Textbook in
Medical Physiology And Pathophysiology
Essentials and clinical problems
Copenhagen Medical Publishers 1999 - 2000

ISBN 87-984078-0-5

Poul-Erik Paulev, M.D., D.Sci

Department of Medical Physiology
University of Copenhagen

http://www.mfi.ku.dk/people/paulev/

Return to content

Preface

This book is for the student of medical physiology. The book contains clinical problems for preclinical students; just following the preclinical test the clinical students face real life problems at the clinical courses. The book is aimed at easing the transfer to the clinic and act as a refresher for medical doctors.

Medical physiology and pathophysiology **integrates basic topics and the patho-physiological mechanisms governing human life.**

I wish to thank the illustrator and designer of the book, Kirsten McCord, who has drawn all illustrations and acted as the prime technical editor.

We acknowledge the kind assistance of: Peter Bie, Jesper Brahm, Per Bredmose, Poul Dyhre –Poulsen, Lars Findsen, Ole Frederiksen, Rolf Gideon, Albert Gjedde, Finn Gustavsson, Niels Henrik Holstein-Ratlou, Jørn Hounsgaard, Jens Ingeman Jensen, Finn Michael Karlsen, Paul P. Leyssac, Margrethe Lynggaard, Joop Madsen, Miroslaw Pokorski, Ole Siggaard-Andersen, Jørgen Skydsgaard, Marek Treiman, Thomas Zeuthen, Jørgen Warberg and others. I apologize to the colleagues and students who have contributed without being mentioned here.

To the student

The student is recommended to read this book in the following way. The first time reader of a Chapter is advised to read first of all *Highlights* The next step is to read only *Study Objectives*, *Definitions* and the running text *Essentials* - Definitions and Highlights are taken from the running text and represent repeated core material. Typically, throughout the book, you will find links to illustrations (Fig.) and equations (Eq.).

After reading the book in this way, it is advisable to deal with the sections *Pathophysiology, Equations,* and *Self-Assessment* just before the final examination. Some of the *Case Histories* in a Chapter present numeric problems.

Poul-Erik Paulev

Copenhagen

Return to content

- Welcome
- Divisions
- Teaching
- People
  - Sorted by division
  - Sorted by last name
  - Poul-Erik Paulev
- Links

Printer friendly version

Poul-Erik Paulev

MD, Dr.Med.Sci

E-mail: ppaulev@mfi.ku.dk

Phone +45 353-27412 or13
Private phone +45 458-13271

Division: **Endocrinology & Metabolism**

- **On-line textbook in physiology**
- **Selected Books**
- **Curriculum Vitae**
- **Selected Bibliography**

Section I: *Cells and Action Potentials - Muscle Cells and Disorders*

Cellular physiology - or *biophysics* - is a discipline covering basic characteristics of most cells. The following 2 chapters do not pretend to include all essential topics; however, it is important to clarify these topics initially, because such concepts (the Na^+-K^+ -pump, radioactive decay etc.) will be referred to throughout the book. An alphabetic list of abbreviations and symbols is present.

Chapter 1.
Cells And Action Potentials

Study Objectives

- To *define* activity, activity coefficient, concentration, diffusion, flux, molality, molarity, normality, osmolality, osmosis, pressure and radioactivity.

- To *describe* the diffusion potential, the equilibrium potential, facilitated diffusion, the Donnan-effect, the resting membrane potential, the action potential, and membrane transport including that of glucose.
- To *calculate* the equilibrium potential, osmotic pressure and other variables from relevant variables given.
- To *draw* the action potential curve.
- To *explain* the colloid osmotic pressure of plasma, hyponatraemia, overbreathing, radioactive decay and the elimination rate constant, the nerve conduction, signal transduction, the function of the action potential, the Na^+-K^+ -pump and the transport proteins.
- To *use* the ideal gas law and the above concepts in problem solving and case histories.

Principles

- *The ideal gas law relates the pressure P, the volume V, and the number of mol of the gas n, to the Kelvin temperature T: ($P \times V = n*R*T$). At standard temperature, pressure, dry (STPD) the volume occupied by 1 mol of any ideal gas is 22.4 litre (l). - By analogy in an ideal solution, one mol of solute will exert an osmotic pressure of one atmosphere, if it is dissolved in 22.4 l of water. Van´t Hoff´s law for ideal solutions is an equivalent to the ideal gas law.*

Definitions

Absolute temperature (T) is measured in Kelvin or K from the *absolute zero point* (-273 ^oC).

Activity is a corrected concentration measure of any species (ie, the free concentration multiplied by the activity coefficient). The activity is often measured with ion selective electrodes. - In diluted solutions - below 10^{-3} molar - there is no correction for uncharged molecules.

Activity coefficient is the fraction of the free ions, which is chemically active. - For sodium the activity coefficient is 0.75 in many biological solutions.

Action Potential (AP) is an *all-or-none electrical signal*, which appears as a positive wave when recording internally. The AP is conducted with the same shape and size along the whole length of a muscle cell or a nerve fibre.

Amphipathic molecules contain both a polar and a non-polar region.

The **membrane potential difference** is defined as the potential inside the cell minus the potential outside the cell – the difference is transiently reversed during an AP.

Becquerel (1 Bq) is the preferred unit for disintegration rate of radioactive decay, namely *one radioactive disintegration per s*. Disintegration rates were previously expressed in Curies (Ci), in honour of Marie Curie, who discovered radium.

Concentration (**C** or brackets around a substance [Na^+]) is the mass or moles (mol) per unit of fluid volume.

Diffusion is a net transport of atoms or molecules caused by their random thermal motion in an attempt to equalise concentration differences (DC).

The **Donnan effect** is the *extra osmotic pressure of protein solutions* caused by impermeable protein molecules resulting in uneven distribution of small, permeant cations and anions (in blood plasma).

The elimination rate constant (k) is the fraction of the total amount of a given substance in the distribution volume of the body eliminated per time unit. Elimination with a constant rate is exponential. The *half-life* for a substance eliminated exponentially is equal to 0.693 k^{-1}. This is just a simple mathematical deduction.

Flux (*J*) is the amount of a substance transported along a pressure gradient through an area unit (A is measured in m^2) of a membrane in moles per second (s). *Convective flux* is the net amount of molecules transported through A per time unit (mol s^{-1} m^{-2}), caused by a pressure gradient and *fluid* (liquid or air) volume transport.

An **ideal semipermeable membrane** is permeable to water only, but impermeable to all solutes. Most real *semipermeable membranes* are permeable to water and to low molecular substances (crystalloids), but not to macromolecular substances (colloids such as proteins).

Molar concentration (molarity) is the number of moles of a substance totally dissolved per litre (l) of solution - often given in mmol per l or mM. One mol of a substance is the amount of that substance containing Avogadros number, $6.022 * 10^{23}$ molecules per mol.

Molality is the number of mol totally dissolved substance per kg of solvent, frequently water. One *equivalent* is the molar mass of all the ions that contain $6.022 * 10^{23}$ single charges or *valences* when fully dissociated.

Motility is the reciprocal resistance of a molecule towards movements.

Normality of a solution is the number of equivalents per l (Eq l^{-1}).

Osmolality is a measure of the osmotic active particles in one kg of water. Plasma-osmolality is given in Osmol per kg of water. Water occupies 93% of plasma in healthy persons.

Osmolarity is the number of osmotically active particles dissolved in a litre of solution.

A **permeable membrane** allows the passage of all dissolved substances and the *solvent* (mainly water).

A **selectively permeable membrane** is permeable to a particular compound (sucrose, Na^+, Ca^{2+}, anions only or to cations only).

- **Pressure** (***P***) is measured as force per area unit - that is in Newton per square m or *Pascal.*

- **Osmosis** is transport of solvent molecules (mainly water) through a semipermeable membrane. **Osmotic pressure** (π) is the hydrostatic pressure, that must be applied to the side of a rigid ideal semipermeable membrane with higher solute concentration in order to stop the water flux, so that the net water flux is zero.
- **Radioactivity.** Some nuclei are *unstable* or *radioactive*, because they release certain particles such as helium nuclei or electrons. Other radioactive substances emmanate gamma-rays with an extremely short wavelength. All radioactive decay processes follow an exponential pattern. If N_o is the initial number of unstable nuclei, the number of nuclei remaining after a time t (N) is given by $N = N_o * e^{-kt}$, where k is a constant characteristic of each nuclide, called the *disintegration constant*. This is the *law of radioactive decay* (Eq. 1-6).
- **Volume** (***V***) in litres (l). *Standard temperature, pressure, dry* (STPD) is an abbreviation for a volume at standard temperature of 273 K, standard pressure of 101.3 kPa or 760 mmHg, and dry air.

Essentials

Three topics are treated here: 1. Transport through membranes, 2. Resting membrane potentials, and 3. Action potentials.

1. Transport through membranes

Membrane transport refers to solute and solvent transfer across both cell membranes, epithelial and capillary membranes.

1a. Membranes

Biological membranes are composed of phospholipids stabilised by hydrophobic interactions into bilayers (Fig. 1-1). The membranes contain approximately 50% lipids and 50% proteins.

Fig. 1-1: Model of a cell membrane built by phospholipids separating receptors, channels, proteins (Pr^-), glycoproteins (receptors, antigens etc) and glycolipids.

Phospholipids are *amphipathic*. One region is polar consisting of charged choline, etanolamine and phosphate head-groups (bullets in Fig. 1-1). The other region is non-polar, consisting of tails of fatty acyl chains (Fig. 1-1). The non-polar regions tend to avoid contact with water by self-association. Any other arrangement with disruption of *hydrogen bonds* (between O and H atoms) of water has a high energy cost. Integral proteins are deeply imbedded in the membrane, and the model shows 3 protein molecules spanning the membrane (ie, transmembrane proteins). Surface proteins are not shown. The proteins carry receptors to which transmitter substance bind. Carbohydrate chains are shown forming glycolipids with antigenic or receptor function, or glycoproteins with other receptor functions.

The *molar concentration* (molarity) is the number of mol totally dissolved substance per litre (l) of solution - often given in mmol per l or mM. Ions in plasma are conventionally measured in mM with flame-photometry by the ability to absorb monochromatic light. Na^+ is mainly dissolved in the water phase of plasma (93%). The $[Na^+]$ in plasma is therefore smaller than the *Na^+-activity* which is recorded in the water phase alone with ion selective electrodes. For conventional reasons ion selective electrodes are calibrated to match the well-known flame photometry values, although the *activity* is the biological important variable. Molality, normality, and flux are described above in *Definitions.*

Mechanical, electrical, thermal, or *gravitational forces* drive *migration* of molecules. These forces move the molecules passively in a direction determined by the vector of the force.

Diffusive flux (J^{dif}) is the movement of molecules by diffusion caused by a concentration gradient (dC/dx) in the direction x. The *diffusion coefficient* (D) is a proportionality constant

that relates flux to the concentration gradient (dC/dx). *Einstein* defined D as (k×T×B), where T is absolute temperature and B is *motility* of molecules. Motility is the reciprocal resistance towards movements (velocity/N or $m*s^{-1}*N^{-1}$). The concept (k*T) is the thermal, molecular energy. A molecule diffuses from higher to lower concentration that is *down its concentration gradient.* Accordingly, dC/dx has a negative slope, when molecules diffuse in the direction x. Then, it is easy to calculate the diffusive flux (per m^2) according to Eq. 1-2. This relationship was first recognized as early as in 1855 by the anatomist and physiologist Fick, and it has since been named after him: *Fick's first law* of diffusion. The flux by simple passive diffusion is directly proportional to the concentration of dissolved molecules (Fig. 1-2).

Fig. 1-2: Two types of passive molecular transport: Simple diffusion and the much larger facilitated diffusion. C is concentration.

Einstein's relation states that for average molecules in biological media, the mean displacement squared, $(dx)^2$, is equal to 2 multiplied by D and by the time (t) elapsed, since the molecules started to diffuse (see Eq. 1-3). For molecules with $D = 10^{-9}\ m^2\ s^{-1}$, the time required to diffuse 1 mm is 0.5 milli-second (ms). To diffuse 10 and 100 mm, the time required increases 100-fold each time: 50 ms and 5000 ms.

Facilitated diffusion takes place through *transport proteins* not linked directly to metabolic energy processes (Fig. 1-2). Facilitated diffusion shows saturation or Michaelis-Menten kinetics, because the number of transport proteins is limited. The *saturation kinetics* is different from the energy limitation in primary active transport. *Amino acids, glucose, galactose and other monosaccharides* cross many cell membranes by facilitated diffusion.

An ideal *semipermeable membrane* is permeable to water only, but impermeable to all solutes. Most real *semipermeable membranes* are permeable to water and to low molecular substances (crystalloids), but not to macromolecular substances (colloids such as proteins).

An ideal *ion-selective membrane* is permeable to anions only or to cations only (Na^+, Ca^{2+}, or to Cl^- and NO_3^-). Ion-selective membranes are used in ion-selective electrodes to measure the activity of selective ions in plasma water (Box 1-1).

Box 1-1: Ion-selective membranes	
Selectivity:	**Either anions or cations**
Specificity:	**Cation membranes distinguish between Na^+, Ca^{2+}, and K^+.**
	Anion membranes distinguish between Cl^- and NO_3^-.

The relation between the potential difference measured with an ion-selective electrode and the activity (ionised or fully dissociated form) for a certain ion (eg, K^+) is given by the Nernst equation (Eq. 1-5).

1b. Osmosis and osmotic pressure

Osmosis is transport of solvent molecules (mainly water) through a semipermeable membrane. The water flows from a compartment of high water concentration (or low solute concentration) to one of low water concentration (or high solute concentration). The greater difference between the solute concentration of the two compartments, the more is water unevenly

distributed between the two compartments. Water diffuses down its chemical potential gradient into the compartment with higher solute concentration, causing the chemical potential gradient to be reduced until solute equilibrium is reached.

Osmotic pressure is the hydrostatic pressure, that must be applied to the side of an ideal semipermeable membrane with higher solute concentration in order to stop the water flux, so that the net water flux is zero.

The *colligative properties of water* are strictly related to the solvent or water concentration alone. Water molecules are bound together by hydrogen bonds in clusters of several hundred molecules, forming a structure looking almost like crystals. Sites between the clusters, where the distance between water molecules are larger than elsewhere are called *bubble nuclei* because these sites seem to initiate formation of gas bubbles in decompression sickness. These sites are also likely locations for substances dissolved in water. With decreasing water concentration, the water vapour tension, and the freezing point is reduced, whereas the boiling point, and the osmotic pressure of the solution is increased as compared with pure water. The size of the osmotic pressure of a solution depends of the number of dissolved particles per volume unit.

The osmotic pressure (π) depends on the absolute temperature (T Kelvin or K) and on the number of dissolved particles per volume unit (N/V equal to the molar fraction).

This relationship was first recognized by *van't Hoff* and applies to ideal solutions only. Real physiological solutions, such as the cytosolic phase and extracellular fluid, differ from the ideal solutions, which are very dilute.

A correction factor called the *osmotic coefficient* (f) corrects for these differences in osmolality. For physiologic electrolytes it is 0.92 - 0.96, and for carbohydrates it is 1.01.

A solution has the *ideal osmolarity one*, when it contains ($6.022* 10^{23}$) osmotically active particles per l. Diluted solutions have an *osmolar concentration* or *osmolarity* (Osmol per l) numerically equal to the *sum total* molarity of all dissolved particles (mol per l). In biological solutions the molarity is different from osmolar concentration. The number of Osmol per l is: (f N/V). - The corrected van't Hoff law is developed in Eq. 1-4.

Osmolality is simply the number of mol per kg of water in the fluid frequently given as mOsmol kg^{-1}. Fully dissociated molecules have twice the osmolality of undissociated molecules. Plasma- osmolality can be approximated by the calculation expressed in Eq. 1-7. Plasma-osmolality is measured by freeze point depression or by boiling point increase. The osmolality of the ICV is approximately 290 mOsmol kg^{-1}, which is simplified to 300 mOsmol kg^{-1} in Fig. 1-4. The osmolality of the extracellular fluid must be the same, since cell membranes are not rigid, so they cannot carry any essential pressure gradient. The total number of mOsmol in the ICV and ECV of a standard person is thus 8400 and 4200, respectively (Fig. 1-4).

The *colloid osmotic pressure* is equal in magnitude to the hydrostatic pressure, which must be applied at the luminal side of the capillary barrier, in order to stop net transport of water caused by uncharged colloids in the blood plasma. *Colloids* are mainly *plasma proteins*.

The *osmotic pressure* is equal in magnitude of a certain hydrostatic pressure. This pressure column must be applied to the solution to restore the free energy or *chemical potential* of its water to that of pure water. The tendency of water to pass a membrane depends on its chemical potential (ie, vapour pressure). The chemical potential of water decreases with solutes present

and with decreasing temperature.

Uniformly distributed substances in diluted solutions behave like gas molecules at atmospheric pressure (atm). The osmotic pressure can be expressed as $P_{osmot} = C \times RT$, which is the equivalent of the ideal gas equation ($P \times V = n$RT). Here C is the concentration of dissolved solutes, and the derivation of the relationship is based on the chemical potential of water. R is the gas constant (= 0.082 l × atm × $Osmol^{-1} \times K^{-1}$). At *standard temperature, pressure, dry* (STPD) the volume occupied by 1 mol of any ideal gas is *22.4 l.* Thus, STPD is an abbreviation for a volume at standard temperature of 273 K, standard pressure of 101.3 kPa or 760 mmHg, and dry air. In an ideal solution, one mol of solute will by analogy be dissolved in 22.4 l of water, and will exert an osmotic pressure of one atmosphere.

In biological solutions at 310 K, such as an ultrafiltrate of plasma (interstitial fluid, ISF), with an osmolality of 0.300 Osmol per kg water, the osmotic pressure must be:

P_{osmot} = 0.3 (Osmol kg^{-1}) × 0.082 (kg × atm × $Osmol^{-1} \times K^{-1}$) × 310 (K) = 7.63 atm (=773 kPa) or the pressure exerted by a column of water *76 m high*.

Only *net gradients* across endothelial and plasma membranes are important, and they depend upon *protein concentration gradients*. This is because all the electrolytes (crystalloids) have diffused to equilibrium across the capillary endothelial membrane, whereas proteins (colloids) cannot.

The average colloid osmotic pressure (π_{coll}) of plasma is approximately *3.6 kPa* (27 mmHg). The dissolved proteins have a molality of 1 mmol per kg water, and a net average of 17 negative charges per molecule (ie, 1 mmol kg^{-1} or 17 mEq kg^{-1}). Milli-Equivalents abbreviates *mEq.* The proteins are directly responsible for *2.4 kPa* (18 mmHg). The remaining *1.2 kPa* (9 mmHg) of the colloid osmotic pressure is due to the unequal distribution of permeable ions, the *Gibbs-Donnan law* or the *Donnan effect* (see below).

1c. The Donnan effect across the capillary membrane

Let us consider a closed system with two compartments separated by a rigid membrane that is permeable to water and to small ions. In the presence of solutions with different NaCl concentrations, water and ions permeate rapidly in both directions across the membrane. Electrical neutrality in each of the two solutions requires that the simultaneous movement of Cl^- match any net movement of Na^+, so the equivalents of anions and cations are the same. The number of times the two ions collide with one side of the membrane is *proportional* to the product of their concentrations: $[Na^+] \times [Cl^-]$. At equilibrium the fluxes of NaCl in each direction are identical, and ultimately the concentrations are the same all over.

Let us now add protein to one compartment ($compartment_p$ modelling streaming plasma), which is separated by a membrane (modelling the capillary endothelial membrane) from the other compartment ($compartment_{ISF}$ modelling interstitial or tissue fluid). The model still contains only Na^+ and Cl^- that can cross the membrane. At equilibrium the product of concentrations of the two ions on either side of the membrane must be equal, and the transmembrane potential corresponds to the equilibrium potential of the small permeant ions (the Nernst equation, Eq 1- 5). Transforming the Nernst equation reveals that the concentration product of any pair of diffusible ions is identical on either side of the membrane at equilibrium:

$$[Na^+]_{ISF} \times [Cl^-]_{ISF} = [Na^+]_p \times [Cl^-]_p.$$

On the plasma side, which contains *impermeant anions* (negative proteinates), the concentration of permeant anion (Cl^- is the model) must always be less than on the interstitial fluid side. The concentration of permeant cation (Na^+ is the model) must always be greater than in the ISF.

The sum of permeant anion and cation concentrations in plasma is always greater than the sum of the same anion and cation concentrations in ISF:

$$[Na^+]_{ISF} = [Cl^-]_{ISF} \; ; \; [Na^+]_p + [Cl^-]_p > [Na^+]_{ISF} + [Cl^-]_{ISF}$$

This is a simple mathematical argument: The sum of unequal sides of a rectangle is greater than the sum of the sides of the square with the same area. It explains why the osmotic pressure in plasma exceeds that of the tissue fluid. This is not due to the plasma proteins alone, but is also due to the *higher* concentration of small, permeant ions in the plasma.

The *Donnan effect* is the *extra osmotic pressure* of protein solutions caused by the uneven distribution of small, permeable cations and anions. The Donnan effect causes a 5% and 10% concentration difference across the capillary barrier between the plasma and ultrafiltrate concentrations of monovalent and divalent ions, respectively.

In the above equations Na^+ and Cl^- are model ions for all the cations and anions. In our body other anions and cations are present, the Na^+ and Cl^- concentrations are not alike, and the *capillary membrane* is far from rigid. Nevertheless, the Donnan equilibrium implies an accumulation of charges on the side with the negatively charged proteins. This potential difference across the *capillary membrane* is termed the *Donnan potential* **–** a potential, which is developed across cell membranes without a sodium-potassium pump.

The Donnan factor at the capillary membrane is 0.95, so a plasma- $[Na^{+}]$ of 150 mmol measured in each kg of plasma-water is in equilibrium with 142 mM in the interstitial fluid between the cells.

Strictly speaking, there is no such thing as a cell with rigid cell walls in the animal kingdom, so the Donnan effect is theoretically unfounded in animal cell membranes – except at the capillary barrier.

1d. The Na^+-K^+-pump.

The Na^+-K^+-pump is a *transmembrane protein* in the cell membrane (Fig. 1-3). The pump contains a channel, which consists of two double subunits: 2 a - and 2 b - subunits. The catalytic subunit (a) is an Na^+-K^+-activated ATPase of 112 000 Dalton, and the b -*subunit* is a glycoprotein of 35 000 D.

Fig. 1-3: The Na^+-K^+-pump consists of 2 α - and 2 β - subunits (Pi = Phosphate).

The pump is a *primary active transporter*, because it uses the cellular energy of the terminal phosphate bond of ATP (Fig. 1-3). The Na^+-K^+-pump transports 3 Na^+ out of the cell and 2 K^+ into the cell for each ATP hydrolysed. This is a net movement of positive ions out of the cell, and therefore called an *electrogenic transport*. The constant influx of Na^+ is shown as well as the leakage of K^+- and Cl^-. In a steady state the net transport of each ion across the resting membrane is zero.

+ +

The Na -K -pump is located in the *basolateral* exit-membrane of the epithelial cell (Fig. 1-3). The primary active ion-transport provides metabolic energy for the secondary water absorption through the luminal membrane. Hereby, the active pump in the exit-membrane drives the luminal transport across the entry membrane. This transport of NaCl and water is surprisingly nearly isotonic. The bulk flow can take place against a large osmotic gradient, and increases in diluted solutions. The *entry membrane* is often highly permeable to water.

The Na^+-K^+-pump builds up a high cellular electrochemical gradient for K^+ and indirectly for Cl^- (Fig. 1-3). The water outflux is coupled to the outward transport of K^+ and Cl^-. The interstitial fluid receives ions and glucose, causing its osmolarity to increase. The osmotic force causes water to enter the interstitial fluid via the cell membranes and the gaps between the cells (tight junctions). This in turn causes the hydrostatic pressure in the interstitial fluid to rise. The hydrostatic force transfers the *bulk of water, ions and molecules* through the thin-walled, tubular capillaries to the blood. When excess of water (solvent) passes through tight junctions, they lose part of their *tightness* and the solvent water drags many Na^+/Cl^--ions out (*solvent drag).*

In a healthy standard person nutrients and oxygen are transported into the cell interior from the extracellular fluid through the cell membrane (Fig. 1-4). The intracellular fluid volume, ICV, is 26-28 l. The *extracellular fluid* volume (ie, ECV of 14 l) consists of the circulating blood plasma (3-3.5 l) and the *interstitial fluid* (ISF) with a volume of 10.5-11 l in the spaces between cells. Total body water (here 42 l) accounts for 60% of body weight. The body is cleared of 24 mol of carbon dioxide by the lungs in 24 hours and of other substances by the kidneys (Fig. 1-4). A yellow tube on the diagram symbolises the gastrointestinal channel, where nutrient molecules are absorbed and waste products are eliminated through the liver bile.

Fig. 1-4: Salt- and water- transport through a cell membrane separating the intracellular and extracellular compartment.

The diagram also shows the Na^+-K^+-pump together with leakage of K^+, Cl- and water (Fig. 1-4). The net transport of each substance is zero in the steady state.

The Na^+-K^+-pump is responsible for maintaining the high intracellular [K^+] and the low intracellular [Na^+]. The energy of the terminal phosphate bond of ATP is used to actively extrude Na^+ and pump K^+ into the cell.

Jens Christian Skou of Denmark initiated the study of the Na^+-K^+-pump already in the 1950ties and received the Nobel Prize for his contribution to basic chemistry and physiology in 1997.

The membrane also contains many K^+- and Cl^--channels, through which the two ions *leak* through the cell membrane.

Intestinal and kidney tubule cells transport substrates, such as glucose and amino acids, in a *substrate-Na^+ cotransport* in the luminal membrane, linked to the Na^+-K^+-pump of the basolateral membrane. This is called a *secondary* active transport of substrate. Such a transport is powered by an actively established gradient (ie, the Na^+-gradient)

The many ion-transporting ATPases form classes or families showing amino acid sequence homology.

1e. Glucose transport proteins (GLUTs) and insulin receptors

A family of homologous carrier proteins that are coded by distinct genes mediates glucose-transport. The transport proteins (GLUTs) show a marked tissue-specificity, which reflects differing transport needs of various tissues. This is facilitated transport (Fig. 1-2).

Five human *glucose-transporters* are cloned and identified (GLUT 1-5). The GLUT 1 resides in placenta, brain, perineural sheaths, red cells, adipose and muscle tissues. GLUT 2 is found in the liver, pancreatic b-cells, proximal renal tubule cells, and the basolateral membranes of small intestinal cells. GLUT 3 is ubiquitously distributed, found predominantly in the brain and in lower concentrations in fat, kidney, liver and muscle tissues. GLUT 4 is confined to tissues with insulin-responsive glucose uptake (muscle, heart and fat stores). GLUT 5 is found in the luminal membrane of small intestinal cells, and also in brain, muscle and adipose tissues. Some of these transporters also allow fructose and galactose to pass.

Fig. 1-5: Insulin, insulin receptors, with D-glucose transport proteins (GLUTs) and their translocation.

In *adipocytes and muscle cells,* glucose transport is profoundly influenced by insulin (Fig. 1-5).

1. As insulin binds to its large T-shaped *insulin receptor*, many intracellular vesicles are stimulated.
2. They contain a high number of membrane penetrating GLUTs, which translocate from the intracellular pool towards the cell membrane.
3. When these vesicles - rich in glucose transporters - fuse with the cell membrane, the number of glucose transporters increases substantially, thereby increasing D-glucose uptake up to ten times.
4. As the insulin-receptor complex dissociates, the GLUTs translocate again to the intracellular stores in the vesicles (Fig. 1-5).
5. The glucose transport ceases.

The *insulin receptor* is a glycoprotein found in the cell membranes. The T-shaped receptor protruding from the cell membrane contains 1370 amino acids forming two a- and two b-sub-units. The two a-subunits are entirely extracellular, whereas the two b-subunits span the membrane. Insulin binding on a-subunits stimulates a *protein kinase* on the intracellular part of the receptor to phosphorylate tyrosine residues on the b-subunit and on endogenous proteins. The exact molecular mechanism linking the receptor kinase activity to changes in cellular enzyme activity and transport processes remain uncertain; but it is shown that the *kinase activity* is essential for signal transduction.

2. Resting membrane potentials

A *membrane potential difference* is conventionally defined as the intracellular (j^i) minus the extracellular (j^o) electrical potential. The ion concentrations (activities) inside the cell and outside the cell are called C^i and C^o, respectively.

When a microelectrode penetrates a membrane, it records a negative potential with respect to an external reference electrode caused by different permeability of anions and cations. This is the *resting membrane potential* (RMP values in Box 1-2). The resting membrane potential is an essential mechanism in storing and processing information in neurons and other cells.

Concentration gradients across cell membranes are present for several ions, whereby they diffuse from one location to another. The ion with the highest permeability and concentration

gradient, such as the potassium ion, establishes a membrane potential. This potential enhances or inhibits the flux of other ions and the ultimate situation is an electroneutral flux.

The chloride ion diffuses extremely rapidly, but otherwise positive ions (cations) diffuse more rapidly than negative ions (anions) through a membrane. However, as an example the permeability for Na^+ is low (0.2 nm $*s^{-1}$) compared to that of K^+ (5-40 nm/s) in neurons.

The *equilibrium potential* for a certain diffusible ion across a membrane that has a concentration gradient over the membrane, is precisely that *membrane potential difference*, which opposes the flux due to the concentration gradient so that the net transport of the ion concerned is zero. The equilibrium potential is simply calculated by balancing the diffusion potential of the ion with the opposing electrical force. The electrical force working on the ion is proportional to the electromotive force of the field. As a consequence, the total driving force on the ion and its *diffusion flux* is *zero*. Nernst introduced this equilibrium potential shortly before year 1900. The *Nernst equation* for the equilibrium potential of Na^+ across a selective permeable membrane at 310 K is found by insertion of the ion activities (concentrations) inside and outside the cell (Eq. 1-5).

Box 1-2: Resting membrane potentials (RMP) and equilibrium potentials (V_{Eq}) in different cells.

	RMP (mV)	**V_{Eq} (mV)**
Resting skeletal muscle	- 80	- 80 for Cl^-
and myocardial cells	- 90	- 94 for K^+
		+ 60 for Na^+
		+ 130 for Ca^{2+}
Smooth muscle cells	- 40 to –60 (oscillations)	Variable
Neurons	- 70	As above

In skeletal muscle cells the resting membrane potential is -80 mV, and the *equilibrium potential* of Na^+ is +60 mV. Hence, the electrical driving force is: (-80 - (+60)) = -140 mVolts. Accordingly, there is a net passive influx of Na^+ into these cells down an electrochemical gradient (Box 1-2). The net influx is small because the resting Na^+ penetration is almost exactly balanced by active extrusion.

The resting membrane potential (RMP) is calculated from the *Millman equation* (Eq. 1-8). The RMP is mainly a diffusion potential (see above).

Neurons typically have four structures: The cell body, dendrites, axon and axon terminals (Fig. 1-6). Dendrites are elaborate branching processes that arise from the cell body, and they are pathways for incoming signals from other neurons to the cell body. Integration of incoming signals occurs mainly in the *axon hillock*. This is the part of the cell body, which gives rise to an elongated tube called an axon, a fibre that can be up to 1.2 m long.

Fig. 1-6: The neuron with cell body, dendrites, axon and axon terminals.

Near its termination each axon divides into fine branches, each of which ends in an *axon terminal* (ie, synaptic button or Bolton terminal). The axon terminals contain mitochondria and

synaptic vesicles filled with neurotransmitter. These presynaptic structures are the sites where electrical signals are converted into chemical messages for transmission to nearby neurons. Unipolar neurons only have a single major process extending from the cell body. Bipolar and multipolar neurons have two or more major processes arising from the cell body. Most neurons have only one axon, a few more than one and some neurons function without an axon. Their location, structure and functional properties (Box 3-1) classify neurons. Communication from an axon to a dendrite is called *axodendritic*, from a dendrite to another is termed *dendro-dendritic*, from a dendrite to an axon is called *dendro-axonal*, from a dendrite to the soma is called *dendro-somatic*, and between two axons is referred to as *axo-axonal*.

Neuronal membranes are composed of lipid bilayers stabilised by hydrophobic interactions, and thus function as barriers to free diffusion for water-soluble molecules. The ability of the neuronal membrane to control the movement and concentration of charged particles generates ion gradients with a charge difference across the membrane. The potential difference across the resting membrane is called the *resting membrane potential* (RMP).

Ion channels and gates are classified according to the gating stimulus to which they respond. Voltage-gated ion channels are located along the axon of a neuron and responsible for the action potential. These ion channels are sensitive to local anaesthetics. Voltage-gated Na^+-K^+- and Ca^{2+}-channels contain membrane spanning helices – often with amino acid sequence homology. Action potentials in cardiac muscle cells have a plateau phase, where Ca^{2+} enters the cytosol via slow Ca^{2+}-channels. This Ca^{2+}-entry plays an important role in excitation-contraction coupling.

Ligand-gated ion channels are responsive to particular neurotransmitters. Ligand-gated ion channels open in response to substances such as acetylcholine. These channels are permeable to small cations (often unselectively: Na^+- K^+- NH_4^+- and Ca^{2+}). These channels are involved in generation of the postsynaptic potential and the endplate potential.

NaCl is found in high concentration outside the neurons, whereas [K^+] is high inside the cell. These ion gradients maintain a constant leakage of NaCl into the cell, and a leakage of K^+ out. The gradients are maintained by the Na^+-K^+-pump, which is thus controlling the resting membrane potential. *Cl^--ions* distributes passively across most neuronal membranes and contributes little to the resting membrane potential, but they are important for the modulation of incoming signals. At rest, many K^+ -channels are open and K^+ moves down its concentration gradient out of the cell, whereby the inside becomes negatively charged (until it is difficult for K^+ to leave the cell, and the K^+-outflux slows down). The RMP approaches the equilibrium potential for K^+.

3. The action potential

Neurons can carry electrical signals along their whole length without any loss of signal strength. This electrical signal is an all-or-none phenomenon, termed the *action potential***.** The incoming signals to dendrites and cell bodies consist of small, graded changes (ie, small *synaptic potentials*) in the resting membrane potential caused by the actions of neurotransmitters and modulators. *Synaptic potentials* are spatially and temporally summated in the axon hillock of the cell body. The synaptic potential is graded according to the stimulus and shows decrement conduction in that its size decreases with increasing distance (wave length several mm). The local synaptic potential cannot in itself initiate an action potential. When the strength of the summated synaptic potentials is sufficient to reduce the resting membrane potential at the axon hillock below the *threshold* it opens Na^+ -channels. These Na^+ -channels are voltage-gated, because the change in voltage opens or closes a gate over each

pore. The Na^+ -channels is usually closed at conditions with a normal resting membrane potential. When the Na^+ -channels open and allow Na^+ to flow into the cell down its concentration gradient, the influx itself depolarizes the neuron further, whereby more voltage-gated Na^+ -channels open. A propagating action potential (approaches the equilibrium potential for Na^+) in the axon is generated with a *positive* voltage overshoot simultaneously with the peak membrane conductance to Na^+ ($g_{Na}{}^+$ in Fig. 1-7). This is followed by the *repolarisation phase* (conductance for Na^+ goes down and up for K^+), when the potential returns toward the resting membrane potential. The potential may overshoot the resting value, causing a transient hyperpolarization known as the *hyperpolarising afterpotential* (Fig. 1-7) close to the equilibrium potential for K^+.

Sustained depolarization inactivates the voltage-gated Na^+ -channels, and shuts off the Na^+ -influx. Opening of voltage-gated K^+ -channels allows an increased outflux of K^+ to counterbalance the influx of Na^+. The membrane conductance to K^+ increases *slowly*, and reaches a peak in the repolarization phase ($g_K{}^+$ in Fig. 1-7). This K^+ -outflux causes the neuronal membrane potential to return to its normal resting value, when the Na^+ -channel is inactivated. The signal conduction is unidirectional, because newly opened Na^+ -channels become refractory for a time, when they are inactivated. As these areas are blocked for further depolarization for a time, the depolarization can proceed only in the forward direction towards resting Na^+ -channels.

Fig. 1-7: Transmembrane potentials and Na^+-K^+- conductance (flux) in a neuron.

The action potential is an *all-or-none electrical signal*, which appears as a positive wave when recording internally. The action potential is conducted with the same shape and size along the whole length of a muscle cell or a nerve fibre.

The refractory periods
During the early part of the action potential the cell membrane is completely refractory. A new stimulus, regardless of its size, cannot evoke an action potential. Almost all Na^+-channels are inactivated, and will not reopen until the cell membrane is repolarized. This is the *absolute* refractory period covering most of the peak and lasting until well into the repolarizing phase (ARP in Fig. 1-7).

During the hyperpolarizing afterpotential, a *suprathreshold stimulus* is able to trigger a new AP, albeit of smaller amplitude than the first action potential. This period is called the *relative* refractory period (RRP in Fig. 1-7). The cell membrane is relatively refractory, because some Na^+-channels are voltage-inactivated and at the same time K^+-conductance is increased.

Nerve conduction
The lipophilic core of the cell membrane is an electrical insulator, but the salt solutions of the cytoplasm and the extracellular fluid act as conductors of electrical current. Opening of many voltage-gated Na^+-channels, whereby the Na^+-conductance is increased about 104-fold, so the membrane is instantly depolarised, causes the action potential. The action potential essentially spread by alterations of the voltage-gated Na^+-channels.

Depolarization spreads along the membrane of excitable cells by local currents flowing to the adjacent segments of the membrane. This is shown in Fig. 1-8A. The phenomenon is called the *local response* or *electrotonic conduction*. The depolarization decreases mono-exponentially

+

from the excitation site. Na -channels will be recruited in all areas of the membrane, where the threshold potential is exceeded. The Na^+-channels behind the peak of the action potential are refractory. This explains why an action potential travels in both directions, when it is evoked in the middle of a nerve.

Fig. 1-8: **Spread of the action potential along an unmyelinated (A) and a myelinated (B) axon. The refractory channels prevent the action potential from proceeding in more than one direction. The action potential (wavelength in cm) essentially jumps from node to node or over several nodes facilitating high-speed conduction.**

The *myelin sheath* consists of 20-300 layers of insulator substance produced by Schwann cells wrapping round the axon. The *nodes of Ranvier* are the lateral spaces (1 mm wide) between adjacent Schwann cells, which stretch 1-2 mm.

The effects of this arrangement are as follows:

Very little current is lost through the electrical insulation of the *myelin sheath*. Thus, the *electrotonic conduction* is rapid with only a small decrement in amplitude. The electrotonic conduction is virtually instantaneous. Because of the insulation the depolarization can spreadmuch faster.

Saltatory or *leaping* conduction occurs, because the action potential is generated only at the nodes (Fig. 1-8). The cell membrane below the myelin sheaths has hardly any Na^+- channels and is therefore inexcitable. Saltatory conduction is up to 50 times faster than the conduction through the fastest unmyelinated axons. The action potential can also jump over a number of nodes to that farthest away, because the action potential wavelength is several m.

The Na^+-channels there are activated by the *electrotonic conduction*.

Since the ionic currents are restricted to the nodes of Ranvier in the myelinated axons, this minimises disturbances in the Na^+- and K^+-gradients, that are restored by an active process in which the Na^+-K^+-pump, driven by ATP, pumps Na^+ out and K^+ into the cell. The main energy cost is to restore the Na^+- K^+- balance.

Myelination of the nerve fibre thus reduces energy cost of maintaining the resting membrane potential following an action potential.

Typical values for normal ion concentrations (intracellularly and extracellularly) are given in Box 1-3.

Box 1-3: Normal ion concentrations in muscle cells and in plasma.

Intracellular osmolality mmol * (kg of water)$^{-1}$		**Plasma concentration ranges mmol * (l of plasma)$^{-1}$**
Na^+	10	135 - 146
K^+	155	3.5 - 5.0
Ca^{2+}	0.0001	1 – 1.2
Mg^{2+}	12	0.7 – 1.1
Cl^-	5	95 – 106

HCO_3^-	10	22 - 27

The total calcium concentration in plasma is 2.2 – 2.7 mM, but only 45% is ionized (Chapter 30).

Pathophysiology

This paragraph deals with simple conditions, where the student needs no prior knowledge, so only hyponatraemia, pseudo-hyponatraemia and overbreathing are described.

1. Hyponatraemia is defined as a plasma-[Na^{+}] below 135 mM. This is a common condition caused by a high water intake (*water intoxication*), reduced water excretion in kidney disease, salt loss or other causes described in Chapter 24.

Hyponatraemia must be distinguished from the rare condition *pseudo-hyponatraemia.* Spuriously low [Na^{+}] are measured in plasma (Na^+ is predominantly confined to the water phase), simply because its concentration is expressed per l of plasma. Normally, 93% of plasma is water, and the non-water fraction is 7% (mainly proteins). In cases with too much lipid, protein, glucose, urea, or alcohol in the blood plasma (ie, hyperlipidaemia, hyperproteinaemia, hyperglucosaemia, uraemia, alcoholaemia etc), the normal [Na^{+}] is reduced by dilution with the increased non-water fraction. Thus, the *calculated* plasma-osmolality (Eq. 1-7) is less than the *freeze-point-measured* osmolality. This discrepancy is called the *osmolality-gap.*

There is no need for treatment with salt-solutions in such spurious conditions. Pseudo-hyponatraemia also occurs, when a blood sample is taken from an arm vein, where a glucose solution is infused. The plasma-[Na^{+}] obtained from such a blood sample is low.

2. Overbreathing is also called *hyperventilation*. Overbreathing is frequently caused by *panic attacks* (ie, hyperventilation tetany). The high ventilation washes out too much carbon dioxide (CO_2)/carbonic acid, whereby the CO_2 tension of the arterial blood decreases simultaneously with an increase in its pH. This is an alkalosis (arterial pH above 7.45). Alkalosis dissociates proteins by mass action and form Ca^{2+}-proteinate (Fig. 17-9). The falling extracellular concentration of free Ca^{2+}-ions reduces the threshold and opens Na^+-channels in neurons, muscle cells and myocardium. The resulting reduction in membrane potential increases the excitability of the tissues, which causes continuous muscular contractions (ie, *tetanic cramps*).

Equations

- **Uniformly distributed substances** in diluted solutions behave like gas molecules at atmospheric pressure (atm). The osmotic pressure can be expressed as:

$$\text{Eq. 1-1: } P_{osmot} = C * R*T,$$

which is the equivalent of the *ideal gas equation* ($P = n/V* R*T$). C is the concentration of dissolved solute. R is the gas constant (= 0.082 l atm mol^{-1} K^{-1}).

- **Fick's first law of diffusion** deals with the diffusive flux per m^2 :

$$\text{Eq. 1-2: } J^{dif} = -D* dC/dx.$$

The dimension of D is found by dimension analysis of the equation for J^{dif}:

$$(J^{dif} \text{ moles } s^{-1}\ m^{-2}) = D\times (dC/dx \text{ moles } m^{-3}\ m^{-1}).$$

Accordingly, D has the dimension: $m^2\ s^{-1}$. D is small, when the molecules are large and when the surrounding medium is viscous. The *permeability coefficient* (ie, permeability) for a membrane is the flux (mol $m^{-2}\ s^{-1}$) divided by the concentration (mol m^{-3}) for a given substance and has the unit m s^{-1}.

Einstein's relation states that the displacement squared, $(dx)^2$, is equal to 2 multiplied by D and by the time (t) elapsed, since the molecule started to diffuse:

$$\text{Eq. 1-3: } (dx)^2 = 2 * D * t.$$

The corrected van't Hoff law:

$$\text{Eq. 1-4: } \pi = T\times R\times f \times N/V \text{ or } \pi = T\times R\times DC$$

where R is the ideal gas constant ($0.082\ l\times atm\times mol^{-1}\times K^{-1}$ or $8.31\ J\ (K\ mole)^{-1}$), and DC is the *concentration gradient*. This is the law for ideal or extremely dilute solutions.

A **membrane potential difference** is conventionally defined as the *intracellular* (j^{i}) minus the extracellular (j^{o}) electrical potential. The *ion activities* (concentrations multiplied by the activity coefficient) inside the cell and outside the cell are here called C^i and C^o, respectively. The *Nernst equation* for the equilibrium potential of Na^+ across an ion-selective membrane at 310 K reads:

$$\text{Eq. 1-5: } j^{i} - j^{o} = (\mathbf{R}\ T/z\ F)\ \ln(C^{o}_{Na^+}/C^{i}_{Na^+}) \text{ Volts (V)}$$

$$V_{EqNa^+} = 61.5 \log (C^{o}_{Na^+}/C^{i}_{Na^+}) \text{ mV}.$$

In the equation above **R** is the *ideal gas constant* ($8.31\ J\ (K\times mole)^{-1}$), T is the absolute temperature, z is valence of the ion with sign (z= +1), and F is the Faraday constant (96 500 coulombs per equivalent). The activity coefficient for sodium is 0.75 and used to convert concentration to activity.

The law of radioactive decay: If N_o is the initial number of unstable nuclei, the number of nuclei remaining after a time t (N) is given by:

$$\text{Eq. 1-6: } N = N_o * e^{-k\ t}.$$

where k is a constant characteristic of each nucleide, called the *disintegration constant*.

Plasma-osmolality is calculated as follows:

$$\text{Eq. 1-7: Plasma-osmolality} = (2* [\ Na^{+}]\ ) + [\ \text{glucose}] + [\ \text{urea}]$$

Normally, the plasma-$[\ Na^{+}]$ is 140 mmol in 1 litre of water, and both plasma-[glucose] and plasma-[urea] are around 5 mmol per l of water. The *plasma-osmolality* is given in mOsmol per kg of water. One l of water is approx. equal to 1 kg of water.

The **Millman equation.** A convenient version to calculate the resting membrane potential (RMP) at body temperature is:

$$\text{Eq. 1-8: RMP} = (g_{K^+} \times Eq_{K^+} + g_{Na^+} \times Eq_{Na^+} + g_{Cl^-} \times Eq_{Cl^-}) / (g_{K^+} + g_{Na^+} + g_{Cl^-})$$

This equation shows that the RMP is determined by the conductance (g) of the membrane to K^+, Na^+ and Cl^-, and by their equilibrium potentials.

Self-assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

A. Positive ions (cations) diffuse more rapidly through a membrane than negative ions (anions).
B. The Na^+-K^+ -pump located in the cell membrane, is responsible for maintaining the high intracellular [K^+] and the low intracellular [Na^+].
C. The permeability for Na^+ in cell membranes is high compared to that of K^+ and Cl^-.
D. In skeletal muscle cells the resting membrane potential is -80 mV, and the equilibrium potential of Na^+ is +60 mV.

A membrane potential difference is conventionally defined as the extracellular minus the intracellular electrical potential.

II. Each of the following five statements have False/True options:

A. The local, subthreshold response is graded according to the stimulus.
B. Accommodation is a progressive decrease in firing frequency despite maintained depolarization.
C. Voltage inactivation of Na^+-channels is involved in the accommodation and in the refractory periods.
D. During the early part of the action potential the cell membrane is relatively refractory.

Voltage-gated Na^+-, K^+- and Ca^{2+} -channels are comprised of subunits with membrane spanning domains.

III. Each of the following five statements have False/True options:

A. Saltatory conduction is up to 50 times faster than the conduction through the fastest unmyelinated axons.
B. Saltatory conduction occurs because the cell membrane beneath the myelin sheaths has a high density of Na^+-channels.
C. At standard temperature, pressure, dry (STPD) the volume occupied by 1 mol of any ideal gas is 22.4 l.
D. Elimination with decreasing rate is exponential.

Facilitated diffusion shows saturation kinetics, and takes place through transport proteins not linked to the metabolic energy processes.

Each of the following five statements have False/True options:

A. Osmolarity is the number of osmotically active particles in each l of solution.

B. Facilitated diffusion does not show saturation kinetics but moves solutes up-hill.

A family of homologous carrier proteins that are coded by distinct genes mediates facilitated diffusion.

C. Glucose-transport. The transport proteins (GLUTs) show a marked tissue-specificity which reflects differing transport needs of various tissues.
D. The relation between the potential difference measured with an ion-selective electrode and the activity is given by the Nernst equation.
E. The Donnan effect is the extra colloid osmotic pressure of protein solutions caused by uneven distribution of small, diffusible cations and anions.

Try to solve the problems before looking up the answers.

Highlights

The ideal gas law relates the pressure P, the volume V, and the number of mol of the gas n, and the Kelvin temperature T: ($P \times V = nRT$).

- *One mol is the amount of a given substance containing Avogadro's number. The volume occupied by any ideal gas is 22.4 l at STPD.*
- *Fick's first law of diffusion relates the diffusive flux per m^2 to the concentration gradient.*
- *The Donnan effect of plasma is the extra osmotic pressure of protein solutions caused by uneven distribution of small, permeable cations and anions.*
- *The membrane potential difference is conventionally defined as the intracellular minus the extracellular electrical potential.*
- *The action potential is an all-or-none electrical signal, which appears as a positive wave when recording internally during activity in a neuron or a muscle cell.*

- *Saltatory or leaping conduction occurs, because the action potential is generated only at the nodes. The cell membrane below the myelin sheaths has hardly any Na^+- channels and is therefore inexcitable. Saltatory conduction is up to 50 times faster than the conduction through the fastest unmyelinated axons.*

- *In cases with too much lipid, protein, glucose, urea, or alcohol in the blood plasma (ie hyperlipidaemia, hyperproteinaemia, hyperglucosaemia, uraemia, and alcoholaemia) etc, the normal plasma-$[Na^+]$ is reduced by dilution with the increased non-water fraction.*
- *Overbreathing is caused by panic attacks. The high ventilation washes out the carbon dioxide (CO_2)/carbonic acid, whereby the CO_2 tension of the arterial blood decreases simultaneously with an increase in its pH. This is an alkalosis. Alkalosis dissociates proteins and form Ca^{2+}-proteinate. The falling extracellular concentration of free Ca^{2+}-ions opens Na^+-channels in neurons, muscle cells and myocardium. The resulting reduction in membrane potential increases the excitability of the tissues, which causes continuous muscular contractions (ie, tetanic cramps).*

Further Reading

Alberts B, Brady D, Lewis J, Raff M, Roberts K, and JD Watson. Molecular biology of the cell. Garland Publ Inc, NY & London, 1994.

Apps DK, BB Cohen, and CM Steel. "Biochemistry." Bailliere Tindall, London, 1994

Chapter 2
Muscle and Cells Disorders

Study Objectives

- To *define* the concepts gap junction, motor unit, synaptic & neuromuscular transfer, isometric and isotonic contraction, plasticity, post-synaptic potentials, and recruitment.
- To *describe* the electromyogram, three types of motor units and three types of muscle tissue (striated, smooth, and myocardial tissue), modulation of neurotransmission with facilitation, potentiation, neurotransmitters and receptors.
- To *explain* the function of the neuromuscular junction, the synapses, the neurotransmitters, and the control of the muscular force by frequency variation and recruitment. To explain disorders of the neuromuscular junction, the skeletal muscles, the smooth muscles and the myocardium.
- To *use* the above concepts in problem solving.

Principles

- *Waller's law of neuronal degeneration: When a motor axon has been severed, the rough endoplasmic reticulum accumulates proteins required for repair of the axon. The axon and the myelin sheath distal to the injury die and are phagocytized. The neuroglial Schwann cells remain alive, proliferate and form long rows along the pathway previously occupied by the dead axon. The severed axon regenerate along this pathway.*
- *Dale's law: A single neuron liberates only one neurotransmitter at all its synapses. Although the law is frequently valid, there are several exceptions, where two or more co-transmitters are released at all the synapses of a single neuron.*

Definitions

- **Excitatory postsynaptic potential** (EPSP) refers to a transientdepolarization of a neuron membrane. The combined effect of EPSPs from hundreds of presynaptic terminals can summate to evoke an action potential.
- **Gap junctions** are transmembrane protein pores between cells. The pores represent a low electrical resistance. Most *electrical synapses* contain many gap junctions allowing free passage of ions and small molecules in both directions when open.
- **Inhibitory postsynaptic potential** (IPSP) is a transient hyperpolarization of a neuron membrane. The negativity of the resting membrane potential increases (normally -70 mV) and summation of IPSPs may result in an effect.
- **Isometric contraction** is a muscular contraction at constant length.
- **Isotonic contraction** is a muscle contraction at constant tension (load).
- A **miniature endplate potential** is probably caused by the spontaneous release of a single acetylcholine vesicle into the synaptic cleft. This is called *quantal release*.
- **Motor unit** refers to one motor neuron and the group of muscle fibres it innervates. All muscle fibres belonging to a certain motor unit are of the same type.
- **Neurotransmission** refers to transfer of signals from one neuron to another mediated electrically or chemically.
- The **neuromuscular endplate** is the *contact zone* between the axons of motor neurons and striated muscle fibres. The acetylcholine containing vesicles of the axon terminals dock on the *release sites* of the presynaptic membrane with high affinity. The muscle cell membrane at the endplate is folded in *junctional crypts*. Nicotinic acetylcholine

receptors are concentrated at the openings of these crypts.

- **Plasticity** refers to mechanical plasticity of smooth muscle tissue or to an amplification produced by synapses, which transmit better when frequently used.
- **Recruitment** refers to the increase in force and contraction velocity of a muscle by activation of more and more motor units.
- **Sarcomere** is a contractile unit of a muscle fibril containing the halves of two I-bands with the A-band in between (ie, the part of the fibril between two neighbour Z-lines).
- **Synaptic transfer** refers to the transmission of signals from one neuron to another, and the site of contact between the two neurons is called the *synapse*.

Essentials

This paragraph deals with 1. Neuromuscular junctions, 2. Synapses, 3. Skeletal muscles, 4. Smooth muscles and 5. Cardiac muscle tissue.

1. Neuromuscular junctions

The neuromuscular endplate is the *contact zone* between the axons of motor neurons and striated muscle fibres. Axon terminals have vesicles containing *acetylcholine* (Fig. 2 -1). The vesicles dock on the *active zones* or *release sites* of the presynaptic membrane with high affinity. The muscle cell membrane at the endplate is folded in junctional folds or *crypts* (Fig. 2-1). *Nicotinic acetylcholine receptors* (Chapter 6) are concentrated at the openings of these junctional crypts. The release sites are located directly over the acetylcholine receptors (Fig. 2-1). The postsynaptic membrane has *acetylcholinesterase* all over its surface.

The *nicotinic acetylcholine receptor* is related to a ligand (acetylcholine)-gated ion channel found not only in the neuromuscular junction, but also at all autonomic ganglia (Chapter 6) and in the central nervous system (CNS). The receptor is fixed into the postjunctional membrane, whereas *acetylcholinesterase* is loosely attached to its surface. The receptor has five integral protein subunits (2a , 1b , 1g , 1d), surrounding a *central ion channel pore* that is opened by the binding of 2 acetylcholine molecules to the 2 a - proteins (Fig. 2-1). Opening of the ion channel increases the conductance for small cations (Na^+ and K^+) across the postjunctional membrane, depolarising the membrane potential of the cell. These ion channels are not voltage-gated (not dependent on changes in membrane potential), like most cation channels in neurons, cardiac and skeletal muscle cell membranes.

Fig. 2-1: The neuromuscular junction and intracellular events. Acetylcholine = ACh. The ACh-receptor to the right is magnified.

The acetylcholine-vesicles are probably already stored close to the *release zones*, awaiting the release signal (Fig. 2-1). When the action potential (AP) reaches the axon terminals, the axon membrane is depolarised, and voltage-gated Ca^{2+}-channels are transiently activated. This causes Ca^{2+} to flow down its concentration gradient from the outside into the axon terminal. The influx of Ca^{2+} at the release zones causes the vesicles to fuse with the axon membrane, and empty acetylcholine into the 50 nm wide cleft by *exocytosis* (Fig. 2-1).

After crossing the synaptic cleft by diffusion, acetylcholine binds to its *receptor protein* on the muscle cell membrane. This binding complex opens the ion channel and increases the conductance for small cations across the muscle cell membrane. The influxes of Na^+ depolarise the *endplate* temporarily, the transient depolarization is termed the *endplate potential* (EPP). The EPP dies away when acetylcholine is hydrolysed to acetate and choline by the enzyme, *acetylcholinesterase*. The EPP has a large safety margin, as a *single action potential* in the motor axon will produce an EPP that always reaches the threshold potential in the muscle fibre.

Rapid contraction of the muscle fibre is achieved by propagation of the muscle action potential along the whole length of the muscle fibre membrane and into the small, transverse tubules, which penetrate all the way through the muscle fibre (T-tubules in Fig. 2-1).

The acetylcholine binding at the motor endplate increases endplate conductance and generates an action potential (AP) in all directions from the end plate (Fig. 2-1). The electrical excitation of the sarcolemma and the transverse tubules (T-tubules) during the AP triggers – by an unknown mechanism - the *sarcoplasmic reticulum* to release a pulse of Ca^{2+} (Fig. 2-1). The Ca^{2+}-channels opens transiently in the vicinity of each sarcomere (Fig. 2-1). The sarcoplasmic [Ca^{2+}] increases from 10^{-7} to 10^{-6} M (which is the threshold). This Ca^{2+} diffuses to the adjacent myofilaments, where they bind strongly to troponin C on the active filament, and end the troponin-tropomyosin blockade. This enables cyclic crossbridges to work as long as the high [Ca^{2+}] is maintained, whereby contraction occurs. A continually active Ca^{2+}-pump returns Ca^{2+} to the sarcoplasmic reticulum, and another Ca^{2+}-pump in the cell membrane also reduces sarcoplasmic [Ca^{2+}] . Then the thin filament is *off duty*, because Ca^{2+} is withdrawn from its troponin C, the troponin-tropomyosin-blockade is re-established and *relaxation* ensues. The terminal cisternae of the sarcoplasmic reticulum contain granules of calsequestrin, a protein that can bind Ca^{2+} and reduce the concentration gradient (Fig. 2-1).

Neurons with *motor* function have the ability to synthesise acetylcholine, because they contain *choline-acetyltransferase*. This enzyme catalyses the production of acetylcholine from acetyl-CoA and choline. Almost all cells produce *acetyl-CoA* and *choline*. Choline is also actively taken up from the extracellular fluid via a mechanism indirectly powered by the Na^{+}-K^{+}-pump. There is a 50% reuptake of choline from the synaptic cleft; hence some choline must be synthesized in the motor nerve.

The postjunctional membrane depolarizes spontaneously - resulting in so-called *miniature endplate potentials* (MEP-potentials). A miniature endplate potential is probably caused by the spontaneous release of a single vesicle into the cleft. This is called *quantal release*.

An endplate potential is prolonged when *cholinesterase-inhibitors* are present in the synaptic cleft. This is because these substances (eserine, edrophonium, malathion, parathion etc.) inhibits the enzyme and thereby protects acetylcholine from being hydrolysed by the enzyme. The life dangerous parathion poisoning is described in chapter 6. Under normal conditions, the endplate potential is terminated by the rapid hydrolysis of acetylcholine by acetyl-cholinesterase.

Acetylcholine is a transmitter in the CNS, in all motor neurons, in all preganglionic neurons of the autonomic nervous system and postganglionic parasympathetic fibres, and in a few postganglionic sympathetic fibres. The cholinergic receptor subtypes are shown in Table 6-2.

2. Synapses

Chemical synapses prevail in humans, but we also have electrical synapses in gap junctions.

A chemical synapse consists of a neuronal presynaptic terminal, a synaptic cleft and a subsynaptic (or postsynaptic) membrane with associated receptor proteins (Fig. 2-2). The chemical synapse is highly developed in the CNS. It conducts the signal one way only, and has a characteristic *synaptic delay*.

The *presynaptic axon terminal* typically broadens to form a *bouton terminaux* (presynaptic terminal).

Fig. 2-2: A synapse between a preganglionic and a postganglionic neuron.

1. The action potential, originating in the CNS, *depolarises* the axon membrane by selective influx of Na^+, which has a large electrochemical gradient. Repolarization follows rapidly by selective K^+-efflux (Fig. 2-2).
2. When the action potential reaches the presynaptic membrane, Ca^{2+} enters the terminal through voltage-gated *Ca^{2+}-channels*.
3. Vesicles containing transmitter, fuse with the presynaptic membrane and release their contents of acetylcholine into the synaptic cleft (Ca^{2+}-induced exocytosis).
4. Transmitter molecules (acetylcholine, ACh) diffuse across the synaptic cleft and bind to specific receptors, which are located into the postsynaptic membrane (Fig. 2-2). This ligand binding elicits a transient opening of *pores***,** which are specifically permeable to small cations. The synaptic cleft of a chemical synapse is about 30 nm.
5. The ACh-receptor opens and allows influx of Na^+, whereby the membrane depolarizes and an action potential is generated which propagates along the length of the postganglionic axon (Fig. 2-2). This is an appropriate response of the postsynaptic cell to the received signal.
6. The effect is rapidly terminated by the highly specific enzyme acetylcholinesterase, which hydrolyses acetylcholine into two inactive products (acetic acid and choline).

Influx of Na^+ or efflux of K^+ through the pores of such receptors changes the postsynaptic membrane potential. If the presynaptic action potential (AP) results in a postsynaptic depolarization, the transient is called an *Excitatory Post-Synaptic Potential* (EPSP). If the AP results in a postsynaptic hyperpolarization, the transient is called an *Inhibitory Post-Synaptic Potential* (IPSP). Excitatory synapses often use *glutamate* as the transmitter. The pores are penetrated mainly by Na^+, which enters the cell, depolarizes the membrane, and produces an EPSP.

The axon hillock on the *cell body* has a high density of voltage-gated Na^+- and K^+-channels. The axon hillock probably integrates the many synaptic potentials, and from here the action potential is generated. The *dendrites* have voltage-gated channels for K^+ and for Ca^{2+}. Recent evidence suggests that dendrites also contain voltage-gated Na^+-channels, which are involved in *electrogenesis* (ie, movement of charge across the membrane).

Each neuron in the CNS is in contact with up to 10^5 presynaptic axon terminals. Synaptic inputs are integrated at the axon hillock by either *spatial* or *temporal* summation.

Spatial summation occurs when inputs from several axons arrive simultaneously at the same postsynaptic cell. Their postsynaptic potentials are additive. EPSPs summate and move the membrane potential closer to the threshold level for firing. Conversely, EPSPs and IPSPs cancel each other out.

Temporal summation occurs when successive APs in a presynaptic neuron follow in rapid succession, so that the postsynaptic responses overlap and summate. Summation is possible because the synaptic potential lasts longer than action potentials by a factor of 10-100 times.

Each individual synapse contains receptors, ion channels, and other key molecules, which are sensitive to the neurotransmitters released at the site. These specific protein molecules are involved in synaptic plasticity and summation.

Electrical synapses. A *gap junction* is a transmembrane pathway of low electrical resistance that connects the cytoplasm of adjacent cells. A gap junction allows the membrane potential of

the adjacent cells to be *electrically coupled*. Gap junctions form *electrical synapses***,** which differ from chemical synapses in that transmission, is instantaneous.

An electrical synapse consists of several protein pores, which close in response to increased intracellular [Ca^{2+}] or [H^{+}] in a cell, thereby increasing their resistance. Open gap junctions exchange ions and small molecules up to a molecular weight of 1000 Dalton.

Gap junctions are found in simple reflex pathways, where rapid transfer of the electrical potential is essential, and between non-neural cells such as epithelial and myocardial cells, smooth muscle cells and hepatocytes.

Neurotransmitters are divided into classical, rapidly acting non-peptides (Box 7-1) and putative, slowly acting neuropeptides (Box 7-2) - all dealt with in Chapter 7.

Here is only described the function of GABA, neuropeptides and dopamine.

The major *inhibitory* transmitters are GABA (gamma-aminobutyric acid) in the brain and glycine in the spinal cord. Binding of GABA to the GABA-receptor opens the pore for Cl^{-} influx, whereby the subsynaptic cell membrane *hyperpolarises* (Fig. 2-3). The increase in Cl^{-} conductance stabilises the membrane potential and decreases the efficacy of excitatory transmission. The GABA-receptor pore is permeable to K^{+} besides Cl^{-}. The GABA-receptor has a major inhibitory role in brain function and is the binding site for barbiturates (used as hypnotics in anaesthesia) and for benzodiazepines (used to relieve anxiety).

Fig. 2-3: A $GABA_A$ -receptor in an inhibitory synapse.

The $GABA_A$-receptor shown here is related to *sedation* and *mood*, whereas the $GABA_B$-receptor controls *spasticity* (Chapter 7). Picrotin blocks the GABA-channel.

Glutamate, aspartate and related acidic amino acids are the most important *excitatory transmitters* in the brain and spinal cord. Excitatory neurons possess *excitatory* amino acid (EAA) receptors. EAA receptors are a family of receptors with at least four different ions channels: The N-methyl-D-aspartate-receptor **(**NMDA), and three so-called *non-NMDA receptors* - one of which is the *glutamate receptor*. The NMDA-receptor operates with K^{+}-efflux, while Na^{+} and Ca^{2+} enters the subsynaptic neuron. Mg^{2+} and many antiepileptic drugs block the NMDA-receptor channel (Chapter 7). Opening of Na^{+}- and Ca^{2+}-channels, which allow an increased influx of Na^{+} and Ca^{2+}, cause the membrane potential to approach the threshold level for excitation. Both a reduced Cl^{-}-influx to the neuron and a reduced K^{+}-efflux move the membrane potential towards the threshold level and possible excitation. The NMDA-receptor has a separate glycine site.

Neuropeptides (Box 7-2) have slow excitatory or inhibitory transmitter actions. Peptides cannot be synthesized locally in the axon terminals, because they do not have ribosomes.

Fig. 2-4: Peptide neurotransmitters

Peptides are water soluble, and act as hormones by binding to specific cell-surface receptors. *Cell-surface receptors* are a family of guanosine triphosphate-binding proteins, so-called GTP-binding or *G-proteins,* which control and amplify the synthesis of second messengers. Cell-surface receptors for neurohormones can function as transport protein and possess enzyme activity (Fig. 2-4).

Neuropeptides are build by a sequence of amino acids. Neuropeptides are synthesized in the

cell bodies of the neurons and transported to the terminal buttons by rapid axonal transport (Fig. 2-4). Some neuropeptides are released together with a non-peptide co-transmitter (Box 7-2).

Some neuropeptides are produced when a large *mother-peptide* is cleaved into several active neuropeptides. Neuropeptides are released from the nerve terminal near the surface of its target cell, and diffuse to the receptors of the target cell. Low concentrations of neuropeptides typically affect the membrane potential by changing the conductance of the target cell to small ions. The action of neuropeptides usually lasts longer than that of enzyme-inactivated transmitters. Following prolonged synaptic transmission, neuropeptides are deactivated by *proteolysis*.

Dopamine and other catecholamines derive from tyrosine via DOPA, which stands for the precursor 3,4-dihydroxy-phenylalanine. Dopamine is the actively accumulated into storage vesicles in the nerve endings together with noradrenaline and ATP. Dopamine activates both presynaptic and subsynaptic D_2-receptors (Fig. 2-5).

Fig. 2-5: Dopamine receptors and the interactions with noradrenaline (NA).

Noradrenaline can be oxidatively deaminated by monoamine oxidase (MAO) located on the external membrane of mitochondria (Fig. 2-5). The enzyme COMT (catechol-O-methyl transferase) can also methylate noradrenaline to nor-metanephrine. MAO and COMT are important in metabolising circulating catecholamines. Re-uptake of noradrenaline is the most important terminator of its actions.

Activation of both D_2-receptors opens K^+-channels and the increased outflux of K^+ hyperpolarizes the membrane. Blockage of the presynaptic D_2-receptors in substantia nigra with antipsychotic drugs reduces K^+-outflux and increases dopamine production and release.

Loss of *dopamine-containing neurons* in substantia nigra results in the lack of dopamine at the D_2-receptors of the striatal neurons. These neurons degenerate in Parkinson's disease causing *muscular rigidity* and *hand tremor* (Chapter 4).

3. Skeletal muscles

Skeletal or striated muscles are attached to a skeleton. Striated muscles are called *striated*, because they have a striking banding pattern. Microscopy with polarised light reveals *dark* (optically anisotropic) *striations* or *A bands* alternating with light or optically isotropic striations or *I bands*. Running along the axis of the muscle cell or *muscle fibre* is the *myofibril bundles* of *filaments* that are visible on electron micrographs. The A band contains the *thick filaments* of myosin, and the I band contains *thin filaments* of actin and tropomyosin (Fig. 2-6). The thin filaments are anchored to a transverse structure termed the *Z disc* (Fig. 2-6). Each contractile unit contains the halves of two I-bands with the A-band in between. This unit is a *sarcomere*. Sarcomeres have a length of 2.3-2.5 mm between the two Z discs at rest. The central A band is a relatively isotropic substance - also termed the *H-band* - with an *M-line* of darkly stained proteins that link the thick filaments into a fixed position. Contraction takes place by sliding of the filaments.

The sliding of filaments against each other is called the *sliding filament hypothesis*, and since contraction works by cycling of millions of crossbridges, it is also called the *theory of crossbridge cycling*.

The thin filaments are 1-1.2 mm long and consist of small globular proteins that form two helic pearl strings. The *double helix of actin* is supported by a long, thin molecule of *tropomyosin*

that is situated along the groove of the double strands of actin (Fig 2-6). Each tropomyosin molecule interacts with 7 actin molecules on each side. *Troponin* is composed of 3 subunits: Troponin-C binds Ca^{2+}, troponin-T reacts with tropomyosin, and troponin-I inhibits the actin-myosin-interaction, when Ca^{2+} is absent. *Dystrophin* is another normally occurring cytoskeletal muscle protein.

The thick filaments are 1.6 mm long, and consist of large myosin molecules. Myosin is a dimer of almost 500 kD. Each monomer consists of one heavy chain and two light chains. The *heavy chain* consists of a helical tail and a globular head (Fig. 2-6). The *light chains* are associated with the head of the heavy chain. Since myosin is a dimer, the *double-helix tail* must end in two *globular heads* (Fig. 2-6). The globular heads contain the ATPase activity and the actin-binding site. The *light chains* control the rate of cross bridge cycling.

Fig. 2-6: Thick and thin filaments. The crossbridge cycle.

The crossbridge cycle theory states that there are *multiple cycles* of *myosin-head* attachment and detachment to *actin* during a muscle contraction. When myosin binds to actin, an actinomyosin complex is formed - with an extremely active ATPase. The interaction between actin and myosin and the hydrolysis of ATP is the basic process that converts chemical energy into mechanical energy.

Each crossbridge consists of two heads. At rest the crossbridge from myosin is not attached to actin. The globular myosin heads are oriented perpendicular to the filament axis (Fig. 2-6), and they have a high standard affinity for actin.

1. Stimulation of a muscle liberates Ca^{2+} in the sarcoplasma, which removes the troponin-tropomyosin blockage of the actin, and actin can react with the binding sites on the globular heads. The crossbridge is now bound to the thin filaments (Fig. 2-6).
2. The binding accelerates the release of ADP and P_i from the actin-myosin complex, and the attached global heads change conformation by 45^o with respect to the filament axis. The head of the crossbridge drags the thick filament *10 nm* along towards the Z-disc or - at constant length - a proportional force is developed. Multiple repetitions of this short sliding process is necessary to result in an appreciable muscle shortening. In the absence of ATP, the crossbridge cycle stops here and the binding is immobile (*rigor link* and *rigor mortis*).
3. The following stage is the binding of ATP to the myosin heads, which weakens the binding to actin and disrupts the *rigor link.*
4. Then ATP is partially hydrolysed on the myosin head, and the resulting energy is stored in the perpendicular head, which has a renewed high standard affinity for actin. If Ca^{2+} is present, a new crossbridge cycle is initiated and may occur 100 times each s. With a cycle movement of 10 nm this is 1000 nm per s for each half of the sarcomere.

Fig. 2-7: Force-length diagram

Force is required to stretch a relaxed muscle, because muscle tissue is elastic, and the force increases with increasing muscle length (Fig. 2-7). The passive blue curve reflects the properties of the elastic, connective tissue, which becomes less compliant or stiffer with lengthening (Fig. 2-7).

A muscle contraction at constant length is termed *isometric.* Force is measured in Newton (N), and *one N* is the force required to accelerate a mass of one kg with an acceleration of **one** m s^{-2}. In muscles, the traditional expression for *force* is stress or *tension* in N per cross-sectional

area of the muscle (N m), which is actually pressure (Pascal, Pa). Here, the ordinate is force expressed as a percentage of the maximal force (Fig. 2-7).

1. The length at which maximum active contractile force is developed is called L_o, corresponding to a sarcomere length of 2.15 m m (Fig. 2-7). L_o is the length of the muscle in the body when at rest. At this length there is a maximum number of active crossbridges (Fig. 2-7). When an isolated muscle in an *isometric force* or *stress-meter* is stimulated, the active muscle force decreases with the decrease in overlap between thin and thick filaments; at a sarcomere length of 3.65 m m the isometric force reaches zero (Fig. 2-7). The force is always proportional to the number of cycling cross-bridges interacting with the thin filament.

2. Force also declines at muscle lengths less than L_o (Fig. 2-7). Thin filaments overlapping, and thick filaments colliding against Z-discs cause this. The isometric force (stress) decreases as the sarcomere length is reduced, as shown with the sarcomere length of less than 2.15 m m (Fig. 2-7).

3. When the active muscle length is stretched beyond any overlapping between the thin and the thick filaments the muscle can only develop a force of zero (see the sarcomere length of 3.65 m m with an A-band of 1.6 m m in Fig. 2-7).

The lengths of the thick and thin filaments of human striated muscles are similar (1.6 and 1.2 m m, respectively). They generate *maximal tension forces* at L_o, corresponding to a sarcomere length of 2.2 m m, namely 300 kN per m^2 or kPa.

Muscle power or *work-rate* (Eq. 2-1) is the product of muscle force (N) and shortening velocity (m s^{-1}). The maximal work rate of human muscles is reached at a contraction velocity of 2.5 m s^{-1}. The maximal work-rate is thus (300 kPa *2.5 m s^{-1}) = 750 kW per square meter of cross sectional area.

Hill developed an equation for the shortening velocity of isotonic muscle contractions (Eq. 2-2). The equation is illustrated in Hills force-velocity diagram (Fig. 2-8).

The maximum force is developed at the initial length (Fig. 2-8 right: 18 g of load). At 18 g there is no shortening – the length is unchanged. Stimulation of the unloaded muscle results in maximum shortening velocity (100%). An unloaded crossbridge can cycle at maximal rate, indicated by maximal shortening velocity (Fig. 2-8 right).

The shortening velocity decreases rapidly as the afterload is increased describing a hyperbola (Fig. 2-8 right). With increasing loads the latency is increased and the shortening is reduced (4 and 9 g in Fig. 2-8 left). The latency depends on the length of thepreceding isometric phase. The maximal velocity of shortening is directly proportional to the *myosin ATPase activity*. We increase the velocity of muscle shortening under a given load by the recruitment of additional motor units.

The long human arm muscles shorten at a rate of 8 m per s. Muscles can bear a load of 1.6 times the maximal force before the crossbridges are broken, but under such extreme conditions the work-rate (power) of the muscle approach zero (no shortening in Fig. 2-8 left). This is also the case when a person attempts to lift a motor car - the speed of shortening is zero (isometric contraction). On the other hand, the speed at which a pocket thief operates is probably impressive, although the force is minimal.

Maximal work-rate occurs at a load of 1/3 of the maximal isometric force of the muscle. Here

the contractile system has optimal efficiency in converting chemical energy into mechanical energy.

Fig. 2-8: Hill's force-velocity diagrams (right) and related shortening curves (left).

A further rise in filament velocity seems to reduce the potential for actin-myosin interaction. The *crossbridge cycling rate* falls as the load on the crossbridges increases (Fig. 2-8 right).

In a muscle, the force of contraction is graded by *increasing* the frequency of action potentials, and by *recruiting* more muscle cells. Prolonged crossbridge contraction results in physiological *tetanus*. This is a prolonged muscle contraction maintained by the prolonged Ca^{2+}-influx caused by repetitive stimulation.

Human skeletal muscles consist of three functional types of motor units. A motor unit is a motor neuron with the muscle fibres it innervates. All muscle fibres belonging to a motor unit are of the same type. The three types of muscle fibres are characterised in Box 2-1.

Box 2-1. Structural, functional and histochemical characteristics of twitch fibres.

Classification	***Red (I)***	***Red (IIA)***	***White (IIB)***
	Slow oxidative (SO)	FOG	FG
	Intermediate	Red	White
	Slow	FR	FF
	Slow-twitch	Fast-twitch red	Fast-twitch white
Myoglobin	High	High	Low
Oxidative enzymes	High	Intermediate	Low
Glycolytic activity	Low	Low	High
Glycogen	Low	High	Intermediate
Mitochondria	Intermediate	High	Low
Mitochond.ATPase	Intermediate	High	Low
Sarcoplasmic retic.	Intermediate	Dense	Dense
Fibre diameter	Small	Intermediate	Large
Contractions	Postural	Endurance	Powerful
Shortening velocity	Low (I)	Intermed. (IIA)	High (IIB)
Recruitment	First	Second	Last

Most human skeletal muscles are a mixture of all three types of motor units, although the proportions vary considerably.

Type I: The *slow* motor units contain *slow-oxidative* (SO) red *slow-twitch* fibres.They are adapted to continuous postural muscle activity. The fibres have many mitochondria and a high content of myoglobin (red fibres). They depend on aerobic metabolism and the glycogen content is high. Slow motor units have weak but long lasting contractions (slow reaction to a signal or twitch). The fibres are small and are first to be recruited. During light work these highly excitable motor units activate red fibres suited for prolonged activity or endurance activities. Endurance training increases the oxidative capacity of the activated motor units, whereas strength training increases cellular hypertrophy.

Type IIA: *Fast*-twitch, fatigue-resistant (FR) motor units have type IIA twitch fibres with a high or intermediate content of mitochondria, myoglobin, and glycogen. These fibres also rely upon oxidative metabolism (fast oxidative glycolytic = FOG) and have a high level of both oxidative and glycolytic metabolism. The motor units provide contractions of intermediate force and duration, and they resist fatigue. FOG fibres are of intermediate size, and they are recruited before the white fibres. This is in accordance with the *size recruitment principle*: Small or intermediate motor units are easier to activate by excitatory postsynaptic potentials (EPSPs) than large neurons.

Type IIB: Fast-twitch fatiguable (FF) motor units produce fast contractions (fast-twitch), and fatigue easily, as the name implies. Their *large white fibres,* with their dense sarcoplasmic reticuli, are adapted to activities requiring large forces with rapid control of contraction and relaxation. The fast-twitch white fibres (also called type IIB due to the highest shortening velocity) have few mitochondria, small amounts of myoglobin (white fibres), and depend on glycolysis (high anaerobic metabolism). They have only small amounts of glycogen (fast glycolytic = FG). The FF motor neuron is large, the axon is thick and it branches so greatly that the FF motor unit innervates more muscle fibres. This is why FF motor units are capable of powerful contractions. The cell body receives type Ia afferents. The FF units are recruited last and mainly during maximal efforts such as sprinting. The production of ATP by glycolysis matches the high rate of ATP consumption.

We have three major *metabolic sources* of ATP:

1. Phosphocreatine, which is an immediate energy source used for intense white fibre activity such as sprinting. Lohmann's creatine kinase catalyses the efficient reforming of ATP from ADP by the conversion of a small phosphocreatine pool to creatine. Following exercise the *oxygen debt* is repaid and the phosphocreatine pool is restored (Chapter 18).
2. The glycogen stores of the muscle produce ATP rapidly but inefficiently by *glycolysis*, with lactate as the end product.
3. Glucose, free fatty acids, triglycerides and amino acids in plasma are substrates for oxidative phosphorylation. This is a most efficient pathway and the slowest source of energy due to the many steps in the process (Chapter 20).

4. Smooth muscles

The same molecules as in striated muscle essentially cause contraction in smooth muscle, but the intracellular organisation and the dynamic characteristics are entirely different (Box 2-2).

Box 2-2: Characteristics of skeletal, cardiac and smooth muscle cells.

	Skeletal	***Cardiac***	***Smooth muscle***
Diameter (m m)	Up to 100	10	Up to 5

Length (m m)	200 000	50	Up to 200
T-tubules	Yes	Yes	No -Simple caveoli
Regular sarcomers	Distinct	Distinct	No -Look smooth
Regular Z-discs	Yes	Yes	No- but dense bodies
Regular myofibrils	Yes	Yes	Irregular myofibrils
Troponin	Yes	Yes	No
Sarcoplasmic reticulum	Yes	Yes	Simple reticulum
Gap junctions	No	Yes	Yes (single-unit)
Extracellular Ca^{2+}	No	Yes	Yes
Refractory period	Short	Long (300ms)	Long
Latency (ms)	10	10	200
Twitch (ms)	10-100	300	3000
Resting membrane pot.(mV)	-80	-90	-50
Force	High	High	Low maintained for days
Energy cost	300-fold	High	Low
Disorders	Atrophy	Cardiac	Asthma, hypertension

Smooth muscles are called so because they lack the distinct *sarcomeric bands* of *striated* muscles. Smooth muscle cells are spindle-shaped and line the hollow organs and the vascular system; the smooth muscle cells are extremely small (Box 2-2). Smooth muscle cells contain a few thick myosin-filaments, and many thin actin-filaments attached to *dense bodies* by a-actin (helical sarcomers). The cells are without *regular* sarcomers, Z disc's, myofibrils and T-tubules. Smooth muscle cells lack troponin. Dense bodies are analogous to Z disc's, and some dense areas are attached to the cell membrane. Smooth muscle cells do not contain a typical sarcoplasmic reticulum, which can store and release Ca^{2+}. Instead some fibres possess an analogous *simple reticular system* located near the *caveoli* of the cell membrane. Caveoli are small invaginations of the membrane, similar to the T-tubules of striated muscles. The more extensive the reticular system is in the smooth muscle fibre, the higher is its shortening velocity due to release of Ca^{2+} mediated by IP_3. Smooth muscle cells maintain large forces almost continually at extremely low energy costs.

The same tension or tone is maintained for days in smooth muscle organs (intestine, urinary bladder, gall bladder) and can be obtained in striated muscle at high energy cost (up to 300

times the smooth muscle rate of ATP consumption).

Smooth muscle cells are extremely sensitive to *extracellular* [Ca^{2+}].

During an action potential the inward flux of ions is not Na^+, but Ca^{2+} through slow Ca^{2+}-channels. They open mainly in response to a ligand binding, but we have also voltage-dependent Ca^{2+}-channels.

The force-length relation is qualitative similar to that of striated muscles, so the *sliding-filament mechanism* is probably analogous (Fig. 2-6).

The smooth muscle mechanism is special, because stimulation results in a maintained isometric force with strongly reduced velocities. Smooth muscle contractions are extremely slow. Ca^{2+} probably regulates the number of active crossbridges in smooth muscle slowly and indirectly.

Smooth muscle cells contain some mitochondria, and they show a slow contraction pattern superimposed on the lasting tonus. Smooth muscle contractions typically last for 3 s, in contrast to striated muscle with total contraction periods of 10-100 ms. Since the energy demand in smooth muscle is extremely low, it is balanced by the *oxidative ATP synthesis*. Smooth muscle cells do not have an oxygen debt as striated muscles do, although they produce *large amounts of lactate*. This is probably because the ATP-synthesising glycolytic mechanism is located in the cell membrane and is linked to the ATP-utilising Na^+-K^+-pump. Smooth muscle contains far fewer myosin filaments than striated muscle. The myosin crossbridge heads of smooth muscle contain an isoenzyme with much less ATPase activity than that of striated muscle. Ca^{2+}-entry through the cell membrane is much slower than internal release of Ca^{2+}.

A contracting smooth muscle fibre releases Ca^{2+} from two pools. The *large extracellular fluid pool* is essential. In the fibre that possesses a sarcoplasmic *reticulum* similar to the sarcoplasmic reticulum of striated muscle, there is a fast intracellular pool. The smooth muscle cell membrane contains a *$3Na^+$-$2K^+$-pump,* a delayed *K^+-channel,* a ligand-activated and a voltage-dependent *Ca^{2+}-channel,* a sarcolemmal *Ca^{2+}-pump,* and a Na^+- Ca^{2+}-exchanger (Fig. 2-9).

1. A stimulatory ligand is bound to membrane receptors for G-proteins and for ligand-gated Ca^{2+}-channels (Fig. 2-9). The major Ca^{2+}-influx takes place through the ligand-gated (noradrenaline) and the voltage-gated Ca^{2+}-channels. The Ca^{2+}-influx depolarizes their membrane, whereby Ca^{2+} further permeates the membranes. The depolarization by ligand binding thus indirectly opens the voltage-gated channels.
2. When a stimulus acts on reticular receptors via a G-protein it activates phospholipase C. *Phospholipase C* hydrolyses *phosphatidyl inositol diphosphate* (PIP_2) into IP_3 and diacyl-glycerol, DAG (Fig. 2-9).

Fig. 2-9: Contraction and relaxation in smooth muscle cells. The ligand is acetylcholine in visceral cells and noradrenaline, ATP and peptide hormones in vascular smooth muscle cells.

3. IP_3 is bound to a receptor on the simple sarcoplasmic reticulum and this second messenger binding elicits a controlled release of Ca^{2+} from the reticulum. Hereby, the sarcoplasmic [Ca^{2+}] rapidly increases above the threshold for contraction (0.1 m M).

4. The crossbridge cycling is regulated by a *myosin light chain kinase* (MLC kinase) dependent upon both Ca^{2+} and calmodulin. The phosphorylation of myosin to myosin-phosphate is drastically accentuated by the binding of 4 Ca^{2+}-calmodulin to MLC kinase forming a complex. The phosphorylated light chain myosin reacts with actin in the thin filaments and contracts. The rate of sliding and of ATP-splitting is up to 1000-fold slower than in striated muscles.
5. Ca^{2+} is actively pumped out of the cell by an ATP-demanding Ca^{2+}-pump and through a Na^{+}- Ca^{2+}-exchanger (antiport). The antiport uses the energy of the Na^{+}-gradient for influx. Reuptake into the poorly developed sarcoplasmic reticulum and the mitochondria is slow compared to cardiac and skeletal muscle tissue.
6. Below the Ca^{2+}-threshold the myosin light chains are dephosphorylated by myosin light chain phosphatase and the contractile structures relax.
7. The Na^{+}-K^{+}-gradient across the cell membrane is maintained by the Na^{+}-K^{+}-pump (Fig. 2-9).

When the high intracellular [Ca^{2+}] during an action potential is lowered again towards the resting level, the cell relaxes. This is accomplished by stimulation of the *sarcolemmal* Ca^{2+}-pump, and by blockade of both Ca^{2+}-input and Ca^{2+}-release.

Metarterioles and precapillary sphincters without nerve fibres can still respond to the needs of the tissue by the action of local tissue vasodilatators. The following factors cause smooth muscle relaxation, and therefore vasodilatation: Adenosine, NO, lack of oxygen, excess CO_2, increased [H^{+}], increased [K^{+}], diminished [Ca^{2+}], and increased [lactate].

Endothelial-derived relaxing factor (EDRF) is recently shown to be *nitric oxide* (NO). Activation of endothelial cells produces NO from arginine, and NO diffuses into the smooth muscle cells. NO stimulates directly the enzyme *guanylate-cyclase*, and by that intracellular [cGMP] elevates.

Circulating acetylcholine *contract***s** the arterial smooth muscles when bound to *cholinergic receptors*.

Smooth muscle cells grow (hypertrophies) as a response to the needs of the body, and they also retain the capacity to divide.

During *hypertension* the lamina media of the arterioles hypertrophies which increases the *total peripheral vascular resistance* in the systemic circulation. These topics are further developed in Chapter 9.

During *pregnancy* the (single-unit, see below) smooth muscles of the myometrium are quiescent and contain few *gap junctions* under the influence of progesterone. At term the myometrium grows and the number of gap junctions *increases*, due to the high oestrogen concentration. Now the myometrium is well prepared for the co-ordinated contractions during *parturition* (see Chapter 29).

Smooth muscle changes length without marked changes in tension. Initially, there is a high tension developed upon stretching; then the tension falls as the myosin and actin filaments are reorganised by slowly sliding against each other. A sudden expansion of the venous system with blood results in a sharp rise in pressure followed by a fall in pressure over minutes. The smooth muscle fibres in the walls of the venous system are highly compliant, because they have accepted a large blood volume without much rise in pressure (*delayed compliance*).

Smooth muscle cells are frequently involved targets in diseases such as hypertension, stroke,asthma, and many gastrointestinal diseases.Smooth muscle cells can be divided into multi-unit smooth muscle and single-unit smooth muscle.

Fig. 2-10: Contraction of multi-unit smooth muscle cells (vascular). A single contraction is elicited by an electrical stimulus and later acetylcholine elicits tetanus. Contraction of multi-unit smooth muscle is controlled by extrinsic innervation or by hormones. Mechanical contact junctions between the cells are not found.

1. In *multi-unit smooth muscle tissues* each cell operates entirely independent of other cells and the cell does not communicate with other muscle cells through gap junctions**.** The discrete cells are separated by a thin basement membrane and often innervated by a single neuron, and their main control is through nerve signals. Thousands of smooth muscle cells belonging to the multi-unit type join by the common innervation in a *functional syncytium***.** Multi-unit smooth muscle is found in the eye (the ciliary muscle and sphincters as the iris muscle of the eye), in large arteries, in the vas deferens, and in the piloerector muscles that cause erection of the hairs. These muscle cells are normally quiescent, insensitive to stretch and they are activated only through their autonomic nerves. Each muscle is composed of *multiple motor units*, hence the name: *multi-unit smooth muscles*. The nerve fibre branches on a bundle of smooth muscle fibres, and form *junctions* with varicosities filled with transmitters. These junctions are analogous to the neuromuscular junctions of striated muscles. The neurotransmitters are acetylcholine and noradrenaline. Multi-unit smooth muscles have developed a *contact junction* with shorter latency than the slowly operating *diffuse junctions* mainly found in the single-unit type.
2. *Single-unit smooth muscle cells* are arranged in bundles such as the arrangement in a viscera eg. intestine, uterus and ureter (Fig. 2-11). These smooth muscle cells communicate through hundreds of *gap junctions*, separating the cell membranes by only 2-3 nm, and from pacemaker tissue of variable location, action potentials are generated initiating a contraction of the muscle. In this respect single-unit cells resemble the cardiac muscle.

Fig. 2-11: Single-unit smooth muscle cells resemble cardiac muscle. Activity propagates from cell to cell through gap junctions forming an electrical syncytium. The dense bodies and dense areas contain alpha-actin.

Action potentials generated in one cell can activate adjacent cells by ionic currents spreading rapidly over the whole organ and securing a co-ordinated contraction as though the tissue were a *single unit* or a *syncytium*. These cells are characterized by their spontaneous motility and by their sensitivity to stretch. The spontaneous activity is usually modified by the autonomic nervous system. *Visceral smooth muscle* undergoing peristalsis, generates propagating action potentials from cell to cell.

Other cell-to-cell contacts are *desmosome's* and *intermediary junctions* subserving structural contact. These intermediary junctions transfer mechanical force from one smooth muscle cell to another on the plasma membrane, causing the single-unit smooth muscle cell to function like a *stretch transducer*.

5. Cardiac muscle tissue

Myocardial cells are built of regular sarcomers just like the skeletal muscles, and they are contracting fast. Myocardial cells form an electrical syncytium in the same way as the single-unit smooth muscle cells. The characteristics of myocardial, skeletal and smooth muscle cells are presented in Box 2-2. Myocardial cells are mononuclear and the myoglobin, enzymatic and mitochondrial content are large just as the red fibres of skeletal muscles. The metabolism of myocardial cells is similar to that of red skeletal fibres, both being designed for endurance

rather than speed and strength. The oxygen supply to the heart muscle must be maintained, if it is to synthesise ATP at a sufficient rate. Myocardial cells deprived of oxygen for 30-s cease to contract.

Myocardial cells most resembles smooth muscle in its auto-rhythmicity and syncytial function. Pacemaker cells in the sinus node determine the normal cardiac frequency, because they send out spontaneous action potentials along the conduction system of the heart with a higher frequency than any other cells in the heart. Vagal stimulation releases acetylcholine at the pacemaker cells. Acetylcholine increases the K^+-permeability, whereby K^+ leaves the cell and hyperpolarizes the cell membrane. This is why the pacemaker (cardiac) frequency is reduced by vagal nerve stimulation. Sympathetic stimulation or adrenaline reduces the K^+-permeability, so the depolarization is shortened, and the pacemaker frequency increased.

The prolonged action potential characteristic for myocardial cells is initiated by an abrupt Na^+-influx (phase 0) through *fast Na^+-channels* just as in the striated muscles. The AP plateau is due to a slow Na^+-Ca^{2+}-channel, which deliver Ca^{2+} for the contraction activation. The action potential releases Ca^{2+} from the sarcoplasmic reticulum to the sarcoplasma. The effect is distributed by the cardiac T-tubule system.

Cardiac contraction by crossbridge cycling depends on the presence of extracellular Ca^{2+} just as in smooth muscle tissue. Therefore, use of Ca^{2+}-antagonists reduces the contractile force of the heart, whereas drugs, which increase Ca^{2+}-permeability across the membrane, improve the contraction. In the heart, Ca^{2+}-influx tends to prolong the depolarization just as in smooth muscle cells. The cardiac glycoside, digoxin, selectively binds to and inhibits the sarcolemmal $3Na^+$-$2K^+$-pump, which leads to an increase in intracellular $[\ Na^{+}]$. Although the Na^+-efflux is inhibited, the redundancy of Na^+ affects the *Na^+-Ca^{2+}-exchanger* (3 Na^+ out for one Ca^{2+} into the cell), leading to an increase in cellular $[\ Ca^{2+}]$ and in the force of contraction. This is the mechanism of the increase in contractile force by digitalis glycosides.

Pathophysiology

This paragraph deals with 1. Disorders of the neuromuscular junctions (myasthenia gravis), 2. Skeletal muscle disorders (dystrophia, dystonia, muscle injuries), 3. Smooth muscle disorders (asthma, hypertension etc) and 4. Myocardial disorders (coronary artery disease, arrhythmias, and chronic heart disease).

1. Disorder of the neuromuscular junction (Myasthenia gravis)

This serious disease is acquired, but the cause is unknown. The development of this *autoimmune disorder* may be related to other diseases. Rheumatoid arthritis treated with D-penicillamine has resulted in myasthenia gravis. More than 50% of the myasthenia patients have *thymic hyperplasia* and some patients have a real *thymoma*.

Many of these patients have an increased blood concentration of antibodies against their own *acetylcholine receptor protein*. There is a *decreased density* of receptor proteins on the postjunctional membrane. This was shown by the use of radioliganded toxins from poisonous snakes (which bind irreversibly to the acetylcholine receptor protein).

The patients are tired and the muscles are extremely weak. This is particularly so for the proximal limb muscles, the extraocular muscles and the neck muscles, whereby the patient has difficulties in lifting the head. Mastication and swallowing is a difficult process.

Fig. 2-12: Neuromuscular junction with antibodies and decreased density of

acetylcholine-receptors in a patient with myasthenia gravis.

As just mentioned the blood of most patients with myasthenia gravis contains *autoantibodies* against *acetylcholine (ACh) receptor proteins* on the cell surfaces of the motor end plates etc. The autoantibody competes for the ACh receptor and inhibits synaptic transmission, so muscular contraction is greatly inhibited. Deposition of immune complexes eventually destroys the ACh-receptor protein.

Intravenous injection of an *anticholinesterase* improves the muscle strength immediately, but the beneficial effect is gone within 3 min.

Thymectomy improves the condition and the prognosis also in the group of patients without thymoma.

Oral anticholinesterase (such as pyridostigmine) has beneficial effect over 2-4 hours. They inhibit the enzyme acetylcholine-esterase, and thereby prolong the effect of naturally occurring acetylcholine on the receptors. In severe cases this treatment is inefficient, and immune-suppressants such as corticosteroids are sometimes favourable.

2. Skeletal muscle disorders

Muscular dystrophy is an inherited disorder of skeletal muscles. *Duchenne muscular dystrophy* is an X-linked recessive muscle disorder characterized by the absence of *dystrophin* in the striated muscles and in the myocardium. The locus is localised to the Xp21 region of the X chromosome. Dystrophin is a normally occurring cytoskeletal muscle protein. The patient is a boy, who has to climb up his legs in order to reach the erect posture. Typically, there is proximal weakness with compensatory *pseudohypertrophy* of the calves. There is no cure and the patient dies from myocardial damage.

Dystonias are *prolonged muscle contractions* leading to muscular spasms. There is a simultaneous action of opposing agonist and antagonist groups that produce abnormal postures. Dystonia is painful and particularly resistant to treatment.

Dystonia musculorum deformans begins in childhood with generalized spasms that affect gait and posture. In most cases the cause is a genetic defect.

Spasmodic torticollis causes the head to turn (torticollis) or change posture. Patients with a trigger zone on the jaw benefit from acupressure here.

Muscle injuries are dealt with in Chapter 18.

3. Smooth muscle disorders

The most important disorders are *asthma* (Chapter 14) and *systemic hypertension* (Chapter 12). Smooth muscles are also involved in a disorder of swallowing (*achalasia*), where the myenteric plexus and the lower oesophageal sphincter fail to respond with *receptive relaxation*, and the food accumulates in the oesophagus. Other disorders of the gastro-intestinal smooth muscles are also treated there.

4. Disorders of the myocardium

Coronary artery disorders (smooth muscles and myocardial disease) and congestive heart disease are treated in Chapter 10, and cardiac arrhythmias in Chapter 11.

Only direct therapeutic uses of the systems developed till now are described here.

Nitro-glycerine, nitroprusside and similar drugs relax smooth muscles by transfer of NO from endothelial cells. NO increases *intracellular* [cGMP] (Fig.11-1), which is the basis for the

beneficial effect of the drugs on cardiac cramps. These second messengers activate protein kinases that phosphorylate effector proteins such as Ca^{2+}-pumps and K^{+}-channels. Such vasodilatators stimulate the *sarcoplasmic Ca^{2+}-pump*, inhibit Ca^{2+}-influx and stimulate K^{+}-efflux through the delayed K^{+}-channel (reduces the excitability). Hereby, the high intracellular [Ca^{2+}] during an action potential is lowered towards the resting level (10^{-7} mM), and the smooth muscle cell relaxes producing vasodilatation.

Equations

- **Muscle power** (or work rate) equals the product of muscle force and shortening velocity

Eq. 2-1: Power (W) = Force (N) * Velocity (m s^{-1}).

- **Hill's equation.** The force-velocity curve is shown in Fig. 2-7. The curve fits Hill's equation:

Eq. 2-2: Initial shortening velocity (v) = (Po - P)*b/(P + a)

where P is the force or load acting on the muscle, Po is the maximal isometric force or load, **a** is a constant with the dimensions of a force, and b is a constant with the dimensions of velocity.

Self-assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

- **A.** Motor neurons synthesise acetylcholine unrelated to their content of choline-acetyltransferase.
- **B.** There is a high density on the subsynaptic membrane of specific acetylcholine receptors.
- **C.** The receptor protein for acetylcholine contains a voltage-gated channel for cations.
- **D.** Binding of acetylcholine elicits a transient opening of ionophores, which are specifically permeable to small ions.
- **E.** Parkinson's disease is possibly caused by loss of dopamine containing neurons in the substantia nigra.

II. Each of the following five statements have False/True options:

A. Nitro-glycerine, nitroprusside and similar drugs contract smooth muscles by transfer of nitric oxide from endothelial cells.
B. All myasthenia patients have a thymoma.
C. During hypertension the lamina media of the arterioles hypertrophies which increases the total peripheral vascular resistance in the systemic circulation.
D. The nicotinic acetylcholine receptor is related to an acetylcholine-gated ion channel found not only in the neuromuscular junction, but also at all autonomic ganglia and in the central nervous system.
E. When the high intracellular [Ca^{2+}] during an action potential is lowered again towards the resting level, the cell contracts. This is accomplished by stimulation of the sarcolemmal Ca^{2+}-pump, and by blockade of both Ca^{2+}-input and Ca^{2+}-release.

Try to solve the problems before looking up the answers.

Highlights

- *Recruitment is the increase in force and contraction velocity of a muscle by activation of more and more motor units.*
- *Synaptic transfer refers to the transmission of signals from one neuron to another, and the site of contact between the two neurons is called the synapse.*
- *A chemical synapse consists of a neuronal presynaptic terminal, a synaptic cleft and a subsynaptic membrane with associated receptor proteins. The chemical synapse is highly developed in the CNS. It conducts the signal one way only, and has a characteristic synaptic delay.*
- *A gap junction or electrical synapse is a pathway of low electrical resistance that connects cytoplasm of adjacent cells. A junction couples adjacent cells electrically and thus allows synaptic transmission without delay.*
- *Neurons with motor function have the ability to synthetize acetylcholine, because they contain choline-acetyltransferase.*
- *GABA (gamma-aminobutyric acid) in the brain and glycine in the spinal cord are inhibitory neurotransmitters. Binding of GABA to the GABA-receptor opens the pore for Cl^- influx, whereby the subsynaptic cell membrane hyperpolarizes. The GABA-receptor has a major inhibitory role in brain function and is the binding site for barbiturates (used in anaesthesia) and for benzodiazepines (used towards anxiety).*
- *Glutamate, aspartate and related acidic amino acids are the most important excitatory transmitters in the brain and spinal cord. Excitatory neurons possess excitatory amino acid (EAA) receptors. These EAA-mediated synapses predominate in the CNS.*
- *Each neuron in the CNS is in contact with up to 10^5 presynaptic axon terminals. Synaptic inputs are integrated by either spatial or temporal summation.*
- *Neuropeptides are built by a sequence of amino acids. Neuropeptides are synthesized in the cell bodies of the neurons and transported to the terminal buttons by rapid axonal transport.*
- *Loss of dopamine-containing neurons in substantia nigra results in lack of dopamine at the D_2-receptors of the striatal neurons. These neurons degenerate in Parkinson´s disease causing muscular rigidity and hand tremor.*
- *Blockade of the presynaptic D_2-receptors in substantia nigra with antipsychotic drugs reduces K^+-outflux and increases dopamine production and release.*
- *The crossbridge cycle theory states that there are multiple cycles of myosin-head attachment and detachment to actin during a muscle contraction. When myosin binds to actin, an actomyosin complex is formed - with an extremely active ATPase.*
- *Muscle power or work-rate is the product of muscle force (afterload in N) and shortening velocity ($m\ s^{-1}$). The maximal work rate of human muscles is reached at a contraction velocity of 2.5 $m\ s^{-1}$. The maximal work-rate is thus (300 kPa *2.5 $m\ s^{-1}$) = 750 kW per square meter of cross sectional area.*
- *Tetanus is a prolonged muscle contraction maintained by the prolonged Ca^{2+}-influx caused by a high stimulation frequency.*
- *Smooth muscle cells are frequently involved targets in diseases such as hypertension, stroke, asthma, and many gastrointestinal diseases.*
- *Smooth muscle cells maintain large forces almost continually at extremely low energy costs. The same tension or tone is maintained for days in smooth muscle organs (intestine, urinary bladder, and gall bladder).*
- *Myocardial cells form an electrical syncytium in the same way as the smooth muscle cells do.*
- *Myocardial cells deprived of oxygen for 30-s cease to contract.*

The most important smooth muscle disorders are asthma and hypertension.

- *The most important myocardial disorders are coronary artery disease, arrhythmias, and chronic heart disease.*
- *Myasthenia gravis is a disorder of neuromuscular contraction. The patients frequently have an increased blood concentration of antibodies against their own acetylcholine receptor protein and thymic hyperplasia.*
- *Duchenne muscular dystrophy is an X-linked recessive muscle disorder characterized by the absence of dystrophin in the striated muscles and in the myocardium.*

Further Reading

- Kupfermann, I. "Functional studies of cotransmission." *Physiol. Rev.* 71: 683, 1991.
- Pollack, G.H. "Muscles and molecules: Uncovering the principles of biological motion." Seattle, Washington, 1990. *Ebner & Sons*.
- Alberts, B. et al. "Molecular biology of the cell." *Sec. Ed.*, 1989, Garland Publishing, Inc., New York & London.

Return to top

Return to Content

SECTION II.

The Nervous System

My colleagues Jørn Hounsgaard and Poul Dyhre-Poulsen have contributed importantly to this section.

The nervous system is the essential control system for the human body. It consists of a central nervous system (CNS) and a peripheral nervous system. All information originates in sensory receptors and, enters the CNS via the peripheral nervous system. The CNS controls all the activities of the human body ranging from contractions of striated and smooth muscles to the exocrine and endocrine secretions.

The nervous system is an extremely rapid signal transduction system and the most important communication network in our body. The main integrative functions allow selection and processing of incoming signals to produce an appropriate response.

The nervous system includes sensory receptors that detect events in the body as well as in the outer world. Signals or action potentials from the sense organs travel through peripheral, afferent nerves to the CNS, where they are processed. The CNS controls the various activities of the body by motor mechanisms that generate movements and glandular secretions through efferent nerves. The afferent and efferent nerve fibres distributed throughout the body form the peripheral nervous system that is subdivided into a somatic and an autonomic part.

Neurons are highly specialised cells that are excitatory, inhibitory and sometimes neurosecretory. Neurons receive and transmit signals (action potentials) to other neurons or effectors. Neuronal networks account for information in a memory, evaluation of available knowledge, decision making, and transmission of response signals to appropriate effectors. The human nervous system contains about 10^{12} neurons forming at least 10^{15} synapses.

Frequently used abbreviations in this section are CSF for cerebrospinal fluid, EAA for excitatory amino acids, ECF for extracellular fluid, ECV for extracellular fluid volume, EEG for electroencephalogram, and REM for rapid eye movements. A complete list of abbreviations is present in Chapter 35.

Chapter 3.

The Somatosensory System And Disorders

Study Objectives

- To *define* adaptation, adequate stimulus, coding, sensory receptors including taste and smell, molecular receptors, receptor potential, stimulus transfer, types of sensory nerve fibres, conduction velocity, and threshold stimulus.
- To *describe* skin receptors, articular receptors, nociceptors and central pathways, the effect of chordotomy, thalamic surgery, and prefrontal lobotomy.
- To *draw* Hills force-velocity curve and the voltage-duration curve for nervous stimulation.
- To *calculate* one variable from relevant information's given.
- To *explain* cortical somatotropic and columnar organisation, the control of taste and smell, the control of nociceptive transmission (gatecontrol), central analysis, central pain, headache, referred pain, allodynia, causalgia, hyperalgesia, trigeminal neuralgia, thalamic syndrome, phantom limb pain, hyperalgesia, and Brown-Sequards syndrome.
- To *use* the concepts in problem solving and case histories.

Principles

- *Critical empiricism. In brain research any scientific observation presupposes a theory that can be falsified. Theories that fail to be falsified in repeated scientific projects are temporarily acceptable. This philosophy is generally applicable.*
- *Sherrington's integration law. The integrative action of the nervous system unifies separate organs to form an individual personality.*

Definitions

Adaptation or *accommodation* of sensory receptors refers to a progressive decrease in firing frequency despite maintained depolarisation.

Adequate stimulus refers to the stimulus, for which the receptor has a lower energy threshold than for other stimuli - ie, the stimulus to which the receptor is most sensitive.

AMPA is an abbreviation of *alpha-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid*. This is an excitatory amino acid (EAA) and one of the EAA-gated ion channels carries AMPA-receptors with high affinity for a subclass of glutamate receptors.

Causalgia means hyperalgesia (hypersensitivity to pain) elicited by a cold stimulus.

Coding: In neurons the stimulus intensity is coded by the frequency of action potentials.

Dermatomes: Every spinal dorsal root is destined to a segment of the skin called a dermatome.

Endogenous opioids are substances in the CNS with opiate-like effects.

GABA is the abbreviation for gamma-amino butyric acid – the most common inhibitory transmitter in the brain.

The **gate-control hypothesis** of pain states that pain transmission is suppressed by innocious signals in thick myelinated afferents (group II), whereas the pain sensation is enhanced by signals in thin afferents.

Headache is pain on the surface of the head, which is actually due to anomalies in intracranial or extracranial structures.

Hyperalgesia means hypersensitivity to pain.

NMDA is the abbreviation for N-methyl-D-aspartate.

Nociceptors or nocireceptors are responsive to stimuli that potentially cause injury.

Plasma membrane receptors consist of a protein or glycoprotein molecule, an ion channel pore or a specific enzyme (G-protein).

Receptor potential: The stimulation of a receptor elicits a *generator potential* that is graded according to the stimulus strength. When the stimulus is strong enough to reach the threshold, an action potential is fired. Sensory receptors translate stimulus intensity to impulse activity in sensory afferents.

Sensory receptors are either *neurons* in the case of vision, smell and cutaneous senses, or *modified epithelial cells* in the case of auditory, vestibular and taste senses.

Threshold stimulus refers to the weakest stimulus to which the receptor will react.

Trigeminal neuralgia (French: *Tic douloureux*) is a condition with daily paroxysms of violent pain in part of the trigeminal area lasting only some seconds. The pain is provoked by eating, washing the face or by cold in the face.

Essentials

This paragraph deals with *1. Sensory receptors and nerve fibres, 2. Blood-brain barrier and CSF, 3. Regeneration of nervous tissue, 4. Sensory pathways, 5. Central opiate receptors, 6. Taste and smell.*

1. Sensory receptors and nerve fibres

Sensory receptors are either *neurons* in the case of vision, smell and cutaneous senses, or

modified epithelial cells in the case of vision, auditory, vestibular, smell and taste senses. The special sensory receptors for vision, hearing and balance are described in Chapter 5.

Some sensory receptors have characteristics similar to the well-known *plasma membrane receptors*. Plasma membrane receptors consist of a protein or glycoprotein molecule, an ion channel or a specific enzyme (G-protein).

The stimulation of a receptor elicits a receptor potential (*generator potential*) that is graded continuously with stimulus intensity. When the stimulus is strong enough to reach the threshold, action potentials (APs) are fired. In neurons the stimulus intensity is coded by the frequency of action potentials.

Sensory receptor systems are *biological transducers* with a dynamic range of *up to* 10^{12} in the most sensitive organ, the *ear*. The *threshold* is the reciprocal of the sensitivity. The *threshold* is the weakest stimulus to which the receptor will react. The *sensitivity* of a sensory receptor is greater the smaller its threshold stimulus is.

The Pacinian corpuscles in the skin are gigantic receptors (almost 1 mm long) consisting of concentric layers like onion-scales in the microscope. There is a single axon in the axis of each corpuscle (Fig. 3-1). Stimulus intensity is coded by the single axon. Compression deforms the axon and depolarises the membrane by opening of Na^+-channels.

Fig. 3-1: The Pacinian corpuscle (1 mm long) is a vibration detector.

The depolarisation generates a graded *receptor potential* forcing a current towards the first node of Ranvier with maintained stimulus. The receptor potential rapidly decreases (rapid adaptation), because the adequate stimulus is alterations in the deformity rate (vibrations in the range 150-300 Hz). At the first node of Ranvier, a propagating *action potential* along the axon is released, provided the generator potential is sufficiently large (Fig. 3-1). The Pacinian corpuscle is located in the deeper layers of the skin and connective tissue. The afferent fibres are thick (Type Ab or II), and they lead the signals to synapses in nucleus gracilis and cuneatus of the spinal cord.

Steven proposed the *power law* that is given as Eq. 3-1. The *interpreted stimulus strength* (ISS) is equal to a constant (k) multiplied by the *actual* stimulus strength (SS) raised to the power n. The situation n equal to 1 describes a linear relation between stimulus and resulting activity in the conducting neuron (Fig. 3-2).

Perception of *taste, heat and angular acceleration* are described by *power functions* or *transfer functions* with n just above 1, whereas *hearing and smell* are described by functions with n lower than 1. The only sensation with a particularly large value of n (about 3) is *pain*, which is why pain is felt so severe with increasing stimulus intensity!

Fig. 3-2: A graphical description of the power law.

However, the power law and other types of curve fitting with transfer functions have hardly improved our understanding of sensory modalities.

Both *conduction velocity* and *size* is used in classification of nerve fibres. The fibres are divided into types A, B and C, based on the three main *conduction velocities* shown in the record of the *compound action potential* from a mixed nerve. Type A, B and C refer to phases in the combined action potential. Type A fibres are the fast conducting myelinated fibres (thick fibres subdivided into a , b , g , and d), type B are preganglionic sympathetic fibres, and type C are the small, unmyelinated fibres. Another classification is based on the *thickness of the axons* (I-IV). The size classification became necessary, when Aa - fibres were separated into two subgroups: Ia and Ib.

In Box 3-1, the velocity classification (A-C) is written first, and the size classification (I-IV) is given in parentheses.

Box 3-1: Classification of nerve fibres

Fibre type	Function	Axon diameter m m	Conduction
		/ Myelin + -	velocity, m per s
Aα (I)	motor a - fibres	9-18/+	70-120
	spindle afferents (Ia)		
	tendon organs (Ib)		
Aβ (II)	touch and pressure	5-12/+	30-75
Aγ (II)	motor to muscle spindles	3-6/+	18-36
Aδ (III)	pain, pressure, temperature	1-5/-	4-30
B (III)	preganglionic	3/-	3-12
C (IV)	pain, touch, heat	1/-	1-2

The Aa (I) fibres are *motor α -fibres* and *proprioceptors* from the annulospiral endings of muscle spindles (Ia) and from Golgi tendon organs (Ib).

The Ab (II) fibres conduct discrete touch and fine pressure signals from cutaneous tactile receptors.

The Ag (II) fibres are motor fibres to muscle spindles. They have their origin in the spinal cord.

The Ad (III) fibres transfer pain sensations, decline in skin temperature as well as crude, passive touch and deep pressure.

- The B fibres (III) are autonomic preganglionic fibres.
- The C fibres (IV) are unmyelinated and lead *pain, touch* and *signals* from heat receptors from the skin. The C fibres have no myelin sheath.

Sensory receptors in the nervous system are classified as *exteroceptors* (located on the body surface), *proprioceptors* (located in muscles, tendons and joint capsules), *interoceptors* (located in the viscera), and *telereceptors* (stimulated by events far from the person).

Cutaneous receptors are exteroceptors (Fig. 3-3). Pacinian and Meissner corpuscles are rapidly adapting (dynamic) touch velocity detectors in glabrous skin. In hairy skin, hair-follicle receptors are velocity detectors (they adapt rapidly). Meissner corpuscles are located in the papillae of the hairless skin such as fingertips, lips and clitoris. Merkels discs and Ruffini-end organs are slowly adapting (static) touch intensity detectors both in hairless and hairy skin. Merkels discs are found in elevated dome corpuscles in hairy skin (up to 50 Merkel discs in a corpuscle of 0.5 mm in diameter). These so-called *Iggo dome receptors* are extremely sensitive and transmit touch signals to a single nerve fibre. In this way, even weak tactile stimuli can

create sensation in the CNS.

Fig. 3-3: Cross section through an area of hairless and a hairy skin showing 6 types of sensory mechanoreceptors. All the mechanoreceptors are supplied by afferent nerve fibres of group II, except the free nerve endings (group IV).

The free nerve endings, with unmyelinated afferents (group IV) conduct signals with low velocity (Box 3-1), and register passive touch, such as, slow strokes with a piece of cotton.

Thermoreceptors are also exteroreceptors. We have cold-receptors just below the skin surface (200 nm deep). Cold receptors respond to changes in temperature. Heat receptors are also located in the skin. The location of certain heat and cold points in the skin is determined by bringing a thin hot or cold object in contact with the skin. *Thermoreceptors* react to temperature changes. *Cold receptors* and *heat receptors* in the skin are located close to the surface. Both types of receptors are also located in the deep tissue and in the CNS. Both types of receptors *discharge spontaneously* at normal temperature, *dynamically* when skin temperature is changing rapidly, and *adapt slowly*.

Proprioreceptors, located in muscles, joints and joint capsules, are mechanoreceptors (muscle spindles, Golgi-receptors, Pacinian and Ruffini corpuscles, and free nerve endings). The Ruffini mechanoreceptors are also called *joint receptors*, because they are located in ligaments, tendons and *articular capsules*. They provide information for the CNS concerning articular movements, movement velocity and joint position. Joint receptors of the proximal joints are particularly sensitive. The static and dynamic receptors inform the CNS about the position and movement of the joint, respectively. These receptors enable us to sense the position of the joint with great accuracy.

Accommodation of sensory receptors or *adaptation* is a progressive decrease in firing frequency despite maintained depolarization. The frequency of action potentials from stimulated receptors fall, although the stimulus is maintained at constant strength.

Accommodation or adaptation occurs, when a proportion of the voltage-gated Na^+-channels is rapidly *inactivated* by depolarisation, which also opens K^+- channels. This makes the cell more refractory to stimulation. Accommodation can also be caused by a hyperpolarization induced by gradual activation of Ca^{2+} -dependent K^+ -channels.

Fig 3-4: Accommodation or adaptation curves from different sensory receptors.

Pain- and cold-receptors, Merkel discs and Ruffini-end organs adapt extremely slowly and incompletely (Fig. 3-4). Joint receptors, smell-taste-receptors, muscle spindles, carotid sinus- and pulmonary stretch receptors, and the optic nerve, all adapt somewhat better (Fig. 3-4).

Hair-follicle receptors, Meissner corpuscles and Pacinian corpuscles adapt rapidly, just as many free nerve endings (Fig. 3-4).

Nociceptors or *nocireceptors* (pain receptors) are responsive to stimuli that potentially cause injury. Nociceptors are free nerve endings of two types. The *fast adapting* Ad fibre mechanical nociceptors (group III) are high-threshold, finely myelinated afferents that originate superficially in the skin. The *slowly adapting* C-polymodal nociceptors (group IV) are unmyelinated afferent fibres that originate in the deeper cutaneous tissue, and respond to various mechanical, thermal and chemical stimuli (Fig. 3-5). In the spinal cord nociceptive afferents synapse with secondary neurons in lamina I and II. These sensory neurons ascent in the spinothalamic tracts.

The fast adapting pain through group III fibres is bearable (acute, sharp, stinging, somatic pain), compared to the slowly adapting unbearable pain (diffuse, burning, prolonged secondary, visceral pain) through group IV fibres.

Fig 3-5: Mechanical, polymodal and visceral nociceptors. A visceral pain afferent synapses in the spinal cord with the neuron of the lateral spinothalamic tract on which the cutaneous group IV pain afferent terminates.

When nociceptors become sensitised (ie, more responsive), their thresholds are reduced, thus causing *hyperalgesia* (ie, hypersensitivity to pain). Many substances such as bradykinin, histamine, leucotrienes, prostaglandins, serotonin, and K+ that are often released near damaged or dying cells sensitise nociceptors. K^+ *activates* the nociceptors. Substance P is also released from polymodal nociceptors through an axon reflex with antidromal signal transduction in afferent group IV fibres, causing hyperalgesia, vasodilatation and increased capillary permeability (Fig. 3-5). Glutamate may be co-released with *substance P* from the polymodal C-fibre terminals.

The *gate-control hypothesis* of pain states that pain transmission is suppressed by innocious signals in thick myelinated afferents (group II), whereas the pain sensation is enhanced by signals in thin afferents. Inhibitory interneurons of the lamina II in the dorsal horn of the spinal cord perform the gate-control through a special type of presynaptic inhibition called primary afferent depolarization (PAD), and the receptors on the cell body of the secondary neuron is the gate. The gate control hypothesis explains why innocuous signals, mediated by large myelinated afferents, can inhibit pain mediated by thin myelinated afferents.

The *adequate stimulus* is the stimulus, for which the receptor has a lower energy threshold than for other stimuli (ie, the stimulus to which the receptor is most sensitive). The adequate stimulus for pain receptors is mechanical deformation, extreme temperature or tissue damage. The *sense impression* depends on the site in the brain which receives the sensory signal (ie, *central analysis*) and on the receptor localisation (ie, *peripheral analysis*). This is how different neurons transmit different types of sensations, even though they may transmit the same electrical signals (see Chapter 8).

The CNS discards more than 99% of all incoming signals as irrelevant.

The visual system is an example of a *specific information line* for a certain modality of sensation. The neurons in the retina, the optic nerve, the lateral geniculate nucleus, and the visual cortex describe just such a dedicated neuronal pathway. The specific information line through which the signal is conducted determines the way in which a suprathreshold stimulus is perceived (eg, pressure applied on the eye will be perceived as light).

The auditory system also forms a *specific or labelled line* all the way from receptor to cortex. In all cases the specific region in the cerebral cortex, where the nerve fibre ends determines the modality of sensation.

Now, where is the sense interpretation localised and what is its intensity?

1. Coding in the sense organ is *peripheral analysis*, which is based on the peripheral location of the receptor. External energy is transformed to a *receptor potential* that triggers APs in afferent nerve fibres. Peripheral analysis depends upon the location and the special structure and sensitivity of the receptor. The pattern of firing of APs is the only possible variable for coding information in a single neuron. Examples of firing patterns are *on-off* patterns with mean frequencies, *off-on* patterns, *transient* patterns or *adaptation*, *long-lasting* patterns, firing with latency etc.
2. *Central location coding in the CNS* is termed *central analysis*, which is related to the sense impression.

2. Blood-brain barrier and CSF

The blood-brain barrier consists of *tight junctions* between the endothelial cells of the capillaries in the CNS and of neuroglia. This barrier only allows extremely small or hydrophobic molecules to pass into the brain. The cerebral microcirculation consists of *strong arterioles* that can constrict to carry a high arterial pressure without brain oedema.

Many large molecules cannot pass from the blood to the cerebrospinal fluid (CSF) across the choroid plexus, a tight junction barrier that is called the *blood-cerebrospinal fluid barrier*.

The *blood-CSF barrier* of the choroid plexus allows some large molecules to pass from the

blood to the CSF.

Fig. 3-6: The blood-brain and blood-CSF barriers showing the daily formation of 500 ml of CSF.

The blood-brain and the blood-CSF barriers exist in all areas of the brain, except in the so-called *circumventricular organs* (hypothalamus, the pineal gland, and the area postrema). These discrete organs have highly fenestrated capillaries that are easily penetrated by large and small molecules as well as ions. The circumventricular organs are located close to essential control centres in the hypothalamus and brain stem regions regulating respiration, blood glucose concentration, and extracellular fluid osmolality.

The two brain barriers are almost impermeable to large molecules such as plasma proteins, but highly permeable to CO_2, oxygen, water, alcohol, anaesthetics, hallucinogens, and other lipophilic substances. The blood-brain barrier is almost completely impermeable to water-soluble molecules, electrolytes such as H^+, whereas CO_2 passes through the barrier to the medullary chemoreceptors (Fig.16-3).

Humans produce 500 ml of CSF daily. The total CSF volume is only 1/3 of the daily production. Most of the 500 ml of CSF is produced in the *choroid plexuses* in the four brain ventricles, and the remaining is produced across the blood-brain barrier.

The ventricular system and the central spinal channel are covered with *ependyma*. The absorption of CSF takes place through the *arachnoidal granulations*, which protrude, into the *sinus sagittalis*. The rate of absorption is directly related to the pressure in the cranial cavity - in particular the CSF- pressure. Proteins can pass through large holes in the endothelial cells. The CSF is separated from the brain cells by the *thin* pia mater. Substances that enter the CSF can easily diffuse into the brain interstitial fluid. Drugs that cannot pass the *blood-brain barrier* can enter the brain through pia mater, when infused into the CSF (Fig. 3-6).

The CSF passes from the lateral ventricles (I and II) through the *foramen of Monroe* into the third ventricle (III), through the *aquaduct of Sylvius*, the fourth ventricle (IV), and out into the subarachnoid space through the *foramina Luschkae & Magendie* (Fig. 3-7).

Fig. 3-7: Anatomical structures involved in CSF-formation and absorption.

The normal CSF-pressure in a supine person is up to 10 mmHg (1.3 kPa) or 136 mm of water.

The secretion of fluid by the choroid plexus depends on the active Na^+-transport across the cells into the CSF. The electrical gradient pulls along Cl^-, and both ions drag water by osmosis. The CSF has lower [K^+], [glucose], and much lower [protein] than blood plasma, and higher concentrations of Na^+ and Cl^-. The production of CSF in the choroid plexuses is an active secretory process, and not directly dependent on the arterial blood pressure. The CSF is separated from the brain cells by the extremely thin *pia mater*. All natural substances that enter the CSF can easily diffuse into the brain extracellular fluid (Fig. 3-6).

CSF leaves the four ventricles through the roof of the 4th ventricle, traverses the subarachnoid space, and is reabsorbed into the blood of the venous sinuses via the arachnoidal villi. The *absorption* here is directly related to the CSF pressure in the cranial cavity. Large holes through the endothelial cells allow proteins to enter the blood.

3. Regeneration of nervous tissue

Severe injury to nervous tissue causes cell death. Neurons are postmitotic cells. For this reason lost neurons cannot be replaced.

There is, however, considerable capacity for regeneration of axons in the peripheral nervous system. Both growth and maintenance of axons require the *nerve growth factors* (NGF). NGF is an essential survival factor for neurons outside the CNS - in particular sensory neurons. NFG binds to receptors belonging to the insulin receptor family (*tyrosine kinase family*).

When a motor axon has been severed, the cell body undergoes **chromatolysis**. This is a

neuronal reaction, where the rough endoplasmic reticulum (the *Nissl bodies*) becomes active. The Nissl bodies accumulate proteins required for repair of the axon. The *axonal reaction* is an attempt to repair the fibre by production of new protein structures that are transported along the axon. Therefore, proteins distend the rough endoplasmic reticulum. The axon and the myelin sheath distal to the injury die and are phagocytized. The neuroglial Schwann cells that had formed the myelin remain alive. This is the so-called *wallerian degeneration* named after Waller.

The Schwann cells proliferate and form long rows along the pathway previously occupied by the dead axon. The severed axon regenerates along this pathway, and *growth cones* may eventually reinnervate the target organ.

Neurological injury probably involves excessive *glutamate receptor stimulation* as a common pathway.

Glutamate is the most important of the *excitatory amino acids* (EAAs) in the spinal cord and the brain. Glutamate stimulates the family of EAA-receptors including AMPA-, NMDA-and metabotropic receptors. NMDA means N-methyl-D-aspartate. - Effective glutamate antagonists are applied in clinical studies of pain.

The inhibitory amino acids, GABA and glycine, and the monamines and endogenous opiods inhibit the second-order neurons of the spinothalamic tract.

Fast axonal transport of organelles in the cytosol occurs as rapidly as 0.4 m per day. At this rate synaptic vesicles can travel along the motor axon from the spinal cord to a patient's foot within three days. Fast axonal transport of enzymes and organelles occurs on microtubuli in the axons, and is not interrupted by resting periods in cell compartments outside the transport system (Fig. 3-8). Oxidation of glucose in the mitochondria provides ATP for the Na^+-K^+-pump and for transport filaments and microtubules embedded in the axonal cytoplasm (Fig. 3-8).

Fig. 3-8: Axonal transport of vesicles, organelles and proteins by microtubuli.

Slow axonal transport occurs as diffusion of cytosolic proteins and organelles such as mitochondria. This transport occurs at a rate 100 times more slowly than fast axonal transport. Organelles or enzymes are stored in different cell compartments on their way or their direction of transport reverses.

Axonal transport can be *anterograde*, when it occurs in the direction from the soma to the axonal terminals. Axonal transport can also be *retrograde*, when it occurs in the opposite direction. Here vesicles are degraded by lysosomes, when returned to the soma. A typical example of slow transport is the transfer of the many mitochondria towards the terminal of an axon.

In the CNS, fast neurotransmission is *inhibitory* or *excitatory*. In the neuromuscular junction, *each signal* is always excitatory and sufficient to trigger a muscular contraction. In the neuromuscular junction, acetylcholine is the only neurotransmitter, whereas in the CNS there is a large variety of neurotransmitters (see Box 7-1 and 7-2).

The sensory system transmits signals from sensory nerve receptors in the body. The nerve receptors are located in the skin, muscles, tendons, joints and viscera. The signals are transferred to the CNS by a pathway of first, second, third, and higher-order neurons. The third and higher order neurons are located in the *thalamus* and the *cortex*. The cell body of the first order afferent neuron is located in the dorsal root or in the cranial nerve ganglia. The signals pass through the spinal cord, the brain stem, and the thalamus before reaching the cerebral cortex.

4. Sensory pathways

Several sensory tracts and pathways synapse in the *nuclei of the thalamus* (the spinothalamic tracts). The *somatosensory thalamus* is a relay station for most sensory modalities. The sensory inputs are processed in somatotopic areas of the thalamus, and are then transferred to

appropriate cortical areas. The *somatotopic organisation* is maintained all the way to the cortex.

The *reticular activating system* (RAS) of the brainstem is involved in arousal acting in concert with the thalamus.

The *spinothalamic tract* conveys pain and temperature (lateral tract), and also crude passive touch (ventral tract). The first-order neurons are afferent Ad fibres (III) which have cell bodies in the spinal ganglia. Second-order neurons cross immediately to the opposite side of the spinal cord, and ascend in the lateral and ventral spinothalamic tract.

Fig. 3-9: The spinothalamic tracts and their sensory function.

Pain and temperature reach the thalamus in the lateral spinothalamic tract (in the lateral funiculus). The second-order axon terminates in the *somatosensory thalamus* (the ventral posterior lateral nucleus and the central lateral nucleus). The third-order neurons pass from the somatosensory thalamus via the *thalamocortical fasciculus* to the *somatosensory cortex* or the *primary sensory cortex* (somatic sensory area I, or area 1, 2, 3 in Fig. 4-2) with the *sensory homunculus.* Some third-order neurons also pass to the somatic sensory area II of both hemispheres.

Proprioception and active tactile signals are transmitted through sensory nerve fibres to the spinal cord. Primary afferent fibres ascend in the *dorsal columns* all the way to the medulla oblongata. These primary axons synapse with second-order neurons in the gracile and the cuneate nuclei. These second order neurons cross the midline in the medulla, and ascend in the medial lemniscus to end in the somatosensory thalamus. The *medial lemniscus pathway* transmits proprioception and fine tactile senses.

The *spinothalamic tract* is the most important pathway for *pain*. The *second order neurons* of the spinal tracts have their cell bodies in the lamina I, II and V of the spinal cord. These cells receive excitatory signals from nociceptors in the skin, muscles and viscera. The action potentials from the nociceptors are conducted along the axon to the spinal cord and release neurotransmitters such as the excitatory amino acid, glutamate, and different neuropeptides. When these neurotransmitters bind to the receptors on the postsynaptic membrane of the secondary neurone, they increase the permeability to small ions, and excite secondary, postsynaptic neurons. The secondary neurons of the spinothalamic tract projects mainly to the *contralateral thalamus* by crossing over immediately through the anterior commissure to the opposite side of the spinal cord within the incoming segment.

***5*. Central opiate receptors**

The *endogenous analgesia system* is a pain control system descending from brainstem to the spinal cord (Fig. 3-10).

As an example, this system may explain why a runner who twists his leg during a competition may finish the run before he really feels the pain. As soon as he has passed the goal and stop running the pain often becomes severe, and he cannot run at all.

The cell bodies of the neurons belonging to this system are located in the *periaqueductal grey area* of the midbrain, the *periventricular areas*, locus coeruleus, and the areas surrounding the *aqueduct of Sylvius* (Fig. 3-10). Signals from these cell bodies reach the medullary *raphe magnus* nucleus and the medullary *nucleus reticularis gigantocellularis* with nucleus reticularis *paragigantocellularis* lateralis. The nuclei transmit signals via the *descending pain-suppressing pathway* in the dorsolateral column to a *pain inhibitory complex*. Stimulation or increased tone of the analgesia system can suppress strong pain signals entering the spinal cord through the dorsal spinal horn. These regions contain *opioid receptors*. There are at least 4 types of *central opiate receptors* and their subtypes: **u** for morphine-like drugs, d and k for enkephalins, and the non-selective s -*receptors*.

Endogenous opioids are substances with opiate-like effects. These substances are naturally occurring in the nervous system (b -endorphin, met-enkephalin, leu-enkephalin, dynorphin and

many others). Endogenous opioids are derivatives of *three* large protein molecules encoded by three different genes. These mother-molecules are *pro-opio-melanocortin* (POMC), *proenkephalin* and *prodynorphin.*

Fig. 3-10: Opiate receptors in the CNS. Enkephalin, b-endorphin and dynorphin (3 peptides) appear in normal human cerebrospinal fluid.

Enkephalins inhibit both type C and type Ad (III) pain fibres presynaptically in the dorsal horns. *Enkephalin* is the endogenous ligand for the d -opiate receptors. Dynorphin has much higher affinity than morphine and is only found in small quantities close to the dynorphinergic k -opiate receptors. b -*endorphin* is present in the hypothalamo-hypophysary system.

Presynaptically located opiates inhibit depolarization of nerve terminals and reduce synaptic transmission. The purpose of pain is to protect the body from further or imminent harm.

A special type of burning pain is provoked by noxious heat or by capsaicin (which contain a vanillyl-group) in chilli, paprica and pepper. These spices and heat stimuli seem to activate a vanniloid receptor subtype 1 in sensory nociceptors with terminals in the dorsal horn of the spinal tract. Activation opens Ca^{2+}-channels and the Ca^{2+}-influx is probably involved in the burning sensation.

Gyrus cinguli has the highest density of *central opiate receptors*. Pyramidal cells are contacted by *opiate secreting interneurons* that inhibit arriving pain signals.

6. Taste and smell

The sensations from the anterior 2/3 of the tongue travel with the trigeminal nerve fibres, through the *chorda tympani* into the facial nerve (VIIth), and eventually reach the *solitary tract* of the *brain stem*. Taste signals from the back of the tongue and surrounding tissues are transmitted through the glossopharyngeal nerve (IXth) into the tractus solitarius. All taste fibres synapse in the *nuclei of the solitary tract* and the axons of these neurons project to the thalamus. From the thalamus third-order neurons reach the lower part of the *primary sensory cortex* in the postcentral gyrus (somatosensory area I = area 1 in Fig. 4-2).

Fig. 3-11: Taste buds and taste pathways from the tongue.

Acids evoke *sourness*, because H^+ stimulates special *H^+-receptors* in the taste buds. *Saltiness* is produced by the anions of inorganic salts. The *Cl^--receptor* is particularly effective in registering saltiness. Our taste buds at the base of the tongue also have *bitter-receptors* stimulated by many long-chain organic compounds. Many alkaloids (quinine, caffeine, and nicotine) also taste bitter. *Sweet-receptors* are stimulated by sucrose, glucose, lactose, maltose, glycerol, alcohol, aldehyde, ketone, and organic chemicals.

In the *upper nasal cavity* the mucous membrane is yellow and termed the *olfactory membrane*. It contains 100 million bipolar neurons called *olfactory cells*. They contain hairs or *olfactory cilia*. The olfactory cells are *smell receptors*. They work as telereceptors, and the smell pathways do not include the thalamic relay station and a neocortical projection area. Instead, the olfactory cells pierce the cibriform plate and synapse in the olfactory bulb. The olfactory tract then transmits the olfactory signals to the olfactory cortex at the surface of the temporal lobe. In the *limbic system* (Fig. 4-3), olfactory information is correlated with feeding behaviour and emotional-motivational behaviour.

Fig. 3-12: The olfactory region, its receptors and pathways.

Pathophysiology

This paragraph deals with *1.The thalamic syndrome, 2. Brown-Sequards syndrome, 3. Special sensory pain disorders, 4. Taste and smell disorders.*

1. The thalamic syndrome

The thalamic syndrome is frequently caused by *thrombotic* blockade of bloodflow to the somatosensory thalamus. The destruction of thalamic neurons in one hemisphere leads to

ataxia and *loss of sensations* from the opposite side of the body. After a few months different types of sensations return, but they are accompanied by pain.

2. The Brown-Sequards syndrome

The Brown-Sequards syndrome or paresis includes all effects of *transection* of only one half of the spinal cord at a certain level. All motor functions on the side of the lesion are blocked in the segments below the level (paresis, spasticity, and loss of vasoconstrictor tone). Sensations of pain and temperature from all lower dermatomes on the opposite side of the body are lost, because of transection of the contralateral spinothalamic tract. The only sensation left on the side of transection is crude touch, because it is transmitted in the opposite ventral spinothalamic tract. The total sensory loss is therefore termed *dissociated anaesthesia*.

3. Pathological pain

Hyperalgesia means *hypersensitivity to pain*. Hyperalgesia is caused by either hypersensitive pain receptors (sunburned skin), or by facilitated transmission. Facilitated transmission is due to abnormal stimulation of peripheral nerve fibres and neurons of the spinal cord or of the thalamus.

A special type of hyperalgesia is present when *herpes virus* infects one or more dorsal root ganglia. The virus excites the neurons and causes pain in the dermatomal segment subserving the ganglion. The segmental pain circles halfway around the truncus on the affected side. The virus is also transported by axonal flow to the cutaneous terminals, where it causes a characteristic rash confined to the dermatome. The disease is called *herpes zoster*.

Causalgia is *hyperalgesia and heat-cold-sensations* accompanied by sweat secretion in a region with nerve lesion. The hyperactivity in sympathetic efferent neurons and in nociceptive afferents running along arteries is unexplained.

Even the lightest touch at sensitised *trigger areas* release - within seconds - severe lancinating pain throughout the affected branch of the trigeminal nerve. The cause is unknown, and therapy is usually unsuccessful.

Referred pain and central pain

Pain that originates from deep organs is poorly localised and often referred coming from superficial structures. This may be explained by the fact that pain signals from viscera are transmitted through neurons in the CNS that also transmit pain signals from a specific area of the skin. Pain due to *myocardial ischaemia* (angina pectoris) is commonly described as pain originating from the inner side of the left arm, and termed *referred pain*.

Dermatome anaesthesia

The mammalian embryo is segmented into so-called *somites*, which are innervated by an adjacent part of the spinal cord. Every spinal dorsal root is destined to a segment of the skin called a *dermatome*. Also muscles (*myotomes*) bone (*sclerotomes*), and viscera are related to specific segments of the spinal cord or brain stem. The dermatomes are drawn with sharp borders, which is unrealistic (Fig. 3-13). There is considerable overlap and several successive dorsal roots must be interrupted to produce *dermatome anaesthesia*. In case of serious spinal cord injury the dermatome map is useful for determination of the level and extent of the lesion.

Fig. 3-13: Dermatomes

Central pain is a sensation of pain in absence of peripheral nociceptive stimuli. Central pain is processed in the *cortical pain areas*, and caused by lesions along the nociceptive pathways (peripheral nerves, the spino-thalamo-cortical tract and the thalamus).

Amputation of a limb is sometimes followed by *phantom limb pain*. The patient suffers from severe pains, and the sensation is projected to the amputated limb. It is not known whether the mechanism is central or peripheral.

Trigeminal neuralgia (*tic douloureux*) is a condition with daily paroxysms of violent pain in

part of the trigeminal area lasting only some seconds. The pain is provoked by eating, washing the face or by cold in the face. The mandibular and maxillary area is involved. The paroxysm finishes with saliva-tears- and sweat secretion. The disease is probably located to the trigeminal Gasserian ganglion, but the cause is usually unknown. Drugs or surgical procedures on the ganglion have variable effect.

Headache

Headache is caused by anomalies in intracranial or extracranial structures.

1. Intracranial headache is released from nociceptors in the meninges or in the arteries and veins at the base of the skull. The brain tissue itself hardly contains pain receptors. The intracranial types cover frontal, occipital, migraine, and meningeal psychogenic and pressure headache.

Nociceptors are stimulated by stretch (dura or tentorium), by dilatation (vessels), or by chemical means (eg. histamine, 5-hydroxytryptamine etc). The pain signals reach the CNS through the 5^{th} and 9^{th} cranial nerve and the cervical sensory fibres.

Stimulation of supratentorial nociceptors is referred to the frontal area via the 5^{th} cranial nerve as *frontal headache*. Subtentorial nociceptors cause *occipital headache* through the 2^{nd} cervical nerve.

Migraine headache

Migraine or hemicrania means *unilateral headache*, which is frequently but not always present. Migraine is defined as *recurrent attacks* of headache associated with gastrointestinal and visual disorders. There is evidence for a genetic aetiology of migraine: an autosomal dominant inheritance with reduced penetrance (ie, expression only in permissive environments).

Classical migraine has prodromal symptoms (aura) with visual disturbances due to ischaemia in the retina. The onset is often in the eye region with spread towards the vertex or towards the other eye and typically accompanied with nausea, emesis, photophobia and *scintillating scotomata*. Sometimes sensorimotor abnormalities occur on one side of the body with ataxia, dysphasia and syncope.

Frequently migraine occurs without aura, which make the diagnosis difficult.

Migraine is of unknown origin, but it is sometimes associated with prolonged psychological stress. A hypothesis claims that prolonged emotional stress in sensitive individuals causes *reflex vasoconstriction* of intra- and extra-cranial arteries. The brain ischaemia explains the prodromal phenomena, and leads to accumulation of vasodilatating substances such as adenosine, ADP, NO etc. At the onset of the aura, the plasma concentration of 5-hydroxytryptamine rises, and it falls during the migraine attack.

After a brief period of aura the vessels dilatate, pulsate forcefully, and the walls become oedematous. These changes are believed to cause the *migraine headache*.

Food containing nitrites and tyramine may precipitate migraine attacks.

Severe attacks of migraine are treated with 5-hydroxytryptamine$_1$ agonists, such as sumatriptan, or with ergotamine tartrat.

Prophylaxis of migraine is carried out with *5-hydroxytryptamine antagonists* (pizotifen, methysergide) and b *-adrenergic blockers*.

Psychogenic headache varies in severity and location. This headache is generally accentuated by conflicts or by anxiety with excessive sweating, tachycardia and hyperreflexia. This type of headache can be the first sign of depression, if the condition is worse in the morning following sleep disturbances.

Meningitis headache is accompanied by contraction of the neck muscles (stiff neck). The dura and the venous sinuses are inflamed, and the headache is severe.

Pressure headache. Intracranial mass lesions (tumours, abscess, bleeding, and traumata) are usually surrounded by brain oedema, whereby the basal vessels and the meninges are displaced and generate pain. This causes a special pressure headache, which is exaberated by supine rest, bending over, straining, sneezing and coughing. Any elevation of intracranial pressure induces this type of headache. Pressure headache is often accompanied by vomiting.

Subdural haematoma must be suspected after head trauma with pressure headache.

Suddenly occurring headache following a trauma may be caused by subarachnoid haematoma.

When combined with fever, neck stiffness, back stiffness, and vomiting the cause may also be meningitis, where a history of *sore throat* is frequently obtainable.

Fig. 3-14: The cutaneous innervation of the head and types of headache.

2. Extracranial headache is common and released from nociceptors in extracranial vessels, in the muscles of the head and neck or by inflamed mucous membranes of the sinuses, *sinus pain* (Fig. 3-14). The pain is felt directly over the frontal or the maxillary sinuses in the case of sinusitis.

Typical is the *muscle contraction headache* located in the frontal or occipital muscles. The frontal and/or occipital-nuchal muscles are tender. Frequently, both the occipital and the cervico-trapezial muscles are tense and tender with specific pain points (*loci dolendi*). Acupuncture or lazer therapy of these points is often effective. Treatment is difficult when depression or accident sequelae is the underlying cause.

3. Taste and smell disorders

The geniculate ganglion is a sensory ganglion for taste, which lies at the genu of the facial nerve. The nerve fibres join the facial nerve in the chorda tympani and carry taste from the anterior two-thirds of the tongue. Cranial lesions involving the petrous temporal bone cause *loss of taste* in this area together with an unpleasant loud distortion of noise called hyperacusis. Hyperacusis is due to paralysis of the stapedius muscle.

The sensory fibres of the glossopharyngeal nerve carry taste from the posterior third of the tongue. Cranial lesions involving the jugular foramen may damage the glossopharyngeal nerve often together with the vagus and the accessory nerve.

Loss of the ability to smell is called *anosmia*. Head injuries involving the cibriform plate or tumours damaging the sensory pathway may cause anosmia. Damage of the olfactory receptors in the nasal mucosa by upper respiratory infections may lead to anosmia.

Equations

- Stevens proposed the *power law* to account for the non-linearity of most physiological mechanisms. The interpreted stimulus strength (ISS) is equal to a constant (k) multiplied by the actual stimulus strength (SS) raised to the power n:
    - Eq. 3-1: $ISS = k*SS^n$.

In his original version, only the exponent **n** differed for each type of sensation. The equation can be modified by subtracting different constants from SS before raising it to the power **n**, or by changing the value of **k**.

Self-Assessment

Chapter 3. Multiple Choice Questions

I. Each of the following statements has True/False options:

- A. The somatosensory thalamus is a relay station for most sensory modalities.
- B. Glutamate is the main inhibitory transmitter in the CNS, whereas GABA is the dominant excitatory transmitter.

C. Pain and temperature reach the thalamus through the lateral spinothalamic tract.

- D. Presynaptic transmission of opiates inhibits depolarization of nerve cell membranes.
- E. The adequate stimulus of the cutaneous mechanoreceptors is deformation of the receptor.

II. Each of the following statements has False/True options:

- A. The Nissl bodies are stacks of rough endoplasmic reticulum.
- B. Taste, heat and angular acceleration follow transfer functions, so the interpreted stimulus strength decreases with the rise in actual stimulus strength.
- C. Sensory receptor systems are biological transducers with a dynamic range up to 10^{12}.
- D. The B nerve fibres are autonomic preganglionic axons with a diameter less than 3 m m and a conduction velocity of 3-12 m per s.
- E. The CSF has higher concentrations of K^+, glucose, and protein than blood plasma, and lower concentrations of Na^+ and Cl^-.

Case History A

A female of 32 years is admitted to a neurosurgical ward with a discrete lesion in the spinal cord caused by a traffic accident. Her vital functions are unaffected. The most important signs are a complete lack of cutaneous temperature sensibility and pain sensibility in the left leg and the lower left side of the trunk below the umbilicus (the navel).

Where in her spinal cord is the lesion localised?

Try to solve the problems before looking up the answers

Highlights

- *The nervous system is a rapid signal transduction system and the main communication network in our body. The integrative functions allow selection and processing of incoming signals to produce an appropriate response.*
- *The nervous system includes sensory receptors that detect events in the body as well as in the outer world. Several sensory tracts and pathways synapse in the nuclei of the thalamus (the spinothalamic tract). The somatosensory thalamus is a relay station for many sensory modalities.*
- *Neuroglia is supportive cells that sheath and protect neurons. Myelinated axons propagate APs up to 50 times faster than unmyelinated with the same diameter. Neuroglia also eliminates transmitters more rapidly from the synapse. The neuroglia constitutes half of the brain volume, and there are about 10^{12} to 10^{13} glial cells in the human brain.*
- *Sensory receptors in the nervous system are classified as exteroceptors (located on the body surface), proprioceptors (located in muscles, tendons and joint capsules), interoceptors (located in the viscera), and telereceptors (stimulated by events far from the person).*
- *Sensory receptors are either neurons in the case of vision, smell and cutaneous senses, or modified epithelial cells in the case of auditory, vestibular, smell and taste senses.*
- *C fibres (IV) are unmyelinated and lead pain, touch and heat signals from the skin.*
- *The gate-control hypothesis of pain states that pain transmission is suppressed by innocious signals in thick myelinated afferents (group II), whereas the pain sensation is enhanced by signals in thin afferents. Inhibitory interneurons in the dorsal horn of the spinal cord perform the gate-control through a special type of presynaptic inhibition*

called primary afferent depolarization (PAD), and the receptors on the cell body of the secondary neuron is the gate.

- *All taste fibres synapse in the nuclei of the solitary tract and the axons of these neurons project to the thalamus. From the thalamus third-order neurons reach the lower part of the primary sensory cortex in the postcentral gyrus (somatosensory area I).*
- *Dermatome anaesthesia. The mammalian embryo is segmented into so-called somites, which are innervated by an adjacent part of the spinal cord. Every spinal dorsal root is destined to a segment of the skin called a dermatome. Also muscles (myotomes) bone (sclerotomes), and viscera are related to specific segments of the spinal cord or brain stem.*
- *Loss of the ability to smell is called anosmia. Head injuries involving the cibriform plate or tumours damaging the sensory pathway may cause anosmia. Damage of the olfactory receptors in the nasal mucosa by upper respiratory infections may lead to anosmia.*
- *The thalamic syndrome is frequently caused by thrombotic blockade of bloodflow to the somatosensory thalamus. The destruction of thalamic neurons in one hemisphere leads to ataxia and loss of sensations from the opposite side of the body. After a few months some of sensations return, but they are often accompanied by pain.*
- *Meningitis headache is accompanied by contraction of the neck muscles (stiff neck). The dura and the venous sinuses are inflamed, and the headache is severe.*
- *Migraine headache begins with prodromal nausea and vision disturbances, often occurring about one hour prior to the headache. The pain is located on one side of the head in classical cases.*

Further Reading

- Caterina, MJ, MA Schumacher, M Tominaga, TA Rosen, JD Levine, and D Julius. The capsacain receptor: A heat-activated ion channel in the pain pathway. Nature 389: 816-24, 1997.
- Kandel, E.R., J.M. Schwartz and Jessop. "Principles of neural science." New York: *Elsevier Science Publ. Co.*, 1991.
- Russell, M.B. and J. Olesen. "The genetics of migraine without aura and migraine with aura." *Cephalalgia* 13 (4): 245-8, 1993.

Return to content

Return to chapter 3

Chapter 4.

Brain Function, Locomotion And Disorders

Study Objectives

- To *define* akinesia, amnesia, aphasia, arousal, coma, rigidity, a motor unit and three different unit types, habituation and non-associative learning, conditioning, and long-term potentiation.
- To *describe* the primary motor cortex, the corticospinal pathways, damage to the corticospinal pathways, the control of nucleus ruber, the symptoms rigidity and spasticity, nerve conduction velocity, monoaminergic transmission, postural control, neck-and labyrinthine reflexes, the control of voluntary movements, the cerebellar cortex and its pathways, cerebellum and motor learning, damage to the cerebellum, damage to the basal ganglia, cause and therapy of Parkinson´s syndrome, the main functions of the brain lobi and the hippocampus with effects of typical lesions, synaptic plasticity in brain growth and brain damage.
- To *draw* a model of the basal ganglia with pathways, a muscle tendon with efferent and afferent pathways, and a model of recurrent inhibition of motor neurons.
- To *explain* the components in a reflex arch, the muscle tendon, the Golgi tendon organ, the flexor reflex, the crossed extensor reflex, reciprocal innervation and inhibition, alpha-gamma-coactivation, the effect of gamma-efferents on muscle length, the effects of a spinal cross sectional lesion, the orientation reflex, exteroceptive and proprioceptive reflexes, and muscular force. To explain the electromyogram, autonomic movements, damage to the capsula interna, damage to the pyramidal system, the techtospinal pathways, nucleus raphe, locus coeruleus, the EEG during different conditions, sensory and motor aphasia, and hemispheric dominance.

- To *use* the concepts in problem solving and case histories.

Principles

The functional unit of the nervous system is the neurone with its cell body, dendrites and axon, which terminates in a synapse.

Action potentials passing down the axon release chemical neurotransmitters at the synapse.

Definitions

- **Agnosia** refers to lack og ability to recognise and interpret a sensory stimulus. Agnosia is related to a lesion of the sensory cortex.
- **Akinesia** or *hypokinesia* or *bradykinesia* means inability to initiate normal movements. Akinesia is a typically finding in Parkinsons disease.
- **Anterograde amnesia** is a lack of ability to learn anything new. This is a consequence of bilateral removal or damage of the hippocampi.
- **Aphasia** is a condition with disorders of the language function. Lesions of the left hemisphere produce deficits in the language function of most people.
- **Apraxia** refers to lack of ability to perform certain practical actions (unbutton a jacket

etc) – often found with parietal-lobe syndromes. The apraxia of gait (failure of skilled walking) is due to frontal-lobe disease.

- **Arousal** is a high level of consciousness also called alertness.
- **Ataxia** refers to uncoordinated movements in particular found as ataxic gait in cerebellar disorders.
- **Athetosis** refers to slow, serpentine, writhing involitional movements of the hands or of most of the body. Athetosis is seen following neonatal insults (cerebral palsy).
- **Coma** is an unresponsive state of unconsciousness from which the patient cannot be awakened even with the most vigorous stimuli.
- **Chorea** refers to rapid involitional hyperkinesia with jerky movements of the limbs. Chorea is found in Huntingtons disease.
- **Dominant hemisphere** is the hemisphere that *controls the expressed language*. Lesions of the left hemisphere produce deficits in language function of most people. These deficits are called aphasia.
- **Electroencephalogram** (EEG) is a recording of a *rhythmic electrical activity* from the surface of the skull. In humans, the EEG is recorded from a grid of standard leads.
- **Flexor withdrawal reflex** is a nociceptive (pain) reflex involving all muscles of a limb in flexor withdrawal. This is an attempt to protect the limb from further damage. The reflex can activate extensor muscles of the opposite limb (the crossed extension reflex).
- **Habituation** refers to a gradual diminution of the response to a repeated stimulus without behavioural consequences.
- **Hemiballism** refers to violent swinging movements of one side of the body almost simulating the throwing of a ball. Hemiballism is caused by lesion of the contralateral nucleus subthalamicus.
- **Learning** is a change of behavior caused by neural mechanisms affected by experience.
- **Long-term memory** (*long term potentiation*) is a prolonged storage and retrieval of new information.
- **Memory** refers to the neural storage mechanisms for experiences.
- **Non-associative learning** means that the learning is unassociated to the stimuli.
- **Orientation reflex** is a fundamental change of behaviour, where the eyes, head and body are turned toward an alarming external stimulus.
- **Prosophenosia** refers to the inability to recognize faces following extensive damage of both occipital and temporal lobes.
- **RAS** is an abbreviation for a large region of the reticular formation of the brainstem termed the *reticular activating system* (RAS). Stimulation of this system causes arousal and an *arousal reaction* in the electroencephalogram.
- **Retrograde amnesia** refers to a condition, where the patient cannot recall information from the *memory*.
- **Rigidity** is a clinical condition with muscle stiffness caused by a high tonus level in the alpha-motor units of limb muscles. The muscle resistance is increased towards slow, passive movements of the limb and it is equal in opposing groups of muscles. This condition is called *lead-pipe rigidity* and it is found in Parkinson´s disease.
- **Sensitisation** is the opposite of habituation. The increased response upon repetition of a stimulus has important behavioural consequences in order to avoid the threat.
- **Sensory aphasia** refers to damage of the Wernicke area with difficulties in understanding written or spoken language, although single words can be heard.
- **Sopor** or *clouding of consciousness* is a term for reduced wakefulness.
- **Spasticity** is a clinical phenomenon following lesion of one pyramidal tract. Loss of the inhibitory effect of the corticospinal pathway increases the spinal reflex activity of the gamma-loop. The muscle tone is increased towards rapid, passive movements of the limbs resulting in a sudden *clasp-knife effect*. Stroke, spinal cord lesions, neonatal insults (cerebral palsy) or multiple sclerosis causes spasticity.
- **Stupor** is a sleepy state from which the patient can be aroused by vigorous stimuli.
- **Tone.** Skeletal muscle tone is a low level contractile activity in some motor units driven

by reflex arcs from muscle receptors. Normally, the muscles feels relaxed and flaccid during passive movements of the limbs. – Increased muscle tone is called hypertonia. Hypertonia is found as spasticity in cerebral palsy and as rigidity in Parkinsons disease. Low muscle tone is called hypotonia and found in cerebellar disorders.

- **Tremor.** Rest tremor with pill-rolling movements of the fingers is found in Parkinsons disease. Intention tremor or action tremor is characteristic of cerebellar lesions.

Essentials

This paragraph deals with *1. The cortex and the reticular activating system, 2. Higher brain functions, 3. The limbic system and the hippocampus, 4. Spinal organization of the motor control, 5. Descending motor pathways, 6. Motor control by the brain.*

1. The Cortex and reticular activating system

Six cell layers are recognized in typical regions of the cerebral *neocortex* and they are numbered layers I-VI (Fig. 4-1). The neocortex first appears with the mammals, and its structure is phylogenetically younger than the *allocortex*.

All six layers contain glial cells and more or less of the three typical neurons: Stellate cells, pyramidal cells and fusiform cells. The superficial layers receive and process information, whereas the deep layers are the sites of origin of most cortical efferents.

The six neocortical layers are as follows:

I. The molecular layer contains numerous dendrites, axons and axon terminals almost without cell bodies.
II. The external pyramidal layer contains mainly densely packed stellate cells, which are GABAergic (inhibitory) interneurons.
III. The external pyramidal layer contains small pyramidal cells. Pyramidal cells use excitatory amino acids (aspartate, glutamate) as transmitters. Layers I, II, III connect adjacent cortical regions and integrate cortical functions.
IV. The internal granular layer resembles layer II with many stellate cells. Most of the sensory signals project to layer IV.
V. The internal pyramidal layer resembles layer III. The pyramidal cell bodies increase in size inwards.
VI. The multiform layer consists of long spindle-shaped or fusiform cells arranged perpendicular to the cortical surface.
VII. The perpendicular collections of neurons, axons and dendrites in the cortical areas form the so-called cortical columns.

Fig. 4-1: The cerebral neocortex with pyramidal, stellate and fusiform cells.

The pyramidal and fusiform cells of layers V and VI provide the output from the cortex. The pyramidal cells have long axons passing to other cortical regions to the brain stem and to the spinal cord.

Thalamocortical afferents mainly project to layers I, IV, and VI, whereas corticothalamic projections have their origin in pyramidal cells in layers V and VI. These connections form a reverberating thalamocortical system, which excite the cortex and contribute to the patterns of the electroencephalogram (EEG).

A large region of the reticular formation of the brainstem is termed the ascending reticular activating system (RAS), which determines our state of consciousness, by its connection with the thalamocortical system (Fig. 4-2). The RAS transmits facilitatory signals to the thalamus.

The thalamus excites specific regions of the cortex, and the cortex then excites the thalamus in a reverberating circuit with fast acetylcholine and long-lasting neuropeptides as transmitters. Such a positive feedback loop is what wakes us up in the morning. External stimuli and internal factors (inhibitory interneurons) serve to create a balance of different activity levels during the day. RAS maintain the ascending thalamic activity, but also a certain descending activity level in our antigravity muscles and reflexes. An inhibitory region in the medulla can inhibit RAS and thus both its ascending and descending activity.

A high level of consciousness is called arousal or alertness, which is recognized in the EEG as a high frequency-low voltage shift (see below). An orientation reflex, a fundamental change of behaviour following an external stimulus, often accompanies arousal. The eyes, head and body are turned toward the external stimulus.

Impaired consciousness is caused by malfunction of the neurons in the RAS, and the impairment has at least three levels. Sopor or clouding of consciousness is a term for reduced wakefulness, stupor is a sleepy state from which the patient can be aroused by vigorous stimuli, and coma is a unresponsive state of unconsciousness from which the patient cannot be awakened even with the most vigorous stimuli. Brain stem compression at the mesencephalon leads to coma and death.

Fig. 4-2: The RAS and the thalamocortical system.

The electroencephalogram (EEG) is a rhythmic electrical activity recorded from the surface of the skull. In humans, the EEG is recorded from a grid of standard leads (Fig. 4-2). During neurosurgery the electrical activity is recorded from the surface of the cortex as an electrocorticogram. In normal adult persons, the dominating frequencies are 8-13 Hz (a-rhythm) over the parietal and occipital lobes, as long as the subject is awake and relaxed with his eyes closed. With open eyes, the EEG becomes desynchronized with low amplitude (10 mV) and the dominant frequency increases to 50 Hz. The theta- (3-7 Hz) and delta rhythms (0.5-2 Hz) are observed during light and deep sleep, respectively.

A thalamocortical rhythm produces coordinated extracellular currents, when the brain is not exposed to external stimuli (Fig. 4-2). The EEG recording is due to large synaptic potentials by whole groups of mainly pyramidal cells. The EEG pattern is desynchronized by sensory inputs through the thalamus. The level of alertness (in RAS) also modifies the EEG pattern.

Each pyramidal cell - as with each Purkinje cell in the cerebellum - posses an extraordinarily large number of synapses (10^6). The potentials recorded on the surface of the skull are 50-100 mV. The large pyramidal cells form a dipole with one pole directed toward the surface of the cortex, and the other toward the white matter.

When an external stimulus evokes an EEG change, the change is termed a cortical evoked potential. The large numbers of synaptic potentials in the cortical region are summated to form an evoked potential recorded on the skull by an electrode placed over the associated cortical area. However, evoked potentials are small, so measurement requires repeated stimulation and signal averaging. The evoked potentials over the auditory, visual, and somatosensory cortex (areas I and II) are used clinically to assess the integrity of the respective sensory pathway.

Circadian periodicities are changes in biological variables that occur daily. The circadian controller is the so-called biological clock, probably located in the suprachiasmatic nucleus of the hypothalamus. The biological clock receives many projections from sense organs including projections from the retina signaling light and darkness. These signals are transmitted further to the pineal gland according to one hypothesis. Darkness probably stimulates melatonin secretion by the pineal gland, which inhibits the secretion of gonadotropic hormones from the

anterior pituitary, and thus reduces sexual drive. Melatonin secretion decreases with age.

Destruction of the biological clock disrupts many biological rhythms, such as oscillations in body temperature, other vegetative functions and the sleep-wake cycle.

The astronomic 24-hour cycle is shorter than the biological sleep-wake cycle (normally 25 hours). When flying east the astronomic cycle is shortened further acerbating the discrepancy between the two cycles. This increases the problems of adjusting the circadian systems, which often require a week to regain their normal phase relation to the biological clock. Problems caused by changes of biological rhythm are summarized in the term jet lag. Melatonin is used clinically to reduce the jet lag.

The endogenous circadian periodicity of the sleep-wake cycle is normally 25 hours - see above. Sleep is divided into four stages based on EEG. The relaxed individual with eyes closed has 8-13 Hz a-rhythm. As he falls asleep, he passes through the four stages of sleep. During these stages the muscles are relaxed, all vital functions are decreased, and the gastrointestinal motility is increased.

Stage 1 is light sleep, where a-rhythm is interspersed with theta rhythm. Stage 2 is somewhat deeper sleep dominated by slow waves and by sleep spindles (periodic spindle-shaped bursts of a-rhythm) and by large, irregular K-complexes. Stage 3 is characterized by delta waves and by occasional sleep spindles. Stage 4 is recognized by the very slow delta waves with frequencies around 0.5-1 Hz. The subject is difficult to wake up.

A different form of sleep with complete loss of muscle tone occurs periodically every 90-min during stage 1 sleep. This is termed rapid eye movement sleep or REM sleep. Eye movement artifacts and a desynchronized EEG (low voltage, fast activity as in the arousal reaction when awake), is characteristic for REM sleep. The subject is difficult to wake up, so the condition is therefore also termed paradoxical sleep.

Fig. 4-3. Differences in sleeping pattern between three age groups.

Spontaneous erection occurs during REM sleep, and an irregular heart rate and respiration are often observed. Dreams occurring during REM sleep are often recalled by the person when awake.

Children and young adults have all 4 stages of sleep and several periods of REM sleep (Fig. 4-3). The depth of the non-REM sleep diminishes through the night and the REM periods increase in duration (Fig. 4-3).

Stage 4 sleep disappears with age, and stage 3 sleep decreases in duration (Fig. 4-3). The REM sleep is also reduced, and wake periods occur in increasing number. This is why elderly people believe that they do not sleep sufficiently.

The passive theory of sleep claim sleep to be caused by reduced activity in RAS. However, transecting the brainstem in the midpontile region produces an animal that never goes to sleep. Stimulation at the nucleus of the solitary tract can induce sleep, suggesting that sleep be an active process related to centres below the midpontile level.

The question is difficult to address. An educated guess is that sleep is an active, energy saving condition, preferable to most animals. The metabolic rate during sleep falls to 75% of the basal metabolic rate.

2. Higher brain functions

Each hemisphere consists of the following four lobes: the frontal - occipital - parietal and temporal lobes.

The frontal lobe, located in front of sulcus centralis (central fissure), is involved in motor behaviour. The frontal lobe contains the primary motor (area 4), the premotor (area 6), and the supplementary motor areas (frontal eye areas 8 and 9 of Fig. 4-4). These cortex areas are responsible for planning and execution of voluntary movements.

The motor speech areas (44 and 45 or Broca's area) are located close to the motor cortex, on the inferior frontal gyrus of the dominant hemisphere in humans (the left hemisphere is controlling the expressed language in most people). Lesions here cause motor aphasia (difficulties with speech and writing). Patients with lesion of Broca's area (in the dominant hemisphere) frequently suffer from paralysis of the opposite side (right) of the body.

Fig. 4-4. The human cerebral cortex of the left hemisphere controls the expressed language.

The frontal cortex is also involved in personality and emotional behaviour - including attention, intellectual and social behaviour.

The occipital lobe is located behind the parietal and temporal lobe, and involved in visual processing and visual perception. Adjustments for near vision are controlled by the primary visual cortex in area 17 and in the cortex around the calcarine fissure occipital lobe. The conscious visual perception takes place in the primary visual cortex. The secondary visual cortex is in area 18 and 19, where visual impressions are compared, interpreted and stored (Fig. 4-4).

The important primary somatosensory area I is located on the postcentral gyrus (area 1,2 and 3 in Fig. 4-4). There is a distinct spatial representation of the different areas of the body in the postcentral gyrus (the sensory homunculus). The secondary somatosensory area II is located in the rostral part of area 40, close to the postcentral gyrus (Fig. 4-2). The somatic association or interpretation areas (areas 5 and 7) are located in the parietal cortex just behind the somatosensory area I (Fig. 4-2).

Each side of the cortex receives information exclusively from the opposite side of the body.

Widespread damage to the somatosensory area I causes loss of sensory judgement including the shapes of objects (astereognosis).

Auditory and vestibular signals are processed and perceived by the superior temporal gyrus (area 41 in Fig. 4-2). Area 42 is the secondary auditory centre, where auditive signals are interpreted and stored.

The medial temporal gyrus helps control emotional behavior in the limbic system and all the functions of the autonomic nervous system.

Signals from the auditive (area 42), visual (areas 18 and 19) and somatic (areas 7 and 40) interpretative areas are integrated in the posterior part of the superior temporal gyrus. This large gnostic area is specially developed in the dominant hemisphere, where it is called the general interpretative or language comprehension area (Wernicke's area). Damage in Wernicke's area causes sensory aphasia (i.e., difficulties in understanding written or spoken language, although single words can be heard).

Learning processes

Learning is a change of behavior caused by neural mechanisms affected by experience. Memory refers to neural storage mechanisms for experiences. The hippocampus is involved in learning and memory.

1. Non-associative learning means that the learning is unassociated to the stimuli.

Habituation refers to a gradual diminution of a response by repetition of a stimulus, because experience show that the stimulus is unimportant. Sensitization is the opposite of habituation. Firstly, a strong threatening stimulus triggers a certain response, but repetitions of the stimulus increase the size of the response in order to avoid the threat. This evaluation is called the reward and punishment hypothesis. The neural processes are probably related to the function of the hippocampus.

In the snail aplysia a facilitating interneuron releases serotonin onto the presynaptic terminal of a neuron. This stimulates adenylcyclase and the formation of intracellular cAMP in the presynaptic terminal. The resulting protein kinase activation causes phosphorylation and blockage of K^+-outflux. The K^+-outflux is necessary for recovery from the action potential. Lack of K^+-outflux prolongs the presynaptic action potential considerably. This causes a prolonged Ca^{2+}-influx into the presynaptic terminal with increased release of neurotransmitter and facilitated synaptic transmission.

2. Associative learning is the process of learning by associations between stimuli. The free radical nitric oxide (NO) modulates learning.

Conditioning refers to a neural process of associative learning, where there is a temporal association (optimum 0.5 s) between a neutral stimulus (eg, a sound before food) and an unconditioned stimulus (food) that elicits a response (gastro-intestinal secretion). Repetition of the sound-food manoeuvre develops into a conditioned reflex, where the sound alone elicits salivary secretion.

In operant conditioning the response is associated with reinforcement, which changes the probability of the response. Positive and negative reinforcement increases the probability of the response, whereas punishment reduces its probability. Learning is highly improved by happiness. Light stress is an advantage in learning something new. However, substantial stress is not helpfull in the recall process, and stress can completely block the memory.

Strategic behavior is the basis for our social life. Strategic or motivated behavior is related to homeostasis in general (defence, reproduction, temperature and appetite control). Previously, strategic behavior was explained by negative feedback with the purpose as a fixpoint, and with the human brain playing a minor role. Today it is generally accepted that the cerebral drive is a dominant determinant for strategic or motivated behavior. The drive that arouses individuals from inactivity originates in the limbic system (including the hypothalamus), that is acting in close relation to the thalamus and the cerebral cortex. The limbic system is connected to the autonomic control functions of the brainstem reticular formation by the medial forebrain bundle. These vital functions are thermocontrol, appetite control and sexual behavior (Chapter 6).

The dominant hemisphere is the hemisphere that controls the expressed language. Lesions of the left hemisphere produce deficits in language function of most people. These deficits are called aphasia. The left planum temporale in the floor of the lateral fissure of Sylvii is larger than that of the right hemisphere in most people - not only right-handed. The right hemisphere is dominant for functions related to language (intonation, body language), and to mathematically related functions. Each hemisphere controls the contralateral side of the body.

Information between the two hemispheres is transferred through the anterior commissure and the corpus callosum. The language centres on the left hemisphere cannot influence the right hemisphere unless the corpus callosum is intact. The two hemispheres can operate relatively independently with language. One hemisphere can express itself through spoken language. The other communicates non-verbally.

If an animal with intact corpus callosum and optic chiasm learns a visual discrimination task with one eye closed, the task can still be performed with the untrained eye alone, even when the optic chiasm is transected before the animal is trained. Therefore, visual information is transferred as long as the corpus callosum is intact.

Surgical transection of the corpus callosum has been performed to prevent epilepsy from spreading. When such a patient fixate his vision on a point on a screen, it is possible to stimulate only one hemisphere by showing an object to one side of the visual field. Similar objects (key, ring, nail, fork etc) can be manipulated (but not seen) through an opening below the screen. Healthy persons can locate the correct object with either hand. Split-brain patients, with the picture of the object transferred to the right hemisphere, can locate the correct object with the left hand (ie, right hemisphere), not with their otherwise preferred right hand.

Jigsaw puzzles are solved with such manipulo-spatial capabilities. Right-handed patients with split brain can solve three-dimensional puzzles, if the visual signals can reach the motor cortex for the hand to explore. The visual and motor cortex are connected to each other only in the same hemisphere, when the corpus callosum is cut.

Memory research has characterized three temporal stages in human memory processes.

- 1. An immediate memory holds sensory information for a few hundred milliseconds to seconds for analysis and further processing. The immediate memory is erased by new incoming signals, so we can only remember a few new telephone numbers at a time. Accumulation of Ca^{2+} in the presynaptic terminals with each signal possibly causes prolonged release of neurotransmitter at the synapse (synaptic potentiation).

- 2. The short-term memory is covering seconds to a few minutes, and the short-term memory receives selected information from the immediate memory. Information is erased as new items displace old data. If a person sees a rapid succession of slides, it is the last slide that remains in the short-term memory. We store recent events in the short-term memory, by a neural activity with improved synaptic efficacy that lasts for seconds to minutes. The improved synaptic efficacy is possibly due to synaptic potentiation, presynaptic facilitation, or impulses circulating in neuronal circuits for a restricted period.

- 3. The long-term memory is a large and permanent memory. The long-term memory receives information from the immediate and the short-term memory. Recycling of information through the short-term memory is termed rehearsal. The likelihood of a successful storage in the long-term memory increases with the number of cycles. When the long-term memory is searched for a certain information, it may take minutes to recall the memory. The long-term memory is subdivided into the intermediate long-term memory, which lasts for days or weeks and can be disrupted, and the long lasting long-term memory, which lasts for years.

- 4 The long lasting long-term memory is the storage in the brain of highly overlearned information as one's own name and address. This memory is difficult to disrupt, and it is seldomly affected in retrograde amnesia (see below).

The long-term memory and consolidation of memory relate to effector protein synthesis at the synapses. Electron microscopy suggests an increased number of vesicular release sites in the presynaptic terminals.

Retrograde amnesia is a term used for a condition where the patient cannot recall information from the immediate and short-term memory. The mild form of retrograde amnesia is typical following head lesion with loss of consciousness (cerebral commotion). The short-term memories have only been rehearsed a few times and probably stored only discretely.

The long term-memories are widespread in the cortex as structurally maintained modifications of the synapses after many rehearsals. Only in severe cases is the long-term memory involved.

3. The limbic system, the hippocampus and emotions

The limbic system is the neuronal network that controls emotional and motivational behavior. Motivational behavior include control of vegetative functions such as body temperature, respiration, circulation, osmolality of body fluids, sexual behavior, smell, thirst, appetite and body weight.

Hypothalamus constitutes the major part of the limbic system, and is located in the middle of the other limbic elements.

Fig. 4-5. The limbic system. The corpus callosum is transected, and we are looking at the medial aspects of the right hemisphere.

The limbic cortex begins in the frontal lobe as the orbitofrontal cortex, extends upward as the subcallosal gyrus, over the corpus callosum and into the cingulate gyrus (Fig. 4-5). The limbic cortex finally passes caudal to the corpus callosum down towards the hippocampus, para-hippocampal gyrus and uncus at the medial surface of the temporal lobe (Fig. 4-5). The fornix connects the hippocampus to the mamillary body. The mamillothalamic tract connects the mamillary body to the anterior nucleus of the thalamus. Thalamus connects to the cingulate gyrus, and its cortex is associated with the hippocampus. Stria terminalis connects the amygdaloid body to the midbrain septum and to the mamillary body (Fig. 4-5).

The limbic paleocortex links the subcortical limbic structures to the neocortex. Hereby, the limbic system relates behavior and emotions to the intellectual cortex functions.

Another important pathway is the medial forbrain bundle, which connects the limbic system to the autonomic control functions of the brainstem reticular formation.

The hippocampus connects with the cerebral cortex, the midbrain septum, the hypothalamus, the amygdaloid and the mamillary bodies and acts both as a store and a recall centre (Fig. 4-5). The hippocampus is the decision-maker, determining the importance of incoming signals. Hippocampus becomes habituated to indifferent signals, but learns from signals that cause either reward (pleasure) or punishment. Hippocampus is the "brain librarian" (helps the cortex to store new signals into the long lasting long-term memory). The signal molecule, nitric oxide (NO), modulates aspartate responses related to hippocampal long-term potentiation.

Bilateral removal of the hippocampi in epileptic patients permanently disrupts the ability to learn anything new (anterograde amnesia). Other lesions of the hippocampi reduce previously learned memory material (retrograde amnesia - see above). Long-term alterations imply a rise in the number of synapses. Cholinergic synapses in the midbrain septum are essential to our memory, and these neurons are dependent upon the nervous growth factor. Repeated activation of a sensory pathway increases the reaction of pyramidal cells. Such a reaction may last for weeks in the hippocampus and be involved in storage and retrieval of new information in the

long-term memory.

Our memory (cortex and hippocampus) works as a filter. Perhaps only 1 per mille of all received signals contain useful or emotional information and are catched in the memory. Unfortunately, we are unreliable witnesses, because we invent emotional "information" concerning a factual experience. The easiest facts to remember are those that make sense. All facts, concepts and acquired skills are stored in a ready-to-use fashion. Feelings play a large role in memory, and strong impressions that are charged with emotion etch themselves into our memory.

A recollection is split up into numerous subunits in different regions of the brain. Later, all subunits are brought together by the hippocampus into a complete memory (eg, a certain smell act as a strong clue to a clear memory from way back). One individuals recollection of a particular incident can trigger off anothers, whereby new associations can be created.

4. Spinal organization of motor control

Motor activity can be voluntary or involuntary. Voluntary movements are planned and started by feedforward control, and when maintained for a while they are regulated by feedback loops. Involuntary movements comprise reflexes, such as the stretch reflexes, and autonomic functions, such as the respiratory muscle movements. We have motor centres in the cerebral cortex, the brainstem, the spinal cord, the cerebellum, and the basal ganglia. Motor centres all receive sensory information in an organized neural structure termed a somatotopic map (see the motor homunculus).

We have 200 different skeletal muscles, which are controlled by more than 300 000 motor units.

A motor unit is comprised of a a-motor neuron, all its axon terminals, and the skeletal muscle fibres it innervates. The number of muscle fibres in a motor unit varies from 2 in highly regulated eye muscles (entirely red fibres) to 2000 in the quadriceps femoris muscle. The motor unit is the final common pathway, because all muscle fibres of the unit contract when a motor unit is activated. Adjacent motor units interdigitate, so they can support each other. The muscle power is increased by recruitment of more motor units and by increased frequency of discharge in each unit.

We have three types of motor units (a-motor neurons) in a mixed muscle such as the gastrocnemius. The three types of motor units are characterized in Chapter 2, Box 2-2.

The myotatic stretch reflex

A spinal reflex is a stereotyped motor reaction to an input signal. The myotatic stretch reflex is the most crucial monosynaptic reflex for the maintenance of the erect body posture in humans.

Fig. 4-6: The phasic myotatic stretch reflex and reciprocal innervation (F-).

The reflex has two components. Firstly, the primary annulospiral endings (group Ia) of the muscle spindles trigger the phasic stretch reflex. Secondly, both primary and secondary endings elicit the tonic stretch reflex.

- 1. The phasic stretch reflex is elicited in the clinic by a light tap on a muscle tendon. When the patellar tendon from the quadriceps muscle is stretched quickly by the tap, a discharge is elicited in the afferent fibres (Ia) from the primary endings of the muscle spindle (Fig. 4-6). This is the phasic myotatic stretch reflex or the so-called patellar reflex. These Ia fibres synapse directly (monosynaptically) on a-motor neurons that

supply the extensor muscles of the knee (E+ in Fig. 4-6). The response elicited is a brief contraction of the latter. Of all the presynaptic terminals arriving to the motor neuron up to 90% are located on the surface of the dendrites. The remaining 10% synapse on the soma of the motor neuron.

- The Ia afferent fibres also synapse with small group Ia inhibitory interneurons in the grey matter of the spinal cord, as the one synapsing with the upper a-motor neuron in Fig. 4-6. This neuron innervates the semitendinosus muscle, which flexes the knee joint (F- in Fig. 4-6). The reflex inhibition of antagonist muscles when synergistic muscles are contracted is called reciprocal innervation. In pathologic conditions, the phasic stretch reflexes may be depressed or hyperirritable.

- 2. Passive bending of a joint triggers the tonic stretch reflex. This elicits a discharge in both groups Ia and II afferents from the muscle spindle. The tonic stretch reflex contributes to the erect body posture and helps maintain posture by increasing the tone of the physiologic extensor muscles (ie, antigravity muscles).

Renshaw inhibition and presynaptic inhibition

Renshaw inhibition. Cajal found that the a-motor axons give off thin recurrent (antidromal) collaterals in the grey matter of the spinal cord (Fig. 4-6). These collaterals synapse with Renshaw interneurons in the ventral horn (Fig. 4-6). The Renshaw cells synapse with a-motor neurons of synergistic muscles, and thus inhibit monosynaptic reflexes (postsynaptic inhibition). Stimulation of each a-motor unit inhibits adjacent motor units (ie, recurrent inhibition). This is also called the principle of lateral inhibition, whereby the motor response is confined to selected units only.

Descending signals from the brain can either amplify the postsynaptic inhibition or reduce its effect. Renshaw cells make it possible for the higher brain centres to influence spinal reflexes by central inhibition or facilitation.

Presynaptic inhibition. Presynaptic terminals contain a large number of voltage-gated Ca^{2+}-channels. Ca^{2+} must enter the presynaptic terminal from the extracellular space before the vesicles can release their neurotransmitter at the synapse. Presynaptic inhibition takes place at presynaptic contact sites on the presynaptic terminals. Activation of these sites closes many Ca^{2+}-channels, and thus inhibits transmitter release.

The Golgi tendon organ

The Golgi tendon organs are the serially located terminals of group Ib fibres wrapped around bundles of collagen fibres in the tendons. Golgi tendon organs monitor the force in the tendon; they are activated either by stretch or by contraction of the muscle. The adequate stimulus is the force developed in the tendon.

The inverse stretch reflex or the Golgi tendon reflex completes the stretch reflex by a force-controlling feedback. The Golgi tendon organs monitor force in the tendons. Golgi tendon organs are in series with the muscle fibres - not parallel as the muscle spindles. If the extensor muscles of the thigh are fatigued, as during standing, the force in their tendons begins to decrease. This reduces the discharge of the Golgi tendon organs. This acts as a compensating feedback, which excites the a-motor neurons and increases the force of contraction. The inverse stretch reflex helps maintain the force of muscular contraction and posture during standing. During the rapid contraction of the myotatic stretch reflex, the inverse stretch reflex reduces the force of contraction. The stretch reflexes regulate the length of the muscle, and

provide a length-force feedback to the CNS.

The muscle spindle

The muscle spindle monitors muscle length and rate of change of length (velocity); they are particularly abundant in muscles that are capable of fine movements and in large muscles that are dominated by slow twitch fibres. The organ is shaped like a spindle, which lies in parallel to the large, regular, extrafusal muscle fibres. Each organ contains two main types of intrafusal muscle fibres: Nuclear bag fibres which swell in the equatorial region due to all the nuclei located here, and thin nuclear chain fibres which have central nuclei arranged in line (Fig. 4-7). The primary afferent fibres (Ia) twine around the equatorial regions of both the bag and chain fibres like a corkscrew or annulospiral; the annulospiral nerve endings signal length and velocity. The secondary afferent fibres originate mainly from the nuclear chain fibres and with a few branches originating from the nuclear bag fibres (Fig. 4-7). They monitor only the length of the muscle.

Two types of g-motor neurons innervate the muscle spindle. The dynamic g-motor axons form plate endings (P_2) on the nuclear bag fibres, while static g-motor axons form creeping trail endings on nuclear chain fibres (Fig. 4-7). The intrafusal fibres receive a Ab-motor fibre, which terminates with P_1 plate endings on both extra- and intrafusal muscle fibres (Fig. 4-7). The Ab-motor fibres may be involved in a-g-coactivation.

When the extrafusal fibres contract, the muscle spindles shorten, whereby the discharge rate of their afferents decreases.

Fig. 4-7: The structure of a muscle spindle with a bag and a chain fibre.

Activity of the g-motor neurons causes the polar spindle regions to contract on either end. This elongates the equatorial regions so that muscle spindles can adjust to stretch (Fig. 4-7).

Descending commands from the brain often cause contraction of both extrafusal and intrafusal fibres simultaneously so that the muscle spindle is sensitive to stretch at all muscle lengths. When the muscle is stretched, the muscle spindles are simultaneously stretched with it, and the discharge rate of the afferents is increased.

The flexion reflexes

The flexion reflexes are triggered by various flexion reflex afferents including nociceptors. The flexion reflexes have a long latency, because it involves polysynaptic interneurons. The afferent discharge causes excitatory interneurons to activate a-motor neurons that innervate ipsilateral flexor muscles. The afferent discharge also causes inhibitory interneurons to inhibit a-motor neurons, supplying the ipsilateral extensor antagonists.

The flexor withdrawal reflex is crucial. This reflex is also called a nociceptive reflex or a pain reflex, and involves all the muscles of a limb in flexor withdrawal in order to protect from further damage. In addition, the reflex can activate the extensor muscles of the opposite limb. This contralateral activity is termed the crossed extension reflex by reciprocal innervation.

The locomotor pattern generator controls flexion reflexes involved in locomotion.

Severe visceral disease can trigger contraction of the chest and abdominal muscles, which reduces pain by limiting movement of the body. When examining the abdomen of such a patient it will be observed that the muscles are tense. This sign is called defence musculaire,

which is a viscero-somatic protective reflex.

Coordination of limb movements

We possess pattern generators or neural circuits in the spinal cord, for every limb and for respiration, chewing etc. The midbrain locomotor centre, via the reticular formation and through the reticulospinal tracts, organizes the commands. Such spinal pattern generators also account for other movement patterns like scratching, dancing etc.

5. Descending motor pathways

Clinical dichotomy traditionally subdivides the descending fibres into the pyramidal and the extrapyramidal pathways; this is based on the fact that the corticospinal tract passes through the medullary pyramids. Therefore, interruption of the corticospinal or pyramidal tract was supposed to cause pyramidal tract disease (see later). The problem, however, is that the loss of the corticospinal tract does not explain all the classical signs of pyramidal tract disease.

The concept of extrapyramidal pathways raises other problems. The concept of extrapyramidal tract diseases is generally used to designate one or more disorders of the basal ganglia. While, extrapyramidal pathways do play a role in basal ganglia diseases (as in cerebellar disease), the main motor pathway involved in basal ganglia diseases is the corticospinal tract!

The descending motor pathways can also be dichotomized based on their endpoint in the spinal cord, and hence which muscles they control and how. Pathways ending in the lateral horn of the spinal cord (on motor neurons or interneurons) are called the lateral descending motor system (the rubrospinal tract and the lateral corticospinal tract). Pathways ending on the medial ventral horn interneurons are termed the medial descending motor system (containing reticulo-, tecto-, and ventriculo-spinal tracts).

The lateral corticospinal, the corticobulbar (to the facial motor and hypoglossal nucleus) and the rubrospinal tracts control the manipulative movements of the limbs and the lower face and tongue muscles. The corticospinal and corticobulbar tracts originate from areas 4, 6, 8, 9, and somatosensory area I (areas 1, 2, 3 in Fig. 4-4). The large and small pyramidal cells and the giant pyramidal cells of Betz are the cells of origin of these tracts. The corticospinal tract descends through the internal capsule and brainstem. At the medullary pyramid 80% of the fibres cross to the opposite side and descend in the dorsal lateral funiculus as the lateral corticospinal tract. The fibres of this tract end on motor neurons and interneurons in the lateral horn of the spinal cord. These motor neurons innervate distal muscle groups. Interruption of the lateral corticospinal tract implies loss of the fine control of the digits. Interruption of the corticobulbar tract to the facial motor and hypoglossal nucleus implies loss of voluntary movements of the lower face and tongue. Interruption of the rubrospinal tract from the red nucleus combined with corticospinal lesions give rise to difficulty in separating finger, hand and arm movements. The red nucleus is closely linked to the deep cerebellar nuclei.

The lateral or dorsolateral descending system allows the primary motor cortex to modify the reflexes and pattern movements at the level of the spinal cord.

The medial or ventromedial descending system involves the ventral corticospinal tract and much of the corticobulbar tract ending in the medial group of brainstem and spinal cord interneurons. The ventral corticospinal tract continues caudally in the ventral funiculus on the same side and ends bilaterally on the medial interneurons. They control the axial muscles and bilateral activity including chewing and wrinkling of the eyebrows.

Other medial system pathways originate in the brainstem:

- 1. The lateral vestibulospinal tract excites motor neurons that innervate proximal postural muscles. It receives input from all compartments of the vestibular apparatus and from cerebellum to the lateral vestibular nucleus.

- 2. The medial vestibular tract receives signals from the semicircular ducts and from cerebellum, and excites motor neurons in cervical and thoracic segments. Thus, it controls the head position in response to angular accelerations of the head.

- 3. The pontine reticulospinal tract excites motor neurons to the proximal extensor muscles to support posture.

- 4. The medullary reticulospinal tracts have mainly inhibitory effects on many spinal reflexes.

- 5. The tectospinal tract from the superior colliculus causes contralateral movements of the head in response to touch and auditory stimuli. This tract allows the integration of hearing and vision with motor performance.

- 6. Pathways from the solitary nucleus and the interstitial nucleus of Cajal are involved in the pharyngeal stage of swallowing. The solitary nucleus receives all sensory signals from the mouth including taste, and is involved in cardiovascular and respiratory control.

The ventromedial system is important for the normal muscle tone and body posture.

Monoaminergic descending pathways

- 1. The neurons of the pontine locus coeruleus and nucleus subcoeruleus contain nor-adrenaline (NA). These nuclei project to and inhibit interneurons and motor neurons of the spinal cord through the lateral funiculi.

- 2. The neurons of the raphe nuclei in the medulla, which are connected to the limbic system also, contain serotonin. The serotonergic nuclei project to and inhibit dorsal horn interneurons reducing pain transmission, and they also project to and excite ventral horn motor neurons of the spinal cord, thereby enhancing motor activity.

- 3. There is also a descending dopamine pathway.

The three monoaminergic pathways function as motor system amplifiers.

6. Motor control by the brain

The primary motor cortex (area 4) on the precentral gyrus controls distal muscles of the extremities. Area 4 is organized parallel to the somatosensory cortex. The face is represented laterally near the Sylvian lateral fissure, and the legs on the medial part of the hemisphere. The cortical representation is somatotopic and disharmonic, as indicated by the motor homunculus.

The premotor cortex helps control proximal and axial muscles.

The supplementary motor cortex is involved in motor planning and in coordination of movements. The frontal eye fields initiate saccadic eye movements.

Corticospinal neurons discharge before voluntary muscle contraction, and the size of the discharge is related to the size of the contractile force. The somatosensory cortex and the posterior parietal association cortex receive feedback from the sensory neurons system, which

helps correct motor feed-forward commands.

The role of the cerebellum

The little brain, also termed the motor autopilot, helps regulate movements and posture, influences muscle tone, eye movements and balance.

Cerebellum is particularly concerned about the timing of rapid muscular activities including the interplay between agonist and antagonist muscle groups. Motor learning is programmed in the cerebellum. Cerebellum compares the proprioceptive input from the actual movements, with the movements intended by the motor control areas of the brain. Cerebellum controls the sequence of movements, and makes corrective adjustments just like an autopilot.

The cerebellar cortex is characteristically folded and consists of three phylogenetically different structures related to three afferent pathways (inputs). The large neocerebellum in higher mammals is also called the pontocerebellum and consists of the hemispheres and vermis caudal to the primary fissure. The paleocerebellum or spinocerebellum consists of vermis of the anterior lobe, pyramis, uvula and paraflocculus. The small archicerebellum or vestibulocerebellum is simply the flocculonodular lobe.

Three important outputs from the cerebellum also divide it into three functional units. The vermis of the cerebellar cortex projects to the fastigial nucleus, the pars intermedia to the globose and emboliform nuclei, and finally the hemisphere, which projects to the large dentate nucleus (Fig. 4-8).

Fig. 4-8: Neuronal connections between the cerebellar cortex and the deep cerebellar nuclei.

The cerebellar cortex is build up of three layers. The superficial molecular layer with axons, dendrites and many synapses, the Purkinje-cell layer and the granular layer (Fig. 4-8). The small granule cells send their axons into the molecular layer, where they divide and send so-called *parallel fibres* in each direction along the folium. These fibres excite the dendrites of the Purkinje and the Golgi cells. The Golgi cells inhibit the granule cells by feedback inhibition. Stellate and basket cells are interneurons that inhibit dendrites and cell bodies of the Purkinje-cells, respectively. Each Purkinje cell is stimulated from a climbing fibre, which projects from the inferior olive. All neurons with cell bodies in the cerebellar cortex are inhibitory except for the granule cells. The cerebellar cortex modulates the activity of the deep cerebellar nuclei.

The incoming pathways to the cerebellum end as *mossy fibres* on the granule cells. Each mossy fibre reach many granule cells. The input signals through the mossy fibres evoke *simple spikes* (single action potentials) in Purkinje-cells. The climbing fibres produce repetitive or complex discharges in Purkinje cells. *Complex spikes* of long duration and low frequency are involved in the cerebellar programming of motor learning. The Purkinje-cell axons terminate in the deep cerebellar nuclei or in the lateral vestibular nucleus.

This is the basis for *cerebellar coordination* and fine, rapid adjustments of complex movements. The cerebellar hemisphere affects movements on the same side of the body, because of its crossed connection to the motor system. The motor system projects contralaterally.

Discrete electrical stimulation of cerebellum does not cause movements or sense impressions, so it is also termed the *silent brain*.

The vestibulocerebellum projects to the vestibulospinal and reticulospinal tracts, which

coordinate balance and eye movements. The vestibulo-ocular reflex produces conjugate eye movements in the direction opposite to that of the head movement. The vestibulo-collic reflex increases the neck muscle tone damping the induced movement.

The spinocerebellum receives proprioceptive input from the spinal cord (the spinocerebellar tracts). The spinocerebellum controls the axial muscles through the medial descending motor system, and the proximal limb muscles through the rubrospinal tract of the dorsolateral system.

The pontocerebellum receives decision signals and motor control signals from the cerebral cortex by way of pontine nuclei. The pontocerebellum is involved in motor planning, and controls the distal limb muscles through the lateral corticospinal tract.

The basal ganglia

The main function of the basal ganglia is to initiate and stop movements. The basal ganglia inhibit the thalamus, and thus reduce the thalamic stimulation of the motor *Cortex*.

Fig. 4-9: The basal ganglia and their interplay. Transmitter stimulation is marked by +, and inhibition by -. The affected cell bodies or axons at disease states are marked with a bar.

The basal ganglia also contribute to cognitive (i.e., intelligence, knowledge, and motor learning) and affective (i.e., emotional) functions.

The basal ganglia include the globus pallidus and striatum. Striatum consists of the nucleus caudatus and the putamen. These deep brain nuclei function in collaboration with several thalamic nuclei, substantia nigra and the subthalamic nucleus (Fig. 4-9).

The striatum receives afferent fibres from the cortex (Glutamate + = glutaminergic excitatory fibres), and dopaminergic (inhibitory) fibres from substantia nigra (Dopamine -). Striatum projects to the globus pallidus and to the substantia nigra. These connections are GABAergic and inhibitory (GABA - in Fig. 4-9). Globus pallidus receives afferent GABAergic fibres from striatum, and projects to the thalamus with GABAergic efferents. In the striatum, there are excitatory cholinergic pathways.

Pathophysiology

This paragraph deals with 1.Pure lesion of the medullary pyramid, 2.Abnormal muscle tone, 3. Spinal transection syndrome, and 4. Cerebellar disease. -

Capsular stroke, Parkinson´s disease, dyskinesias and epilepsy are all dealt with in Chapter 7, which is a systematic description of neurological and psychiatric disorders. Read chapter 7 before trying to solve the case histories.

1. Pure lesion of the medullary pyramid

The control of fractionated finger movements is absent. There is a positive sign of Babinski. Flexion reflexes are not found, and neither is spasticity. On the contrary, muscle tone is decreased. In summary, a pure interruption of the corticospinal tract alone does not show the same signs as capsular stroke.

The main deficits caused by medial lesions are reduced muscle tone in the physiologic extensors, loss of balance during walking and standing, and loss of rightening reflexes (they tend to restore head and body position). However, fine finger movements are quite normal.

2. Abnormal muscle tone

Spasticity is used in clinical neurology to describe muscles resisting fast, passive movements of the limbs, especially in extreme articular positions. When the limbs are moved in extreme articular positions, the increased muscle resistance suddenly disappears. Spasticity includes hyperactive stress reflexes and foot clonus. The resistance dominates in the physiological extensors (antigravity muscles). Spasticity is typical for stroke, where the capsula interna is damaged. The resulting disruption of the lateral descending system is extended by damage of other cortical efferents to the basal ganglia, the thalamus and pons (see Chapter 7).

Rigidity is muscle stiffness caused by prolonged activity in the motor units. The muscle resistance is increased towards passive movements of the limbs in any direction (lead pipe rigidity). This condition is found in Parkinson's disease (see Chapter 7).

3. Spinal transection syndrome

The spinal shock is immediately recognized by several characteristic symptoms: flaccid paralysis with loss of stretch reflexes, areflexia, loss of autonomic functions, and of all sensation below the level of transection. After a few weeks the spinal shock fades away and the reflexes return and become hyperactive (foot clonus), including mass reflexes and flexion reflexes. A spastic paralysis or paresis replaces the flaccid paralysis.

4. Cerebellar disease

Cerebellum can suffer from damages at two locations: 1. Damage to the flocculonodular lobe causes nystagmus and difficulties in gait and balance (i.e., resembling lesion of the vestibular apparatus). 2. Damage to the vermis or the intermediate region and hemisphere, results in motor disturbances of the trunk and limbs, respectively.

Cerebellar disorders include cerebellar incoordination, dysequilibrium, and loss of muscle tone.

Cerebellar incoordination comprises ataxic gait, as seen in alcohol intoxication and in disseminated sclerosis. Another type of ataxia is dysmetria, where there is an inability to move the limbs to the desired position. Many patients manifest their ataxia as dysdiadochokinesis, which is a disturbance of the normal ability to make repeated supinations and pronations of the lower arms. Complicated muscle function is stepwise - not smooth. Intention tremor is seen when the patient is asked to touch a target. Speech is slow and slurred, a defect termed dysarthria or scanning speech.

Dysequilibrium results in balance problems, and the patient falls to the affected side. Gyratoric vertigo is a genuine rotational or merry-go-round vertigo with the associated loss of equilibrium. This cerebellar vertigo is similar to that following lesion of the vestibular apparatus.

Loss of muscle tone is called hypotonia. The hypotonic lack of damping causes the leg to swing back and forth, when the patellar reflex is triggered - so-called pendular knee jerk.

Cerebellar nystagmus is involuntary movements of the eyeballs around their natural position - often accompanied by rotational vertigo, when the flocculonodular lobe is damaged.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. The frontal cortex is involved in motor and emotional behavior.

B. The somatic association or interpretation areas (area 5 and 7) are located in the temporal cortex.

C. Recycling of information through the primary memory is termed rehearsal.

D. Retrograde amnesia following brain commotion is a loss of the short-term memory.

E. The limbic system relates behavior and emotions to the intellectual cortex functions.

II. Each of the following five statements have False/True options:

A. The EEG arousal reaction is a low frequency-high voltage shift.

B. Circadian periodicities are changes in biological variables occurring once a day.

C. N-methyl-D-aspartate-(NMDA)-receptors bind aspartate, dopamine and glutamate.

D. Dreams occur during REM sleep, and the person always reproduces them when awake.

E. Dominating EEG frequencies of 8-25 Hz are characteristic of light sleep.

III. Each of the following five statements have True/False options:

A. Fast fatigable motor units consist of type IIB twitch fibres with few mitochondria and small amounts of myoglobin.

B. The Renshaw cells synapse with a-motor neurons of antagonistic muscles, and thus inhibit monosynaptic reflexes.

C. The cerebellar hemisphere affects movements on the opposite side of the body.

D. The Golgi tendon organs are the serially located terminals of group Ib fibres wrapped around bundles of collagen fibres in the tendons.

E. The ventromedial descending system involves the ventral corticospinal tract and much of the corticobulbar tract ending in the medial group of brainstem and spinal cord interneurons.

4. Case History A

An outstanding Russian composer, 63 years of age, recovered from a cerebral insult. However, he could no longer understand spoken or written language, although his speech was fluent. The composer also maintained his ability to compose excellent music.

- *1. What is the name of this deficit in language function?*
- *2. Where in the brain is the lesion* ***localized and in what side of the brain?***

4. Case History B

A male of 65 years suddenly falls and is found in deep coma by the doctor. There is a left-sided hemiplegia with short arm-long leg as a flexion reflex. The paralysis and areflexia turns into spastic hemiparesis with a positive sign of Babinski. The deep stretch reflexes (patellar- and Achilles-tendon reflexes) are enhanced. There is loss of superficial reflexes (the abdominal and cremasteric reflexes). When the Achilles-tendon reflex is triggered it releases foot clonus. When the patient is awake from coma his facial nerve paresis is examined. He can knit his brows and turn his eyes upwards.

- *1. What is the pathophysiologic basis for this condition?*
- *2. What are spasticity and foot clonus?*
- *3. Is the facial nerve paresis central or peripheral?*

Try to solve the problems before looking up the answers .

Highlights

- *The reticular activating system (RAS) transmits facilitatory signals to the thalamus. The thalamus excites the cortex, and the cortex then excites the thalamus in a reverberating circuit. Such a positive feedback loop is what wakes us up in the morning. During the day external stimuli and internal factors including inhibitory interneurons balance the different activity levels.*
- *Impaired consciousness is caused by malfunction of the neurons in the RAS, and the impairment has at least three levels. Sopor or clouding of consciousness is a term for reduced wakefulness, stupor is a sleepy state from which the patient can be aroused by vigorous stimuli, and coma is a unresponsive state of unconsciousness from which the patient cannot be awakened even with the most vigorous stimuli.*
- *Circadian periodicities are changes in biological variables that occur daily. The circadian controller is the so-called biological clock, probably located in the suprachiasmatic nucleus of the hypothalamus. The biological clock receives many projections from sense organs including projections from the retina signaling light and darkness.*
- *Children and young adults have all 4 stages of sleep and several periods of REM sleep each night. The depth of the non-REM sleep diminishes through the night and the REM periods increase in duration.*
- *The motor speech areas (44 and 45 or Broca's area) are located close to the motor cortex, on the inferior frontal gyrus of the dominant hemisphere in humans (the left hemisphere is controlling the expressed language in most people). Lesions here cause motor aphasia (difficulties with speech and writing). Patients with lesion of Broca's area (in the dominant hemisphere) frequently suffer from paralysis of the opposite side (right) of the body.*
- *The medial temporal gyrus helps control emotional behaviour in the limbic system and all the functions of the autonomic nervous system.*
- *The hippocampus is involved in learning and long lasting long-term memory. This is what makes hippocampus the decision-maker.*
- *Motor centres all receive sensory information in an organized neural structure termed a somatotopic map (motor homunculus).*
- *The motor unit is the final common pathway, because all muscle fibres of the unit contract, when a motor unit is activated. Adjacent motor units interdigitate, so they can support each other. The muscle power is increased by recruitment of more motor units and by increased signal frequency in each unit.*

- *Renshaw cells make it possible for the higher brain centres to inhibit or facilitate spinal reflexes.*
- *Cerebellum or little brain is also termed the motor autopilot, because it helps regulate movements and posture, and influences muscle tone, eye movements and balance.*
- *Cerebellar disorders include cerebellar incoordination, dysequilibrium, and loss of muscle tone.*
- *The main function of the basal ganglia is to initiate and stop movements. Disorders of the basal ganglia, such as lack of dopamine in substantia nigra, result in a clinical syndrome with rigidity, hand tremor, and akinesia (Parkinson´s disease).*

Further Reading

Schuman, E.M., and D.V. Madison. "Nitric oxide and synaptic function." *Annu. Rev. Neurosci.* 17: 153-183, 1994.

Thomson, R.F. "The brain. A Neuroscience Primer." *Second Edition*. Freeman, 1993.

McIntosh, A.R., C.L. Grady, L.G. Ungerleider, J.V. Haxby, Rapoport, S.I., and B. Horwitz. "Network analysis of cortical visual pathways mapped with position emission tomography." *J. Neurosci.* 14 (2): 655-666, 1994.

Return to top

Return to content

Chapter 5.
Special Senses And Disorders

Study Objectives

- To *define* adequate stimuli for vision, hearing and balance. To define astigmatism, cataract, glaucoma, hypermetropia, myopia, presbyopia, colour blindness, hemianopsia, scotoma, strabismus, and visual acuity and agnosia. To define nerve deafness, conduction deafness, presbyacusis, nystagmus, and transportation sickness.
- To *describe* the anatomy of the eye, including the retina and sensory pathways and cortical visual organisation. To describe the inner ear including hair cells, and the vestibular system. To *describe* the eye movements, receptive fields, the colour vision, the sound transfer, the mechanical-electrical transduction in hair cells.
- To *calculate* the correction of refractive disorders, and the hearing loss from relevant variables given.
- To *explain* the function of photoreceptors, dark adaptation, the pupillary reflex, the sensitivity to sounds, the travelling wave and the frequency theory. To explain the diagnosis and pathophysiology of the following disorders: astigmatism, hypermetropia, myopia, presbyopia, colour blindness, hemianopsia, visual agnosia, strabismus, nerve deafness, conduction deafness, presbyacusis, nystagmus, and transportation sickness.
- To *use* the concepts in problem solving and case histories.

Principles

- *The human eye as light perceives electromagnetic radiation at wavelengths between 400 and 725 nm. Electromagnetic radiation ("waves") do not require a medium for propagation. This spectrum of wavelengths is seen in a rainbow. Light reflected from a star traverses the empty space. Electromagnetic radiation in any narrow band within this spectrum is termed monochromatic light.*
- *The camera obscura is a simple model of the eye. The camera obscura is a device in which a small aperture creates a reversed image on the receptive surface. The aperture can be extended if a convex lens is mounted in front of it. The image then produced on the receptive surface (retina) is reversed and reduced (see textbooks in physics).*
- *A mechanical wave is a wave that propagates by disturbing the particles of a medium. Sound waves are such mechanical or matter waves. The laws of quantum mechanics determine their behaviour.*

Definitions

- **Anopsia** is a visual field defect reaching the periphery of the field (see glaucoma).
- **Astigmatism** is a refractive disorder of the eye in which the curvatures of the cornea or lens are different along different meridians. The different meridians therefore have different focal distances.
- **Cataract** is an eye disease, where the vision is blurred by an opaque lens.
- **Colour blindness** is a group of recessive inherited sex-linked (X chromosome) disorders. Colour blindness is inherited from the father - with the daughter as a carrier - to her son.
- **Conduction deafness** is caused by impairment of the mechanical conduction of sound into the cochlea.

Far-point (F or *punctum remotum*) for the eye is the fixation point in the un-accommodated eye.

- **Glaucoma** is a term used for eye disorders with loss of optic nerve fibres and of the visual field. Frequently, the cause is an increased intraocular pressure (above 25 mmHg) due to reduced fluid outflow at the irido-corneal junction.
- **Hypermetropia** or *far-sightedness* is a refractive disorder with insufficient refractive power, whereby the far point is always located behind the eye.
- **Hemianopsia** means loss of vision in half of the visual field of both eyes. The loss of vision refers to the visual filed, and thus to the contralateral half of each retina.
- **Myopia** or *near-sighted* is a refractive disorder, where the patient can only foveate diverging light waves (both from the near- and the far point). The patients usually have elongated eyeballs.
- **Near-point** (N or *punctum proximum*) of the eye is the fixation point for the maximally accommodated eye (ie when the lens is in its most spherical configuration).
- **Nerve deafness** is caused by damage of the cochlea, the auditory nerve or nucleus. There is hearing loss by both air and bone conduction.
- **Nystagmus** is a disorder with abnormal involuntary movements of the eyeballs.
- **Presbyacusis** refers to the decline with age in the capacity of hearing high tones.
- **Presbyopia** is called old mans sight, because the lens loses its elasticity and hence its ability to assume a spherical shape, so the patient cannot accommodate for near vision.
- **Receptive field** is an area of the visual field from which light is perceived through a certain ganglion cell in the retina.
- **Scotoma** or *localised blindness* is an island-formed visual defect caused by a lesion of the retina in one eye, or by partial interruption of the optic nerve (see glaucoma). The visual field defect may be absolute or relative.
- **Strabismus** *(squint)* or *cross-eyed* is an eye disease, where the visual axes of the two eyes do not converge on the fixation point of the object simultaneously.
- **Transportation sickness** or *kinetosis* is a disorder with vertigo, nausea and vomiting due to rapid changes in the direction of motion.
- **Visual accommodation** is the rise in refractive power of the lens, obtained as the lens rounds up because of contraction of the ciliary muscle and relaxation of zonule fibres.
- **Visual acuity** is the resolution capacity of the eye. Cones have a high-resolution capacity and hence a high visual acuity, because the light is focused on the fovea, where the cones are concentrated.
- **Visual agnosia** or mental blindness, is lack of the ability to combine the seen object into a concept.

Essentials

This paragraph deals with *1. The visual system, and 2. The auditory and vestibular system.*

1. The Visual System

This system detects, transmits and interprets photic stimuli. Photic stimuli are electromagnetic waves with wavelengths between 400 and 725 nm. This is *visible light* or the adequate (effective) stimulus for the eye.

The eyes can distinguish brightness and colour. The photoreceptors are rods and cones located in a specialised epithelium called the retina. In each eye the *retina* contains about 6 million cones and 120 million rods. In the peripheral region of the retina both rods and cones converge on *bipolar cells*. The bipolar cells converge on ganglion cells giving rise to the one million nerve fibres in each optic nerve. In addition, there are horizontal cells and amacrine cells in the retina. They conduct impulses laterally.

Rods and cones

Rods are most sensitive in the dark (*scotopic vision*). More than hundred rods converge on each ganglion cell. There are no rods at all in the fovea.

Cones operate best in light (*photopic vision*). Cones have a *high-resolution capacity* and hence a *high visual acuity*, because the light is focused on the fovea, where the cones are concentrated. The high resolution is also due to the small *convergence* of cones to bipolar cells in the fovea (approximately a 1:1 relationship). Cones are responsible for colour vision. Cones are surrounded by pigment, except where the light enters.

The eye contains chamber fluid, which is produced by filtration and secretion in the ciliary processes. The intraocular pressure is normally 1.3-2.6 kPa (10-20 mmHg). Increased resistance to fluid outflow at the iridocorneal junction leads to *increased intraocular pressure* with loss of optic nerve fibres- or *glaucoma*. In this condition, the retinal artery is compressed at the optic disc, where it enters the eye. This causes retinal and optic nerve atrophy which eventually results in blindness.

A *diopter* is the unit for the refractive power of a lens. The *diopter* (D) equals the reciprocal value of the focal length of the lens in metre (m).

Visual accommodation is the *rise in the refractive power of the lens*, obtained as the lens rounds up, because of contraction of the ciliary muscle and relaxation of the zonule fibres.

Each object we look at has a special target point (the *fixation point*), from which light passes un-refracted through the nodal point of the eye and focuses on the fovea, creating the sharpest possible image. The nodal point in the eye is precisely the point through which a light beam passes un-refracted. The *far point* (F or punctum remotum) for the eye is the fixation point in the un-accomodated eye (Fig. 5-1). The *near point* (N or punctum proximum) of the eye is the fixation point for a maximally accommodated eye (ie, when the lens is in its most spherical configuration). The refractive power of the lens can vary between 12 and 26 D. The *accommodative power of the eye* is the rise in refractive power from the un-accomodated to the maximally accommodated condition (see Eq. 5-1). A child of 10 years has 12-14 D, a 20 year old person 10 D, and a 60 year old person only 1 D in accommodative power.

The optical distance convention defines all distances measured from a light source to the eye, to be positive. Thus, all distances from the eye to the light source are negative. Hence, the distance from the *nodal point* of the eye to a point in front of the eye is negative. Convex refractive media bend (convergence) in-falling light behind the media and thus have a positive diopter. Concave lenses have refractive powers with negative Diopters, because the focal point is in front of the lens.

Convergence or near vision occurs when the eye focuses on an object closer than 6 m from the eye. Near vision - even with only one eye - triggers *accommodation* and *pupillary constriction*. The ciliary muscle and the pupillary sphincter muscle are innervated by the parasympathetic oculomotor nerve, and the two muscles contract simultaneously for near vision.

The visual fields of both eyes are perceived as only one continuous visual fields (the Cyclops eye effect). This is *fusion* or the illusion that we are looking at the world with only one eye.

In a healthy eye, the light from an object in the visual scenario is focused sharply on the retina by the cornea and the lens. Both of these refract (bend) light. The cornea has a refractive power of *43 D*, and the healthy lens has a refractive power that varies between *12-26 D*. Thus the total refractive power is *56-69 D*. The lens allows the eye to accommodate, so that both near and distant objects can be focused on the retina and thus clearly seen. When we look at distant objects with normal eyes and relaxed ciliary muscles, the object foveates automatically. However, when we look at nearby objects, the light is initially focused behind the retina. The lens then rounds up, by contraction of the ciliary muscles and relaxation of the zonule fibres (i.e., accommodation), to focus the image on the fovea.

The normotropic eye has the ideal refractive power. *Parallel light* from the far point (F in the upper part of Fig. 5-1) *foveates* on the retina in the un-accomodated eye. Light from the near

point (N in Fig. 5-1) in the totally accommodated eye also foveates.

Fig. 5-1: Hypothetical light rays for emmetropic, myopic, facultative and absolute hypermetropic eyes.

The coloured space in front of each eye is the fraction of the three dimensional space, which can be focused on the retina for a given visual axis (Fig. 5-1).

Eye movements

Conjugate movements are movements of both eyes in the same direction and magnitude, so that the relation between the visual axes is maintained. When focusing on far away objects, the parallel axes are maintained during conjugate movements. Likewise, conjugate eye movements maintain the convergence angles of the eye required for focusing on nearby objects.

Saccadian or *jumping movements* are rapid eye movements. Saccadian eye movement is an instantaneous reposition of the eye that occurs when reading or when focusing on a flash of light in the peripheral visual field. The velocity of the movement is up to 500^{o} per s. The latency period is 250 ms, and the contraction time is 50 ms. The compensatory eye movement involving the vestibular system, occurs when the head rotates. This is also an example of Saccadian eye movement.

In contrast, *pursuit movements* are smooth eye movements that allow the eye to track a moving object. They have a velocity of up to 30^{o} per s.

These two movements work together in *optokinetic nystagmus*. This is a shift between smooth pursuit movements and correcting jumps. The direction of nystagmus is by convention indicated by the rapid correcting phase.

Even during foveation of an object the eyes are not totally still. The eyes are continuously performing *miniature eye movements*, which occur at a rate of 3 microsaccades per s, with mean amplitude of 0.1^{o}.

Photoreception

The number of photoreceptors in a human eye is estimated to be 110-130 million rods and 5-7 million cones.

Each *photoreceptor cell* includes an outer and an inner segment, which are united by a thin cilium. The outer segments are directed towards the pigment epithelium of the peripheral retina, and contain stacks of *disks* that are rich in photo-pigment molecules. The inner segments contain the cell nucleus and numerous mitochondria. The *rods* are predominant outside the fovea, and they contain much more pigment (10^{8} rhodopsin or molecules per rod) than do cones. Rods are so sensitive that a single photon can trigger a rod response. Rods are therefore well suited for night vision. Rhodopsin or *visual purple* has two absorption maxims: 350 and 500 nm. The spectral extinction curve for rods corresponds to that of rhodopsin, suggesting that rhodopsin is the chemopigment in rods. Rhodopsin consists of a glycoprotein (opsin) and a chromophore group (11-cis-retinal). Retinal is the aldehyde of vitamin A_1 (retinol).

The fovea only contains *cones*. Cones function in the daytime with maximal visual acuity and colour vision. The human eye possesses *three* types of cones, each with a specific pigment related to the three basic colours: red (erythrolab), green (chlorolab) and blue (cyanolab). The cones in the fovea do not contain cyanolab.

When the human eye is fully adapted to darkness, its rods have open Na^{+}-channels, and the resulting influx of Na^{+} maintains depolarised rods with a resting membrane potential of *-40 mV*. The rod cell synapses with bipolar and horizontal cells, and releases *glutamate* as long as the dark depolarisation is maintained. Na^{+} is continuously removed from the rod by the Na^{+}-K^{+} pump.

Inside the rod a special amplification takes place. Light absorption by a *single rhodopsin molecule* activates thousands of G-protein molecules (transducin), which then activate large

quantities of cGMP phosphodiesterase in the discs. Each of these enzyme molecules catalyses the hydrolysis of cGMP to 5'-GMP at a rate of thousands per second. The reduction in [cGMP] closes the Na^+-channels, and hyperpolarises the cell. The amplification mechanism is probably why the eye is capable of detecting a single photon.

A similar cascade of reactions takes place in cones, when they are stimulated. Cones are so small that the hyperpolarization occurs rapidly.

Each ganglion cell has a receptive field in the retina that is comprised of a number of photoreceptors.

The fraction of a receptive field belonging to each photoreceptor is added to neighbour areas in order to obtain the receptive area of a bipolar or a horizontal cell. An on-bipolar cell is depolarised by white light, whereas an off-bipolar cell is hyperpolarised. Signals are transmitted from the photoreceptors to the ganglion cells as a graded response. These small receptive areas are summated to form a circular receptive field for each ganglion cell (Fig. 5-2). Ganglion cells can generate action potentials and transmit signals to the brain.

1. One type of ganglion cell has a centrally located excitatory area, surrounded by an inhibitory annular area (Fig. 5-2). Together these form an on-centre off-surround receptive field. Here, an on-response is triggered in the bipolar cell that is connected to the on-ganglion cell.
2. Another type of ganglion cell has a centrally located inhibitory area (inhibited by light), surrounded by an excitatory annular area (Fig. 5-2). These form an off-centre on-surrounding receptive field.
3. A third type of ganglion cell is connected to both on- and off-bipolar cells, so its centre is both stimulated and inhibited by white light.

Fig. 5-2: Ganglion cell receptor fields in the retina.

The ganglion cells can also produce *transient* or *sustained* reactions. These reactions are due to adaptation to light (decreased sensitivity with exposure) and lack of adaptation, respectively (Fig. 5-2).

Ganglion cells in the fovea are connected to few or only one cone. Some ganglion cells are excited by blue light and inhibited by its opponent colour yellow. Other cells are excited by green and inhibited by the opponent colour red. This mechanism is the so-called *colour contrast analysis* of the retinal ganglion cells. Colour opponent neurons are found not only in the ganglion cells but also in the lateral geniculate nuclei.

Retinal signals pass through the main visual pathway: the optic nerve, the lateral geniculate nucleus, the optic radiations (the geniculostriate tracts), the primary visual cortex, the pretectal area, the Edinger-Westphal nucleus, the oculomotor nerve and the ciliary muscle.

Each point of the retina has a corresponding location in the dorsal lateral geniculate nucleus and in the visual, striate cortex (area 17). The nerve fibres in the optic nerve run so the upper quadrants of the retina are represented in the upper half of the nerve, and the lower quadrants in the lower half.

Such a *retinotopic map* is present in the *lateral geniculate nucleus* and maintained throughout the visual pathways and in the visual cortex. The receptive field in the retina is maintained all the way to the cortex. This is the basis of fusion. The consequence is that the *right striate cortex* receives information about objects located in the left side of the visual field, and the striate cortex in the left hemisphere receives information about the right side of the field of vision. In general, each hemisphere of the brain is connected to sensory and motor activity of the opposite side.

The lateral geniculate nucleus has three different pairs of neuronal layers (1-2, 3-4, 5-6). Ganglion cells from the ipsilateral (same side) eye projects to layers 2, 4 and 6, whereas ganglion cells from the contralateral eye projects to layers 1, 3, 5. The lateral geniculate

nucleus is involved in integration and registration of pictures formed in corresponding areas of the retinal surfaces. Some neurons react to white light (with circular receptor fields), while other neurons react to opponent colours. When we jump from one highlight to another in the visual field, each jump is called a saccade. Selection of visual stimuli may be located in the lateral geniculate neurons (possibly performing gate control).

Most of the neurons in the geniculate nucleus projects to the striate cortex by way of the optic radiations (geniculostriate tract). Neurons in a certain column of the lateral geniculate nucleus project to precisely the same part of the striate cortex (area 17). The lateral geniculate nucleus also receives information from the cortex (in particular the visual cortex) that is essential for selection of signals of particular interest.

The striate cortex (area striata, area 17) is located around the calcarine fissure on the medial side of each occipital lobe. The optic radiation ends mainly in synaptic contact with simple cells in layer 4 of the striate cortex. Simple cells have on- and off-fields. Complex cells receive inputs from several simple cells, and hypercomplex cells receive inputs from several complex cells.

Axons from one eye terminate in millions of functional units called ocular dominance columns consisting of about 10^3 neurons. Cortical neurons are arranged in orientation columns showing orientation selectivity for lines edges or bars. Other cortical neurons are arranged in direction columns showing direction selectivity. Colour blobs are interspersed among the other columns (see later).

A large area at the occipital pole represents the macula, and the upper and lower half of the visual field is represented below and above the calcarine fissure. The upper layers of the superior colliculus perform visual processing. The deep layers produce eye movements.

The cortical area V4 contains colour-sensitive neurons, and the visual association areas 18 & 19 (Fig. 4-4) contain many cells with complex functions.

The absolute sensitivity depends upon the adaptive condition of the retina, the pupillary diameter, and the source of light (spectral composition, exposure time, and light source dimensions). The threshold for the completely dark-adapted eye is (7×10^{-11}) Watts/m^2.

The Trichromatic Theory

Light adaptation is a decrease in visual sensitivity during constant stimulation. This occurs rapidly because the rhodopsin bleaches readily. Hence, in daylight (photopic cone vision) we are dependent on cones for vision. Night vision (scotopic rod vision) is extremely sensitive to light, because of dark adaptation. It takes at least 20 min in dark surroundings before the rods become fully adapted. In a dark movie theatre, we have scotopic vision with low visual acuity and colour blindness. As soon as the film is projected we experience partial light adaptation, so that the photopic cone vision is resumed.

The trichromacy theory postulates that an appropriate mixture of the three basic colours can produce any colour: red, green and blue. The three types of cone pigments have different opsins, and opsins that differ from that in rhodopsin. Groups of cortical neurons called cortical *colour blobs* respond specifically to colour signals, and also receive signals from adjacent columns of the visual cortex. Cortical *colour blobs* are probably the primary stations for perception of colour, and they are found both in the primary and the secondary visual cortex areas. Perception of spectral opponent colour pairs is located in discrete colour blobs of the visual cortex.

The three cone pigments are Erythrolab for red (maximal sensitivity at 555 nm), chlorolab for green (525 nm), and cyanolab for blue (450 nm). The absorption spectra of the photopigments overlap considerably. The three cone types are uniformly distributed in the retina, except in the fovea. *Fovea has no cyanolab cones* and no rods. This gives the fovea partial physiologic scotoma (ie, no blue vision and no scotopic dark vision). The real *physiologic scotoma* is the *dark spot* corresponding to the *optic papilla*. Inhibition of neighbour ganglion cells from on centre field ganglion cells is called lateral inhibition; it occurs also in the lateral geniculate

nucleus or in the visual cortex. Lateral inhibition provides simultaneous contrasts and enhancement. Each colour-contrast neuron is excited by one colour and inhibited by the opponent colour. Opponent colours are red-green, yellow-blue, and green-purple.

Contrast analysis begins already in the retina and is elaborated centrally in the lateral geniculate nucleus, the thalamus and the visual cortex. If there is a multilevel neural system for the analysis of colour mixing, we also need to assume the existence of a neural system for colour brightness, depending upon the intensity of the light. Healthy people are trichromats, because they have all three cone pigments.

Spatial resolution or minimum separabile is the capacity of the eye to see two stimulated retinal areas as separated. In healthy young humans the spatial resolution is about 1/60 degree, depending upon luminosity, exposure time, patterns and opponent colours in the visual scenario. The most important factor limiting this capacity is the cerebral integration.

Temporal, visual resolution is the capacity of the eye to see consecutive light stimuli as separate. Intensity is directly related to duration of perception of light. Contrast further decreases temporal resolution, a flash of light in the dark is perceived for longer than in bright surroundings. Temporal resolution is also determined by the wavelength of light. The eye is maximally sensitive at the absorption maxims of the three cone pigments and rhodopsin.

The positive after-picture is a visual impression lasting longer than the stimulus. It is visible on a dark background following exposure of the eye to intense light. The negative after-picture follows the positive afterpicture as a dark shadow or as the opponent colour. The negative after-picture is due to adaptation of the area in the retina related to the picture.

A flickering source of light liberates successive flashes so rapidly that they fuse, and appear to be continuous. In the darkness of a movie theatre we do not sense the flickering frequency of 24-48 frames each s, or those of a television screen with 50-60 frames per s. With increasing intensity of illumination the critical fusion frequency increases abruptly. This is why young persons can look directly into a neon light and see its flickering character even with 60-100 flashes each s. Accordingly, the cones of the healthy human eye have a critical fusion frequency around 60-100 flashes per s with optimal illumination. The photopic cones are much more sensitive to rapid alterations of light intensity than the rods.

Movements in the visual scenery are depicted as opposite movements on the retina. Convergent inputs from the eyes result in depth perception (ie, stereopsis or stereoscopic vision). Stereopsis depends upon the medial, longitudinal fasciculus and the corpus callosum. These structures co-ordinate the movements of the two eyes. The two eyes are 7-8 cm apart, which causes slight disparities between their retinal images. Disparate receptive fields and thus excitation of specific cells in the secondary visual cortex probably exhibit the perception of depth.

Distance evaluation requires high visual acuity and experience with objects of known size.

Essential for the development of the baby’s brain is human milk proteins and long chain fatty acids in the mother milk. Protein deficiency from birth reduces formation of brain neurons and thus limits brain development including the development of visual capacity. Many vitamins and key proteins have hormonal and transmitter function in the brain, and lack of such substances in the critical growth period just after birth, results in irreversible damage. The action of endogenous nerve growth factor is necessary for the normal functional and anatomical development of the visual system. In the critical period of visual development, which is the first two years of life, the child must be exposed to a multitude of visual stimuli. This is necessary for the development of neurons and key substances that can record future visual stimuli. The ability to fuse the two optic fields is a process that has to be practised. This fact is an important basis for the treatment of cross-eyedness (strabismus).

Cross-eyedness or squint (strabismus) is an eye disease, where the visual axes of the two eyes do not converge on the fixation point of the object simultaneously. Thus the retinal images do not fuse on corresponding areas on the two retinas. Since the fixation line only foveate in one eye, the patient can learn to suppress the other picture in the brain. Hereby, double vision is

avoided at the expense of visual acuity.

2. The Auditory And The Vestibular System

The two systems share the labyrinth, and transmit signals to the brain through the 8th cranial nerve. The two systems record fluid movements and use the so-called hair cells as mechanical transducers.

Sounds are sense impressions that consist of complex mixtures of compression and decompression waves that can be broken down to pure tones by Fourier analysis. Pure tones are sinusoidal waves of a specific frequency (cycles per s or Herz = Hz) and amplitude.

Sinusoidal waves can change phases. The normal human ear is sensitive to pure tones with frequencies between 10 and 30 000 Hz, in a young person.

As people age, their capacity to hear high tones declines. This condition is termed presbyacusis.

Sound propagates at 343 m/s in air at 20^{o}C, although each single air molecule only moves a few mm in the direction of propagation. The unit of sound pressure (p) is Pascal (Pa).

According to international convention the sound pressure level (SPL) is expressed in decibel (dB) - see Eq. 5-2.

Any rise in the SPL of 10 dB implies a rise in sound pressure by a factor of 3, since the log of 3 is 0.5: 10 dB = 20 log 3 (Eq. 5-2).

Speech has an intensity of 60-65 dB, and sounds that exceed 100 dB can damage the ear. A constant sound stimulation only results in minor adaptation. The human ear has the largest sensitivity around 1000-4000 Hz, the range for normal speech.

The sound pressure waves in air are converted into sound pressure waves in the fluid column within the cochlea. The pressure wave in the air is transmitted via the tympanic membrane and the ossicles (malleus, incus and stapes), to the fluid of the cochlea. The foot plates of the stapes inserts in the oval window, and separates the middle ears from the fluid of the cochlea. The ratio of the effective surface area of the tympanic membrane to that of the oval window is 14:1, and the pressure is increased further by the differing lengths of the lever arms in the chain of ossicles. By this area-pressure amplification, hearing is improved by more than 25 dB.

When the external ear is filled with water during diving, hearing is seriously reduced.

Two muscles are found in the middle ear. They dampen movements of the ossicular chain when the ear is exposed to extremely high pitch sounds that can be anticipated. These muscles are the tensor tympani muscle supplied by the trigeminal nerve, and the stapedius muscle supplied by the facial nerve. Exposure to sounds above 90 dB elicits reflex contractions.

The cochlea is composed of three tube systems coiled together to form a pyramid: scala vestibuli, scala media and scala tympani (Fig. 5-3). The part of cochlea beneath the oval window is called scala vestibuli, and it is filled with a fluid column termed perilymph.

Fig. 5-3: A cross section through one of the turns of the cochlea.

The perilymph conducts the pressure wave to the basilar membrane, which is displaced within the endolymph together with the whole organ of Corti, which contains the hair cells.

Each hair cell has 40-100 hairs (*stereocilia*). The hairs have different heights, and when the pressure wave displaces the hairs towards the tallest hair, the hair cells are depolarised. When the basilar membrane moves upward towards the scala media, the reticular lamina shifts upward and inward (Fig. 5-3), causing the hair cells to depolarise. Downward movement of the basilar membrane towards the scala tympani moves the reticular lamina downward and outward (Fig. 5-3). This movement hyperpolarises the hair cell membrane.

The endolymph in the scala media has a potential difference of +80 mV with the perilymph as reference. The inside of the hair cell is -60 mV compared to the perilymph; this is a resting membrane potential about the same size as in most neurons. Thus the total potential difference between the inside of the hair cell and the endolymph in the scala media is -140 mV. This

resting membrane potential is maintained by Na -K -pumps in the Stria vascularis (Fig. 5-3).

Bending of the hair change the conductance of K^+-ions through the apical hairy membrane, and this is how the resting membrane potential is changed. A current flow is produced through the hair cell from apex to base, which is resting on the basilar membrane (Fig. 5-3). This current flow or receptor potential can be recorded extracellularly with microelectrodes as the cochlear microphone potential (ie, the sum of receptor potentials from many hair cells). This potential has the same frequency as the acoustic stimulus, and the potential is analogous to the output voltage of a microphone. The cochlear microphone potential follows the sound stimulus without latency, without measurable threshold, and without fatigue in contrast to neuronal action potentials.

Stimulated the hair cells release neurotransmitters (glutamate, aspartate) that excite the cochlear nerve fibres. Thus, the propagating action potentials are generated in the cochlear nerve fibres.

A high frequency tone produces travelling waves along the basilar membrane. High tones travel only a short distance from the stapes along the basilar membrane to their resonant point, where the displacement amplitude of the basilar membrane is maximal (Fig. 5-4). Low frequency tones travel all the way to the apex of the cochlea (Fig. 5-4). The higher the tone frequency, the more basal located in the cochlea is the resonant point and its potential.

Fig. 5-4: Displacement of the basilar membrane illustrates the travelling wave theory (von Bekesy).

The existence of such a maximum of the travelling wave is termed *frequency dispersion.* Since different frequencies excite differently located hair cells the argument is called the place analysis theory. The brain also utilises the temporal structure of the sound stimulus. This is the so-called periodicity analysis.

The receptor potentials generate action potentials in the cochlear nerve (8^{th} cranial nerve) that travel to the cochlear nuclei. Secondary neurons transfer the signals from here to the superior olivary nuclei that co-ordinates the two ears, or directly to the inferior colliculus through the lateral lemniscus (representing both ears). Axons from the inferior colliculus ascend to the medial geniculate nucleus of the thalamus. Axons from this thalamic nucleus form the auditory radiation, which terminates in the auditory cortex in the superior temporal gyrus (areas 41 and 42 in Fig. 4-4). High frequencies are projected to the rostral auditory cortex, and low tones to the caudal section.

The duration of a sound stimulus is encoded in the duration of the neural signal, and its intensity by the level of neural activity.

Projections from the auditory cortex also descend to the medial geniculate nucleus and the inferior colliculus. The oligocochlear bundle controls several sound impressions. Efferent stimulation through these pathways inhibits the sensitivity of these nuclei for sounds, while increasing their tone selectivity. This phenomenon, and a high degree of motivation, explains how a mother can hear her baby cry in spite of noise, and also how we can hear an individual in a crowd (the cocktail party effect).

Localisation of a sound source depends upon the difference in time between the arrival of a low frequency sound signal to the left and right ears (time delay). The sources of low frequency sounds (below 2000 Hz) are localised by this time delay. The source of high frequency sounds is localised by the difference in sound amplitude arriving at each ear caused by the dampening of the sound intensity.

Sounds in the region of 2000 Hz cannot be detected by either mechanism. On average, the distance between the two organ's of Corti is about 0.16 m. Then, a sinusoidal wave or pure tone with exactly the same wavelength coming from one side of the head, is in phase when they reach the ears. This wavelength corresponds to the frequency of 2144 Hz (343 m/s divided by 0.16 m). In this instance the subject will be unable to determine the source of the sound.

The vestibular system detects if the body is in balance. The sensory unit of the auditory-vestibular system is the membranous labyrinth, located in the petrous portion of the bony labyrinth. The membranous labyrinth contains endolymph and is surrounded by perilymph; it is composed of the auditory cochlear duct or scala media, and the balance regulating the Vestibular system. The vestibular system consists of three semicircular ducts and two otolith chambers (the utricle and the saccule). Each semicircular duct has a swelling termed an ampulla (Fig. 5-5).

Fig. 5-5: The spatial orientation of the three semicircular ducts in the upright person (left). The horizontal duct is not drawn. The membranous labyrinth is shown to the right.

The semicircular ducts consist of a horizontal duct, a superior and a posterior duct at right angles to each other, so that they cover all three planes in space. The semicircular ducts all communicate with the utricle. The utricle joins the saccule, which receives new endolymph from the cochlear duct.

The sensory organ of each utricle and saccule is called a macula. The sensory organ of each semicircular ampulla is the crista ampularis.

Each macula contains thousands of hair cells. Vestibular hair cells each have many stereocilia (hairs) on their apical surface just as cochlear hair cells do; however, they also have a large stereocilium called kinocilium. The hairs are imbedded in a gelatinous substance, the otolithic membrane that also contains earstones or otoliths. These otoliths increase the specific gravity of the otolithic membrane to twice that of the endolymph. Thus their hair cells are sensitive to linear acceleration such as gravity and to static equilibrium control, but not to angular accelerations of the head. The macula of the utricle is located in the horizontal plane, and the macula of the saccule in the vertical plane.

Each crista ampularis consists of many hair cells. Here the hairs are imbedded in a large gelatinous substance termed a cupula. The cupula occludes the lumen of the ampulla completely, and its material has the same specific gravity as the endolymph. The cupula is concerned with equilibrium control during motion and with angular acceleration (rotation of the head), but is unaffected by linear acceleration.

When the stereocilia are bent toward the kinocilium, the conductance of the apical cell membrane increases for positive ions, and the hair cell becomes depolarised. Bending the stereocilia in the opposite direction hyperpolarizes the cell. The depolarised hair cell releases glutamate or aspartate and increases the discharge rate of the nerve fibre with which it synapses.

The utricles and saccules are sensitive to linear accelerations. When we suddenly thrust our body forward, the otolithic membranes fall backwards on the cilia of the hair cells until the thrust stops. Then, the otolithic membranes fall forwards. The signals to the brain make us feel as if we were falling backwards. Therefore, we lean forward until the otolithic membranes are in balance.

Pathophysiology

This paragraph deals with 1. Refractive disorders, 2. Colour blindness, 3. Visual field defects, 4. Mental blindness, 5. Deafness (hypacusis), 6. Nystagmus, and 7. Kinetosis.

1. Refractive Disorders

(myopia, hypermetropia, astigmatism, presbyopia, and cataract).

Near-sighted (*myopic*) patients usually have *elongated eyeballs*. More rarely, myopia can be caused by too high refractive power in the lens system. Myopic persons can only foveate diverging light waves - both from F and N (Fig. 5-1). The images of distant objects are focused in front of the retina, and the image is blurred on the retina. Both F and N are located in front of the eye. Concave lenses (-D) accomplish correction. The weakest concave lens compatible with optimal visual acuity is the best correction, as the accommodation is eliminated.

Hypermetropic or far-sighted persons usually have *shortened eyeballs,* and F is always behind

the eye. In rare cases hypermetropia can also be caused by insufficient refractive power in the un-accomodated eye. The *absolute hypermetropic eye* can only focus images of distant objects behind the retina (Fig. 5-1). The *facultative hypermetropic eye* can focus converging light beams on the retina without accommodation (rest in Fig. 5-1). This patient can read the Snellen letters without problems; they also foveate diverging light beams by accommodation, but then the patient gets eyestrain due to fatigued ciliary muscles. Convex lenses (+D) correct hypermetropia. The strongest convex lens compatible with optimal visual acuity is the best correction, as the accommodation is eliminated.

Astigmatism is a refractive disorder of the eye, in which the *curvatures of the cornea or lens are different along different meridians*. The different meridians therefore have different focal distances. Therefore, astigmatism can be corrected with *cylinder lenses* that correct the curvature differences.

Presbyopia is called *old man's sight*. The far point remains where it is, so the un-accomodated refraction is unaltered. The ability to accommodate is changed, so that N approaches F.

With age, the lenses of most people loses its elasticity, and hence their ability to assume spherical shape. The lens is the organ in our body with the highest protein concentration. Alterations of lens proteins probably cause *progressively increasing stiffness of the lens*. The accommodative power decreases from 14 D in a child to less than 2 D at the age of 50. The patient's eye becomes incapable of accommodation for near vision and reading. Convex lenses correct presbyopia.

Cataract is an eye disease, where the *vision is blurred by an opaque lens.* Precipitation of lens proteins can occur in several ways. It is often due to oxidative processes. The lens needs oxygen, but strong sun light or radiation can *oxidise* lens proteins in unprotected eyes. The oxidation is enhanced by hyperbaric oxygen therapy and by high blood [glucose] in diabetics. Oxidants in the food may be the cause in some patients. Antioxidants, such as vitamin A and D, seem to protect against the loss of transparency in long-term studies. Today, excellent surgical techniques are used to eliminate the opaque lens and re-establish normal refraction.

2. Colour blindness

The three colour genes are located on an *X chromosome*. Females have two X-chromosomes, and colour blindness is rare among females. Colour blindness is inherited from the *father* - via the daughter - to her *son*. The trait is recessive and sex linked. The total incidence is about 8% of the male and 0.5% of the female population. *Monochromats* lack all three or two cone pigments, an extremely rare disorder. *Dichromats* lack one of the three cone pigments. Proteus is the first or red component, so *protanopic people* are blind for the red part of the spectrum. They cannot separate red and yellow signals in traffic. *Deuteranopic patients* are blind for the second or green colour, and *tritanopics* are blind for blue - the third basic colour. *Abnormal trichromats* have a reduced amount of one cone pigment: *Protanomalous trichromats* lack erythrolab, *deuteranomalous* (the most frequent type) lack chlorolab, and *tritanomalous* lack cyanolab.

3. Visual field defects

Visual field defects are caused by interruptions of the visual pathways. *Hemianopsia* means loss of vision in half of the visual field of both eyes. The loss of vision refers to the visual field, and thus to the contralateral half of each retina (shown with two colours in Fig. 5-6).

Homonymous hemianopsia means that the same side of the visual field for each eye is defective. Corresponding halves of each retina has lost vision (black in the illustration). Homonymous hemianopsia occurs from lesions of the entire optic tract, the lateral geniculate body, the optic radiation, or the entire visual cortex of the contralateral hemisphere (Fig. 5-6). A lesion of the striate cortex often spares the large macular area at the occipital pole. This results in a disorder termed *homonymous hemianopsia* with macular sparing (Fig. 5-6). Partial lesions may cause *quadrant-anopsia.*

Heteronymous hemianopsia can be bitemporal or binasal (Fig. 5-6). Bitemporal hemianopsia

results from damage of the optic nerve fibres as they cross the optic chiasm (Fig. 5-6).

An expansively growing pituitary tumour, perhaps related to acromegaly, can damage crossing fibres, originating from ganglion cells in the nasal halves of each retina. Expansion of the tissues surrounding both carotid arteries is a rarity, which can damage nerve fibres from the temporal halves of each retina, and cause binasal hemianopsia.

Fig. 5-6: Visual field defects. Lesions are shown with bars.

As long as the patient sees with both eyes, he may not experience any visual defect caused by damage to non-corresponding areas of the retina.

Localised blindness or *scotoma* is caused by a lesion of the retina in one eye, or by partial interruption of the optic nerve. Interruption of the entire optic nerve results in complete blindness or anopsia (Fig. 5-6).

Ophtalmoscopy is an important diagnostic tool able to establish both eye disorders and systemic diseases.

Fig. 5-7: The eye background (fundus) of the right eye in a healthy person. – A typical hypertensive eye background is shown in Fig. 9-6.

The normal ophtalmoscopic picture of the fundus is seen in Fig. 5-7. The papilla is clearly visible with the central artery and vein, and the cone-filled fovea is located to the left. The papillo-macular nerve bundle connects to the cones of the macula, but this bundle is invisible. This person has a small pigmented area along the lateral side of the papilla (Fig. 5-7).

4. Mental blindness

Bilateral temporal lobe lesions can lead to the Klüver-Bucy syndrome. In this condition, the temporal cortex, hippocampus and the amygdaloid body are damaged. The Klüver-Bucy syndrome includes *mental blindness* (visual agnosia). Mental blindness is the inability to recognise objects seen. Besides mental blindness, the syndrome consists of loss of short-term memory, and hypersexual behaviour incompatible with normal social adaptation. This hypersexual behaviour is related to the visual agnosia.

Damage to visual areas of the temporal cortex alone causes isolated *visual agnosia*. Visual agnosia or mental blindness, is lack of the ability to combine the seen object into a concept. This visual agnosia can be *colour*-agnosia (acromat-agnosia) or *face*-agnosia (prosop-agnosia).

5. Deafness (hypacusis)

Nerve deafness is caused by impairment of the cochlea, the auditory nerve or the nucleus. Chloramphenicol, kinin and streptomycin can damage the cochlea. These drugs can cause hearing loss or deafness for *all* sound frequencies. Deafness to specific frequencies is caused by localised damage of the basilar membrane. This is typical for rock and beat musicians, soldiers, and airline pilots. The nerve deaf patient has a hearing loss when tested both by *air conduction* through the middle ear, and *bone conduction* through surrounding bone structures. A certain type of nerve deafness for high tones develops among older persons (*presbyacusis*).

Conduction deafness is caused by impairment of the mechanical conduction of sound into the cochlea. A hereditary disease called otosclerosis is due to fixation of the faceplate of the stapes to the oval window. Otosclerosis, blockade of the external ear with ear wax, otitis media, damage of the tympanic membrane, and of the ossicles all cause conduction damage to hearing. Persons with conduction damage have normal bone conduction.

6. Nystagmus

Nystagmus is a disorder with *abnormal involuntary movements of the eyeballs*. Opto-kinetic nystagmus occurs when travelling in a train or a car. The eyes remain fixed on an object long enough in order to gain a clear image. The semicircular ducts cause the eyes to rotate in the direction opposite to the direction of travel. Optokinetic nystagmus involves the vestibular nuclei, the *medial longitudinal fasciculus*, and the oculomotor nuclei.

Post-rotatory nystagmus is observed in a person sitting in a rotating chair. This is the physiologic adequate stimulus for nystagmus.

Caloric nystagmus refers to the horizontal reflex movement of the eye when the external ear is flushed with hot or cold water. The fast phase of the nystagmus is directed away from the ear flushed with cold water, and towards the ear flushed with hot water. The caloric nystagmus test is preferable to the post-rotatory test for testing the nystagmus reflexes, because it examines one ear at a time, and is more convenient.

7. Kinetosis or transportation sickness

Many types of transportation, which subjects passengers to rapid changes in the direction of motion, elicit *kinetosis*. Kinetosis is a disorder with vertigo, nausea and vomiting. The disorder is triggered from the *vestibular system*, provided that the cerebellar function (the flocculo-nodular lobe) is intact. The flocculo-nodular lobes are linked to the equilibrium control of the semicircular system. Persons with destroyed semicircular canals or with destroyed flocculo-nodular lobes can be completely protected from kinetosis at the expense of lost equilibrium during motion.

Equations

The **accommodative power of the eye** is the rise in refractive power from un-accomodated to the maximally accommodated condition:

Eq. 5-1: Accommodative power = 1/F - 1/N.

F and N are the far- and the near point, respectively.

A child of 10 years has the high accommodative power of 12-14 D, a 20-year-old person 10 D, and a 60-year-old person 1 D.

According to international convention the *sound pressure level* (SPL) is expressed in decibel (dB):

Eq. 5-2: Sound pressure level (dB) = 20 log p/p_o.

The actual pressure is p, and the threshold for sound pressure is p_o . The threshold for sound pressure is *20 μPa* in air (p_o) at 1000 Hz in a sound tight chamber for a healthy person. This pressure corresponds to a *sound effect* of 10^{-12} *Watts/m*^{2}.

Self-Assessment

Multiple Choice Questions

I. The following five statements have True/False options:

- **A**. The fovea has no cyanolab cones and no rods.
- **B.** Nystagmus is a disorder with abnormal voluntary movements of the eyeballs.
- **C.** The cornea has a refractive power of 43 D, and the healthy lens has a refractive power which varies between 12-26 D. Thus the total refractive power is 56-69 D.
- **D.** Light absorption by a single rhodopsin molecule activates thousands of G-protein molecules (transducin), which then activate large quantities of cGMP phosphodiesterase in the discs. Each of these enzyme molecules catalyses the hydrolysis of cGMP to 5'-GMP at a rate of thousands per second.
- **E.** Parallel light from the far point foveates on the retina in the fully accommodated eye.

II. The following five statements have True/False options:

A. Nerve deafness is caused by damage of the cochlea, the auditory nerve or nucleus. There is hearing loss by air conduction only.

B. Kinetosis is a disorder with vertigo, nausea and vomiting due to rapid changes in the direction of motion.

C. Drugs, such as chloramphenicol, kinin and streptomycin, can cause hearing loss for all sound frequencies.

D. The organ of Corti contains hair cells, each with 40-100 stereocilia. When the pressure wave displaces the hairs towards the tallest stereocilia, the hair cell is depolarised.

E. Low tones travel only a short distance from the stapes along the basilar membrane to their resonant point, where the displacement amplitude is maximal.

5. Case History A

A 30-year old female complains of eyestrain and frontal headache during reading - sometimes followed by nausea and vomiting.

The patient is placed 6 m (20 feet) from the Snellen test chart. She is able to read line 6, which is the letter size read by a normotropic eye. Now thin convex lenses are placed in front of her eyes, but she can still read line 6. The diopter of the strongest convex lens with which she can still read line 6 is +4 D for both eyes. These converging light rays must be directed against the far-point (F). F must be located behind the eyes at a distance of ¼ m. Examination with concave lenses reveals the strongest concave lens by which she can read line 6 to be -3 D or N = -1/3 m (in front of each eye).

1. *1. What is the refractive anomaly of the patient?*
2. *2. A normotropic person of 30 years has an accommodative power of 7 D. Compare the accommodative power of the female patient to that of the normal person?*
3. *3. Is it possible for the patient to read a fine text 0.2-m in front of her?*
4. *4. The saggital diameter of the patient's eyes is typical. Describe its characteristics.*

- 5. *This patient has a reduced outflow of chamber fluid at the iridocorneal junction. She has an increased risk of developing an eye disease, which is the most common course of blindness in the world. Describe the most likely condition and the relation to her eye anatomy.*

5. Case History B

A patient with a hearing loss of 26 dB is working in a power station, where the daily sound intensity is 100 dB and the air temperature is 20^{o}C.

- ***1**. Calculate the ratio between the sound pressure in the powerhouse, and the sound threshold pressure for a healthy person.*

- ***2.** Calculate the threshold pressure for the patient.*

5. Case History C

A 9-year-old girl suffers from facultative hypermetropia. She is placed 6 m from the Snellen test chart and asked to read line 6, which is the letter size read by a normotropic eye. When thin convex lenses are placed in front of her eyes, she can still read line 6. The diopter of the strongest convex lens with which she can still read line 6 is +5 D for both eyes. Examination with concave lenses reveals the strongest concave lens by which she can read line 6 to be -4 D.

1. *Where are the far point (F) and the near point (N) located?*
2. *Calculate her accommodative power and compare it to the normal value of 14 D.*
3. *Calculate the correction needed.*

5. Case History D

A 40 year old male diabetic has an accommodative power of 10 D. His near point (N) is located 0.05 m in front of the eye (- 0.05 m).

- *1. Calculate the location of F and the necessary correction.*
- *2. What is the name of the refractive disorder?*
- *3. Is the patient capable of driving a car without corrective glasses? Is the vision mildly or seriously reduced?*

Try to solve the problems before looking up the answers.

Highlights

- *Near vision- even with only one eye - triggers accommodation and pupillary constriction.*
- *Rods are most sensitive in the dark (scotopic vision). More than hundred rods converge on each ganglion cell. There are no rods at all in the fovea.*
- *Cones operate best in light (photopic vision). Cones have a high-resolution capacity and hence a high visual acuity, because the light is focused on the fovea, where the cones are concentrated.*
- *The cornea has a refractive power of 43 D, and the healthy lens varies between 12 and 26 D.*
- *The accommodative power of the eye is the rise in refractive power from the un-accomodated to the maximally accommodated condition.*
- *Conjugate movements are movements of both eyes in the same direction and magnitude, so that the relation between the visual axes is maintained.*
- *Saccadian or jumping movements are rapid eye movements. Saccadian eye movement is an instantaneous reposition of the eye that occurs. when reading or when focusing on a flash of light in the peripheral visual field. The velocity of the movement is up to 500^{o} per s. The latency period is 250 ms, and the contraction time is 50 ms. The compensatory eye movement involving the vestibular system, occurs when the head rotates. This is also an example of Saccadian eye movement.*
- *The fovea only contains cones. Cones function in the daytime, with maximal visual acuity, and colour vision. The human eye possesses three types of cones, each with a specific pigment related to the three basic colours: red (erythrolab), green (chlorolab) and blue (cyanolab).*
- *The 3 cone types are uniformly distributed in the retina except in the fovea. The fovea has no cyanolab cones and no rods.*
- *Abnormal trichromats have a reduced amount of one cone pigment: Protanomalous trichromats lack erythrolab, deuteranomalous (the most frequent type) lack chlorolab, and tritanomalous lack cyanolab.*
- *Movements in the visual scenery are depicted as opposite movements on the retina. Convergent inputs from the two eyes result in depth perception (ie, stereopsis or stereoscopic vision). Stereopsis depends upon the medial, longitudinal fasciculus and the corpus callosum. These structures co-ordinate the movements of the two eyes.*
- *A flickering source of light liberates successive flashes so rapidly that they fuse and appear continuous.*
- *The normal human ear is sensitive to pure tones with frequencies between 10 and 30 000 Hz in a young person.*
- *Speech has an intensity of 60-65 dB and sounds that exceed 100 dB can damage the ear.*
- *A high frequency tone produces travelling waves along the basilar membrane. High tones travel only a short distance from the stapes along the basilar membrane to their resonant point, where the displacement amplitude of the basilar membrane is maximal.*
- *Low frequency tones travel all the way to the apex of the cochlea. The higher the tone frequency, the more basal located in the cochlea is the resonant point and its potential.*

- *The auditory and the vestibular systems share the labyrinth, and transmit signals to the brain through the 8th cranial nerve. The two systems record fluid movements and use the so-called hair cells as mechanical receptors.*
- *The medial geniculate nucleus and the inferior collicle can increase its tone selectivity by dampening other sound signals. This explains the cocktail party effect.*
- *The cupula is concerned with equilibrium control during motion and with angular acceleration (rotation of the head).*
- *The utricles and saccules are sensitive to linear acceleration.*

Further Reading

- Berardi, N., A. Cellerino, L. Dominici, M. Fagiolini, T. Pizzorusso, A. Cattaneo, and L. Maffei. "Monoclonal antibodies to nerve growth factor affect the postnatal development of the visual system." *Proc. Natl. Sci.* , USA, 91 (2): 684-688, 1994.
- Stryer, L. "Cyclic GMP cascade of vision." *Annual Rev. Neuroscience* 9: 87, 1986.
- Brodal, A. "Neurological anatomy in relation to clinical medicine." *Edition 3.* New York, 1981, Oxford University Press.
- Von Bekesy, G. "Experiments in hearing." *New York, 1960.* Mc-Graw-Hill.

Return to content

Return to chapter 5

Chapter 6.
The Autonomic Nervous System And Disorders

Study Objectives

- To *define* receptors, autonomic neurotransmitters and blocking drugs, homeostasis, receptors and related concepts.
- To *describe* the anatomy and the physiology of the sympathetic and the parasympathetic nervous system, the visceral afferent system, the enteric nervous system, transmitter mechanisms in autonomic ganglia and at peripheral receptors, the bladder emptying, the pupillary reflexes.
- To *explain* the central autonomic control, the autonomic control of temperature, appetite, thirst, the subsynaptic autonomic mechanisms, emotional disorders, the Kluver- Bucy-syndrome.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The autonomic nervous system mediates neural control of the internal milieu despite substantial environmental changes.*
- *Cannons law: The peristalsis in the small intestine is polarised, so it always proceeds in the oral-aboral direction.*

Definitions

- **Autonomic neurotransmitters** are adrenergic and cholinergic substances (Box 6-1 and 6-3).
- **Autonomic blocking drugs** (sympatholytics and parasympatholytics) block the normal effect of sympathetic and parasympathetic neurotransmitters.
- **Cholinergic receptors** are *nicotinic* (with a fast EPSP within ms) and *muscarinic* (with a slow EPSP lasting several seconds). Both cholinergic receptors are transmembrane proteins and both open an ion channel in the protein.
- **Homeostasis** refers to all processes helping to keep in internal milieu of the body constant despite environmental alterations.
- **Mydriasis** refers to dilatation of the pupil by sympathetic stimulation of the dilatator muscle.
- **Miosis** refers to contraction of the sphincter muscle (parasympathetic) resulting in a small pupil.
- **Receptors for neurotransmitters** are specific cellular components, who react with a neurotransmitter, a hormone or a drug (agonist) to produce a biological response in the cell.
- **SIF cells** are *small intensity fluorescent cells*, which possess muscarinic receptors and contain vesicles filled with dopamine. Adequate stimulation releases dopamine, which interacts with dopamine receptor (D_2) on the postsynaptic cell body and modulates the effect of acetylcholine.

  The modulation takes place through a permeability increase for small ions (K^+ out and Cl^- into the cell), hyperpolarizing the cell membrane.
- **The nicotinic receptor** responds to acetylcholine with a rapid influx of Na^+, whereby the membrane is depolarised.
- **The muscarinic receptor**. In the muscarinic M_1 receptor, IP_3 is second messenger and increases cytosolic Ca^{2+}. Activation of the M_2 receptor implies activation of an inhibitory G-protein, which inhibits adenylcyclase. The result is reduced concentration of cAMP, which operates in smooth muscle contraction, with secretion from glands or with a slow EPSP.

Essentials

This paragraph deals with *1. The autonomic system in general, 2. The sympathetic system, and 3. The parasympathetic system.*

1. The Autonomic System In General

The autonomic system directly influences *smooth muscles, glands and the heart* through its two subdivisions, the sympathetic and the parasympathetic system. The two subdivisions function in a dynamic balance aiming at homeostasis.

The *enteric nervous system* is lying within the walls of the gastrointestinal tract and includes neurons in the pancreas, liver and gallbladder, thus being an entity in itself. However, the enteric nervous system is clearly an important part of the autonomic nervous system that controls gastrointestinal motility, secretion and bloodflow.

The central autonomic system

The central autonomic nervous system outflow arises in the hypothalamus, the brainstem, and the spinal cord. The motor and premotor cortex, the cingulate gyrus and the hypothalamus can modulate the function of the autonomic medullary control neurons in the lateral horn of the grey matter. Circulatory changes during exercise and in various stressful situations are influenced or governed by the cortex and deeper brain nuclei. The central autonomic system also modulates release of certain peptides and catecholamines that affect both blood volume as well as the total peripheral vascular resistance.

The cerebral cortex assimilates all inputs of visual, olfactory, labyrinthine, locomotor origin, as well as from other specialised sensors (stretch receptors, chemo-, baro-, osmo-, and thermo-receptors).

The integration of these inputs into an appropriate response takes place in the hypothalamus and in the ponto-medullary centres. From here the efferent signals pass to the periphery via the sympathetic and the parasympathetic pathways.

The primary afferent projections from the baroreceptors reach the solitary tract nucleus (STN), and from here we have connections to the important dorsal motor nucleus of the vagus (DMNV in Fig. 6-4). A high baroreceptor activity stimulates the DMNV, so that the vagal inhibition of the heart is increased. More importantly, the high baroreceptor activity inhibits the sympathetic drive to the heart and vessels thus reducing blood pressure (Fig. 6-4).

The central autonomic structures co-operate in situations of survival character: Fright, flight or fight-response, feeding and drinking in starvation, reproduction and sexual satisfaction for continuation of life, thermoregulation at extreme temperatures and emotional behaviour in crises.

The Fright -Flight Or Fight Response

Aggression and defence responses are elicited in emergency situations. The sympatho-adrenergic system gives rise to the fright, flight or fight-reactions in acutely stressful situations. The sympathetic reactions dominate over the parasympathetic and the subject is aggressive or anxious. The brain releases corticotrophin-releasing factor to the hypothalamic-pituitary portal system. The hypothalamic-pituitary axis secretes adenocorticotropic hormone, the cardiac rate and contractile force increases, the blood is distributed from viscera to the active skeletal muscles by visceral vasoconstriction and preferential vasodilatation. The subject hyperventilates, the gastrointestinal activity is reduced, and there is increased glycogenolysis and lipolysis. The airways dilatate, and the adrenal medullary (catecholamines) and cortical secretion (cortisol) increases. This response is seen in humans exposed to psychological-emotional stress. Stress in general is comprised of severe emotional and physical burdens (fear, pain, hypoxia, hypothermia, hypoglycaemia, hypotension etc).

Cannons emergency reaction is an immediate sympatho-adrenergic response to life-threatening situations, with both sympatho-adrenergic and parasympathetic overactivity. The last phenomenon includes vagal cardiac arrest with involuntary defecation and urination.

Feeding and drinking

Bilateral destruction of the ventromedial hypothalamic nuclei leads to hyperphagia and failure of body weight control. Such animals become obese, and they have high plasma [insulin].

Bilateral lesions in the lateral hypothalamic regions cause a temporary hypophagia.

The cells of the ventromedial nuclei have a special affinity for glucose, and these cells are responsible for insulin secretion from the pancreatic b-cells. Signals from the dorsal motor nucleus of the vagal nerve increase insulin secretion, and sympathetic stimulation inhibits the release of insulin. The ventromedial nuclei seem to function like a glucostat.

Stimulation and ablations of the limbic system affect food intake. Obviously from clinical practice, psychological factors, emotional disturbances, motivations and conditioned behaviour are all affecting our drive for food intake.

Concerning the control of food intake see also Chapter 20 and 27.

Sexual behaviour

Hypothalamic and other limbic system co-operation are responsible for a wide variety of autonomic and somatic phenomena associated with emotions. Stimulation of the midbrain septum yields pleasurable sensations and sexual drive in patients. The dorsomedial nucleus of the hypothalamus is probably a major sex centre responsible for the sexual act. Stimulation of the ventromedial and preoptic regions also releases sexual activities. See also Chapter 29.

The Thermocontrol

Thermoreceptors can initiate generalised reactions to heat and cold. The signals from both superficial and deep thermoreceptors must act through the hypothalamus to arouse appropriate, generalised reactions.

Cooling or heating the denervated lower extremities of spinal men evoked vasoconstriction and shivering or vasodilatation and sweating of the innervated upper body shortly after cooled or warmed arterial blood reached the brain. The anterior hypothalamus is responsible for sensing blood temperature variations. The anterior hypothalamus, in particular the preoptic area, has been shown to contain numerous heat-sensitive cells and less cold-sensitive receptors. Such central thermoreceptors are also found at other levels of the CNS. After destruction of the hypothalamus, the midbrain reticular formation takes over the temperature control. Sections eliminating both the hypothalamus and the mesencephalon leave the medulla and spinal cord to control temperature. The posterior hypothalamus does not contain thermoreceptors. Concerning thermocontrol see also Chapter 21.

The brain and the immune defence system

Internal and external stress affects the prefrontal cortex, whereby the limbic system with the hypothalamus is activated. Hypothalamic nuclei release corticotropin-releasing hormone (CRH) to the portal blood. The blood reaches the adenohypophysis, where CRH triggers the release of adenocorticotropic hormone (ACTH), endorphins and met-enkephalin. ACTH works through different pathways in order to protect the body. ACTH stimulates the adrenal cortex to release corticosteroids, which produce immuno-suppression. Immuno-suppression reduces the number of inflammatory effector cells, including helper T cells and killer cells.

On the other hand, cancer therapists assume that relaxed lifestyle and positive reinforcement may have stimulated the immune defence in some patients with malignant diseases, and explain miraculous remissions. Higher brain centres may even affect the reticuloendothelial production of killer cells through the peripheral nerves to the lymph nodes and bone marrow. See also Chapter 32.

Fig. 6-1: The peripheral autonomic nervous system. β -receptors stimulate glycogenolysis in the liver and lipolysis in lipid tissues.

Autonomic nerves are composed of two neurons termed the preganglionic and the Postganglionic neuron based on anatomical location relative to the ganglion. A preganglionic neuron has its cell body in the spinal cord or brainstem and is modulated by higher centres and by spinal reflexes (Fig. 6-1). The preganglionic axon leaves the CNS from the cranial, thoracic, lumbar or sacral regions and synapse in the autonomic ganglia with the cell body of the postganglionic neuron. The postganglionic neurons innervate the effector organs (Viscera).

Viscera function involuntarily and their activity must be modulated by the autonomic nervous system with excitatory or inhibitory signals. All autonomic nerves have ganglia outside the CNS in contrast to the somatic nervous system, where neural connections are located entirely within the CNS. Most somatic nerves that control motor function are myelinated and have a high conduction velocity, whereas most postganglionic neurons are unmyelinated with a low conduction velocity. However, the preganglionic neurons are mostly myelinated with a high conduction velocity (Fig. 6-1).

Receptors for neurotransmitters are specific cellular components, whose interaction with the neurotransmitter, a hormone or a drug produces a biological response in the cell.

Acetylcholine (ACh) is the transmitter between the pre- and the post-ganglionic neurons, not only in the sympathetic nervous system, but also in the parasympathetic system. The cholinergic receptors are nicotinic or muscarinic. The cholinergic receptors of the ganglia and in the somatic motor endplate are *nicotinic*. Nicotine and acetylcholine activate nicotinic cholinergic receptors. When the action potential arrives at the preganglionic fibre, acetylcholine is released from its terminals and diffuses across the synaptic cleft to bind to the specific nicotinic receptors on the membrane of the postganglionic neuron. Nicotinic receptors are linked to cation channels lined with negative charges.

These channels open enough to allow mainly hydrated Na^+ to enter the cell rapidly (for about 1 ms) and depolarise the membrane (Fig. 6-2).

Fig. 6-2: The nicotinic cholinergic receptor

The resulting current elicits an *excitatory postsynaptic potential* (EPSP). Repolarisation is also fast (ms).

Acetylcholine is also the neurotransmitter for the sympathetic innervation of sweat glands, and they are completely blocked by *atropine*. The acetylcholine receptors of the sweat glands are *muscarinic*, since acetylcholine and muscarine (Fig. 6-3) activate them.

Fig. 6-3: The muscarinic cholinergic receptor

These slowly working surface-receptors are linked to a long lasting cascade of events starting with binding of the hormone to the receptor, activation of G-proteins (see Chapter 7), enzyme activation,

production of second-messengers, protein kinase activation, and phosphorylation of specific proteins such as channels. All these processes are simplified in Fig. 6-4, and the result is opening of K^+-channels, with efflux of K^+, so the membrane is hyperpolarised. In this example, acetylcholine is an inhibitory transmitter.

2. The Sympathetic Nervous System

The preganglionic sympathetic nerve fibres originate in small multipolar neurons in the lateral horn of the grey matter in the thoracic and lumbar spinal cord. The central sympathetic outflow converges on these preganglionic neurons. Their axons are thin myelinated fibres that leave the spinal cord through the ventral root. The preganglionic fibres then leave the spinal nerve forming myelinated white rami communicantes, through which they reach the nearest ganglion in the paravertebral ganglia of the paired sympathetic trunk. Typically, each fibre will end here forming synapses with up to 20 postganglionic neurons. A few preganglionic fibres pass the sympathetic trunk without interruption to form the splancnic nerves that reach the three unpaired prevertebral ganglia (coeliac = solar plexus, superior mesenteric and inferior mesenteric) of the lower intestinal and urinary organs. Most sympathetic ganglia are remote from the organ supplied. The postganglionic fibres are all unmyelinated, and they leave the sympathetic trunk through the grey rami communicantes and thus reach the effectors supplied by the sympathetic system. The effectors are the smooth muscles of all organs (blood vessels, viscera, lungs, hairs, pupils), the heart and glands (sweat glands, salivary and other digestive glands). In addition, the sympathetic postganglionic fibres innervate adipocytes, hepatocytes and renal tubular cells.

The sympatho-adrenergic system is a functional and phylogenetic unit of the sympathetic system and the adrenal medulla. The adrenal medulla is a modified sympathetic ganglion. Any increase in sympathetic activity increase the secretion of adrenaline and noradrenaline from the medulla into the circulation. The preganglionic fibres to the adrenal medulla pass all the way to the special postganglionic cells in the adrenal medulla. The synapse is cholinergic (nicotinic) as it is for all preganglionic synapses. The postganglionic cells of the adrenal medulla have developed to cells filled with chromaffine granules, and are called chromaffine cells. These cells do not conduct signals, but synthesise adrenaline (and noradrenaline) which is released into the blood. Sympathetic stimulation triggers the conversion of tyrosine to dihydroxyphenylalanine (DOPA). A non-specific decarboxylase catalyses the conversion of DOPA to dopamine, which is taken up by the chromaffine granules in the cells. The granules contain the crucial enzyme, dopamine b-hydroxylase. This enzyme is activated by sympathetic stimulation, and catalyses the formation of noradrenaline from dopamine.

A few granules store noradrenaline (NA), while the remaining granules liberate NA to the cytosol, where NA is methylated by phenylethanolamine N-methyltransferase to adrenaline. Adrenaline is taken up by chromaffine granules and stored as the predominant adrenal hormone.

Adrenergic Receptors

The sympathetic system exerts either excitatory or inhibitory actions through adrenergic receptors. Adrenergic receptors are *membrane-receptors*. The dual response to adrenergic stimulation was known before Ahlquist in 1948 proposed that adrenergic receptors could be divided into two groups, a- and b- receptors, on the basis of *blocking drugs* (Box 6-1). The basic idea of Ahlquist is that noradrenaline (NA) act predominantly on *vasoconstricting α-receptors,* and isoprenaline (Iso) predominantly on *vasodilatating β-receptors*. Both types of receptors are stimulated by *adrenaline* (Ad).

Box 6-1: Adrenergic receptor subtypes.
The symbol > indicates the rank order of sensitivity.

	Adrenergic receptors			
	α-receptors	**β-receptors**		
Stimulated by:	NA>Ad		Iso>Ad ³NA	
Blocked by	**Phenoxybenzamine**		**Propranolol**	
	α-receptors	**β-receptors**		
a -receptors	a -receptors	b -	b -receptors	

a_1	a_2	b_1 receptors	b_2	
Stimulated by	NA>Ad	NA>Ad	Iso>Ad=NA	Iso>Ad>NA Salbutamol
Blocked by	**Prazosin**		**Metoprolol**	**Butoxamine**

The rank order of sensitivity of a series of chemically similar compounds for activating a receptor (agonists) or inhibiting the receptor response (antagonists) is considered diagnostic of the receptor subtype. More and more closely related subtypes are distinguished, so there is already three subtypes for each of the following receptors: a_{1ABC}-receptors, a_{2ABC}-receptors, and b_{123}-receptors.

1. The a-receptors are blocked by Phenoxybenzamine and Phentolamine.

The a_1-receptors are located on the surface of target cells (vascular smooth muscle, sphincter muscles of the gastrointestinal tract and bladder, and radial iris muscles). They are highly sensitive to NA, less sensitive to Ad, and almost insensitive to isoprenaline (Box 6-1).

The a_1-receptors act through phospholipase C and through intracellular [Ca^{2+}] elevation. Ca^{2+} binds to calmodulin in the cytosol. The complex activates protein kinase, which catalyses the phosphorylation of proteins. They become enzymatically active, and trigger vasoconstriction.

In contrast, the presynaptic a_2-receptors are located on the presynaptic membrane (sympathetic end bulbs). NA released into the synaptic cleft diffuses to the a_1-receptors on the target cells, but part of the NA diffuses back to the a_2-receptors on the presynaptic nerve terminals. Here, NA activates membrane adenylcyclase, reducing [cAMP] in the cells, and thus inhibiting release of more NA from the vesicles by negative feedback. Hence, a function of a_2-receptors is auto-inhibitory feedback.

These receptors are also found in gastric smooth muscle cells and the b-cells of pancreatic islets. Stimulation decreases gastric motility and attenuates insulin secretion.

2. The b-receptor is blocked by propranolol (Box 6-1).

The b-receptors are located on effector cells that are most sensitive to isoprenaline, but less so to Ad and NA. All b-receptors act through activation of adenylcyclase and cAMP. b_1-receptors are equally sensitive to NA and Ad, whereas b_2-receptors are more sensitive to Ad than to NA (Box 6-1).

b_1-receptors are located in the myocardium - primarily on pacemaker cells. The b_1-receptors of the heart are stimulated by NA which increases cAMP production with increased chronotropic (increased heart rate) and inotropic effect (increased force). Heart patients use Cardioselective b_1-blockers such as Metoprolol, because Metoprolol decreases cardiac arrhythmias and tachycardia.

b_2-receptors are found primarily on bronchiolar smooth muscle cells, vascular smooth muscle, uterine smooth muscle, salivary glands, the intestine and the liver. When NA binds to b_2-receptors, it causes inhibition of the target organ. Therefore, NA causes vasodilatation, bronchodilatation and uterine relaxation. Similarly, sympatomimetics such as b_2-stimulators (salbutamol) increase cAMP production, resulting in bronchodilatation, increased salivary secretion, uterine relaxation and enhanced hepatic glucose output. b_2-stimulators are used to eliminate bronchial asthma attacks.

Butoxamine is a selective b_2-blocker.

Box 6-2: Responses elicited in effector organs by sympathetic and parasympathetic activation

Effector organ	**Adrenergic response**	**Cholinergic response**
Heart		
Rate of contraction	Increase, b_1	Decrease, M2
Force of contraction	Increase, b_1	Decrease, M2

Arteries and arterioles		
in myocardium	Vasodilatation, b_2 (a 1constr)	Vasodilatation, M
in skeletal muscles	Vasodilatation, b_2	
in lungs	Vasodilatation, b_2	
Bronchial muscles	Bronchodilatation, b_2	Bronchoconstriction, M
Gastrointestinal		
motility	Decrease, a_2 ($b_{2,b3}$)	Increase, M
sphincters	Contraction, a	Relaxation, M
secretion	Decrease, a	Increase, M_1
Exocrine glands		
Salivary	Small secretion, a_1	Secretion, M_2
Lacrimal		Secretion, M_2
Digestive	Decreased secretion, a	Secretion, M_2
Airway		Secretion, M
Sweat	Secretion, a 1	Secretion, M
Pancreatic acini	Decreased secretion, a	Secretion, M
Langerhans islets	Decreased secretion, a 2 Increased secretion, b 2	
Lipid cells	Lipolysis, b_1 b_3	
Liver glycogenolysis	Increase, $a_{1\,b\,2}$	
Eye		
Ciliary muscle	Relax., b (far vision)	Contraction, M (near vision)
Dilatator muscle of pupil	Contract., a_1 (Mydriasis)	
Sphincter muscle of pupil		Contraction, M (Miosis)
Kidney	Renin secretion, b_2	
Ureter-motility	Increase, a_1	

Urinary bladder		
detrusor	Relaxation, b	Contraction, M
sphincter	Contraction, a_1	Relaxation, M
Genital organs		
male	Ejaculation, a_1	Erection, M
uterus (pregnant)	Contraction, a_1	
Adrenal medulla		Secretion, N

The near-vision response is also called the convergence response. Near vision -even with only one eye - triggers accommodation and pupillary contraction (Box 6-1). The ciliary muscle and the pupillary sphincter muscle are innervated of the parasympathetic oculomotor nerve, and the two muscles (with M-receptors) contract simultaneously for near vision. This leads to increased refractive power or accommodation, and to pupillary contraction (miosis).

When a person closes his eyelids the pupils enlarge, and when he opens the pupils again the pupils become smaller. This is due to the pupillary light reflex, where retinal ganglion cells are stimulated by light, send signals through the optic nerve to the olivary pretectal nucleus neurons. These light-sensitive neurons are connected to the parasympathetic preganglionic neurons in the oculomotor Edinger-Westphal nuclei on both sides. The light reflex contracts the pupillary sphincter muscle. Argyll-Robertson's pupillary syndrome refers to small, light-refractive pupils with maintained convergence response to near vision. The syndrome is seen in neurosyphilis, when the pupillary light reflex is spoiled by interruption of the fibres from brachium to the olivary pretectal nucleus neurons.

Sympathetic preganglionic neurons in the intermediolateral cell column of segment T1-T2, send ascending axons to the superior cervical ganglion. Postganglionic axons follow the ciliary nerve into the eye. The nerve terminals end on a 1- receptors on the dilatator pupillae muscle, and noradrenaline is neurotransmitter. The sympathetic fibres also contain vasoconstrictors to the facial skin and stimulate facial sweat glands (see Horners syndrome).

Catecholamines are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The synthesis is described in Chapter 29.

Catecholamines increase heart rate and cardiac output by stimulation of the adrenergic b_1-receptors in the myocardium. Catecholamines, released by the adrenal medulla, support the sympathetic system by modifying the circulation during exercise. During exercise the blood is directed to the working muscles from other parts. Noradrenergic nerve fibres innervate blood vessels all over the body. Sympathetic innervation accounts for vascular tone and vasoconstriction.

The most important exercise response in humans is a tremendous vasodilatation in the vascular bed of muscles. The vasodilatation is probably due to a decrease in the a-adrenergic tone of muscular arterioles, and to the action of adrenaline on b_2-receptors.

Catecholamines dilatate the bronchial airways by stimulating adrenergic b_2-receptors. They increase both tidal volume and respiratory frequency. The result is increased ventilation. Catecholamines acting on b_1-receptors cause increased cardiac output. Catecholamines relax the smooth muscles of the digestive tract (b_2-receptors), but contract the sphincters. Catecholamines stimulate metabolism (by activation of the thyroid hormone, T_3) and lipolysis. Adrenaline stimulates hepatic glycogenolysis via b_2-receptors.

Finally, adrenaline stimulates the ascending reticular system (ie, the reticular activating system or RAS) in the brain stem, thus keeping us alert and causing arousal reactions with desynchronisation of the EEG (Chapter 10).

The resistance vessels of the striated muscles in hunting predators (and perhaps in humans) are also innervated by another system. This is the cholinergic, sympathetic vasodilator system. It is capable of a rapid and appropriate bloodflow response during hunting.

Acute stress activates the splancnic nerves and liberates large amounts of adrenaline from the medulla. Diabetics who are developing acute hypoglycaemia, secrete large amounts of catecholamines. Acute muscular activity starts a large catecholamine secretion in exercising persons.

Besides catecholamines, ACTH is also released during stress by increasing hypothalamic signals. ACTH stimulates the glucocorticoid and to some extent the mineralocorticoid secretion through cAMP. Small amounts of glucocorticoids are permissive for the actions of catecholamines.

Plasma catecholamines are rapidly removed from the blood and have a half-life in plasma of less than 20 s. This is the combined result of rapid uptake by tissues and inactivation in the liver and vascular endothelia (see Chapter 29).

The Autonomic Control Of The Cardiovascular System

The brainstem is the primary site for the autonomic cardiovascular control.

High-pressure baroreceptors are distension-activated stretch receptors located in the walls of the carotid sinus and the aortic arch. Increased arterial blood pressure increases the signal frequency in the sensory baroreceptor neurons that project into the medullary cardiovascular centre (ie, the solitary tract nucleus, nucleus ambiguus and the dorsal motor nucleus of the vagus, DMNV). Impulses generated in the baroreceptor neurons with increasing blood pressure, activate the vagal efferents to the heart and inhibit the sympathetic tone towards the heart. As a consequence the heart rates and force of contraction decreases. Impulses generated in cardiac baroreceptors by cardiac filling, also activate the force of contraction (Fig. 6-4).

The postganglionic fibres from the 3 upper cervical ganglia of the sympathetic trunk pass to the heart as the *cardiac nerves* to the *cardiac plexus*.

Fig. 6-4: The autonomic control of blood pressure and heart rate.

The sympathetic effect is dominant. Increased sympathetic activity constricts the veins, which increases cardiac output by augmenting cardiac filling. Arteriolar constriction reduces cardiac output by increasing the arterial blood pressure (ie, afterload). Other sensory inputs from skeletal muscles, lungs, gastrointestinal viscera, hypothalamus and forebrain help to co-ordinate the autonomic cardiovascular responses related to exercise, respiration, and feeding and temperature control. Hormones, such as angiotensin II, can also modulate the autonomic responses through neurons in the circumventricular organs of the brain, These organs (such as the area postrema) lack the blood-brain barrier.

The sympathetic system innervates the sinus node, the coronary vessels and the myocardial syncytium. Each fibre ends in many terminals, and from the terminals the transmitter noradrenaline is released to the b_2 -receptors of the smooth muscle cells of the coronary vessels and of the myocardium (Fig. 6-4). As a result of increased sympathetic tone, the contractility of the myocardium is increased. Thus, the *end systolic volume* falls from its usual volume- as an example from 70 ml to 40 ml, and the *end diastolic volume* increases due to increased venous return of blood from 140 to 180 ml. Hereby, the *stroke volume* is increased from *70* to *140 ml* of blood in the example. A combination of the doubling of stroke volume with a threefold increase in heart rate, results in a *6-fold* rise in cardiac output.

Sympathetic stimulation depolarizes the sinus node, so that the threshold potential is reached faster than normal. Hereby, the heart rate is increased, and may reach 220 beats/min in young persons. Such a high frequency is due to a *maximal sympathetic activation* of the heart combined with a *reduction* of the vagal tone.

Sympathetic Activation

Activation of noradrenergic fibres leads to *peripheral sympathetic vasoconstriction*, so that blood is shunted to central areas. The heart is stimulated through b_1-receptors so that its frequency and contractility is increased. Other organs are also stimulated to make the person fit for *fight or flight* in any stressful situation.

The postganglionic sympathetic fibres have noradrenaline and ATP containing vacuoles in their nerve terminals. Hence, they release noradrenaline and ATP. The noradrenaline is produced in the chromaffine granules of the neuron.

Fig. 6-5: A sympathetic ganglionic synapse with a small intensity fluorescent cell (SIF cell).

Acetylcholine is released from the preganglionic cell and binds to nicotinic receptors on the postganglionic cell. Acetylcholine also binds to small SIF cells (Fig. 6-5) with muscarinic receptors and vesicles that contain dopamine. Dopamine interacts with dopamine receptors (D2 and D4) on the postganglionic cell and modulates ganglionic transmission by increased permeability to small ions and *hyperpolarisation*.

Liberation of noradrenaline and ATP to the blood does not only lead to constriction of arterioles and

arterial vessels, but also constriction of veins and venules. Without venous constriction, the large venous compliance would cause an inordinate amount of blood to be stored in the veins upon sympathetic arteriolar constriction. The consequence would be decreased venous return, which decreases cardiac output and perfusion of vital organs.

Activation of presynaptic purine receptors by adenosine inhibits adrenaline release from the postganglionic terminals innervating the blood vessels. This results in massive vasodilatation.

Exercise and stress demand mobilisation of energy to muscles and heart. Activation of b_2-receptors in the arteriolar wall by circulating catecholamines from the medulla also contributes to vasodilatation in the striated muscles. The total peripheral vascular resistance is reduced during exercise to 20-30% of resting values.

During stress the cutaneous circulation is reduced at first, but then the cutaneous bloodflow rises due to the increased heat production. The brain vessels are only modestly constricted by sympathetic stimulation.

3. The Parasympathetic System

The parasympathetic system has two subdivisions. The cranial division in the brainstem innervates the blood vessels of the head and neck and of many Thoraco-abdominal viscera. The sacral division in the sacral cord innervates the smooth muscles of the walls of the viscera and their glands (the large intestine, liver, kidney, spleen, the bladder and the genitals).

The parasympathetic system only innervates a small percentage of the resistance vessels. Only arteries in the brain and of the penis, the clitoris, and the labia minora receive parasympathetic innervation. Hence, the parasympathetic system has a minimal effect on the arterial blood pressure.

Parasympathetic fibres travelling in the vagus nerve are of utmost importance in affecting the cardiac rate. Vagal fibres innervate the sino-atrial- and the atrio-ventricular-nodes as well as the atrial muscle walls.

The parasympathetic system also innervates the tear and the salivary glands, and the muscles within the eye.

Excitation of the vagus decreases heart rate and atrial contractile force, increases intestinal motility, contracts the gall bladder and bronchi, and relaxes the sphincters of the gastrointestinal tract. The vagal decrease in heart rate is due to the rhythm shift to special P cells, which have a slow rate of depolarisation. Acetylcholine (ACh) is liberated on the cardiac cell membranes, ACh-activated K^+ -channels are opened (via cholinergic receptors and G-regulatory proteins), and K^+ leaks out of the cells, thus opposing the pacemaker current. Vagal stimulation slows down the AV-conduction, causing the co-ordination of atrial and ventricular rhythm to be disrupted. Vagal stimulation can lead to death. Thus external massage of the carotid sinus can cause collar death by greatly increasing vagal stimulation.

The effect of acetylcholine released in the autonomic ganglia can be simulated by nicotine. Conversely, the effect of acetylcholine released by parasympathetic nerve terminals at the target organs can be simulated by muscarine. These observations suggest the presence of two different types of cholinergic receptors. Cholinergic receptors are activated by ACh and by metacholine (MeCH).

Box 6-3: Cholinergic receptor subtypes. ACh stands for acetylcholine, CCh for carbacholine, MeCH for metacholine, DMPP for dimethylphenylpiperazine, and HHSD for hexahydrosiladifenol

Cholinergic receptors				
	Nicotinic		Muscarinic	
Stimulated by:	Nicotine, ACh, MeCH, **DMPP**		Muscarine, ACh, **CCh** MeCH	
Blocked by:	**Hexa**- and **deca**methonium d-tubocurarine		**Atropine**, scopolamine	
Two types of nicotinic receptors			Three muscarinic subtypes	

	Ganglionic	Neuromuscular		$M_1, M_2 M_3$
Stimulated by	Nicotine, ACh **DMPP**	Nicotine, ACh		
Blocked by:	**Hexa**methonium d-tubocurarine	**Deca**methonium Atropine	Pirenzepine Gallamine HHSD	Atropine Dicyclomine

The most important ganglionic blocking drug for blockade of both sympathetic and parasympathetic transmission is *hexamethonium* (Box 6-3).

Cholinergic receptors are located in all autonomic ganglia (nicotinic type), in postganglionic terminals at target organs with parasympathetic innervation (muscarinic type), and in the motor endplate (nicotinic type).

Nicotinic receptors are those activated by acetylcholine, nicotine and nicotinic agonists (ex. dimethylphenylpiperazine, DMPP). Nicotine stimulates all autonomic ganglia simultaneously. Hence, sympathetic vasoconstriction in the limbs and viscera is accompanied by increased gastrointestinal activity and slowing of the heart via the vagus. Nicotinic receptors are blocked completely by d-tubocurarine, and hexa- or decamethonium (Box 6-3). The motor endplate has a different type of nicotinic receptor than the ganglions, since its receptors are not blocked by hexamethonium, but are blocked by d-tubocurarine and decamethonium (Box 6-3).

Acetylcholine, muscarine and muscarinic agonists (pilocarpine and carbacholine, CCh), activate muscarinic receptors. At least 5 different muscarinic receptor molecules have been identified (M_1, M_2, M_3 ..). Activation of the M_1 type is illustrated in Fig. 6-3. Activation of the M_2 type activates an inhibitory G-protein, which inhibits adenylcyclase. Muscarinic receptor activation is linked to G-protein activation and second-messenger systems.

Muscarinic receptors are blocked completely by atropine, and by antimuscarinic drugs such as homatropine and scopolamine (Table 6-3). These drugs do not block the nicotinic effect of ACh on the postganglionic neurons or on the motor endplate.

1. The sympathetic system consists of short preganglionic and long postganglionic nerve fibres. The parasympathetic system contains long preganglionic and short postganglionic fibres.
2. The chemical transmitter at the target organ is noradrenaline in the sympathetic and acetylcholine in the parasympathetic system.
3. The sympathetic system contains adrenergic receptors (a and b), whereas the parasympathetic system has cholinergic receptors (muscarinic or muscarinergic and nicotinic or nicotinergic).
4. Activation of the cholinergic system serves anabolic functions (ie, stay and play), whereas activation of the noradrenergic system serves catabolic functions (ie, fight, fright or flight).
5. Activation of a_1-receptors increases intracellular [Ca^{2+}], which leads to phosphorylation of protein kinases and thus to a response. Activation of a_2-receptors triggers an inhibition of the membrane adenylcyclase, reducing [cAMP] in the cells. b_1- and b_2-receptors activate adenylcyclase, which increases cAMP production in the cell. Muscarinic receptors are completely blocked by atropine. Activation of M_1-receptors increases intracellular [Ca^{2+}]. Activation of M_2 inhibits adenylcyclase, and through an inhibitory G-protein reduces the formation of cAMP.

Pathophysiology

This paragraph deals with 1. Mushroom poisoning, 2. Carbamate and organo-phosphate poisoning, 3. Xerophtalmia and xerostomia, 4. Tachycardia and bradycardia, 5. Smoking, 6. Phaeochromocytoma, 7. Primary or essential hypertension, 8. Horners syndrome, 9. The Kluver- Bucy-syndrome and emotional disorders. - Altzheimers disease and Parkinson's disease are also related to the autonomic nervous system and described in Chapter 7.

1. Mushroom Poisoning

Among the poisonous mushrooms at least two are related to the autonomic nervous system:

Amanita palterina (false blusher) contains a substance with atropine-effect, which completely blocks the muscarinic cholinergic receptors. Atropine-effects are described below.

Amanita muscaria (red fly agaric) contains atropine-like substances, muscarine and other hallucinogens. Atropine-effects are prominent: Motor unrest, delirium (red fly agaric was used by the Vikings to run berserk), mouth dryness, pupillary dilatation, and tachycardia. Some cases are dominated by muscarine-effects: Glandular secretion (sweat, saliva, tear-flow), miosis, eye pains, bradycardia, cramps, respiratory failure, lung oedema, and coma. Muscarinic symptoms are treated with slow intravenous injection of atropine (1 mg in 1 ml). - Ventilation with the mouth-to mouth-method may become fatal for the rescuer.

2. Carbamate And Organo-Phosphate Poisoning

These substances (carbaryle, dimethoate, melathion, parathion etc) are used as insecticides in agriculture. They are anti-cholinesterases so they accumulate acetylcholine in the tissues. The clinical picture is due to the muscarinic and nicotinic effects of acetylcholine. The muscarine-effects are described above. Other symptoms are nausea, vomiting, muscle weakness, paresthesia, bronchospasm, shock and respiratory arrest. Atropine (1 mg iv.) is given repeatedly to obtain complete atropine blockade beginning with pupillary dilatation. Artificial ventilation is often imperative. Cholinesterase-reactivators are tried in desperate situations.

3. Xerophtalmia and xerostomia

Xerophtalmia and xerostomia is dryness of the conjunctiva and the cornea, and dryness of the mouth, respectively. Most of these disorders are of unknown origin (ie essential or primary), although an autonomic cause may be suspected. Xerophtalmia and xerostomia occurs in connection with anaesthesia due to the use of atropine-like substances in order to reduce secretion.

4. Tachycardia and bradycardia.

A balance between the parasympathetic and the sympathetic system normally determines cardiac rhythm. The parasympathetic system predominates. A relative dominance of the sympathetic tone towards the heart leads to tachycardia (ie, a heart rate above 80 beats per min), and a further relative dominance of the parasympathetic system leads to bradycardia (ie, a heart rate below 50 beats per min). Fluctuations of the autonomic tone leads to phasic changes of the sinus node activity. During inspiration, the parasympathetic dominance falls and the heart rate becomes rapid, whereas during expiration parasympathetic dominance increases and the heart slows down. This phenomenon is found in children, and it can even be found in healthy subjects of high age. The phenomenon is called sinus arrhythmia.

5. Smoking

Cigarette smoking is common among teen-agers, and more girls than boys of 15 years smoke cigarettes. Activation of *nicotinic cholinergic receptors* is fast and does not require G proteins.

Nicotinic receptors stimulated by acetylcholine open a Na^+-channel and depolarise the cell membrane. Nicotine stimulates nicotinic receptors on the postganglionic neuron of all autonomic ganglia.

The number of cigarettes smoked and the number of years smoked seem to increase the number of nicotinic cholinergic receptors in brain tissue. Smoker's dependency may depend upon the number of nicotinic receptors. Smokers, deprived of the usual daily dose of nicotine in cigarettes, become depressed in mood, they feel stress and they are un-concentrated. This is called abstinence and the cause is lack of nicotine. The mood of the smoker is immediately improved by smoking. Although smoking is nicotine addiction, nicotine seems to be one of the less dangerous molecules in smoke. Polycyclic aromatic hydrocarbons and nitrosamines are potent carcinogens and mutagens. Such substances release proteinases from granulocytes and macrophages, whereby elastin is destroyed resulting in multi-site lung degeneration or emphysema. The life-threatening dangers are due to cancer and atherosclerosis: Lung cancer, chronic bronchitis and emphysema, cerebral stroke, ischaemic heart disease, peripheral vascular disease, bladder cancer, and memory problems. Male smokers have an increase in the number of abnormal spermatozoa, and pregnant female smokers have an increase in neonatal mortality. Female smokers above 30 years using anti-pregnancy pills have a 40-fold higher risk of cerebral stroke than their non-smoking control group.

6. Phaeochromocytoma

Some patients suffer from attacks of severe hypertension due to adrenaline hypersecretion. The hypertension is caused by release of large amounts of adrenaline from a *medullary tumour* of chromaffine cells. The attacks are sometimes fatal. The diagnosis is important, because the patient can be cured by surgical abolition of the tumour.

7. Primary or essential hypertension

in general begins as a condition with sympathetic overactivity (see Chapter 9).

8. Horner's syndrome

refers to a condition with miosis, facial vasodilatation and loss of facial sweating and enophtalmus

due to damage of the sympathetic nerve supply from the T1-T2-segments to the eye and facial skin.

9. The Kluver- Bucy-syndrome

refers to an *emotional disorder* with bilateral temporal lobe lesions. The temporal cortex, hippocampus and the amygdaloid body are damaged. Mental blindness (visual agnosia) is the inability to recognise objects seen. Besides mental blindness, the syndrome consists of loss of short-term memory, and hypersexual behaviour incompatible with normal social adaptation. The hypersexual behaviour is related to the visual agnosia.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Altzheimers disease is a primary cortical brain atrophy with premature ageing of the brain. Lack of the acetylcholine-producing enzyme choline acetyltransferase and of acetylcholine is characteristic.

B. Mushroom poisoning can be dominated by muscarine-effects: Glandular secretion (sweat, saliva, tear-flow), miosis, eye pains, bradycardia, cramps, respiratory failure, lung oedema, and coma.

C. Cannons law: The peristalsis in the small intestine is polarised, so it always proceeds in the aboral-oral direction.

D. Cannons emergency reaction is an immediate sympatho-adrenergic response to life dangerous situations, with both sympatho-adrenergic and parasympathetic overactivity. The last phenomenon includes vagal cardiac arrest with involuntary defecation and urination.

E. The sympathetic system contains adrenergic receptors (a and b), whereas the parasympathetic system has cholinergic receptors (muscarinic or muscarinergic and nicotinic or nicotinergic).

II. Each of the following five statements have True/False options:

- **A**. Mydriasis refers to contraction of the pupil by sympathetic stimulation.
- **B**. SIF cells are small intensity fluorescent cells, which possess muscarinic receptors and contain vesicles filled with dopamine. Adequate stimulation releases dopamine, which interacts with dopamine receptor (D_2) on the postsynaptic cell body and modulates the effect of acetylcholine.
- **C**. Apraxia refers to a condition with lack of ability to recognise and interpret a sensory stimulus.
- **D**. Phenoxybenzamine and phentolamine block the a-receptor.
- **E**. Adrenaline stimulates the ascending reticular system in the brain stem, thus keeping us alert.

III. Each of the following five statements have True/False options:

A. Argyll-Robertson's pupillary syndrome with small, light-refractive pupils and maintained convergence response to near vision is a typical sign in acute syphilis.

B. Besides catecholamines, ACTH is also released during stress by increasing hypothalamic signals. ACTH stimulates the glucocorticoid and to some extent the mineralocorticoid secretion through cAMP.

C. All autonomic nerves have ganglia outside the CNS in contrast to the somatic nervous system.

D. When stimulated nicotinic receptors work through a slow cascade of events.

E. Catecholamines dilatate the bronchial airways.

6. Case History

A 19-year-old female is in hospital with a cranial lesion caused by a fall from her horse. The following clinical signs are found: 1) speech troubles and hoarseness, 2) swallowing problems and paresis of the soft palate, 3) rapid heart rate, and 4) dilatation of the stomach with vomiting.

Lesion of a certain cranial nerve can explain all symptoms and signs.

- *1. What is the name of the nerve?*

- *2. What is special about this particular lesion?*

Try to solve the problems before looking up the answers .

Highlights

- *The autonomic nervous system mediates neural control of the internal milieu despite substantial environmental changes. The autonomic system directly influences smooth muscles, glands and the heart through its two subdivisions, the sympathetic and the parasympathetic system. The two subdivisions function in a dynamic balance aiming at homeostasis.*
- *The sympathetic system consists of short preganglionic and long postganglionic nerve fibres. The parasympathetic system contains long preganglionic and short postganglionic fibres.*
- *The chemical transmitter at the target organ is noradrenaline in the sympathetic and acetylcholine in the parasympathetic system.*
- *Carbamate and organo-phosphate poisoning. These substances (carbaryle, dimethoate, melathion, parathion etc) are used as insecticides. They are anti-cholinesterases, so they accumulate acetylcholine. The clinical picture of poisoning is due to the muscarinic and nicotinic effects of acetylcholine. Atropine (1 mg iv.) is given repeatedly to obtain complete atropine blockade beginning with pupillary dilatation.*
- *The sympathetic system contains adrenergic receptors (α and β), whereas the parasympathetic system has cholinergic receptors (muscarinic or muscarinergic and nicotinic or nicotinergic).*
- *Activation of the cholinergic system serves anabolic functions (i.e., stay and play), whereas activation of the noradrenergic system serves catabolic functions (i.e., fight, fright or flight).*
- *Activation of α_1-receptors increases intracellular [Ca^{2+}], which leads to phosphorylation of protein kinases and thus to a response.*
- *Activation of α_2-receptors triggers an inhibition of the membrane adenylcyclase, reducing [cAMP] in the cells.*
- *β_1- and β_2-receptors activate adenylcyclase, which increases cAMP production in the cell.*
- *Muscarinic receptors are completely blocked by atropine. Activation of M_1-receptors increases intracellular [Ca^{2+}]. Activation of M_2 inhibits adenylcyclase, and through an inhibitory G-protein reduces the formation of cAMP.*
- *The intrinsic enteric nervous system consists of two sets of nerve plexi. The submucosal (Meissner) plexus mainly regulates digestive glands, whereas the myenteric (Auerbach) plexus, located between the longitudinal and the circular muscle layers, is primarily connected with gut motility.*
- *Sympathetic activity (excitement or pain) causes a large pupil, and parasympathetic activity (light or near- sight) causes a small pupil.*
- *Parasympathetic activity controls salivation, gastrointestinal functions (with the enteric nervous system), emptying of the bladder and defaecation.*
- *The limbic system including the hypothalamus controls vital autonomic functions and emotional behaviour.*

Further Reading

- Brodal, A (1981) *Neurological anatomy.* Oxford University Press, New York.
- Loewy, A.D. and K.M. Spyer (editors). "Central Regulation of Autonomic Functions." *Oxford University Press*, N-Y., 1990.

Return to content

Return to chapter 6

Chapter 7.

Neurological And Psychiatric Diseases

Study Objectives

- To *define* anorexia and bulimina nervosa, astrocytes, bradykinesia, delirium, delusions, dementia, depression, dystonia, Huntington‘s disease, mania, manic-depressive psychosis, meningitis, microglia, multiple sclerosis, myoclonus, oligodendroglia, paranoid delusions, Parkinson‘s disease, schizophrenia, sleep apnoea, sleeplessness, and status epilepticus.
- To *describe* or *explain* the pathophysiology of brain injuries, brain inflammations, brain tumours, Parkinson‘s disease, manic-depressive psychosis, schizophrenia, epilepsy, status epilepticus, and common nervous and stress-related psychiatric disorders.

Principles

- *Glutamate is a major neurotransmitter, which act on several classes of excitatory amino acid receptors.*

- *Many neurological disorders are mediated by neuronal injury involving excessive stimulation of glutamate receptors. Glutamate antagonists are used in clinical trials.*

Definitions

- **Akathisia** is an extrapyramidal defect with swaying and twisting body dyskinesia.

- **Anorexia nervosa** is an eating disorder in adolescent females and males resulting in severe malnutrition. The patient has an intense wish to be thin. Biological and psychological factors are involved. In some cases there is a regression into childhood, where the girl tries to escape from the problems of puberty and adolescence.
- **Astrocytes** are specialised neuroglia, which separate nerve pathways, buffer extracellular potassium and repair nerve injuries.
- **Bradykinesia** is a term for slowing of voluntary movements found in Parkinsons disease.
- **Bulimina nervosa** refers to a condition, where the patient is preoccupied with food and periodically eats excessively. The patient sometimes avoids overweight by self-induced vomiting just after binge eating.
- **Chorea** refers to rapid involitional hyperkinesia with jerky movements of the limbs.
- **Cognitive brain functions** are intellectual processes such as calculation, judgement, language, learning ability, memory, orientation and thinking.
- **Computed tomography** (CT) is a technique using X-rays moving across a slice of brain (tomia, to cut). Normal brain tissue absorbs X-rays differently from tumour tissue, infarcted tissue, and coagulated blood and brain oedema. The radiation passing in a certain direction is recorded with scintillation detectors, and a computer processes the signals.
- **Conversion** refers to Freud’s hypothesis that mental energy can be converted into physical symptoms and signs (abdominal pain, blindness, double vision, deafness, muteness, fits with dramatic movements, artistic gait disturbances, hysterical paresis with normal muscle tone and deep reflexes, crude tremor, sensory loss, stigmatisation).
- **Delirium** is an acute impairment of consciousness also called *toxic confusion*.
- **Delusion** is an abnormal belief arising from distorted judgement.
- **Dementia** (*senility* or ageing of the brain) disturbs almost all cognitive brain functions, whereby the personality of the patient is completely changed.

- **Depression** is characterised by early morning waking with unresponsive sadness, guilt feeling, suicidal feelings, and lack of a precipitating factor.
- **Dissociation** refers to an apparent dissociation between different mental activities. An example is a mentally protective cover of enjoyment (euphoria) in terminal cases of painful cancer or AIDS (French: belle indifference).
- **Dystonia** are abnormal involitional muscle contractions that produce abnormal movement patterns and postures.
- **Hallucinations** are sense impressions experienced in the absence of external sense stimuli.
- **Hemiparesis** means weakness of the limbs of one side – frequently occurring in upper motor neuron lesions.
- **Hemiplegia** means total paralysis of the limbs of one side of the body.
- **Huntington's disease** is a chorea-condition with hypotonia, dementia and involuntary movements. This is an autosomal genetic defect on chromosome 4.
- **Magnetic resonance imaging** (MRI) is a scanning technique, where protons in a strong magnetic field are bombarded with radiofrequency waves in order to produce images. MRI scanning can picture brain tumours, multiple sclerosis lesions, and syringomylia among others. MRI scanning can even separate white from grey matter. MRI scanning is replacing myelography, because it can visualise spinal cord compression, spinal cord tumours and malformations.
- **Mania** refers to a psychiatric disorder with periodic elevations of mood with overactivity, restlessness, fast talk, excessive energy, increased sexuality, overwhelming self-confidence, and insomnia.
- **Manic-depressive psychosis** covers severe abnormalities of mood. Mood ranges from severe depressive psychosis over moderate and minor depression, sadness, normal mood, happiness, euphoria, hypomania and severe mania.
- **Meningitis** refers to inflammation of the meninges.
- **Microglia cells** proliferate and move to the site of nerve injury, where they transform to large phagocytes, which remove debris.
- **Multiple or disseminated sclerosis** refers to a common neurological disease caused by inefficient myelin production in the oligodendroglia.
- **Myoclonus** often occurs at night and refers to brief contractions or jerks of one or more muscles. Myoclonus is often related to metabolic or drugs toxicity.
- **Neurosis** refers to a psychiatric disorder in which the personality as a whole is unimpaired and without psychotic symptoms. Neurosis is an amplified, more than normal reaction to mental stress such as anxiety, depression and irritability.
- **Oligodendrocytes** produce myelin sheaths around axons in the CNS just like Schwann cells do in the peripheral nervous system.
- **Paranoid delusions** (*paranoia*) are abnormal beliefs dominated by fear of persecution.
- **Parkinsonism** is a dopamine-deficiency state of the forebrain with bradykinesia, tremor, and rigidity.
- **Psychosis** refers to a psychiatric disorder impairing the whole personality and functioning of the individual (insight, sense of reality, delusions, and hallucinations).
- **Repression** means exclusion of memories, impulses, and emotions from consciousness, because these elements would cause anxiety and stress.
- **Sleep apnoea** is periodic breath holding during sleep. Sleep apnoea often occurs with snoring and airway obstruction in obese patients or in patients with chronic obstructive lung disease.
- **Schizophrenia** means splitting of the mind or disconnection of cognitive and emotional psychic functions. Schizophrenia is a psychosis with hallucinations, dissociation of ideas, intense fear, and paranoid delusions.
- **Status epilepticus** is an emergency condition, where consciousness is not regained between grand mal seizures lasting more than half-an-hour.
- **Tics** refer to focal myoclonus with repeated twitching of facial or neck muscles. Tics may even begin in childhood for unknown reasons and they are extremely resistant to therapy.
- **Tremor** or *shaking* can be caused by hyperthyroidism and by Parkinsonism, but it is also a

typical side effect of alcohol, narcotics and drug abuse.

Essentials

This paragraph deals with *1. Nerve cells, 2. Ion channels, 3. Neurotransmitters, and 4. Signal transduction.*

1. Nerve cells

The cells of the Central Nervous System (CNS) consist of neuroglia and of neurons.

Neuroglial cells outnumber all the neurons in the CNS and they constitute half of the brain volume. Glial cells are known to sheath and protect neurons. Glial cell membranes contain receptors and ion channels. They help control the environment of neurons and thus contribute to the function of neurons. We have three types of neuroglial cells. *Microglia cells* are small cells scattered throughout the nervous system. Microglia proliferate after injury and move to the site of injury. Here they transform to large phagocytes, which remove the debris. *Oligodendrocytes* produce myelin sheaths around axons in the CNS just like Schwann cells do in the peripheral nervous system. *Astrocytes* separate nerve pathways, buffer extracellular [K^{+}] , and repair nerve injuries.

2. Ion-channels

Two classes of proteins span the cell membrane and control ion transfer. The first class is Na^{+}-K^{+}-pumps and other ATP-demanding pumps that actively move ions across the membrane against their electrochemical gradient (Fig. 7-1). The second class is *channels* or pores through which specific ions can pass. Ions traverse such an open channel along the electrochemical gradient. The small ion permeation through the cell membrane at rest is referred to as *leak current* (Fig. 7-1). The typical Na^{+}-channel *opens promptly* in response to *depolarisation* (voltage-gated opening) and also closes rapidly, although the cell is still depolarised. The channels then remain inactivated for a short period. Opening of Na^{+}-channels increases the flux of Na^{+} into the neuron, and depolarizes the membrane, so the effect is excitatory.

Fig. 7-1: Ion channels in a neuronal membrane, where the Na^{+}-channel is closed and the K^{+}-channel is open at rest.

Closure of K^{+}- or Cl^{-} -channels decreases the flux of K^{+} out of the neuron or decreases the flux of Cl^{-} into the cell. These events also depolarise the membrane, and again the effect is excitatory.

Obviously, closure of Na^{+}-channels or opening of K^{+}- or $Cl^{—}$channels have an inhibitory effect by hyperpolarisation.

Voltage-gated Na^{+}-, K^{+}-, and Ca^{2+}-channels comprise subunits with membrane spanning domains (Fig. 7-1). There is amino acid sequence homology in the transmembrane helices of these channels. The *channel protein* includes a charged group, which is sensitive to the electric field across the membrane. During depolarisation the gate opens, which changes the whole channel, rendering it much more conductive to specific ions. Each channel continues to open, close and reopen several times during depolarisation. The *fast* Na^{+}-channels close rapidly and are inactivated during depolarisation due to a *channel polypeptide* located on the cytosolic side (Fig. 7-1).

Opening of Na^{+}-channels requires or results in a rapid change of potential. Partial and slow depolarisation, inactivate a critical fraction of the Na^{+} -channels. This is called *voltage-inactivation*. Voltage inactivation of Na^{+}-channels is involved in the accommodation and in the

refractory periods.

3. Neurotransmitters

Neurotransmitters are signal molecules used by neurons to communicate with each other and with target cells. Chemical synapses are specialised. The presynaptic terminal contains mechanisms for production and storing of neurotransmitters that are released in response to depolarisation.

The postsynaptic membrane carries protein receptors that can detect and identify different neurotransmitters and initiate appropriate responses to stimulation. Finally, there are adequate mechanisms for degradation and reuse of transmitters to ensure rapid onset and offset of arriving signals. Chemical synapses are the sites of action for many drugs.

Neurotransmitters can be divided into two groups: Classical rapid acting non-peptide neurotransmitters (Box 7-1) and putative, slowly acting peptides (Box 7-2).

During development some process of differentiation determines the type of neurotransmitter that a given neuron will synthesise, store, and release. Thus, a single neuron releases the same neurotransmitter from all its synapses - an assumption, which has been generally accepted for years as *Dale's law*.

Recent advances indicate that some neurons can release more than one neurotransmitter. Up to 4 neuropeptides have been localised to a single neuron.

Two or more transmitters released together are called *co-transmitters*. One member of each pair of transmitters appears to be a peptide. Perhaps these peptides act by enhancing the message transferred with the rapid neurotransmitters.

Classical neurotransmitters are substances such as acetylcholine, noradrenaline, dopamine, GABA (gamma-aminobutyric acid), glycine etc (Box 7-1). Their diffusion pathway is short, and they have no other function than neurotransmission.

Catecholamines (dopamine, noradrenaline, and adrenaline) are neurotransmitters both in the sympathetic system and in the CNS. Noradrenaline is the transmitter for most postganglionic sympathetic fibres (some of these fibres use acetylcholine).

In the CNS catecholamines are found in several brain nuclei: *Dopaminergic neurons* are found in the substantia nigra, *noradrenergic neurons* in locus coeruleus, and *serotonergic neurons* in the raphe nuclei and in many midbrain structures.

Serotonin (5-hydroxytryptamine) is a transmitter in brainstem nuclei (in particular the Median raphe) concerned with wakefulness and behaviour. Adrenaline, noradrenaline, dopamine, and serotonin serve as fast neurotransmitters in the CNS in the same way as the Enzyme- inactivated acetylcholine. The most important excitatory amino acid (EAA)-receptors are the glutamate receptors, the N-methyl-D-aspartate (NMDA)-receptors, and the alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-receptors. The glutamate receptor is a typical non-NMDA-receptor. NMDA-activated ion channels are only active, when the membrane is depolarised, and they are specific for Ca^{2+}- Na^{+}- and K^{+}-penetration. AMPA-activated ion channels are specific for Na^{+}-and K^{+}-permeability and depolarise the cell membrane.

Box 7-1: Classical, rapidly acting non-peptide transmitters with their cell-surface receptor type and action/location

Substance	Receptor type	Action/ location
Acetylcholine	**Cholinergic**	**Excitatory/Autonomic ganglia etc**

Adrenaline	**Adrenergic**	**Excitatory/Locus coeruleus etc**
Noradrenaline	**Adrenergic**	**Excitatory mainly (awake, mood)**
Dopamine	**Dopaminergic**	**Inhibitory/Substantia nigra**
Histamine	**Histaminergic**	**Excitatory/Hypothalamus**
Serotonin	**Serotonergic**	**Inhibitory/Median raphe of brain stem**
GABA	**GABA-receptor**	**Inhibitory/CNS**
Glycine	**GABA-receptor**	**Inhibitory/Spinal cord**
Aspartate	**NMDA-receptor**	**Excitatory/CNS**
Glutamate	**non-NMDA-receptor**	**Excitatory/Cortex**
Nitric oxide (NO)	**NO diffuses into cells**	**Behaviour and memory/CNS** **Long-term potentiation**

Neuropeptides

Several of the neuropeptides are well-known hormones. They are synthesised in the soma of the neurons and reach the axon terminals by fast axonal transport.

The secretion of both gonadotropins (LH and FSH) from the anterior pituitary is controlled by the hypothalamic luteinizing hormone-releasing hormone (LHRH). The thyrotropin-releasing hormone (TRH) and somatostatin from the hypothalamus control the pulsatory secretion of thyrotropin or thyroid stimulating hormone (TSH) from the anterior pituitary. The secretion of growth hormone (GH) and prolactin from the anterior pituitary is controlled by two hypothalamic hormones: GH-inhibiting hormone (GHIH or somatostatin) and GH-releasing hormone (GHRH).

Many of these peptides are cut off from a big mother molecule: *pro-opio-melanocortin* (POMC). Cleavage of POMC in the anterior pituitary lobe releases adrenocorticotropic hormone (ACTH) and b -lipotropin. Cleavage of ACTH in the intermediate pituitary lobe releases melanocyte-stimulating hormone (a -MSH) and corticotropin-like intermediate lobe peptide (CLIP). Cleavage of b -lipotropin releases b -MSH and b -endorphin. Endorphins and enkephalins bind to opiate receptors and are called endogenous opiates.

Cleavage of a pre-pro-hormone produced in the hypothalamus releases the two octapeptides, oxytocin and vasopressin together with two neurophysin molecules. The two octapeptides are moved to the neurohypophysis by axoplasmatic transport.

Box 7-2: Putative, slowly acting peptide neurotransmitters are water- soluble and binds to cell-surface receptors. Some neuropeptides have non-peptide co-transmitters (-co).	
Substance/co-transmitter	**Action/location**
LHRH	**Gonadotropin release/anterior pituitary**
TRH	**TSH release/ "**

Somatostatin/noradrenaline (-co)	**GH-inhibition/ "**
ACTH	**Stimulates secretion of adrenal cortex hormones/ "**
β -lipotropin	**Lipolysis/Fat cells**
β -endorphin	**Pain release/opiate receptors**
α -MSH	**Melanocyte stimulation/skin**
β -MSH	**" " / "**
Prolactin	**Development of mammary gland/breasts**
Luteotropin (LH)	**Rupture of follicle/ovaries**
Thyrotropin	**Activates adenylcyclase/thyroid follicles**
Growth hormone (GH)	**Regulates growth/the body as a whole**
Oxytocin	**Stimulates myoepithelial cells/milk ducts and uterus**
Vasopressin (ADH)	**Vasoconstrictor. Stimulates renal water reabsorption**
Enkephalins/adrenaline (-co)	**Pain release/opiate receptors**
Substance P/serotonin (-co)	**Smooth muscle contraction/neurons, endocrine cells**
Gastrin	**Gastric acid secretion/neurons, endocrine cells**
CCK/dopamine (-co)	**Bile and pancreatic enzyme secretion/neurons**
VIP/acetylcholine (-co)	**Smooth muscle relaxation, secretion/neurons**
Insulin	**Reduces blood glucose/pancreatic β -cells**
Glucagon	**Increases blood glucose/ pancreatic α -cells**
Angiotensin II	**Aldosterone secretion. Arteriolar constriction**
Bombesin	**Pancreatic enzyme secretion. Synaptic transfer**
Bradykinin	**Vasodilatator. Synaptic transfer**
Calcitonin	**Bone is remodelling. Synaptic CNS transfer**

CCK stands for cholecystokinin and VIP for vasoactive intestinal polypeptide.

The *gut-brain peptides* are described in Chapter 22, and the *hypothalamo-pituitary peptides* in Chapter 26.

4. Signal transduction

Signal transduction is a cascade of processes from the *receptor-hormone* binding to the final cellular response. Many hormones and neurotransmitters raise the concentration of a second messenger in the target cell via guanyl triphosphate (GTP) and act through it. The receptor-

hormone complex activates a GTP-binding protein (so-called G-protein) which controls and amplifies the synthesis of the second messenger. Hereby, each hormone molecule can produce many molecules of second messenger such as cAMP or cGMP. Furthermore, each protein kinase unit can phosphorylate many molecules of its substrate, resulting in a great amplification factor.

G-proteins function as molecular switches, regulating many cellular processes, such as activation of intracellular enzymes (protein kinase, phosphorylase), activation of membrane enzymes and channels, and activation of gene transcription.

G- protein-linked receptors form a family, which has evolved from a common ancestor. Most G-proteins are membrane bound heterotrimers (a, b, g) and exist in an activated state, where it has high affinity for GTP, and an inactive state, where the molecule prefers GDP.

Hydrophilic (lipophobic) hormones, such as acetylcholine and many peptides, bind to membrane receptor proteins, and the hormone-receptor binding activates the enzyme phospholipase C via active G-protein.

Multiple receptor subtypes can co-exist on a single cell. The b-adrenergic receptors are both stimulated by noradrenaline and both activate a stimulatory G-protein (Gs in Fig. 7-2). Gs activates adenylcyclase, which increases the production of the second messenger cAMP.

Fig. 7-2: A single cell with both β_1- and β_2 -adrenergic receptors. Both receptors activate adenylcyclase through stimulatory G-proteins (Gs).

The result is an additive cellular response.

A single cell with b_1-adrenergic receptors activating adenylcyclase through a stimulating G-protein, and a_2 -adrenergic receptors, inhibiting adenylcyclase via an inhibitory G- protein, results in opposite signals when stimulated by noradrenaline (Fig. 7-3).

Fig. 7-3: Antagonistic reactions to noradrenaline in a single cell.

Noradrenergic stimulation of another single cell with b_1-adrenergic receptors activating adenylcyclase through a stimulating G-protein, and a_1 -adrenergic receptors which activate phospholipase C leads to production of two phosphorylated derivatives of phosphatidylinositol (PI): PI-phosphate (PIP) and PI-diphosphate (PIP_2). Phospholipase C cleaves (PIP_2) into inositoltriphosphate (IP_3) and diacylglycerol (DAG) (see Fig. 7-4).

Fig. 7-4: Independent reactions to the stimulation of two subtypes of adrenergic receptors on a single cell.

IP_3 is a second messenger that binds to Ca^{2+}-channels in the endoplasmic reticulum (ER), so that Ca^{2+} is released to the cytosol. DAG and Ca^{2+} are second messengers that activate *protein kinase C*, which is involved in the regulation of cellular metabolism, growth and many other processes. Inactive cytosolic protein kinase C is activated by Ca^{2+}, and binds to the inner surface of the membrane, where DAG activates it. Ca^{2+} and proteinkinase C catalyses the transfer of phosphate from ATP to the effector proteins. Independent reactions are generated by the presence of these two subtypes of adrenergic receptors.

Specific receptor-ligand bindings also activate phospholipase A_2 via a G-protein. Phospholipase A_2 cleaves membrane phospholipids, and releases arachidonic acid (AA) in the cells. AA activates a precursor to platelet activating factor (PAF) termed lyso-PAF. AA is also the

precursor for the synthesis of endoperoxides, prostacyclin, thromboxanes (mediates platelet aggregation and vasoconstriction) and leucotrienes.

Insulin and related growth factor peptides bind to membrane receptors that are glycoproteins protruding from the membrane. The insulin receptor is typical for this receptor family. Peptide binding to the outer receptor subunit stimulates a protein tyrosine kinase on the inner receptor subunit. This phosphorylates tyrosine residues, both on the receptor itself and on other proteins. The tyrosine kinase activity is essential for signal transduction.

Examples of growth factors are: EGF (epidermal growth factor), FGF (fibroblast growth factor), IGF-II (insulin-like growth factor-II), NGF (neural growth factor) and PDGF (platelet-derived growth factor).

Protein tyrosine kinase activity is abnormally high in certain types of cancer and cellular modification. This can be caused by growth factors or by a mutation of the tyrosine kinase part of the trans-membraneous receptor. Mutations of one gene localised on chromosome 10 can lead to four different syndromes: Multiple endocrine neoplasia, Hirschprung's disease, medullary thyroid carcinoma, and Phaeochromocytoma (se Chapter 28).

The final step is often phosphorylation or dephosphorylation of a particular key or effector protein. Protein kinases and dephosphorylation accomplish phosphorylation by protein phosphatase. Second messengers (cAMP, cGMP, IP_3, DAG, and Ca^{2+}) control the activities of protein kinases such as cAMP-dependent protein kinase A, cGMP-dependent protein kinase, calmodulin-dependent protein kinase, and protein kinase C. Calmodulin binds 4 Ca^{2+}.

The phosphorylation level of an enzyme or an ion channel determines and triggers the physiological response.

Protein phosphatase reverses the effect of protein phosphorylation. The phosphatase dephosphorylates the key proteins, and thus opposes or stops the physiological response.

The free radical gas nitric oxide (NO) is a neuronal messenger in both the central and the peripheral nervous system. The NO gas is membrane permeant, and can bypass normal signal transduction in synapses.

Two types of NO synthase (NOS) have been identified: constitutive Ca^{2+}- calmodulin dependent enzyme, and inducible Ca^{2+} - independent enzyme. Both enzymes are flavoproteins containing bound flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). Both enzymes require the cofactors NADPH and tetrahydrobiopterin (BH_4). NOS catalyses the conversion of L-arginine to citrulline and NO in two step when activated by the Ca^{2+} - calmodulin complex, muscarinic agonists, or other activators (Fig. 7-5).

Fig. 7-5: The biosynthesis of NO with cell-cell effects on target cells, such as smooth muscle cells etc. Non-adrenergic non-cholinergic (NANC) relaxation of the gastrointestinal tract (GIT) and the genito-urinary (GU) system is shown.

NO diffuses to the target cell, where it activates *guanylcyclase* resulting in the formation of cGMP (Fig. 7-5). NO is labile. Hence, a carrier for NO has been postulated. The biological effect of NO is mediated by an increase in cGMP levels, and the effects on target cells are shown below (Fig. 7-5). *Nitric oxide effects* are further developed in Chapters 2, 3, 4, 9, 12, 22, 25 and 30.

The *NO biosynthetic pathway* can be interfered with at several points. Nitrovasodilatators, such as nitro-glycerine, have been used for over a century to treat cardiac cramps or angina pectoris (see Chapter 9). Nitrovasodilatators act by releasing NO and thereby causing coronary vasodilatation. Nitric oxide synthase is inhibited by *L-arginine analogues* (Fig.7-5).

One of the effects of NOS-inhibitors is an increase in blood pressure. NOS-inhibitors are effective in treating endotoxic shock. This is a condition caused by increased NO synthesis by inducible NO synthase, where the sympathetic vasoconstrictors are often ineffectual. The cofactors, like BH_4, can also be manipulated, e.g., by *anticancer drugs*.

Pathophysiology

The first part I concerns neurological disorders of cerebrovascular and brain parenchymal origin (eg, brain trauma, epilepsy, movement disorders, multiple sclerosis, inflammations), and end-up with the differential diagnosis between dementia and delirium - a situation of life-threatening consequences.

The second part (II) is confined to psychiatric disorders (eg, neuroses and psychoses).

I. Cerebrovascular Disorders

These disorders include parenchymal brain damage, and the main groupings are *1. Stroke and minor stroke*, *2.Brain lesions*, *3.Epilepsy*, *4.Movement disorders*, *5.Multiple sclerosis*, *6. Inflammations* and *7.Intracranial tumours*. *8. Dementia contra delirium*

1. Stroke and minor stroke

Thrombo-embolism of the middle cerebral artery is a common cause of stroke

(suddenly occurring unconsciousness with hemiplegia). The *middle cerebral artery* is the artery most often occluded by thrombo-embolism or by atherosclerotic material (Fig. 7-6). Risk factors for stroke are related conditions such as inactivity, obesity, hypertension, smoking, hypercholesterolaemia, hypertriglyceridaemia and oral contraception (see also Chapter 10).

Thrombo-embolism causing a stroke usually leads to *cerebral infarction*. Occlusion of the internal carotid artery or the middle cerebral artery causes infarction of the internal capsule with *aphasia* (lesion of the dominant hemisphere), contralateral *hemiplegia*, and *areflexic flaccid limbs*. The lesion blocks the corticospinal tract as it traverses the internal capsule. Glutamate is released in the ischaemic tissue.

The *lateral descending system* consists of the corticospinal, the corticobulbar and the rubrospinal tracts. Interruption of the lateral descending system to the brainstem and spinal cord causes contralateral paresis, weakness of the finger muscles with loss of fine movements, and loss of the abdominal and cremasteric reflexes. Following the initial spinal shock, a series of release signs are found: *Positive* sign of Babinski, *spasticity* (eg, a motor condition dominated by increased tonic and phasic stretch reflexes), *foot clonus*, and abnormal flexion reflexes. This syndrome is termed the *upper motor neuron disease* or the *pyramidal tract syndrome*. The positive sign of Babinski is a slow dorsiflexion of the big toe and fanning of the other toes, when the sole of the foot is stroked laterally from the heel and forward.

Fig. 7-6: A stroke patient with thrombo-embolism of the right middle cerebral artery.

A pure interruption of the *corticospinal tract alone* (the medullary pyramid) causes *decreased muscle tone* and loss of finger movement control, but it does not induce spasticity and flexion reflexes as the lesion of the lateral descending system. Lesions of the medial descending system (ie, the vestibulospinal, reticulospinal, and tectospinal tracts) causes impaired control of the axial muscles, loss of balance during walking, and loss of rightening reflexes. The fine finger movements are normal.

A patient with thrombo-embolism in the right hemisphere suffers from left-sided hemiplegia. Head and eyes (conjugated eye deviation) are typically turned toward the lesion.

A coma patient, who lacks conjugated eye deviation, is likely to suffer from brainstem injury with severe damage of the reticular activating system. Such a condition has a grave prognosis.

Arteriosclerotic brain arteries, micro-aneurysms and larger aneurysms rupture and it bleed into the brain tissue. This primary intracerebral haemorrhage is seen in patients with hypertension. The clinical picture is the same as in *thrombo-embolic stroke*, although cortical tissue damage with unconsciousness is more common here. Coma is the deepest stage of unconsciousness, where the patient is completely without reactions.

Rupture of an arteriosclerotic brain artery or an aneurysm causes bleeding. Bleeding can interrupt the corticospinal tract as it traverses the internal capsule. Such a block of the excitatory pathways to the spinal cord results in severe contralateral paresis, weakness of the finger muscles with loss of fine movements, and loss of superficial reflexes (the abdominal and cremasteric reflexes).

This is a typical result of interruption of the lateral descending system, and often termed the *upper motor neuron disease* or the pyramidal tract syndrome. The capsula interna damage interferes with other cortical efferents to the basal ganglia, the thalamus, and the pons. Therefore, the symptoms and signs are much broader than those after injury of the corticospinal system only are.

The stroke patient can slide into deep unconsciousness termed *coma*. Coma is the deepest stage of unconsciousness. The comatous patient is completely without reactions to even the strongest stimulus. The EEG is dominated by delta waves. When coma proceeds into *brain death*, the EEG trace shows no electrical activity.

Glutamate is released during cerebral ischaemia, such as the ischaemia occurring after a stroke. In animals, N-methyl-D-aspartate-receptor (NMDA) antagonists can prevent ischaemia-induced neurodegeneration.

Microembolism or fall in cerebral perfusion may cause a syndrome called *transient ischaemic attacks*. Small emboli (clots, atherosclerotic material, air or fat) occlude the small arterioles and the brain capillaries.

Hypertension can lead to fibromuscular hyperplasia of the walls of parenchymal brain arteries and arterioles. The proliferation reduces the calibre of the arterioles and leads to microinfarction. Multiple microinfarcts impair cognitive functions and lead to dementia.

2. Head injuries

If the retinal artery is temporarily blocked by a microembolus, the patient experiences a sudden transient loss of vision (amaurosis fugax). Temporary bloodflow reduction in the posterior cerebral artery to the medial surface of the temporal lobe causes transient amnesia (memory loss). Transient aphasia is caused by bloodflow reduction to the language comprehension area (Wernicke’s area) of the dominant hemisphere. Transient hypoperfusion of this area causes sensory aphasia (ie, difficulties in understanding written or spoken language, although single words may be recognised). When this occurs in the non-dominant hemisphere, the result is apraxia such as dressing apraxia.

The principal causes of head injuries are road traffic accidents and alcohol abuse. Head injury often results in a simple brain concussion, but due to intracranial bleeding the condition sometimes becomes life threatening. The bleeding is located extradurally (epidurally), subdurally or subarachnoid with different degrees of brain parenchymal damage.

Computerised tomography (CT) scanning has revolutionised the diagnostic work with head injuries. CT scanning reveals non-invasively the location of blood, skull fractures and cerebral contusion. The location of blood is epidural, subdural, subarachnoid, intraventricular or intracerebral.

Head injuries are divided into *2a. Simple concussion and brain contusion, 2b. Epidural haematoma, 2c. Subdural haematoma, 2d.Subarachnoid haemorrhage, and 2e. Intracranial mass lesion.*

2a. Simple concussion and brain contusion

Simple concussion is defined as a *transient loss of consciousness* followed by complete recovery. A short period of amnesia is often related to the loss of consciousness. This is a migraine injury, where the duration of the unconsciousness indicates the severity of brain damage.

Brain contusion refers to brain damage with prolonged coma, amnesia and focal signs. Later on such patients often suffer from chronic impairment of higher cerebral functions and hemiparesis. *Post-traumatic* epilepsy is frequently caused by head injury with coma following depressed skull fractures, brain contusion or intracranial haematoma. Actually, depressed skull fracture causes a high incident of post-traumatic epilepsy.

Traffic accident victims with severe brain damage may develop the so-called *punch-drunk syndrome* (dementia with extrapyramidal signs), which typically is found among professional boxers (Chapter 18) and alcoholics.

2b. Epidural haematoma

The *middle meningeal artery* and its branches are located in the temper-parietal region. Skull fractures in this region or in regions traversing a dural sinus often cause bleeding into the epidural space.

Extradural or epidural haematoma is caused by rupture of the middle meningeal artery due to a skull fracture (Fig. 7-7) or by tearing of dural veins to the sigittal sinus. The skull fracture sometimes is accompanied by CSF loss (eg, rhinorrhoea and otorrhoea).

Following the head injury with a period of unconsciousness, the patient may appear in a good condition, but suddenly he looses consciousness or develops hemiplegia. Early surgical drainage is lifesaving.

A blow to the *temporo-parietal region* may lead to fracture with transection of one or more branches of the middle meningeal artery (Fig. 7-7). *Pulsate bleeding* at the high systolic pressure can dissect the dura mater from the calvarium and form an *epidural compartment* filled with blood. This is a gradual process, because the dura adheres firmly to the bones, and there is usually an asymptomatic interval of 5-6 hours. Since the supratentorial volume is fixed, the expanding haematoma displaces an equal volume from the supratentorial compartment, firstly by reducing the CSF volume, secondly by pressing brain tissue through the orifice as a *trans-tentorial* and possibly also as a *subfalcine hernia* (Fig. 7-7).

Fig. 7-7: Development of epidural haematoma with herniation and displacement of the falx cerebri. The unilateral pupillary dilatation is caused by oculomotor nerve palsy.

As the rising intracranial pressure exceeds the pressure in the large venous sinuses, the veins are compressed and the venous stasis secondly impedes the arterial bloodflow to the brain. The result is *cerebral ischaemia* (hypoxia and hypercapnia), that occurs even with a high systolic arterial pressure. The high arterial pressure elicits a decline in heart rate via the arterial baroreceptors. Hereby, the ventricular filling is increased, whereby myocardial contractility is increased. The cortical impairment is recognised clinically as confusion and disorientation.

The brain tissue is displaced by the growing haematoma, and the tissues of the uncus of the hippocampus are pressed through the tentorial orifice as a *trans-tentorial herniation*. Hereby, the oculomotor nerve (III. cranial nerve) is pressed against the edge of the tentorium and the resulting

nerve palsy is shown as a fixed, dilatated pupil (Fig. 7-7).

The clinical picture is that of a *simple concussion* with a brief period of unconsciousness followed by recovery for some hours. Simple concussion is defined as transient loss of consciousness. The loss is due to traumatic malfunction of neurones in the *reticular formation* of the brainstem. After 4-8 hours of seemingly recovery, the patient suddenly looses consciousness and develops hemiplegia. The transtentorial herniation is recognised, when the patient is found with an ipsilateral dilatated pupil (Fig. 7-8), but terminally both pupils are fixed and dilated. In the terminal phase tetraplegia develops.

The herniated brain tissue also compresses and displaces the brain stem (midbrain, pons and medulla) resulting in venous stagnation of blood and *ischaemia*. Impaired function of the neurones in the *reticular formation* and the *cardio-respiratory control centres* leads to unconsciousness and cardio-respiratory failure. Lack of oxygen for even a short period results in neuronal damage and necrosis, which is irreversible.

An *epidural haematoma* must be recognised and evacuated as soon as possible. Otherwise the bleeding progress until death ensues.

2c. Subdural haematoma

Subdural haematoma is an *accumulation of blood in the subdural space* caused by venous bleeding. The cause is head injury with *latency* between the event and the symptoms (headache, confusion, stupor, coma, delirium, hemiparesis, epilepsy etc). CT scanning confirms the diagnosis. The latency is sometimes so short, that the clinical picture resembles that of extradural haematoma.

The arachnoid is bound to the cerebral hemispheres, but unattached to the dura mater. Veins from the cerebral hemispheres cross the subarachnoid space, penetrate the arachnoid and the dura, and finally enter the dural sinuses. Injuries applied to the frontal or occipital regions initiate shock waves through the liquid brain tissue, whereby the cortical veins are cleaved just before the blood has reached the saggital sinus.

Fig. 7-8: Development of subdural haematoma with transtentorial and subfalcine herniation.

In this way blood accumulate in the subdural space (Fig. 7-8). The bleeding may be from only one vein, and the development may be slow. Latency between the time for injury and the occurrence of the first symptom can be weeks or months. Headache and confusion are unspecific indications in the elderly. *Cognitive functions* are often impaired by bilateral subdural haematoma. Manifest dementia is sometimes misinterpreted as senility. CT, MRI or arteriography confirms the diagnosis. Surgical drainage is performed.

New bleeding may develop acutely with terminal transtentorial herniation. Some types of subdural haematoma resolve spontaneously.

2d. Subarachnoid haemorrhage

Subarachnoid haemorrhage is a *spontaneous arterial bleeding* into the subarachnoid space. The clinical picture can be that of delirium. Diagnosis is confirmed with CT scanning, and neurosurgical closure of the aneurysm is sometimes possible.

Bleeding into the subarachnoid space is most often spontaneous rather than traumatic. The circle of Willis and adjacent vessels is the most frequent site for *saccular* or *berry* aneurysms (Fig. 7-9).

Fig. 7-9: The circle of Willis with saccular aneurysms (black).

The aneurysms *rupture spontaneously*, often at rest and the patient experience a sudden, devastating headache followed by loss of consciousness. The *neck is stiff* and the back is stiff as well.

Subarachnoid or intraventricular blood is clearly demonstrated by CT scanning, and in such cases lumbar puncture is unnecessary (Fig. 7-10).

Fig. 7-10: Subarachnoid and intraventricular bleeding in the left lateral ventricle (left). The blood-filled lateral ventricle is also projected to the base of the brain (right).

Previously, the diagnosis was confirmed by the presence of blood in the CSF. Today, the lumbar puncture is often avoided, as a spinal tap causes a sudden pressure differential between the supra- and infra-tentorial compartments. This may elicit transtentorial herniation with brainstem compression and death. Angiography is performed on patients fit for neurosurgical closure of the bleeding site.

2e. Intracranial mass lesions

located *supratentorially* (above the tentorium cerebelli) can compress the brain towards the tentorium as to block the upward flow of CSF and thus its absorption. Such mass lesions are brain tumours, encephalitis, meningitis, haemorrhages, aneurysms, brain abscesses, and the effect is similar to the effect of brain contusion just analysed.

Hereby, the CSF-pressure below the tentorium cerebelli increases. A rise in CSF-pressure below the tentorium results in *papiloedema*, because it creates a high pressure inside the optic nerve sheath and thus pushes fluid into the optic disc or papilla. Ophtalmoscopy reveals blurring of the edges of the papilla and dilated retinal veins without the normal pulsation. The papilla looks like the top of a champignon. Lesions of the vessels result in visible, retinal haemorrhages. Continuous intracranial pressure monitoring is important during treatment of comatous patients with severe head trauma.

Intracranial mass lesions are almost always surrounded by *cerebral oedema*. Mass lesions that include the cerebral cortex often lead to *epilepsy*.

Cerebral oedema is caused by increased pressure in the brain capillaries or by lesions of their walls.

A rise in cerebral arterial pressure above the upper limit for *autoregulation* (ie, almost constant bloodflow despite rising driving pressure) results in brain oedema. *Brain oedema* compresses intracranial vessels, whereby brain bloodflow is reduced and *brain ischaemia* develops. This is the start of a vicious cycle, because the hypoxia increases the capillary permeability and dilates also the arterioles. Hereby, the brain oedema develops further. Hypoxia also blocks the Na^+-K^+-pump, whereby the brain cells swell (eg, intracellular overhydration). Intravenous infusion of a concentrated osmotic solution such as mannitol drags oedema fluid from the brain tissue, and benefit the patient. A patient in coma, suspected of increased intracranial pressure, can be treated with 1- 2 g mannitol per kg iv., while further procedures are carried out.

Intubation and hyperventilation should also be instituted to any comatous patient. Reduction of P_{aCO2} to 25 mmHg (3.3 kPa) will rapidly reduce intracranial pressure by decreasing cerebral bloodflow and blood volume.

Brain stem compression occurs when intracranial mass lesions above the tentorium damage the ascending reticular activating system (RAS). The high pressure pushes the basal parts of the temporal lobes through the incissura tentorii, and the cerebellar tonsils through the foramen magnum. Brain tissue incarceration with brain stem compression is a serious cause of coma, which must be diagnosed and treated immediately. Suspicion of increased intracranial pressure is

contraindication of lumbar puncture, because it may cause brain stem compression by transtentorial herniation. Accordingly, ophtalmoscopy with the exclusion of *papillary oedema* is necessary before any lumbar puncture.

3. Epilepsy

Epileptic seizures are partial or general. Epilepsy is an *abnormal paroxystic discharge from cerebral neurones* resulting in a condition with clinical consequences.

The normal EEG waves are due to synaptic potentials by groups of neurons including pyramidal cells. An epileptic seizure is characterised by *high voltage-high frequency discharge* (100-200 μV) from large groups of neurons or from the entire cortex.

Partial or focal seizures can be caused by an *epileptic focus* anywhere in the cortex. The causes of focal seizures are acquired lesions such as cysts, tumours, scar tissue, infections, ischaemic lesions. The epileptic discharge causes involuntary muscular contractions on the contralateral side. Foci in the somatosensory cortex produce sensory hallucinations called an epileptic *aura*. These hallucinations precede the epileptic seizure. The aura varies and is particular for a certain patient. Epileptic foci in the visual cortex cause visual auras, while epileptic foci in the vestibular cortex produce an aura-feeling of spinning. *Psychomotor epilepsy* originates in the limbic system and causes emotional hallucinations and muscle contractions. Focal seizures are characterised by high *epileptic spikes* in the EEG. *Motor seizures* originate in the motor cortex of the opposite side, and they follow a specific pattern in each patient. They are called *Jacksonian seizures* and often precede the generalised types.

Generalised epileptic seizures involve most of the brain and imply *loss of consciousness*. *Generalised absence* (petit mal) is a transient loss of consciousness. These short attacks are recognised by *spike-doom waves* in the EEG.

Primary generalised tonic-clinic seizure (grand mal) is characterised by an extreme and widely distributed electrical activity, with tonic-clinic convulsions of the entire body. Presumably, a basic neuronal circuit activates the cortex of both hemispheres in generalised seizures. The hyperactive nerve cells release K^+ and glutamate during a seizure.

Small children with high fever (purexia) often react with generalised epileptic seizures called *febrile convulsions.* Diabetics in hypoglycaemia (ie, a blood glucose concentration below 3 mM) may develop generalised convulsions.

Epileptogenesis. The genesis and spread of epileptic discharges are poorly understood. The increased cortical excitability, with *high voltage-high frequency discharge* over the entire cortex, is not explained. Several mechanisms are probably involved:

1. The GABA-receptor complex contains the receptor site for not only GABA, but also for anti-epileptic drugs, including barbiturates and benzodiazepines, that potentiate or mimic the GABA-effect. GABA is the major inhibitory neurotransmitter in the brain. GABA opens chloride-channels, whereby the neurons are hyperpolarised, which reduces the likeliness of epileptic firing.

2. Epileptic seizure activity is either initiated or propagated through N-methyl-D-aspartate-(NMDA)-receptors binding glutamate or aspartate. The NMDA-receptor bind glutamate and the result are a high-frequency neuronal discharge. The hyperactive neurons release K^+ and excitatory amino acids or EAA's (glutamate and aspartate). NMDA-receptors and their ionic pores often work with a Mg^{2+}-sensitive glycine-receptor as a co-transmitter to glutamate. NMDA-receptors are the only ligand-gated channels that are also voltage-gated and Ca^{2+}-permeable. Drugs that effectively block the NMDA-receptor can reduce the abnormal excitatory spread of transmission and the focal epileptogenesis.

3. Block of Na^+-channels. The Na^+-K^+-pump and other Na^+-channels normally re-establish the ionic distribution after a discharge, and thus allow the cell to depolarise again. Some antiepileptic drugs do not alter the first action potential but reduce the repetitive firing-pattern. Carbamazepine and phenytoin block Na^+-channels, which prolongs the relative refractory period, and reduce repetitive firing of neurons.

During seizures, the extracellular [K^+] increases substantially, so the resting membrane potential is reduced. This makes the neurons more excitable and promotes the spread of the discharge. Fortunately, phenytoin blocks better at high K^+-concentrations around the neurons.

Adenosine inhibits the initiation of seizures in experimental animals. Carbamazepine promotes the adenosine-inhibition and thus blocks or reduces epileptogenesis.

Emergency therapy of a seizure is to keep the airways patent, apply diazepam suppositories to patients with prolonged seizures, and intravenous glucose in case of hypoglycaemia. The patient must be protected from harming himself during the few minutes of generalised cramps.

Long-term therapy of primary generalised tonic-clonic epilepsy (grand mal) and partial epilepsy is frequently made by use of carbamazepine or phenytoin. Generalised absence (petit mal) is frequently treated with sodium valproate.

Status epilepticus is life threatening due to cardio-pulmonary insufficiency and must be treated immediately with cardio-pulmonary support and diazepam intravenously.

4. Movement disorders

Disorders of the neurotransmission in the extrapyramidal system results in movement disorders of two types. Loss of movement with increase in muscular tone is termed akinetic-rigid syndromes, whereas disorders with involuntary movements are called dyskinesias.

4a. Parkinson´s idiopathic disease

or *shaking palsy* is characterised by tremor at rest, rigidity, akinesia or bradykinesia and postural changes.

Parkinson‘s disease is characterised by well-preserved cholinergic activity, but a reduction of the dopamine content in putamen and substantia nigra, of noradrenaline and 5-hydroxytryptamine in putamen, and of the GABA-synthesising enzyme glutamic acid decarboxylase in substantia nigra and in the cerebral cortex.

The main pathological mechanism is degeneration of dopaminergic neurons in substantia nigra. The consequence is a severe lack of dopamine in the striatum. The lack of dopamine in substantia nigra hyperactivates the GABA pathways to the thalamus, which activates the motor cortex neurons. This increases the discharge of alpha-motor neurons in the spinal cord resulting in plastic rigidity, akinesia and dyskinesia.

L-DOPA is a precursor of dopamine that is capable of crossing the blood-brain barrier. Administration of this drug, which is transformed to dopamine in the brain, relieves much of the rigidity and akinesia by inhibiting the striatum. Transplantation of dopaminergic neurons into the striatum has been explored.

Increased dopamine activity relieves rigidity. Too much dopamine causes chorea (see below).

Increased acetylcholine activity or reduced dopamine activity causes rigidity and bradykinesia in healthy persons. This is why reserpine, which depletes neurons for dopamine, and butypherones, which block the secretion from dopaminergic neurons, all causes so-called drug-induced

Parkinsonism. Drug-induced Parkinsonism is a common side effect in patients treated neuroleptic drugs or in patients given metochlopramide. Akathisia refers to a condition with restlessness and uncontrolled repetitive movements in patients with drug- induced Parkinsonism. Drug-induced Parkinsonism is refractory to usual drug therapy. The patients respond immediately upon seponation of the inducing drug.

The tremor is often pill rolling between the fingers and thumbs, and not necessarily bilateral. The stiffness or rigidity is called lead-pipe or plastic rigidity, because the high muscle tone is equal throughout the range of passive movements and the same by flexion and extension. Sometimes the resistance to passive movements is jerky, so-called cog-wheel-movements.

Akinesia (hypokinesia, bradykinesia) means inability to initiate normal movements. The face is mask-like with rare blinking and monotonous dysarthria.

Postural changes include a forward posture with short step gait and no arm swinging. The patient easily loose balances and falls stiffly to the ground.

Fig. 7-11: Neurotransmitters in the basal ganglia. ACh stands for acetylcholine.

Dementia or cognitive disturbances are present in some patients. Dementia is ageing of the brain with a resulting loss of mental powers.

A balance between the effects of dopamine and glutamate is a necessity for the normal functioning striatum.

Patients are treated with an amino acid precursor of dopamine: L-dihydroxyphenylalanine (L-DOPA). This molecule can cross the intestinal-blood barrier and the blood- brain barrier easily. Hereby, L-DOPA reach the inside of neurons (carrier-mediated transport).

L-DOPA is normally synthesised in dopaminergic neurons from dietary L-tyrosine. Exogenous L-DOPA is methylated by catechol-O-methyl transferase (COMT) to 3-O-methyl DOPA, or it is decarboxylated by decarboxylase to dopamine. Finally, dopamine is catabolised to homovanillic acid by COMT and monoamine oxidase (MAO). Most of the exogenous L-DOPA is lost by decarboxylation in patients where carbi-DOPA is not administered concomitantly. Carbi-DOPA inhibits the decarboxylase and reduces the side effects of L-DOPA.

When Parkinson patients are treated with too much of the dopamine precursor L-DOPA, they may develop hallucinations, fear and paranoid delusions (schizophrenia-symptoms), because excess dopamine causes schizophrenia or schizophrenic symptoms and signs. Chloropromazine and haloperidol decrease the dopaminergic effects and are called anti-schizophrenics. These drugs block both the D_1 and the D_2 dopamine receptors, so they reduce the dopaminergic effects, but also cause extrapyramidal side effects on the top of cholinergic and adrenergic blockade.

Amantidin increases the dopamine release from nerve terminals in the striatum, so this drug is effective mainly in the early phases of Parkinson´s disease, before all dopaminergic neurons are degenerated (Fig. 7-11).

Anticholinergic drugs, which block muscarinic, cholinergic receptors, are still in use to treat early Parkinson symptoms such as tremor. The side effects are dry mouth, bladder weakness, constipation, and confusion and memory loss.

Parkinson patients develop DOPA resistance following years of dopaminergic medication. The use of antagonists for the glutamate receptor in Parkinson patients decreases the activity of the glutamate pathway to the cortex, and reverses the akinesia and rigidity.

4b.Wilsons disease (hepatolenticular degeneration)

is an *autosomal recessive anomaly* of the copper metabolism. The abnormal gene is located on chromosome 13. Normally, copper is absorbed from the gastrointestinal tract and in the liver it is incorporated into caeruloplasmin. Normally, copper is excreted into the bile.

Wilson-children have low serum caeruloplasmin and serum copper. They fail to excrete copper. Copper is accumulated in the basal ganglia of the brain, the liver (liver cirrhosis), and the cornea/lens. Wilson-children show an *akinetic-rigid syndrome* with liver cirrhosis, haemolysis, anaemia and visual disturbances. Early diagnosis with long-term treatment (penicillamine) has improved the prognosis.

4c. Dyskinesias (Chorea and Myoclonus)

Chorea is a type of involitional hyperkinesia, with jerky movements of the limbs. The movements look almost like voluntary, purposeful movements.

A special type of chorea is *athetosis* (aethymology: "not fixed"), which refers to slow, twisting involuntary movements of the fingers. Athetosis is seen in children following brain damage with bilateral damage of the nucleus subthalamicus.

St. Vitus dance or *Sydenhams chorea minor* is a complication following rheumatic fever with encephalitis. The movements are usually unilateral. Recovery typically occurs spontaneously.

Hemiballismus or *hemichorea* describes violent, throwing movements of one arm or one side of the body - as if the patient tried to throw a ball. The cause is partial lesion of the contralateral *nucleus subthalamicus* - often due to thrombosis. Ballism means flailing movements of the limbs.

Huntington's chorea

Huntington's chorea is an *autosomal dominant genetic defect on chromosome 4.* Its characteristic disorders are hypotonia, dementia and involuntary hyperkinesia. Huntington's chorea is due to a defect in GABAergic and acetylcholinergic interneurons of the striatum and the cerebral cortex. The two transmitters are normally synthesised by the enzymes, glutamic acid decarboxylase and acetylcholine transferase, but their concentrations in the interneurons is markedly reduced. The diagnosis is confirmed with genetic testing. The prognosis is bad with rapid mental deterioration in particular in young patients. The dementia of the Huntington's chorea is caused - as for most types of dementia - by cortical degeneration.

GABA receptors are usually inhibitory. When GABA no longer inhibits the globus pallidus from the striatum, this leads to a stronger inhibition of the thalamus, which is probably the cause of the involitional choreiform movements. Chorea is opposed to rigidity. Low doses of a dopamine agonist may reduce the choreiform movements of patients with Huntington's disease.

Myoclonus refers to brief contractions or jerks of one or more muscles. Myoclonus is often called nocturnal myoclonus, because it occurs at night. Generalised myoclonus resembles epileptic seizures, and is related to epilepsy following brain damage by hypoxia. Myoclonus is related to drug toxicity or metabolic toxicity from renal or hepatic insufficiency.

Tics are repeated twitching of facial or neck muscles. Tics are also called mimic or focal myoclonus. The tics may begin in childhood for unknown reasons. Tics are extremely resistant to any therapy.

Tremor can be caused by hyperthyroidism and by Parkinsonism, but it is also a typical side effect of alcohol, narcotics and drug abuse. Some cases of essential tremor can be reduced by beta-blockers.

5. Multiple sclerosis

Multiple or disseminated sclerosis refers to a common neurological disease caused by inefficient

myelin production in the oligodendroglia.

The cause is unknown, but the acquired defect in the oligodendroglia cells results in demyelinized areas or plaques in the CNS. The prevalence increases progressively with the distance from the equator, and the patients have an abnormal immune response with large concentrations of antibodies to virus infections.

The demyelinized plaques are mainly localised to the brainstem, cerebellum, periventricular region and optic nerves. Motor neurons of the spinal cord and peripheral nerves are rarely affected by demyelinisation.

Blurring of vision in one eye is usual with disc swelling of the optic nerve at ophtalmoscopy. There is also diplopia, vertigo, nystagmus, and dysphagia, when the brainstem is affected. Later paraparesis and tetraparesis develops. Death ensues by lung infections or uraemia.

Magnetic resonance imaging (MRI) with scanning of the brain and CNS can visualise demyelinized plaques in the periventricular white matter or elsewhere. MRI is an expensive technique, where protons are activated with radiofrequency waves to create images. In the CNS, the white and the grey matter are distinguished.

This is a disabling disorder for which there is no cure. Interferon has been tried with some effect on the lesions visualised with MRI. Palliative treatment necessitates teamwork.

6. Inflammations

Meningitis refers to inflammation of the meninges. Clinically, the meningitis syndrome is characteristic. The meningitis syndrome is a patient with high fever, headache, photophobia and vomiting. The patient - often a child - is placid and inactive, consciousness may be impaired and neck stiffness develops.

Bacteria, viruses, fungi, chemicals or drugs cause meningitis, or unusual organisms in immuno-compromised patients.

Immediate administration of intravenous benzylpenicillin is life saving in cases of acute meningococcal or other bacterial meningitis together with urgent investigations.

Encephalitis is inflammation of the brain tissue caused by the same organisms as meningitis. Herpes simplex encephalitis is treated with acyclovir intravenously. Acyclovir inhibits DNA synthesis and thus the proliferation of the virus. Japanese B encephalitis is avoided by vaccination of travellers to the Far East. Other causes to acute viral encephalitis are Coxsackie virus, Echo virus and mumps virus.

AIDS in the CNS is caused by the HIV itself or the CNS disease is caused by other infectious agents - in particular fungi, TB, or Escherichia coli which damage brain cells. The clinical picture is meningitis, myelitis or encephalitis.

Neural syphilis may occur as *tabes dorsalis*. Tabes dorsalis is caused by demyelinisation of the dorsal roots of the spinal cord. Lancinating pains, ataxia, loss of reflexes, muscle wasting, neuropathic joints, Argyll Robertson’s light-stiff pupils and optic atrophy.

Tertiary syphilis can be avoided if the primary syphilis infection is treated correct. Usually, injection of 1 g i.m. Benzylpenicillin for 2 weeks is enough.

Rubella encephalitis caused by rubella virus may progress following some years, because of antibody production against rubella viral antigen.

Creutzfeld-Jacobs Disease (CJD) and *KURU* (among cannibals in Papua, New Guinea) are

known from the spongioform encephalopathy seen at autopsy - the brain looks like an Emmenthaler cheese.

The pathology is similar to that of bovine spongiform encephalopathy of cattle and sheep ("SCRAPIE"). CJD is inherited of transmitted to man with a *prion*. The prion is an *abnormal neuronal membrane protein*, which can mutate like a virus. The prion is resistant to usual sterilisation procedures, and the incubation period is not always for years. Prions are transferred when eating neural tissue from sick cows or sheep and in other ways. The first signs in humans are various neurological insults such as sudden blindness, difficulties in gait and balance, memory and concentration disturbances, and slowly progressing dementia until a rapid death. – The hereditary form of *CJD* is caused by mutation of the human gene (PRNP) for the neuronal membrane protein.

Today, *KURU* is history. The high incidence in New Guinea 30 years ago was due to cannibalistic rituals. *Gadjusek* showed that *KURU* was infectious, and received the Nobel Prize in 1976.

7. Intracranial tumours

Most intracranial tumours are *primary* tumours (75%) and approximately one-quarter are *secondary* (metastases).

The primary malignant tumours are *gliomas* (eg. astrocytomas and oligodendrogliomas), and the primary benignant tumours are *meningeomas* and *neurofibromas*.

The metastases originate from primary tumours in the breasts, bronchi, kidneys, prostate, stomach, and thyroid etc.

Magnetic resonance imaging (MRI) is a scanning technique, where protons in a strong magnetic field are bombarded with radiofrequency waves in order to produce images. MRI scanning can picture brain tumours, multiple sclerosis lesions, and syringomylia among others. MRI scanning can even separate white from grey matter. MRI scanning is replacing myelography, because it can visualise spinal cord compression, spinal cord tumours and other malformations.

Gliomas originate in the neuroglia. *Astrocytomas* are gliomas originating from astrocytes. Astrocytomas are usually located in the cerebrum in adults, and in the cerebellum in children

Oligodendrogliomas originate from the oligodendroglia and grow slowly in the cerebral tissue

Meningeomas originate from the arachnoid matter usually along the venous sinuses above the tentorium. They are benign and grow slowly.

Neurofibromas arise from Schwann cells usually around the 8^{th} cranial nerve (acoustic Schwannomas).

Symptoms and signs of brain tumours are treated already in *2e. Intracranial mass lesions.*

8. Dementia contra delirium

Dementia (*senility* or ageing of the brain) disturbs almost all cognitive brain functions, whereby the personality of the patient is completely changed (cognitive functions as calculation, comprehension, judgement, language, learning ability, memory, orientation, and thinking).

Dementia develops *slowly* and has no diurnal variation. Cortical atrophy is found by using CT or MRI scanning of the brain. The clinical differential diagnosis to delirium and depression is sometimes difficult to establish, but it is consequential to the delirious patient, if the diagnosis is misinterpreted as dementia (Box 7-3). Dementia is an exclusion diagnosis.

Box 7-3: Differences between the syndromes dementia and delirium			
Syndrome	**Dementia**	**Delirium**	**Depression**
Attention and cognition	Variable	Globally impaired	Variable
Consciousness	Normal	Impaired	Normal
Diurnal variation	None	Worst at night	Morning worst
Development	Insidious	Sudden	Variable
Hallucinations	None	Often visual	Auditory
Speech	Perseveration	Difficulty finding words	Normal
Delusions	Absent	Fleeting	Systematised
Primary causes	Cortical trophia/	Illness -intoxication	Loss of NA
Therapy	Palliative	Causative treatment	SSRI

Alzheimer's disease

is a possible cholinergic system disease. Altzheimers disease is a primary (ie, unknown aetiology) cortical brain atrophy. Lack of the acetylcholine-producing enzyme choline acetyltransferase, and of acetylcholine has been demonstrated by neurochemical studies. Alzheimer's disease is a form of presenile dementia or premature ageing of the brain (ie. occurring before the age of 70). The disease is rapidly progressing to complete loss of mental powers, in particular loss of memory and normal emotional behaviour. CT scan shows *cortical atrophy* and excludes brain tumours. At autopsy argentophilic plaques filled with *amyloid protein A4* are found in the hippocampus, basal ganglia, thalamus and the cortex. The gene defect causing familial Alzheimer disease is located on *chromosome 21*, close to the *pro-A4 gene.*

The cholesterol transport to the tissues is also affected. Alzheimer's disease is probably caused by neuronal degeneration in the nucleus basalis close to the globus pallidus, and possibly also to lack of somatostatin and substance P in deep brain centres. Normally, cholinergic axons from the nucleus basalis project to the cortex, and their functions relate to memory and to the limbic system functions.

There is no specific treatment of dementia. Anxiety and depression is treated symptomatically (Box 7-3).

Delirium

Delirium is an acute impairment of consciousness also called *toxic confusion*. Sense impressions are misinterpreted, the mind and memory work incoherently, and the patient is frightened and suspicious because of hallucinations. Relatives often mention senility, but they may also inform about an *acute start*, so delirium is recognisable. Besides being acutely developing, delirium is also worst at night with visual hallucinations and incoherent speech and perseveration.

In contrast, the dementia patient is conscious, cannot find the right words and the development has been slow.

The basis for the delirious pattern is organic brain disease caused by *intoxication* (eg. alcohol, drugs, poisons), brain damage by infections, lesions, subarachnoidal haemorrhage or tumours, systemic infections (malaria, septicaemia, TB), and *metabolic brain damage* (eg, hepatic or renal

failure, hypoxia, vitamin B_2, B_6, B_{12} deficiency).

The treatment of delirium concentrates firstly on the *underlying* disease (including electrolyte disorders, ischaemia etc), and secondly on the symptomatic aspect.

II. Psychiatric Disorders

The international Classification of Disease and Related Health Problems (ICD 10, WHO) is used.

The most serious disorders are *psychoses*, and the most common disorders are *neuroses*.

The description concentrates on the two classical psychoses schizophrenia and manic-depressive psychosis. The following personality disorders (neuroses) are described: Phobic anxiety neurosis, obsessive-compulsive disorders, dissociative-conversion disorders (hysteria), and eating disorders (anorexia nervosa and bulimina nervosa). The pathophysiology of affect and stress is also considered.

1. Schizophrenia

means splitting of the mind or disconnection of psychic functions (emotional and cognitive). Schizophrenia is a *psychosis with hallucinations, dissociation of ideas, intense fear, and paranoid delusions* (paranoia). Schizophrenia is possibly caused by hypersecretion of *dopamine* or by blockage of the *glutamate producing neurons* from the cortex to the striatum. The balance between these two neurotransmitters in the striatum is seriously disturbed.

The clinical syndromes covered by this term include - according to WHO - auditory hallucinations (eg, hearing voices), thought withdrawal with abnormal posture, delusional perceptions with paranoia and external control of emotions with persecution from the outside. The patients' feel that their thoughts and emotions are broad casted and they not only hear voices commenting their lives, but also their own thoughts are spoken aloud.

The cause is sometimes clarified as a biochemical brain damage with hypersecretion of dopamine from neurons in the *mesolimbic dopaminergic system* close to substantia nigra or by blockade of the glutamate-producing neurons from the cortex to the striatum. An imbalance between the effects of dopamine and glutamate spoils the normal function of the striatum. A special gene located in chromosome 5 increases the risk of schizophrenia. Dopamine agonists such as amphetamine, and other psychotic drugs (LSD, mescaline, and ecstasy) can cause schizophrenic psychosis.

The genetic involvement is demonstrated by a 50% risk for the monozygotic twin of an affected person. There is a 40% risk for two affected parents for having a schizophrenic child. The gene is probably located on *chromosome 5*. Some schizophrenics have limbic dysfunction of the left hemisphere.

Schizophrenia begins in young adults of both sexes, and may be more than one entity. Schizophrenics are frequently vulnerable to highly expressed emotions.

Chronic schizophrenics are characterised by lack of drive, underactivity, social withdrawal, and emotional emptiness. Catatonia (stupor, stereotypes, and automatic obedience) was previously seen in many institutional patients, and may still be seen among understimulated patients.

Dopamine blockers - blocking D_1 and D_2 dopamine receptors - are the drugs of choice in acute schizophrenia. These drugs belong to the phenothiazine family (chlorpromazine, trifluoroperazine and promazine).

Side effects are unavoidable as the drugs block both D_1 and D_2 receptors, as well as adrenergic and cholinergic receptors. The side effects are *extrapyramidal* (acute dystonia, Parkinsonism,

akathisia and tardive dyskinesia), *autonomic* (hypotension and ejaculation failure), and *anticholinergic* symptoms (dry mouth, urine retention, constipation and blurred vision).

2. Manic-depressive psychosis

covers severe abnormalities of mood. Mood ranges from severe depressive psychosis over moderate and minor depression, sadness, normal mood, happiness, euphoria, hypomania, and severe mania.

The diagnosis manic-depressive psychosis describes patients with periodic attacks of mania or depression, separated by periods of normal behaviour.

The diagnosis also includes patients with depressive periods alone, or with only manic periods.

Endogenous depression is characterised by early morning waking with unresponsive sadness, guilt feeling, suicidal feelings, and lack of a precipitating factor. Severe depression disturbs mood, talk and initiative. One type of mental depression is related to reduced formation of noradrenaline in the locus coeruleus, and of serotonin in the midline raphe nuclei of the brainstem, which seriously damage the limbic system. - Other types of depression are called exogenous or reactive depressions, because they are considered to be due to exogenous or environmental factors.

Medical drugs that inhibit the production of noradrenaline and serotonin often cause depression.

Hypomania is mild mania with euphoria, overactivity and disinhibition.

The *genetic aetiology* of manic-depressive psychosis is confirmed by the concordance of two thirds of monozygotic twins, and by the fact that more than 20 % of dizygotic twins are concordant. There is a clear overweight of females. *Winter depression* from autumn to spring is frequent in areas with lack of light in the winter months. Up to 20% of the population north of the polar circle suffers from winter depression. Light therapy several hours daily are so effective that overdoses may release manic phases.

Monoamine-neurotransmitters are depleted in depression but increased in manic phases.

Stressful social life events (marriage, divorce, moving house, loss of job, vacation, etc.) often precipitate depression.

Selective serotonin reuptake inhibitors (SSRI) are often preferred in treatment of depressive states, because of rapid effect and lower rate of serious side-effects including addiction. These substances inhibit serotonin reuptake within the synaptic cleft, and are named "happiness pills" in the media. Happiness pills do not exist.

Depression was previously treated with *monoamine oxidase inhibitors* (MAO-inhibitors). They inhibit the enzyme monoamine oxidase A&B and thus the breakdown of monoamines. Hereby, adrenaline, dopamine and 5-hydroxytryptamine are accumulated in the brain. Tricyclic antidepressants block reuptake of monoamines and are likewise effective in the treatment of depressive patients.

Electroconvulsive therapy (ECT) is a physical treatment with rapid effect, often used for cases with suicidal or other deep depressions.

Therapeutics (such as *lithium compounds*) that inhibits the action of noradrenaline or serotonin is effective prophylactic agents against manic phases. In this model mania is caused by overproduction of monoamines, and depression by reduced formation of monoamines in the brain nuclei mentioned above. Actually, *lithium carbonate* is used in the prophylactics of manic phases. A plasma-lithium concentration of 0.5-1 mM is necessary to obtain an acceptable result.

Psychoses are treated with antipsychotics often supplemented with benzodiazepines. The typical antipsychotics are traditionally divided into high-, medium-, and low-potency drugs that are blocking the D2 and D1 dopamine receptors, with secondary blocking of the serotonine – histaminergic- adrenergic – and cholinergic receptors. The main drugs in this category are fluopentixole, haloperidole, zuclopenthizole, chlorpromazine and levopromazine.

Side effects are unavoidable as the drugs block a range of receptor types. The side effects are extrapyramidal (dopaminergic – acute dystonia, parkinsonism, akathasia and tardive dyskinesia), serotonine related (weight gain), histaminergic (sedation), autonomic (hypotension, ejaculation failure, salivation), and anticholinergic (dry mouth, urine retention, constipation and blurred vision).

A new class of antipsychotics which do not fit into the high/low potency classification have been introduced and are gaining world wide use because of their low degree of side effects. Drugs without cholinergic activity do not lead to extrapyramidal side effects. Some drugs have less adrenergic activity and they are generally more limbic than pyramidal in their selectivity compared to the classical antipsychotics. The main drugs in this category are: Amisulpride, risperidone, sertindole, closapine and olanzapine.

Nervous and stress-related personality disorders (Neuroses)

Phobic anxiety neurosis

Anxiety neurosis is a chronic condition or it occurs as attacks of panic. *Acute overactivity* of the sympathoadrenergic system results in precardial pain and palpitations (cardiac neurosis = neurocirculatory asthenia), chest constriction, flatulence and frequent defecation and urination, lack of libido, dizziness, headache, and sleep disturbances.

Attacks of panic anxiety occur in young, nervously sweating persons, who feel that they are dying from cardiac disease or from hyperventilation with tetany and carpopedal spasms. Hyperventilation reduces the carbon dioxide tension in the alveolar air (decreased P_{ACO2}) and thus the Ca^{2+} -concentration in the ECV, which opens Na^{+}-channels, reduces the membrane potential and increases the neuromuscular irritability (see *tetany* in Chapter 17).

Symptoms and signs of anxiety (ie, sweating, palpitations, tremor, tachycardia, flatulence, and urination) are caused by increased release of adrenaline and noradrenaline from the adrenal medulla. Drugs containing b -*adrenergic blockers* are of benefit to the anxious patient, because they block the sympathetic nerves and adrenergic synapses in the CNS. Cognitive-behavioural therapy is also applied with effect – sometimes combined with selective serotonin reuptake inhibitors (SSRI). Actually, SSRI substances are the first choice in anxiety conditions.

Obsessive-compulsive disorders

These patients have *obsessional thoughts* and perform *compulsive actions* to the extent that their social lives are seriously impaired. The patient feels an irresistible obsession to perform a given act - such as washing the hands or superstitious check of a closed door - again and again. To the patient the behaviour is often quite meaningless, but still it is necessary to carry on the ritual. A frequent complaint is that dirt and excretions are nasty, and many obsessions concern excretory processes.

In *cognitive-behaviour therapy* the patients learn not to perform the compulsive rituals.

Eating disorders

Anorexia nervosa is an eating disorder in adolescent females resulting in severe malnutrition. The patient has an intense wish to be thin. Biological and psychological factors are involved. In a few cases there is regression into childhood. The girl or boy tries to escape from the problems of

puberty and adolescence.

The increased concordance in monozygotic twins indicates a genetic aetiology.

This disorder occurs among amenorrhoic teen-age girls, who express an abnormal fear of being fat. Typically, they have an extremely low body weight. The girl is usually bright and knowledgeable, and her parents are overprotective. The patient sometimes realises the presence of problems in accepting the role as a maturing female.

Low plasma Gonadotropin levels with impaired response to LHRH are frequently found.

Positive reinforcement for even small weight gains is sometimes of help. The basic psychological problems must be treated with cognitive-behavioural or other psychological treatment. Tricyclic antidepressants are beneficial in cases of *depression.*

Bulimia nervosa is diagnosed in persons who are preoccupied with food and periodically eats excessively. They may avoid overweight by self-induced vomiting just after *binge eating.* Bulimia nervosa may be associated with anorexia nervosa. *Behavioural therapy* is sometimes successful – sometimes combined with selective serotonin reuptake inhibitors (SSRI).

Dissociative-conversion disorders (hysteria) is characterised by psychologically mediated psycho-somatic disorders. These disorders have no physical pathology; they are not sympathetic overactivity and are produced without the consciousness of the patient.

Freud believed that mental energy was converted into physical disorders such as abdominal pain, blindness, double vision, deafness, muteness, fits with dramatic movements, artistic gait disturbances, hysterical paresis with normal muscle tone and deep reflexes, crude tremor, sensory loss, stigmatisation, - all with *secondary gain*. The disorders are explained as the result of repression, dissociation and *conversion* of mental energy into physical disorders. *Repression* means exclusion of memories, impulses, and emotions from consciousness, because these elements would cause anxiety and distress. *Dissociation* means an apparent dissociation between different mental activities. An example is a protective mental cover of enjoyment in terminal cases of painful cancer (French: *Belle indifference*).

The classical *hysterical triad* include mydriasis (large pupils), lack of pharyngeal reflexes and lack of plantar reflex.

The mental disorders are *amnesia* for long periods, *sleep walking* or *somnambulism* (see below), *imitation* with multiple personalities, globus hystericus and *pseudo-dementia*.

Psychotherapy is a causal treatment, although sometimes impossible to carry through.

Sleep disturbances

include insomnia, somnambulism and sleep apnoea.

Insomnia is *subjective sleep deficiency*. The patient complains that he sleeps too little, or has the impression that he cannot sleep. Such patients sleep more than they think, when studied in sleep laboratories, and their health is not impaired. There is a natural decline of the sleep duration with age, and the use of drugs should be restrained. Monotonous sounds such as music or the sounds from ocean waves have proven to be an optimal "sleeping drug" for many individuals. The common complain of insomnia among the elderly is often curable by regular motion passes (eg, walking, swimming, jogging etc).

Insomnia as *early morning waking* is a sign of depression, but it is also seen as nocturnal confusion in dementia and in delirium, where the cause may be organic brain damage or drug abuse.

Sleepwalking or *somnambulism* is a form of personality dissociation with unknown aetiology. Some of the patients have hysterical patterns (see above).

Sleep apnoea often occurs with snoring and airway obstruction in obese patients or in patients with chronic obstructive lung disease.

Affect and stress reactions to psychological or physical stress are seen in otherwise healthy individuals. One example is described above as *panic attacks*.

Self-Assessment

Multiple Choice Questions

- **I. Each of the following five statements have True/False options:**

- **A.** Cellular responses, mediated by drug receptors linked to ion-channels, are rapid compared to responses mediated by G-protein systems.

- **B**. Drugs acting via receptors have side effects, because they are bound to several receptors, distributed in several tissues, and the receptors are linked to different secondmessengers, which produce different cellular responses.

- **C**. Akathisia is an extrapyramidal defect with swaying and twisting body dyskinesia.

- **D**. The circumventricular organs (ie, hypothalamus, the pineal gland, and the area postrema have a tight blood-brain-barrier.

- **E..** Dopamine agonists. such as amphetamine and other psychotic drugs (LSD, mescaline, ecstasy), can cause schizophrenic psychosis.

II. Each of the following five statements have True/False options:

A. Dopaminergic neurons are found in the substantia nigra, noradrenergic neurons in locus coeruleus, and serotonin sensitive neurons in the raphe nuclei.
B. Catecholamines are neurotransmitters both in the sympathetic and the parasympathetic nervous system as well as in the motor endplate.
C. A single neuron releases only one neurotransmitter from all its synapses.
D. Insulin and related growth factors bind to membrane receptors that are glycoproteins protruding from the membrane.
E. The free radical gas nitric oxide (NO) is a neuronal messenger in both the central and the peripheral nervous system. The NO gas is membrane permeant and can bypass normal signal transduction in synapses.

III. Each of the following five statements have True/False options:

A. A constant small ion permeation through the cell membrane at rest is referred to as leak current.
B. The typical Na^+-channel opens promptly in response to repolarisation.
C. Opening of Na^+-channels increases the flux of Na^+ into the neuron, and depolarizes the membrane, so the effect is excitatory.
D. Immediate administration of intravenous benzylpenicillin is life saving in cases of acute meningococcal or other bacterial meningitis.
E. Most intracranial tumours (gliomas, meningeomas and neurofibromas) are primary tumours and only 25% are metastases (secondary tumours).

Case History A

A professor in linguistics, 59 years old, consults his doctor because of speech and movement problems. The patient is intellectually well functioning, but his speech has changed from motivating to a slow monotonous sequence of words. His gait is slow with small steps, and the standing position is difficult for this previous long distance runner. His facial expression is motionless, and he seems to have difficulties in initiating normal movements. There is tremor of the hands and fingers of the pill-rolling type. When the doctor examines the patient for rigidity, he finds high tonus (plastic rigidity) and cogwheel-movements.

1. *What is the main pathological mechanism of this disease?*
2. *What is it called?*
3. *Why is the muscle tone so high?*

Case History B

A female, 26 years of age, suffers from an epileptic seizure during her work as a nurse on a neurological department. A colleague saw that the nurse suddenly stopped while walking, her eyes and head turned left, her left hand moved in a curious way, and she uttered a cry and felled. The whole body became rigid for a minute, during which time she developed cyanosis. Then the muscles started to jerk rhythmically for a few minutes. She was unconscious during the seizure and remained so for an hour after the seizure. An EEG was taken during and after the seizure. A blood sample was taken and the blood glucose was determined to 5 mM.

- *1. Describe the type of epilepsy starting the seizure and the development into a second type of seizure.*
- *2. Describe the most likely EEG findings during and after the seizure.*
- *3. What is the pathophysiological basis for a grand mal seizure?*
- *4. Was hypoglycaemia involved in the seizure?*

Case History C

A professor in economics, 56 years of age, finds it increasingly difficult to concentrate during his work. His wife and two adult children find him totally different from his normal personality; he is with- drawn, depressed and forgetful.

Two months on antidepressants prescribed by his GP does not improve the condition, which is dominated by lack of memory. Psychiatric and neurological examination disclose no evidence for depression or increased intracranial pressure due to focal brain damage. CT scan shows cortical atrophy and excludes brain tumours. The mental powers are rapidly deteriorating. The total cholesterol concentration in blood plasma is increased.

- *1.What is the most probable diagnosis (two must be considered)?*
- *2.Is there a definite criterion for one of these in this patient?*
- *3.What are the prognoses for these two disorders?*
- *4.Is the disorder of this patient inherited?*
- *After a year the patient passes away.*
- *5.What are probable findings at autopsy?*

Try to solve the problems before looking up the answers.

Highlights

- *Opening of Na^+-channels increases the flux of Na^+ into the neuron, and depolarizes the membrane, so the effect is excitatory.*
- *Closure of K^+- or Cl^- -channels decreases the flux of K^+ out of the neuron or decreases the flux of Cl^- into the cell. These events also depolarise the membrane, and again the effect*

is excitatory.

- *Obviously, closure of Na^+-channels or opening of K^+- or Cl^-channels have an inhibitory effect by hyperpolarization.*
- *Two types of NO synthase (NOS) have been identified: constitutive Ca^{2+}- calmodulin dependent enzyme, and inducible Ca^{2+} - independent enzyme. Both enzymes are flavoproteins containing bound flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD).*
- *The free radical gas nitric oxide (NO) is a neuronal messenger in both the central and the peripheral nervous system.*
- *Nitrovasodilatators act by releasing NO and thereby causing coronary vasodilatation in patients with angina pectoris. Nitric oxide synthase is inhibited by L-arginine analogues.*
- *Signal transduction is a cascade of processes from the receptor-hormone binding to the final cellular response. Many hormones and neurotransmitters raise the concentration of a second messenger in the target cell via guanyl triphosphate (GTP) and act through it.*
- *The receptor-hormone complex activates a GTP-binding protein (so-called G-protein), which controls and amplifies the synthesis of the second messenger.*
- *G-proteins function as molecular switches, regulating many cellular processes, such as activation of intracellular enzymes (protein kinase, phosphorylase), activation of membrane enzymes and channels, and activation of gene transcription.*
- *G- protein-linked receptors form a family, which has evolved from a common ancestor. Most G-proteins are membrane bound heterotrimers (α β γ) and exist in an activated state, with high affinity for GTP, and an inactive state, where the molecule prefers GDP.*
- *Hydrophilic (lipophobic) hormones such as acetylcholine and many peptides bind to membrane receptor proteins, and the hormone-receptor binding activates the enzyme phospholipase C via active G-protein.*
- *Protein tyrosine kinase activity is abnormally high in certain types of cancer and cellular modifications. This can be caused by growth factors or by a mutation of the tyrosine kinase part of the transmembraneous receptor. Mutations of one gene localised on chromosome 10 can lead to four different syndromes: Multiple endocrine neoplasia, Hirschprung's disease, medullary thyroid carcinoma, and Phaeochromocytoma.*
- *Stroke is commonly caused by thrombo-embolism of the middle cerebral artery.*
- *Simple concussion is defined as a transient loss of consciousness followed by complete recovery. A short period of amnesia is often related to the loss of consciousness. This is a migraine injury, where the duration of the unconsciousness indicates the severity of brain damage.*
- *Brain contusion refers to brain damage with prolonged coma, amnesia and focal signs. Later on such patients often suffer from chronic impairment of higher cerebral functions and hemiparesis.*
- *Post-traumatic epilepsy is frequently caused by head injury with coma following depressed skull fractures, brain contusion or intracranial haematoma. Actually, depressed skull fracture causes a high incidents of post-traumatic epilepsy.*
- *Epidural haematoma is caused by skull fractures traversing a dural sinus in the temper-parietal region, resulting in bleeding into the epidural space.*
- *Subdural haematoma is an accumulation of blood in the subdural space caused by venous bleeding.*
- *Subarachnoid haemorrhage is a spontaneous arterial bleeding into the subarachnoid space, often with an acute clinical picture of acute delirium. The circle of Willis and adjacent vessels is the most frequent site for saccular or berry aneurysms that rupture.*
- *Intracranial mass lesions located supratentorially can compress the brain towards the tentorium as to block the upward flow of CSF and thus its absorption.*
- *Epilepsy is an abnormal paroxystic discharge from cerebral neurones resulting in a condition with clinical consequences. Epileptic seizures are partial or general.*
- *The normal EEG waves are due to synaptic potentials by groups of neurons including pyramidal cells. An epileptic seizure is characterised by high voltage-high frequency*

discharge from large groups of neurons or from the entire cortex.

- *Partial or focal seizures can be caused by an epileptic focus anywhere in the cortex. The causes of focal seizures are acquired lesions such as cysts, tumours, scar tissue, infections, and ischaemic lesions. The epileptic discharge causes involuntary muscular contractions on the contralateral side. Foci in the somatosensory cortex produce sensory hallucinations called an epileptic aura.*

Further Reading

- *The Journal of Neuroscience*. Semi-monthly journal published by the Society for Neuroscience, 11 Dupont Circle, NW, Washington DC 20036, USA.

- Hopkins AP (1993) *Clinical Neurology, a Modern Approach.* Oxford University Press, Oxford.

- Sims ACP and DW Owens (1993) *Psychiatry*. 6th edition. London: Bailliere Tindall.

Return to chapter 7

Return to content

Chapter 8.

Cardiovascular Physiology & Disorders

Study Objectives

- To *define* afterload, anaemia, aneurysms, arterial pressure amplitude, diffusion-filtration- and permeability- coefficients, filtration capacity, preload, vascular compliance, and capillary protein permeability.
- To *describe* the circulatory system, distribution of the total blood volume, capillary variability, capillary exchange-perfusion-permeability, venous system, venous pump, venous volume and pressures at different conditions.
- To *describe* Laplace´s law, the law of conservation of matter for determination of volume and flow, the Starling equation, net filtration with lymph formation, oedema protection and formation, lymphatic oedema.
- To *draw* a model of the paracapillary circuit and of the two circulatory systems.
- To *calculate* one cardiovascular variable from selected variables given.
- To *indicate* normal levels of cardiac output, oxygen uptake, arteriovenous oxygen content difference, oxygen binding capacity, haematocrit, haemoglobin concentration, mean corpuscular volume (MCV), mean corpuscular haemoglobin concentration (MCHC), and perfusion coefficients.
- To *explain* the exchange between blood and cells, the control of erythropoiesis, Poiseuille´s law, total peripheral vascular resistance and organ resistance, vascular compliance and specific compliance, viscosity-related factors, and the Fåhræus-Lindquist effect.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The law of conservation of matter states that mass or energy can neither be created nor destroyed (the principle of mass balance). This principle is used to measure physiological blood volumes and bloodflow.*
- *Poiseuille's law is used both in the circulatory and the respiratory system (see Eq. 8-3).*

Definitions

- **Afterload** is the force against which the ventricle contracts. A good index of the maximal afterload tension is the *peak* intraventricular pressure during systole.
- **Anaemia** is defined as a clinical condition with an insufficient oxygen carrying capacity of the blood. A blood haemoglobin concentration below 130 g per l (8 mM) implies a measurable reduction of the working capacity for both sexes.
- **Arterial elastance** or *stiffness* is (DP_t/ DV) or the reciprocal of arterial compliance.
- **Arterial pulse amplitude** or the *pulse pressure* is the difference between the systolic and the diastolic arterial pressure at a certain level.
- **Arteriovenous oxygen content difference** is the difference between the oxygen concentration in arterial blood and that of the mixed venous blood ($C_{aO2} - C_{vO2}$).
- **Bloodflow** is the flow of whole blood to an organ per time unit. A practical index is the relative bloodflow measured per 100 g of tissue. Thus, the bloodflow is expressed in ml of blood per min per 100-g tissue, which is abbreviated as flow units (FU).
- **Bulk flow** is convective transport of fluid with its content.
- **Capillary protein reflection coefficient** (s) is the fraction of plasma protein molecules reflected off the capillary wall following collisions.
- **Cardiac output** is the volume of blood leaving the left ventricle (or the right) each min.
- **Central venous pressure** (CVP) is the pressure in the right atrium and caval veins close to the right atrium.
- **Compliance** of a vessel is the increase of volume per unit of transmural pressure increase (DV/DP_t). Transmural pressure refers to the intravascular pressure minus the extravascular pressure.
- **Contractility** is a measure of the cardiac performance at a given preload and

afterload.

- **Driving pressure** is the mean arterial pressure minus the atrial pressure or CVP.
- **Ectopic focus** is a *pacemaker focus* located in other regions of the myocardium than the sinus node. Active ectopic foci cause abnormal contraction patterns in the related regions of the heart.
- **Embolism** refers to the process through which a thrombus is dislodged from its attachment and travels with the blood until it is lodged in a blood vessel too small to allow its passage.
- **Erythrocyte sedimentation rate** (ESR) is the rate of fall of erythrocytes in a column of anticoagulated blood. ESR is increased, when the plasma is rich in large sticky protein molecules (fibrinogen, immunoglobulins etc) which agglutinate red cells, so they fall rapidly. Severe anaemia, immune reactions, infections, ischaemia, malignancy and trauma increases ESR.
- **Fibrinogen** is a dissolved plasma protein that can be transferred to a blood cell trapping fibrin network by the proteolytic enzyme, thrombin.
- **Filtration:** Transport across a barrier by means of a hydrostatic pressure gradient.
- **Haemolysis** refers to disruption of the red cell membrane with liberation of the cellular content to the plasma of whole blood.
- **Haemostasis** refers to the arrest of bleeding.
- **Hypocoagulability** refers to a condition with a prolonged coagulation time.
- **Jaundice** *(icterus)* is pigmentation of cell membranes, plasma and secretions with yellow bile pigments.
- **Mean arterial pressure** (MAP) at a certain level equals diastolic pressure plus 1/3 of the pulse amplitude as an approximation.
- **Microcirculatory unit** is a collection of vessels that originate from one arteriole, which is characterised by well-developed smooth musculature in its walls.
- **Oedema** is an abnormal clinical state characterised by abnormal accumulation of interstitial or tissue fluid.
- **One atmosphere**. By definition, one atmosphere equals 760 mmHg or 101.3 kPa.
- **Pinocytosis** is a process by which fluid and large molecules can pass the capillary wall in vesicles formed by the cell membrane.
- **Preload** is the end-diastolic filling pressure of the ventricle just before contraction.
- **Plasma viscosity** is measured instead of erythrocyte sedimentation rate (ESR), because it is dependent of the same large protein molecules as ESR, but independent of haemoglobin concentration and obtainable within 15 min.
- **Pressure** is **force per area unit**. The international unit is Newton per m^2 or Pascal (Pa).
- **Pressure Resistance Units** (PRU) are measured as Pascal seconds m^{-1} of blood (or as mmHg seconds ml^{-1} of blood).
- **Serum** refers to plasma that has undergone coagulation and thus is devoid of fibrinogen and many other coagulation factors.
- **Serum ferritin concentration** (Chapter 22) reflects the mass of *stored iron* in the body (normal range 12-140 nM). Most of the ferritin is stored in the tissues and not in the blood serum.
- **Serum iron concentration** (Chapter 22) is Fe^{2+} bound to transferrin. The normal range is 7-36 mM with a mean value around 22 for both sexes. Iron deficiency leads to anaemia of the microcytic, hypochromic type (small, pale red cells).
- **Small-diameter phenomenon** *(Fåhræus-Lindquist):* The viscosity of blood decreases in tubes with a diameter less than 0.5 mm, because the packed cell volume here is relatively low.
- **Solvent drag** refers to transport of solvent, which can also draw solutes across a barrier.
- **Stroke volume** is the volume of blood ejected from a heart ventricle with each beat.
- **Thrombosis** refers to the formation of multiple thrombi or clots within the vascular system.
- **Total peripheral vascular resistance** (*TPVR)* is the resistance of the systemic

circulation. *TPVR* can be calculated as the driving pressure, divided by bloodflow ($Q \circ s$ in ml per s): $TPVR = DP / Q \circ s$. During exercise *TPVR* is reduced to approximately 30% of the level at rest.

- **Transferrin** (Chapter 22) is a plasma protein vehicle with 2 binding sites for Fe^{2+} (normally 35% of the plasma globulin is saturated with iron). *Transferrin saturation* is the serum iron concentration divided by the total iron binding capacity. See iron deficiency.

- **Viscosity** of blood is the inner friction, which is due to interaction between molecules and particles in the blood. The viscosity (η) one Pascal sec (1 Pa s) is the tangential force, working on 1 m^2 of surface area, when dv/dx is 1 (s^{-1}).

Essentials

This paragraph deals with 1. Circulatory organisation, 2. Haemopoiesis, 3. The red cells, 4. Viscosity, 5. Blood coagulation, 6. Vascular compliance and stiffness, 7. Wall tension, 8. Microcirculation, 9. Transcapillary fluid exchange and 10. The lymphatic system.

Circulatory organisation

General arrangement

The cardiovascular system consists of two pumps arranged in *series* (Fig. 8-1). They are the right ventricle that pumps blood into the pulmonary circulation, and the left ventricle, which pumps blood into the systemic circulation. Each of these pumps delivers blood through an efferent tube system (the arteries) and each pump receives blood through an afferent tube system (the veins). In the pulmonary system, blood is pumped from the right ventricle through the lung capillaries and is temporary collected in the left atrium (Fig. 8-1). The coronary arteries are the first arterial branches that arise from aorta just above the aortic valve. Aorta and the elastic arteries are *conductance vessels*; the muscular arteries are *distribution vessels*; the arterioles are *resistance vessels*; the capillaries are *exchange vessels*; venules and veins are *capacitance vessels*. The arterio-venous anastomoses in fingers and toes are *shunt vessels*.

Fig. 8-1: Design of human circulation with the right heart before and the left heart after the lungs.

The principal function of the bloodflow in the cardiovascular system is to provide oxygen (O_2) and nutrients to the tissues of the body and to remove carbon dioxide (CO_2) and waste products. The flow of blood through the cardiovascular system follows physical law's known from fluid mechanics (see principles).

Strictly speaking, *Poiseuille's law* (Eq. 8-3) has validity in a circulatory system, only when the fluid flow is laminar and non-pulsating in horizontally situated cylindrical vessels of constant dimensions. The resistance for laminar flow of a Newtonian fluid is only dependent on the dimensions of the vessel and the viscosity of the fluid. Resistance varies inversely as the fourth power of the radius of the vessel.

For resistances in *parallel*, the total resistance is less than that of any individual resistance (Fig. 8-1 and Eq. 8-4). Although the total cross sectional area of all arterioles is much larger than that of all arteries, their resistance to bloodflow is much greater than that of the arteries. The number of *daughter vessels* is not high enough to balance the decrease in vessel diameter. The resistance is highest in the capillaries and it diminishes as the vessels increase in radius.

For resistances in *series*, the total resistance equals the sum of the individual resistances (Eq. 8-5).

In contrast to Poiseuille's conditions, the bloodflow in the human circulation is pulsating and sometimes turbulent, and its blood vessels are not horizontally located, cylindrical or inflexible. Neither is the blood viscosity constant nor independent of vessel diameter and flow.

At rest the mean red cell velocity in the capillaries is observed to be approximately 1 mm in one s; this provides ample time for gas exchange. Since the circulating blood moves continuously, the cardiac output must pass a cross section of all open capillaries. At rest a cardiac output of 5000 ml per min is a reasonable estimate; when changed into volume rate per s, the cardiac output is equal to 10^{-4} m^3 s^{-1}. Hence, it is possible to calculate the large cross sectional area of all open capillaries in a resting person according to Eq. 8-1 (see Fig. 8-1). The total blood volume is approximately 5 l in a healthy adult.

The *right atrium* receives venous blood from the caval veins, and the left atrium receives oxygenated blood from the pulmonary veins. The two atria function as thin walled reservoirs and conduit organs for the blood (Fig. 8-2). On average, atrial systole contributes only about 15 % of the total ventricular filling, but in cardiac insufficiency the atrial contribution may increase importantly. The left and right ventricles provide most of the energy needed to transport the blood through the circulation. The left ventricle accelerates the blood into the systemic or peripheral high-pressure system, and its walls are thick in contrast to the thin, weak right ventricle, which pump blood into the low-pressure pulmonary system.

The *left ventricle* consists of cardiac muscle fibres originating from the fibrous rings at the base of the heart and the fibres are twitching towards the apex. The orifice between the left atrium and the left ventricle carries two valve cusps, and this valve is called the *bicuspid* or mitral valve. Three cusps form the *tricuspid* valve closing the orifice

between the right atrium and ventricle during systole. Strong filaments (chordae tendineae) arise from the papillary muscles of the ventricles. These chordae are attached to the free edges of the atrioventricular valves and normally prevent the valves from bulging into the atria during ventricular systole. The two *atrio-ventricular valve* systems prevent the leakage of blood backward from the ventricles into the atria (Fig. 8-2). Two other valve systems are interposed between the left ventricle and the aorta (the aortic valves) and between the pulmonary artery and the right ventricle (the pulmonary valves).

The conduction system

The normal heart is characterised by an electrical insulation between the atria and the ventricles mainly due to the fibrous ring (*annulus fibrosus*). However, the heart possesses a specialised electrical system, the cardiac *conduction system* that leads the electrical signal from the atria to the ventricles. The conducting system consists of modified myocardial cells. An optimal timing of atrial and ventricular pumping allows the emptying of the atria to be completed before the ventricular contraction. This allows the heart to pump the required cardiac output.

The heart normally has a self-firing unit, located in the right atrium, called the *sinoatrial node* or sinus node (Fig. 8-2). The *sinus node* contains round cells (*pacemaker cells*), elongated intermediary cells and ordinary atrial cells. The electrical signal that automatically originates from the sinus node has the highest frequency, and the sinus node is thus the *natural pacemaker* of the heart. Even a cardiac transplant patient (the heart is totally denervated) adapts to the altered needs for cardiac function and of course initiates new heart beats as long as the transplant is functioning.

The electric signal from the sinus node activates the atrial walls to contraction, and then reaches the main conduction system at the level of the *atrioventricular node* (AV node). The AV node consists of the same cell types as the sinus node. The impulse is delayed in the AV node, and this delay is allows the atrial systole to squeeze extra blood into the ventricles just before the ventricular systole occurs (see above).

Fig. 8-2: The cardiac conduction system (left) is the only electrical connection between the atria and the ventricles of the normal heart. The anatomy of the four heart chambers is shown to the right.

From the *bundle of His*, the signal is transmitted down a rapid conduction pathway, composed of the right and left *bundle branches*, to stimulate the right and the left ventricle and cause them to contract. The right bundle branch proceeds down the right side of the ventricular septum, and the large left bundle branch perforates the septum and divides into an anterior and a posterior division. These bundle branches divide into a network of conducting *Purkinje fibres* just below the endocardial surface. Purkinje fibres are large diameter cells without T-tubules, and with a long refractive period, so they can block premature depolarisation waves from the atria. The propagation wave spreads in the septum from both branches with the thick left bundle branch being dominant. The spread along the Purkinje fibres is rapid, whereas the spread from the endocardium to the epicardium is slow (Fig. 8-2).

Ectopic foci become pacemakers, when the normal dominant pacemakers fail by blockade or depression: In the AV node, the atria, and the Purkinje fibres or in ischaemic ventricular fibres.

These topics are developed further in Chapter 11.

Distribution of blood and flow

The total blood volume (5 l) is distributed with 60-75% in veins and venules, 20% in arteries and arterioles, and only 5% in capillaries at rest. Of the total blood volume only 12% are found in the pulmonary low-pressure system.

The distribution of the cardiac output to the main organ systems of the body in a healthy person at rest and during maximal exercise is given in Box 8-1.

Box 8-1: Distribution of flow in % of the cardiac output, arteriovenous oxygen content difference, oxygen uptake and absolute bloodflow at rest. The same variables are given for maximal exercise (in brackets).

Organ system	*Distribution*	*A-v difference*	*O_2 uptake*	*Bloodflow*
	Flow%	*ml STPD* l^{-1}*	*ml STPD*min^{-1}*	*ml*min^{-1}*
Splanchnic	*27 (2)*	*40 (80)*	*60 (40)*	*1500 (500)*
Kidneys (300 g)	*22 (2)*	*12-14 (28)*	*16 (17)*	*1200 (600)*
CNS	*14 (1)*	*60 (120)*	*45 (36)*	*750 (300)*
Myocardium (250 g)	*4.5 (6.7)*	*140 (190)*	*35 (380)*	*250 (2000)*
Muscle (35 kg)	*19 (88)*	*50 (160)*	*53 (4200)*	*1050 (26 250)*
Other organs	*14 (1-2)*	*50 (100)*	*38 (35)*	*750 (350)*
Total body	*100 (100)*	*50 (150)*	*250 (4500)*	*5500 (30 000)*

A top athlete can show a 6-fold increase in cardiac output from 5 to 30 l of blood each min, when going from rest to maximal dynamic exercise. The heart rate increases from 60 to 180-200 beats per min. The muscle bloodflow can rise from 3 to 75 ml per min per 100 g of muscle tissue (FU) or factor 25 in a total muscle mass of 35 kg. The muscular arterio-venous-O_2 content difference can rise from the resting level (200 - 150) = 50 ml STPD per l of blood to (200 - 40) = 160 ml STPD per l.

At rest the athlete typically has an oxygen uptake of 250 ml STPD per min. The total muscle bloodflow at rest is (35 000/100) ×3 = 1050 ml of blood per min. The total

muscular oxygen uptake at rest is (1050 50/1000) = 53 ml per min (Box 8-1).*

During maximal dynamic activity the total muscle bloodflow is: (35 000/100)×75 = 26 250 ml/min or 26.25 l per min. The total muscular oxygen uptake is increased to (160 × 26.25 l per min) = 4200 ml STPD per min (Box 8-1).

Accordingly, the total muscular oxygen uptake rises by a factor of (4200/53) almost 80 from rest to exercise.

At the start of exercise, signals from the brain and from the working muscles bombard the cardiopulmonary control centres in the brainstem (see Chapter 18). Both cardiac output and ventilation increase, the α-adrenergic tone of the muscular arterioles falls abruptly, whereas the vascular resistance increases in inactive tissues. The systolic blood pressure increases, whereas the MAP only rises minimally during dynamic exercise. The total peripheral vascular resistance (TPVR) falls during exercise towards 30% of the level at rest, because of the massive vasodilatation in the muscular arterioles of almost 35 kg muscle mass (Eq. 8-3). This is why the major portion of the cardiac output passes through the skeletal muscles (Fig. 18-1) and why the diastolic pressure often decreases during exercise. At moderate exercise the skin bloodflow and heat dissipation is increased (Chapter 21).

The coronary bloodflow increases from rest to exercise (Fig. 10-7 A to B).

2. Haemopoiesis

Haemopoiesis is the formation of blood cells. All blood cells are derived from the multipotent stem cells. Stem cells produce erythroid cells, granulocytes, lymphoid cells, megacaryocytes and monocytes by a number of differentiation steps. Stem cells maintain normal cell populations in a healthy bone marrow controlled by haemopoietic growth factors, and stem cells have the capacity for self-renewal. Haemopoietic growth factors include erythropoietin, interleukins, glucocorticoids, sex hormones and thyroid hormones.

Stem cells and red cell precursors contain ribosomal RNA along with cell organels. The cells lose organels during maturition. Pronormoblasts, normoblasts and reticulocytes at each stage contain less RNA and increasing amounts of haemoglobin. Reticulocytes can still synthesise haemoglobin, have lost the nucleus, and remain in the bone marrow a few days before they enter the peripheral blood. Here, they lose their RNA after a couple of days and become mature red cells. *The reticulocyte count* is normally less than 2.5% of the red cell count, but following haemorrhage or haemolysis the reticulocyte-% increases reflecting increased erythropoiesis. When the bone marrow fails to respond to anaemia, the reticulocyte count may fall below 0.5%.

The normal haematological ranges are given in Box 8-2, together with other values of interest.

Box 8-2: ***Normal haematology values. The normal range varies from one laboratory to another.***

Red cell count	4-6*10^{12} l^{-1}
Leucocyte count	4-11*10^{9} l^{-1}.
Reticulocytes	0.5-2.5% of red cells
Platelet count	150-400*10^{9} l^{-1}
Mean Corpuscular Volume (MCV)	80-96 fl
Mean Cell Haemoglobin Concentration	320-350 g l^{-1}
Mean erythrocyte lifespan	120 days
Haemoglobin (mol. weight monomer)	16 115 Dalton
Haemoglobin concentration (mean)	9.18 mM (149 g l^{-1} = 100%).
Packed cell volume (PCV, haematocrit)	40-50%.
Oxygen binding capacity (haemoglobin)	1.34 ml g^{-1} (60 mmol kg^{-1})
Oxygen concentration in arterial blood	200 ml STPD l^{-1}
Erythrocyte sedimentation rate (ESR)	Less than 20 mm in the first hour
Osmolality of plasma	290 mOsmol (kg water)$^{-1}$

When normal kidneys are perfused with hypoxaemic blood, the peritubular interstitial cells release large amounts of the glycoprotein hormone, *erythropoietin*, with a strong effect on the haemopoietic stem cells in the red bone marrow. The stem cells are stimulated to produce proerythroblasts, which speed up the production of new red cells after a few days. The increased *erythrogenesis* improves tissue oxygenation, which decreases erythropoietin production and the balance is re-established.

Chronic renal failure leads to erythropoietin deficiency, and thus to anaemia, which is of the normochromic, normocytic type.

3. The red cells

Haemoglobin is synthesised in the mitochondria of the maturing red cells. Vitamin B_6 is a co-enzyme for the formation of d-amino-laevulinic acid (ALA) by ALA-synthetase. The reaction is stimulated by erythropoietin. One haemoglobin molecule binds 4 oxygen molecules at most. Haemoglobin consists of *globin* (2 a and 2 b polypeptide chains) and 4 prostetic *haem*-groups (Fig. 8-3). Haemoglobin A (for Adult) has a molecular weight of 64 460 g per mol (Dalton). Haemoglobin A comprises almost all haemoglobin in adults, supplied with only a minimum of haemoglobin A_2.

The polypeptide chains are not covalently linked but are held together by hydrophobic forces. Each haem group is connected to one polypeptide chain, which contain a ring of 4 imidazol-groups. In the centre of the porphyrin ring the one iron atom is coordinated by 6 ligands, four of which bind the metal to the porphyrin chain, one to histidin on either the a- or the b-chains. The last is an open binding, which is able to bind either O_2

or carbon monoxide (CO).

In the lung capillaries haemoglobin is saturated with oxygen at high tensions, where the affinity of (oxy)haemoglobin for more oxygen is high (Fig. 8-3). The affinity between oxygen and haemoglobin is defined and described in Chapter 15, where P_{50} is introduced as an affinity index. A low P_{50} equals a high standard affinity and vice versa. The successive change in affinity during binding of the 4 oxygen molecules to each haemoglobin is caused by molecular interactions among the 4 haem groups. This explains the sigmoid shape of the oxygen dissociation curve (Fig. 15-3). Oxygen is released at the low tensions of the tissues, where the affinity of (deoxy)haemoglobin for oxygen is low. The oxygen tension in the tissue mitochondria may reach extremely low values (zero to 1 mmHg or 0.133 kPa).

Red cells do not contain mitochondria, so they survive on anaerobic metabolism (glycolysis) and the anaerobic intermediate, *2,3-diphosphoglycerate* (2,3-DPG), is produced by the help of a red cell enzyme. As the 4 haem units successively unload oxygen, the b-chains of deoxyhaemoglobin are pulled apart, and 2,3-DPG binds strongly to the 2 b-chains of deoxyhaemoglobin (Fig. 8-3). This electrostatic binding substantially reduces the affinity between oxygen and haemoglobin. – Individuals with high arterial pH (chronic alkalosis) or with low arterial oxygen tension (hypotonic hypoxaemia) increase their concentration of 2,3-DPG in their red cells. Storage of blood reduces the 2,3-DPG concentration with time.

Fig. 8-3: Model of oxyhaemoglobin (relaxed binding structure) and deoxyhaemoglobin (tight binding structure). The circular disc with Fe is haem.

When haem is bound to O_2 or CO, it has a cherry-red colour, and haem is dark red when it is in the deoxygenated form. The breakdown of haemoglobin liberates CO and produces bilirubin that is yellow in colour. Bilirubin is normally excreted with the bile. Failure of bile excretion leads to accumulation of bilirubin in the body. *Jaundice* (icterus) is a yellow pigmentation of the skin, plasma, cell membranes and secretions with accumulated bilirubin and other bile-pigments. Bilirubin and other pigments are also found in the blue-yellow skin-spots following lesions with subcutaneous bleeding.

Notice that when blood is saturated under the normal, ambient O_2 partial pressure (20 kPa = 150 mmHg), the *oxygen capacity of haemoglobin* is 1.34 and not 1.39 ml STPD g^{-1} (Fig. 8-3). The latter holds only for extremely high partial pressures (above 45 kPa), when breathing pure oxygen or oxygen enriched air, where the oxygen capacity is equal to the theoretical.

The rate of fall of red cells is called the *erythrocyte sedimentation rate* (ERS). The ERS is measured in a glass column of whole blood with anticoagulant. ERS is measured in mm as the cell free yellow zone above the red cells following 60 min of sedimentation. ERS is an estimate of the acute phase response. The acute phase response produces high levels of large *sticky proteins* (C-reactive protein, immunoglobulins, fibrinogen) that form rapidly falling piles of red cells. ERS is abnormally increased (above 20 mm) in infections, immunology reactions, ischaemia, malignancy or traumas. Normally, the level is only a few mm per first hour, 15-20 with a common cold, and 50-100 during pregnancy.

4. Viscosity of blood

Viscosity is the *inner friction in the fluid*, which is due to the interaction between molecules and particles in the blood passing a cylindrical vessel. Telescope cylinders (laminae) of blood sliding against each other (Fig. 8-4) can illustrate this inner friction. The outermost blood cylinder rests against the vessel wall (velocity is zero), and the central cylinder moves (laminar flow) with the greatest velocity (v). The velocity profile is parabolic. The *velocity gradient*, with the distance x from the centre of the blood vessel towards the outermost blood cylinder, is called the *shear rate* (dv/dx). The tangential force (F) between these blood cylinders depends upon the area (A) sliding against each other, and the relation to viscosity (η) is given by the equation in the legend to Fig 8-4.

Fig. 8-4: Blood vessel with red cells and arrows shoving different velocity (v).

***F/A* = $\eta \times dv/dx$. The viscosity (η) one Pascal sec (1 Pa s) is the tangential force, working on 1 m^2 of surface area, when dv/dx is 1 (s^{-1}).**

This simplified description is valid for water, gas, and other homogenous fluids that are Newtonian fluids. *Newtonian fluids* are defined as *fluids with a viscosity that is independent of the shear rate*. Newtonian fluids move streamline or with so-called ideal laminar flow.

The viscosity of non-Newtonian fluids decreases with increasing shear rate, according to the equation above. Blood is namely not homogenous with a viscosity that is independent of shear rate. On the contrary, at low shear rates (low bloodflow), the viscosity of blood can be ten-fold higher than normal. The typical normal viscosity of body warm blood is 5 centiPoise equal to 5 milli-Pascal seconds or 5 (mPa*s).

Blood viscosity depends upon the concentration of red cells (the haematocrit).

Fig 8-5: Haematocrit (PCV) and relative viscosity varies along the green line. A normal PCV of 45% is shown with the normal absolute viscosity of body-warm blood.

A patient with anaemia and a PCV of 30% has a low blood viscosity and a poor oxygen transport capacity (Fig. 8-5). On the contrary, a patient with polycythaemia and a PCV of 60% has a high oxygen transport capacity, but the blood viscosity is dangerously high and he may develop thrombosis and emboli (Fig. 8-5).

With increasing bloodflow (and shear rate), an increasing fraction of red cells is being pulled into the axial stream of small vessels, so that friction is being minimised. At high shear rates in large vessels, blood therefore mainly behaves like a Newtonian fluid, with a low and almost constant viscosity, as well as a linear relation between bloodflow and the driving pressure.

The viscosity of blood apparently decreases in tubes with a diameter less than 0.5 mm (the small-diameter effect - or the Fåhraeus-Lindqvist phenomenon - see Fig. 8-6).

Fig. 8-6: The viscosity of blood decreases abruptly in tubes with diameters decreasing from 0.5 mm (Fåhræus-Lindquist effect).

This is because the packed cell volume (PCV) is low in small vessels, since red cells have a tendency to accumulate and pass as a single plug in the fast axial stream, where there is a negligible friction. The slower layers along the vessel wall are passed mainly by plasma. This falling viscosity in the small resistance vessels and in the precapillaries and capillaries reduces the work of the heart. This is why the bloodflow frequently rises *linearly* with the driving pressure and thus actually follows *Poiseuille´s law*, as if blood was a Newtonian fluid.

Bloodflow tends to become turbulent in irregular vessels, where the flow velocity is high and the viscosity is low. Turbulence means irregular movements of the fluid elements - an energy demanding transport process.

Plasma viscosity is sometimes measured instead of erythrocyte sedimentation rate (ESR), because it is dependent of the same large sticky protein molecules as ESR, but is independent of the haemoglobin concentration and obtainable within 15-20 min.

5. Blood coagulation

Whole blood consists of a fluid (plasma) in which blood cells and platelets are suspended. Blood cells consist of red cells (erythrocytes) and white cells (leukocytes). A small amount of anticoagulant to a blood sample blocks the coagulation process, and whole blood sediments into three layers: Below the heavy red cells, then a thin grey-white layer of white cells, and above a yellow fluid (plasma) with an invisible content of most of the platelets. A blood sample without anticoagulants normally sediments with coagulation (fibrin formation) within 5 min. A firm red mass is formed, and after some time it retracts and forms a red cone (a fibrin clot of blood cells and fibrin) surrounded by yellow serum.

Healthy humans possess both a fast extrinsic and a slow intrinsic clotting system. The coagulation process involves at least 3 systems all contributing to the haemostasis. Firstly, a *vasoconstriction* occurs following release of serotonin from damaged endothel cells. Secondly, the *fast extrinsic system* goes into action, and thirdly, the *slow intrinsic system* contribute. Finally, the 2 coagulation systems operate together and converge for common reactive steps in order to produce thrombin (Fig. 8-7).

Disruption of the endothelial barrier by injury initiates a cascade of catalytic events through either or both clotting systems. At each reaction in the chain of events, a proenzyme coagulation factor is activated to its enzymatic form, which can activate the next reaction in the chain. The letter **a** stands for the active form. The enzymes are all endopeptidases (proteases), and their catalytic sites include a serine moiety. By these many steps in the cascade, the process escalates until large amounts of thrombi are released. - Factor IV (Ca^{2+}), factor V (proaccelerin), kininogen, kallikrein, and factor VIII are coagulation co-factors without enzymatic activity.

Thrombin is a protease that is responsible for the formation of fibrin monomers, and thus for formation of a fibrin clot. Its parent molecule is prothrombin (factor II), which is present in normal plasma. Thrombin formation from prothrombin goes through certain cleavage stages, the first of which is by activated factor Xa (Stuart). These reactions are augmented by factor IV (Ca^{2+}), factor V (proaccelerin), and phospholipid (see green oval in Fig. 8-7). Thrombin initiates blood platelet aggregation, and disintegrates the plasma membrane of the platelets so phospholipid is provided. The coagulation factors are synthesised mainly in the liver. - An exception is the large Von Willebrands factor (vWf) complex, which is synthesised in the vascular endothelial cells and in megakaryocytes.

The *fast extrinsic* thrombin formation is initiated by the contact of blood with injured cells (Fig. 8-7). The damaged cells liberate a clot-promoting agent, factor III or *tissue thromboplastin*. Factor III interacts with a plasma protein, factor VII, to start a cascade of reactions by prothrombin activators leading to formation of thrombin within seconds (Fig. 8-7).

Clotting of blood implies conversion of a soluble plasma protein, factor I (or fibrinogen), into an insoluble network of fibrin. First, fibrinogen undergoes limited proteolysis by thrombin. The formed fibrin monomers polymerise into insoluble strands of fibrin polymers (Fig. 8-7). Finally the monomers of the fibrin strands are cross-linked by the enzyme activated (a) fibrin-stabilising factor (XIIIa).

Fig. 8-7: Blood coagulation and fibrinolysis. The roman numbers were originally introduced as a short-cut.

When venous blood is drawn in silicone coated tubes and centrifuged for the separation

of cells and plasma, the isolated plasma clots readily, due to the negative surface charge of glass.

In the absence of thromboplastin, thrombin is formed via the *intrinsic clotting system*. Negatively charged surfaces on damaged cells generate thrombin, trigger inflammatory and immune responses and even activate fibrinolysis.

The first step is that negatively charged surfaces (artificial or injured endothelial barrier) activate factor XII (Hageman) to XIIa, which can activate factor XI in the presence of kininogen. The factor XIa activates the vitamin-K-dependent protein, Christmas factor (IX). Christmas factor is synthesised under the control of a gene on the X-chromosome. Activated Christmas factor (IXa) converts factor X to its activated state (Stuart factor Xa). Stuart factor is a plasma proenzyme - also vitamin-K-dependent. The Xa is the enzyme immediately responsible for the release of thrombin, and the final steps of the two clotting systems are identical (Fig. 8-7). Hepatocytes produce factors X, IX, VII, and II only when vitamin K is present. Insufficient synthesis of these coagulation factors can lead to serious bleeding.

When the endothelial surface of the vascular system is disrupted, platelets normally adhere instantly to exposed structures (collagen and other fibres). Adherent platelets discharge ADP and other substances. Adherent platelets become spherical and send out spicules that look like the legs of a spider. The platelet plug grows and forms a firm haemostatic plug that stops the bleeding. Platelets provide substances that enhance thrombin production, such as phospholipid, the important cofactor in the clotting process.

Blood has the ability to dissolve clots. Fibrinolysis is the dissolution of fibrin. The hepatic plasma glycoprotein proenzyme, plasminogen, is activated to the serine protease, plasmin (Fig. 8-6). Streptokinase, staphylokinase and urokinase convert plasminogen to plasmin. The tissue plasminogen activators are serine proteases. Stress, muscular activity and emotional crises enhance fibrinolysis. Plasmin digests fibrin, fibrinogen and other clotting factors. If plasmin is formed in blood plasma devoid of clots, it is irreversibly inhibited by a_2-antiplasmin (Fig 8-7).

The coagulation process is normally modulated to the needs of the person by inhibitors within the blood. Antithrombin III is the main inhibitor of thrombin and factor Xa, and its effect is potentiated by heparin. Heparin is a negatively charged mucopolysaccharide from mast cells. Heparin binds to antithrombin III forming a complex that rapidly binds serine proteases such as thrombin, thus functioning as a potent anticoagulant. Heparin alone does not inhibit the coagulation process significantly.

Fibrinolysis is inhibited mainly by a_2-antiplasmin, because plasmin combines with antiplasmin in an irreversible link (Fig 8-7).

Vitamin C or ascorbic acid cannot be synthesised in humans, but the vitamin is present in all fresh fruit and vegetables. Hydroxylation of proline to hydroxyproline is necessary for the formation of collagen and thus of the normal tissue including blood vessels. Lack of vitamin C (scurvy) leads to defective blood vessel walls with spontaneous haemorrhage and blue spots.

6. Vascular compliance and stiffness

Distensibility or compliance is the increase of volume per unit of transmural pressure increase (DV/DP_t). The specific compliance is the relative increase in volume per unit of pressure increase. The elastance or stiffness is the reciprocal value of the compliance. The compliance of the venous system can be 30 times as large as that of the arterial system.

The venous system can be expanded to contain more than 75% of the total blood volume. The veins function as capacitance vessels, and become very distended when blood is given in transfusions, in heart insufficiency, or during a heart attack. Severe exercise and loss of blood cause an increase in venous tone, which for a period actually can increase the circulating blood volume. During hard work the muscular venous pump provides up to 1/3 of the energy required for blood circulation (the peripheral venous heart). The venous system also plays an important role by its graded venous return to the heart.

Fig. 8-8: Recording of the arterial blood pressure at rest. Dashed, horizontal lines depict MAP. The pulse pressure of the abdominal aorta is calculated from the arterial compliance of a young person at rest.

The recording shows a systolic peak pressure, a dicrotic notch as the aortic valves close, and a falling diastolic pressure. The yellow area under the dashed green line equals the yellow area above the line (Fig. 8-8).

The mean arterial pressure, MAP, is usually being defined as the diastolic pressure plus 1/3 of the pulse pressure (Fig. 8-8). The mean arterial pressure (MAP) is about 12 kPa (= 90 mmHg) in the arteries. Notice the *fall* in MAP from the abdominal aorta to the femoral artery, whereas the systolic pressure *increases.* The arterial mean pressure falls to a mean value around 2.4 kPa (18 mmHg) in the capillaries.

The *arterial pulse pressure* is the difference between the systolic and the diastolic arterial pressure. At a heart rate of 75 beats/min at rest, the cardiac cycle length is 0.8 s with 0.3 s systole and 0.5 s diastole. A stroke volume of 70 ml is deposited in the aorta and the larger elastic arteries during systole. During the systolic period 26 ml of blood (70×3/8) is streaming through the resistance vessels, leaving the arterial system, so the systolic *volume expansion* is 44 ml of blood. A young healthy subject has an arterial

distensibility or compliance of 1 ml of blood per mmHg, which creates a pressure rise during systole (pulse amplitude) of (70-26) = 44 mmHg (Fig. 8-8). With a diastolic pressure of 70 mmHg, this implies a systolic pressure of 114 mmHg, conventionally written 114/70 mmHg or 15.2/9.3 kPa.

Aging and arteriosclerosis increase the stiffness (reduce the distensibility) of the elastic arteries, causing the arterial compliance to fall from 1 (one) to 0.5 ml of blood per mmHg. In this case, a systolic volume expansion of 44 ml of blood increases the pulse pressure amplitude to 88 mmHg (44/0.5=88), and the blood pressure to perhaps 180/92 mmHg. This is a likely process in an otherwise healthy person of advanced age. Typically, the average diastolic pressure will rise with age.

7. Wall tension

For a *thin-walled organ* with two main radii, Laplace predicted that thetransmural pressure at equilibrium (DP_t), was identical with the fibre tension in the wall (T) divided by the two main radii: $DP = T/(r_1 + r_2)$. This model has often been used (withmodifications for wall thickness, w) for the relaxed ventricle (Fig. 8-9A).

Fig. 8-9: Laplace models for the relaxed ventricle (A), the spherical alveole (B), and the cylindrical blood capillary (C).

For a *thin-walled spherical organ* ($r_1 = r_2 = r$), another Laplace equation can be developed from the equation above (Fig. 8-9B). This model is often used for both alveoli and the spherical ventricle. When the left ventricle becomes more and more spherical by diastolic filling, the T will rise with the transmural pressure. The radius increases with the end-diastolic volume. The more the end-diastolic pressure and the fibre tension rises, the higher is the energy demand and the more O_2 is consumed per heart beat during contraction of the dilatated ventricle.

For an *infinitely long thin-walled cylinder*, like a true capillary or a preferential channel (see below), the r_2 approaches infinity and has no influence on the transmural pressure. Hence, the Laplace equation can be approximated by eq.C in Fig. 8-9C. This model is used for a thin vessel wall, since T/r_2 approaches zero. Surprisingly enough, the thin endothelial barrier (0.3 mm) of a capillary easily carries a pressure of 4.3 kPa (32 mmHg) or more. This is because the capillary radius is so small. According to eq. C in Fig. 8-9C, a small radius (5-10 mm) must imply a small wall tension (T).

In hypertension, the arterial walls hypertrophy, so the wall tension is minimised and hence the risk of vessel rupture (use Eq. 8-9C with correction for **w**all thickness: $T = DP^* r/\mathbf{w}$).

8. Microcirculation

The microcirculation is responsible for the transport of nutrients and oxygen to the tissues, and for removal of cellular waste products and CO_2. The arterioles control the flow of blood to each tissue unit, and the metabolic conditions of the tissue cells determine the diameters of the vessels. Hereby, the tissue unit often controls its own blood flow by local mechanisms.

A *microcirculatory unit* is a collection of vessels that originate from an arteriole, which is characterised by well-developed smooth musculature in its wall (Fig.8-10). Arterioles of the face, fingers and toes often branches into an *arteriovenous anastomose*, which functions as a shunt vessel, but which also can be closed completely. In certain tissues the arteriole branches into metarterioles (with so-called *precapillary sphincters* of smooth muscle fibres without nervous supply), which continue into large capillaries termed *preferential channels* (or thoroughfare channels). These channels shunt the blood to the veins. The small true capillaries have only a thin endothelial cell layer making the wall ideal for exchange

Fig. 8-10: A microcirculatory unit.

The diameter of *true capillaries* is only 5-10 mm, barely enough for erythrocytes to squeeze through. The average length of capillaries is 1 mm, and the linear red cell velocity at rest varies around 1 mm each s. The capillary density is high in cardiac and striated muscle tissue and low in subcutis and in cartilage. Endothelial cells contain actin and myosin. It is uncertain whether capillaries may be able to alter their shape according to the needs of the tissues.

Important *exchange vessels* are thin-walled vessels with a large surface area. Exchange vessels comprise true capillaries, parts of preferential channels, and venules (Fig. 8-10). The number of pores is high in the venous ends of capillaries and in venules. Exchange vessels are any blood vessels, which allow transport of substances through its wall in both directions. The velocity of the bloodflow in capillaries varies, sometimes with rhythmic pulsation, at other times random.

At rest the intracapillary pressure varies from arteriole to venule between 3.3 and 1.6 kPa (25 and 12 mmHg), during arteriolar vasoconstriction between 1.6 and 1 kPa (12 and 8 mmHg), and during vasodilatation between 5.3 and 1.6 kPa (40 and 25 mmHg). Arterial pressure fluctuations have been recorded even in the most distal parts of the capillaries. In venules and veins, however, the flow is smooth without fluctuations.

The capillary wall consists of a layer of endothelial cells (0.1 - 1 mm of thickness) resting on a basement membrane. At least three types of capillaries are present in humans:

1. Continuous capillaries are the most abundant. The distance between endothelial cells

is 5-30 nm (Fig. 8-11). Tight junctions with narrow clefts are difficult to pass for the dissolved molecules and ions. In the continuous capillaries, the water filled pore surface area comprises only 10^{-4} of the total surface.

The continuous capillaries in the brain are low permeable to ions and most hydrophilic molecules, because their tight junctions are really tight (*the blood-brain barrier*).

2. Fenestrated capillaries contain tight junctions and pores or *fenestrations*, which are fluid filled channels with a diameter of 50-100 nm. These are formed by two adjacent cell membranes that have fused during removal of the lipid bilayers, so only a diaphragm of protein lattice is left allowing bulk flow without colloids (Fig. 8-11). Fenestrations are round windows found in the capillaries of organs that transport lots of water (the bowels, glomerular capillaries of the kidneys, pancreas and salivary glands). In each fenestration a bush-like filaments can be demonstrated by electron micrography (Rostgaard). The filaments are composed of a protein core with glycosaminoglycan side chains. The filaments and the protein lattice in the fenestrae keep plasma proteins back (Fig. 8-11). In the glomerular capillaries, water filled fenestrations cover 20% of the surface.

Fig. 8-11: Three types of capillary walls.

3. Sinusoid capillaries have very broad openings between the endothelial cells (Fig. 8-11). These large fenestrations have no diaphragm. Sinusoid capillaries are often found in tissues that are bathed in plasma (liver, spleen and bone marrow).

The *circumventricular organs* of the brain contain lots of fenestrations in the walls. The circumventricular organs are located close to the control centres of the hypothalamus and the brainstem. Any penetration of signal molecules in the neighbourhood of these control centres is of physiological importance. - In other areas with continuous capillaries, most substances cannot bypass the blood-brain barrier and reach the brain cells.

9. Transcapillary fluid exchange

Starling hypothesised that the fluid exchange across the capillary wall was determined by the hydrostatic (P_c) and the colloid osmotic pressure (p_c) in the capillary (Fig. 8-12).

Fig. 8-12: Transcapillary fluid exchange (Starling) is shown over a capillary wall. The pressures are in mmHg. The capillary filtration coefficient is in ml*min^{-1}*kPa^{-1}*(100 g tissue)$^{-1}$. One mmHg equals 133.3 Pa.

The flux of substance (J) over the capillary membrane is determined by ($P \times DC$). Actually, this is nothing but an extension of *Fick's law for diffusive transport* (Eq. 1-2).

Fluid moves out of the arterial end of the capillary by *filtration*, because the net hydrostatic pressure (35-5 = 30 mmHg) is higher than the colloid osmotic pressure (p_c= 26 mmHg), and most of the fluid (9/10) passes again into the blood by *reabsorption* in the venous end (Fig. 8-12). Here, the colloid osmotic pressure (26 mmHg) supersedes the hydrostatic pressure (15-1 mmHg equals14 or 1.9 kPa).

The net diffusion of water molecules across the capillary wall is approximately zero. Instead, the transvascular exchange is caused by a combination of an *outward ultrafiltration* and an *inward colloid osmotic* force. Ultrafiltration is caused by a hydrostatic pressure gradient created by the heart. The hydrostatic pressure gradient is a net outward force, moving water through pores in the capillary wall. Plasma contains dissolved protein, which cannot pass the small pores in capillary walls readily. The plasma proteins create a *colloid osmotic pressure* of about 3.3-3.7 kPa (25-28 mmHg). This pressure is much larger than the interstitial colloid osmotic pressure, so that the colloid osmotic gradient across the capillary wall is a net inward force, which draws water into the capillaries.

Starling in Eq. 8-7 described the transvascular water flow as early as in 1896. The driving forces are the so-called *Starling forces* (see Eq. 8-7). The *capillary protein reflection coefficient* is symbolized s. s is the fraction of plasma protein molecules reflected off the capillary wall. The protein reflection coefficient is 0.9-1.0 for many capillaries, expressing that the colloid osmotic pressure gradient is not reduced over time by diffusion of proteins over the capillary wall.

The *capillary filtration coefficient* (Cap_f) corresponds to the *permeability* of the capillary wall. In the legs Cap_f is around 0.075 ml of fluid per min per kPa in 100 g of tissue (at body temperature). The combined pressures in the Starling equation ($[(P_c - P_t) - s(p_c - p_t)]$) determine, if there is a *net pressure* for water movement across the capillary wall (Eq. 8-7).

In conclusion, water moves out of the arterial end of the capillary by *filtration*, and near the venule end, water moves into the blood by *reabsorption*. This transport along the capillary is called *Starling´s paracapillary circulation*. Thus there is normally a net filtration of water and some proteins into the interstitial space. This water and protein, returns to the blood via the lymphatic system (1/10 of the total filtration in Fig. 8-12). The lymph volume amounts to approximately 3-5 l daily, and is mainly produced in the liver and intestine. Starling presumed – erroneously - that proteins were unable to leave the blood in the capillaries (Fig. 8-13: A).

Fig. 8-13: Two models of transcapillary fluid exchange. The capillary pressure (Pc) is protected from large changes in MAP, but is sensitive to changes of venous

pressure including the central venous pressure.

This assumption is wrong. The capillaries are almost universally permeable to proteins and macromolecules that resemble proteins.

Another physiologist Drinker found protein in lymphatic fluid. Drinker developed a model, which presumed that capillaries to a variable degree were *permeable* to proteins (Fig. 8-13: B). Within a single capillary, the *protein permeability increases from the arterial towards the venous end*.

Let us assume that the heart is pumping out about 9000 l of blood every day. With a packed cell volume of 45% there is 55% plasma. This means that 4950 l is plasma. With a 6% protein concentration there is a total of 297 kg of protein. If less than 0.1 per cent (1/1440) of this protein is filtered into the interstitial fluid and lymph, it amounts to 206 g of protein daily. This amount of protein leaves the blood in the capillaries, and returns almost completely to the blood through the lymph and not the veins (Fig. 8-13: B). Hence, *Starling's paracapillary circulation* obviously plays a dominating role in the transport of *crystalloids* (small molecules of nourishment and waste products) through the capillary wall.

The *capillary hydrostatic pressure* (P_c) varies from tissue to tissue. It is low in the lungs and intestine (1 kPa) and particularly high in the renal glomerular capillaries (6-8 kPa). In resting skeletal muscle capillaries, the pressure is 4.3 kPa (32 mmHg) at the arterial end and 1.6 kPa (12 mmHg) at the venous end. In general, P_c *increases* whenever the mean arterial pressure (MAP) increases, venule pressure (P_v) or resistance (R_v) increases, or when arteriolar resistance (R_a) decreases, according to the formula: $P_c = [(R_v/R_a)\ MAP + P_v]$ developed in Fig. 8-13. Normally, R_v/R_a is approximately 1/10. Thus P_c is *protected* from large changes in MAP, but is *sensitive* to changes in venous pressure including the central venous pressure (CVP).

In tissues, where the perfusion pressure is reduced to a value below a so-called *critical closing pressure*, the bloodflow ceases due to vessel collapse. This is explained by the Laplace model (Fig. 8-9C).

The *myogenic response* also causes an important deviation from Poisseuille´s law. The myogenic response covers reactions where the vascular smooth muscle contracts in response to increased transmural pressure and vice versa. A decrease in transmural pressure (intravascular minus extravascular pressure) of the precapillary vessels elicits precapillary relaxation. A rise in transmural pressure elicits precapillary contraction. Perhaps the stretch of smooth muscle cells opens Ca^{2+}-channels, whereby a Ca^{2+}-influx increases the intracellular Ca^{2+} concentration sufficiently for contraction.

10. The lymphatic system

Macromolecules do penetrate the capillary wall and the content of lymph derives from plasma. Less than 0.1 per cent of all the plasma proteins that are being ejected from the heart in 24 hours, escapes from the capillaries. The venous end of the capillaries is permeated by pores of 40 - 60 nm. Here, macromolecules can pass by filtration in a pressure determined fluid transport. Passage as a whole plasma portion (bulk flow) through fenestrations is also possible.

Transepithelial solvent transport can also draw solutes by *solvent drag*. Gradient dependent transport concepts such as filtration, bulk flow and solvent drag are used by different groups of scientists. When large amounts of lymph is being produced, solvent drag dominates over diffusion. At low lymph production, half of the protein transport is caused by diffusion. Fluid pass through the cell by pinocytosis.

Capillary filtration predominates over capillary *reabsorption* resulting in an overshoot (a net filtration) of interstitial fluid. Most of the net filtration is reabsorbed into the blood of end-capillaries or venules (Starling´s paracapillary circulation).

The lymphatic vessels drain the remaining filtered fluid (Fig. 8-12). The lymphatics are composed of endothelium-lined vessels similar to blood capillaries. Some lymphatics are equipped with one-way valves, so rhythmic activity in nearby skeletal muscles returns the lymph to the circulation via the thoracic duct. Lymph vessels originate as *blind-ended sacs* close to the blood capillaries. Lymph vessels are permeable to proteins, macromolecules and even to cells from the interstitial fluid. The lymphatic drainage is particularly important for transporting *chylomicrons* absorbed from the intestine, and to return plasma proteins that leaks from several blood capillary systems. Lung tissue has no lymphatics, because the lymphatic vessels end at the terminal bronchioli. The lymph from the liver provides us with 50% of the daily lymph produced.

Lymphatic fluids from liver and kidney have a protein concentration equal to plasma's (6-8 g per 100 ml), and lymphatic fluid from the bronchial tree has a similar concentration of protein.

Lymphatic fluids from skin and muscles contain only 2% protein, and brain lymph contains no protein at all.

Pathophysiology

This paragraph deals with *1. Anaemia*, *2. Oedema*, *3. Thrombosis/Embolism*, *4. Haemophilia* *5. Aneurysms* *and* *6. Valvular diseases*.

1. Anaemia

Anaemia is defined as a condition with an insufficient oxygen carrying capacity of the patients blood. For both sexes and all age groups a blood haemoglobin concentration

below 130 g per l (8 mM) implies reduced working capacity and thus a consequential clinical condition. Reference levels for age and sex are also available, but they differ from laboratory to laboratory.

Mean corpuscular volume (MCV) expresses the mean volume of each red cell. MCV is calculated from the packed cell volume (PCV) by division with the red cell count. An example with normal values provides the following: 0.45 (l/l)/5*10^{12} (red cells/l). Thus MCV is equal to 90*10^{-15} l per red cell. One femtolitre (1 fl) equals 10^{-15} l. The *normal range is 80-96 fl*. The MCV index is used to classify anaemia's into microcytic (MCV<80 fl), normocytic (MCV 80-96 fl) and macrocytic forms (MCV >96 fl), but the classification is not causal.

Mean corpuscular haemoglobin concentration (MCHC) provides the mean concentration in each red cell. MCHC is calculated from the haemoglobin concentration by division with the packed cell volume (PCV). An example with normal values provides the following: 150 (g/l)/ 0.45 (l/l). Thus, normal *MCHC is 333 g per l of red cells*. Since the concentration of haemoglobin in a normal red cell is maximal, the maximal value (380 g/l) is the highest occurring. *Normochromic anaemia's* have MCHC values in the range 320-380 mostly within 320-350 g/l. Anaemia with MCHC below 320 g/l is called *hypochromic*, and they are often also microcytic such as in iron deficiency anaemia.

Anaemias are classified into two groups based on their cause. The first group is *deficiency anaemias* with insufficient haemoglobin production due to dietary/ absorptive defects or to bone marrow hypoplasia from cell destruction by chemicals or radiation (Box 8-3).

Deficiency anaemias are caused by defect haem synthesis (iron deficiency, anaemia of chronic disease, sideroblastic anaemia) or by defect globin synthesis (thalassaemia).

The second group is *waste anaemias* with waste of red cells (Box 8-3). The waste of red cells is caused by bleeding (haemorrhage) or by haemolysis.

Box 8-3: Classification and causes of the two major types of anaemia**.**

A	**Deficiency anaemias cause defect synthesis of haem or globin**	
	A1:Iron-deficency anaemia	(insufficient iron for haem synthesis)
	A2: Anaemia of chronic disease	(defect synthesis of haem).
	A3: Sideroblastic anaemia	(defect synthesis of haem).
	A4. Macrocytic anaemia with megaloblasts in the bone marrow	(due to vitamin B_{12} deficiency or folate deficiency)
	A5. Macrocytic anaemia without megaloblasts in the bone marrow	(pregnancy, newborn, hepatic disorders, hypothyroidism, aplastic anaemia).
	A6. Aplastic anaemia	(too few stem cells in the bone marrow).
	A7. Thalassaemia	(defect globin synthesis).
B.	**Waste anaemias: Waste of red cells**	
	B1. Acute bleeding	(loss of red cells).
	B2.Haemolytic anaemias	(increased destruction of red cells).

A1. *Iron deficiency anaemia* is caused by chronic bleeding, growth, endurance exercise, pregnancy and nursing, poor intake, malabsorption). *Iron deficiency* is characterised by low serum-iron, high total iron binding capacity (TIBC), and a transferrin saturation below 19%.

A2. *Anaemia of chronic disease* (defect synthesis of haem):

1.Chronic bacterial, viral, fungal, protozoal, and helminthic infections (see Ch. 33).

2.Chronic inflammatory diseases (eg, rheumatoid arthritis, polymyalgia etc, see Ch. 32).
3.Malignant disorders. This anaemia is characterised by low serum-iron as well as low total iron binding capacity.

A3. *Sideroblastic anaemia (*defect synthesis of haem) with ring sideroblasts, is genetic or acquired. The genetic type is X-linked and transmitted by the mother. The acquired types are caused by alcohol, drugs, lead, other disorders or the cause is unknown (primary type). Sideroblastic anaemia is characterised normal total iron binding capacity, raised serum-iron and raised serum-ferritin.

A4. *Macrocytic anaemia* with megaloblasts in the bone marrow is due to folate deficiency or to vitamin B_{12} deficiency.

Folate deficiency anaemia is recognised when the folate concentration in red cells low. This deficiency is due to poor intake, malabsorption, antifolate drugs and excess utilization. Since the folate stores of the body are low the anaemia develops rapidly (over months) compared to years for pernicious anaemia.

Folate polyglutamates are synthesized in human cells. These compounds are biologically active, as coenzymes in amino acid metabolism and in the DNA synthesis. The synthesis of the biologically active form of folate is dependent of vitamin B_{12}. Lack of folate inhibits the purine-pyrimidine-DNA-synthesis, and without new DNA cell division is seriously reduced. The typical patient appears with glossitis and a megaloblastic anaemia is found. The amount of folate in red cells is below 160 mg ml^{-1}. The normal range is 160-640mg ml^{-1}.

Pernicious anaemia is the most common cause of vitamin B_{12} (cobalamin) deficiency. Pernicious anaemia is characterised by a low serum-[vitamin B_{12}] (below 160 ng l^{-1}). Megaloblastic anaemia with lack of gastric HCl confirms the diagnosis.

Pernicious anaemia is caused by atrophy of the gastric mucosa, resulting in insufficient synthesis of *intrinsic factor*. The stomach cannot secrete intrinsic factor, hydrochloric acid and pepsin.

Pernicious anaemia occurs in three forms: 1) most patients have an *autoimmune* disorder, with plasma antibodies against their own parietal cells; 2) rarely, new-born babies suffer from *congenital* intrinsic factor deficiency with normal pepsin and acid secretion; and 3) finally as vitamin B_{12} *malabsorption*, because of a defect in the intrinsic factor-B_{12} *receptors* in the terminal ileum.

Vitamin B_{12} malabsorption in adults is caused by one of two intrinsic factor antibodies. One antibody blocks the binding of intrinsic factor to B_{12}, so the protease-resistant complex is never formed. The other intrinsic factor antibody blocks the binding of the intrinsic factor- B_{12} complex to the intrinsic factor-B_{12} receptors of the terminal ileum. The result is vitamin B_{12} malabsorption.

Parietal cell *antibodies* are present in the plasma of 90% of all patients with pernicious anaemia. The parietal cells of the gastric glands fail to secrete HCl and intrinsic factor. Intrinsic factor is a glycoprotein, which combines with vitamin B_{12} of the food. This combination normally makes vitamin B_{12} available for absorption in the ileum. The site of red cell production is the red bone marrow, which is normally one of the most proliferative tissues.

The lack of vitamin B_{12} in the liver and the red bone marrow inhibits the *methyl-malonyl Co-A mutase* and also spoils the *purine-pyrimidine-DNA-synthesis*. The inhibition of these and other processes leads to the neurological and the haematological disorders in pernicious anaemia.

The *neurological* features are *progressive polyneuropathy* with degeneration of the posterior and lateral column of the spinal cord and peripheral nerves (eg, optic atrophy, symmetrical paraesthesia, weakness, dementia and ataxia).

Haematological disorders. Lack of vitamin B_{12} in the bone marrow turns the normal erythroblasts into abnormal megaloblasts. The erythrocyte production is inhibited, and the cells synthesise much more RNA than normal and much less DNA. Besides, the formation of leucocytes and platelets suffer causing leucopenia and thrombocytopenia. Instead of normal erythrocytes, the megaloblasts deliver megalocytes to the circulation. *Megalocytes are fragile and only have an average life of 40 days*, as compared to 120 days for adult erythrocytes.

Cobalamine is the chemical name of vitamin B_{12}. Pernicious anaemia is treated with intramuscular injections of hydroxycobalamin storage, followed by 1 mg every 3 months as long as the patient lives.

A5. *Macrocytic anaemia* without megaloblasts in the bone marrow is a physiological anaemia in pregnancy and in new-born babies. This anaemia is also found in patients with alcohol abuse, hepatic disorders, hypothyroidism, and aplastic anaemia. The concentration of vitamin B_{12} and folate in the plasma is normal. The relative number of reticulocytes and the MCV is increased. – In some cases there is fat accumulation in the red cell membrane, but the pathogenesis of these conditions is not clarified.

A6. *Aplastic anaemia* refers to a condition of bone marrow failure with only few pluripotent stem cells in the bone marrow. This is due to immune suppression of stem cells by T suppressor cells, or to direct destruction of the stem cells caused by chemicals, drugs, infection or radiation. Pancytopenia, absence of reticulocytes and an aplastic bone marrow is characteristic.

A7. *Thalassaemia* (see Chapter 33).

B1. *Acute bleeding* (loss of red cells). Normochromic normocytic anaemia occurs following an acute bleeding with plasma dilution, before the iron stores are depleted. - Lack of vitamin K can change the development of even a simple tooth bleeding to a serious condition.

B2. *Haemolytic anaemias* (increased destruction of red cells): They are inherited or acquired. Inherited are hereditary spherocytosis or ellipsocytosis, thalassaemia (defect synthesis of globin -see Chapter 33), Sickle syndromes, etc. Acquired haemolytic anaemias are caused by immune destruction of red cells, membrane defects (paroxysmal nocturnal haemoglobinuria, mechanical destruction of cell membranes, haemolysis caused by renal, endocrine or liver disease. Haemolytic anaemia is characterised by osmotic fragility, reticulocytosis, increased serum-bilirubin, and erythroid hyperplasia of the bone marrow.

General for anaemia

In most cases of anaemia the fall in transport capacity develops slowly, whereby there is time for physiological adaptations to minimise symptoms and signs. A rise in 2,3-DPG improves the release of oxygen to the cells. Unspecific symptoms such as fatigue, headaches and faintness have varying origin and are not always recognised as a disease. Dyspnoea, palpitations, cardiac cramps, and intermittent claudication are also difficult to interpret. The signs of anaemia are tachycardia, systolic murmur over the

heart, and cardiac failure. Drumstick fingers with spoon-shaped nails are seen in chronic anaemia with hypoxia such as in chronic iron deficiency. Jaundice suggests the possibility of haemolytic anaemia.

The falling red cell count reduces the oxygen delivery but also leads to falling viscosity of the blood. The reduced viscosity can reduce the total peripheral vascular resistance (TPVR) to less than half of the resting value, which is an appropriate event, since it easens the cardiac work and improves the bloodflow. A slight fall in systemic arterial pressure reduces the stimulus of the arterial baroreceptors, and causes a rise in heart rate and cardiac output. The low oxygen capacity of haemoglobin is compensated by an increased coronary bloodflow at rest. The myocardial anoxia results in cardiac failure (Fig. 10-10) with oedema, large liver, and stasis of the neck veins. Severe anaemia increases respiration, metabolic rate, and temperature due to the large cardiopulmonary work.

2. Oedema

Oedema is an abnormal clinical state characterised by accumulation of interstitial or tissue fluid. Cutaneous oedemas can be diagnosed by the simple test: pitting on pressure. Theoretically, oedemas are caused by *three* different mechanisms:

1. A hydrostatic pressure gradient, which is too great (so-called *high pressure oedema* or *cardiac oedema* at heart failure with increased venous and central venous pressure),
2. A colloid-osmotic pressure gradient, which is too low and caused by too low concentrations of plasma proteins (so-called *hunger oedema* and *renal oedema*), and
3. Leakage in the capillary endothelium (so-called *permeability oedema* with too much protein in the oedema fluid). Burns cause increased capillary permeability for proteins, by infections or by allergy.

Cardiac oedema develops in the dependent parts of the human body, where the hydrostatic gradient is greatest (see congestive heart failure, Fig. 10-10).

Renal oedema is frequently found in loose tissues, such as the subcutaneous tissue around the eyes (see Chapter 25).

Lymphatic oedema is special form of oedema that can be congenital or acquired. A child born with insufficient development of the lymphatic system will suffer from gradual swelling of the affected body part as a result of accumulation of interstitial fluid. Surgical destruction of lymphatic vessels can result in acquired, lymphatic oedema (eg, following mastectomy).

Inflammative processes, cancer cells or filarias (elephantiasis) also can obstruct lymphatic vessels, so the limbs swell and become oedematous “elephant limbs.”

3. Thrombosis and embolism

Thrombosis refers to a condition with formation of multiple thrombi or clots within the vascular system. The cause can be damage of the vessel wall, reduced bloodflow, increased viscosity and hypercoagulability of the blood.

Embolism refers to the process through which a thrombus is dislodged from its attachment and travels with the blood until it is lodged in a blood vessel too small to allow its passage. The flowing blood carries emboli from thrombus material in the deep pervic or leg veins to the lungs, where they block the bloodflow as life-threatening *pulmonary emboli.*

Venous thrombosis is is frequently related to peripheral artery disease or to immobilisation. Bed rest or long immobilisation as during long flights can result in deep venous thrombosis, presenting with pain in the calf and ankle oedema. Anticoagulation therapy and elastic support stockings are used to reduce the risk of pulmonary embolism.

4. Haemophilia

The bleeding disorders known as *haemophilia* are relatively seldomly occurring, but vitamin K deficiency must be recognized as a common and serious bleeding disorder, which can give rise to acute bleeding anaemia (B1).

Haemophilia A is the most frequent *genetic* disorder of the intrinsic clotting system, characterised by a low coagulant concentration of *antihaemophilic factor (VIII).* This disorder is linked to the X-chromosome, and haemophilia affects only males, who transfer the abnormal gene to their daughters, all of whom are carriers. The female carrier of the abnormal gene is usually without symptoms and signs of disease.

Haemophilia B (Christmas disease, Factor IX deficiency) is not as common.

Most haemophiliacs suffer episodes of spontaneous bleeding. Repetitive joint bleeding (haemarthrosis) leads to *crippling arthritis.*

The *activated* partial thromboplastin time tests the competency of the slow intrinsic clotting pathway. The contact factors are maximally activated by first mixing citrate plasma with powdered glass. Then partial thromboplastins (V, cephalin, and inosithin) are added. After addition of phospholipid and Ca^{2+}, the time it takes for coagulation to occur is measured. This is a preferential test of the intrinsic clotting pathway, because factor III (tissue thromboplastin from injured cells) is not available to trigger the extrinsic clotting pathway (Fig. 8-6). Normal values are 35-45 s; the time is prolonged in blood from patients with circulating anticoagulants. The time is also prolonged in

haemophilia and in other disorders with defective intrinsic pathway factors.

Von Willebrands disease. In most forms of Von Willebrands disease the plasma is deficient in both factor VIII and Von Willebrands factor. The disease affects both sexes, which is similar to mild haemophilia. The disorder is *inherited* as an autosomal dominant trait. The *bleeding time* tests the capacity of platelets to form plugs. A blood pressure cuff is applied to maintain venous pressure at 5.3 kPa, and a standardised incision is made on the volar surface of the forearm. Bleeding stops when a proper plug of platelets has aggregated. The incision is blotted with filter paper at 30 s intervals. The normal bleeding time is 4.5 min. The bleeding time is prolonged to at least 10 min in Von Willebrands disease.

5. Aneurysms

Aneurysms are abnormal dilatations on a vessel typically due to degenerative processes in the wall. Aneurysms on brain or coronary arteries may rupture (leading to sudden death), because of their high lateral pressure (Eq. 8-2).

Aortic aneurysms are usually due to arteriosclerosis with large atheromas in the wall. Aneurysms are found as pulsatile dilatations of the abdominal or thoracic aorta (CT scanning or ultrasound examination). Rupture of an aortic aneurysm presents as shock with epigastric pain, and requires immediate surgery. Bleeding inside the wall of the aorta obstructs the lumen (so-called *dissecting* aortic aneurysm), and also here emergency surgery is required.

Left ventricular aneurysm is a complication to ischaemic heart disease often diagnosed by echocardiography (a case is drawn in Fig.10-8).

Saccular aneurysms are found on the circle of Willis and its adjacent branches. Pulsations cause pressure on surrounding structures, and spontaneous rupture often causes sudden death.

6. Valvular disease

Opening and closure of cardiac valves is studied with echocardiography. This is a versatile non-invasive technique used by cardiologists. When valvular diseases cause the valves to open too little (*stenosis*) or not close firmly enough (*insufficiency*), the function of the heart is severely impaired.

Valvular disorders are treated in Chapters 10 and 12.

Equations

- *A geometrical argument:* The relationship between linear mean velocity (v^-) and the bloodflow in one s ($Q^\circ s$) is determined by the cross sectional area (A):

  Eq. 8-1: $Q^\circ s = v^- \times A$.

- *Bernoulli's equation* (see Chapter 13) states that the total driving energy, applied to a continuously flowing, small, ideal fluid volume (d*V*), which is flowing frictionless and laminar, equals the sum of 3 types of energy the kinetic energy ($1/2\ \rho\ v^2$ - fluid density (ρ) multiplied by the squared velocity) , the potential energy at the height (h) and the gravity (G), and the laterally directed energy (ie, the lateral pressure, *P*, directed towards the walls).

  Eq. 8-2: Total energy zero = $dV\ (1/2\ \rho * v^2 + h * \rho * G + P)$.

  The lateral pressure is highest, where the velocity is lowest (eg, aneurysm ruptures). The equation of continuity states that the velocity varies inversely with the cross-sectional area of the tube. Consequently, the lateral pressure is highest where the cross sectional area of the tube is largest.

Poiseuille´s law: The volume rate (*Vdot*) is equal to the driving pressure (D*P*) divided by the resistance: V° = D*P*/Resistance. For the left ventricle, the bloodflow is actually cardiac output (Q°), so the equation reads:

Eq. 8-3: $Q^\circ = DP/TPVR$ ($l\ min^{-1}$).

The driving pressure (D*P*) is the mean arterial pressure (MAP) minus the atrial pressure, and *TPVR* is the total peripheral vascular resistance.

TPVR is directly related to the blood viscosity (h) and to the length (L) of the vascular system, and inversely related to its radius in the 4^{th} power:

Eq. 8-3a: $TPVR = 8\ h\ L/r^4$.

Doubling the length of the system only doubles the resistance, but halving the radius increases the resistance sixteen-fold. - *Poiseuille´s law is an approximation!*

- *Vascular Resistance in parallel organs*. In the systemic or peripheral circulation the resistance in the single organs are mainly placed in parallel, and the resistance of all organs (R_1 toR_n) are related to the total (TPVR) by the following relation:

  Eq. 8-4: $1/TPVR = 1/R_1 + 1/R_2 + ...1/R_n$.

- *Vascular resistance in portal circulations*. There are only a few *serially* connected elements (portal circulation): Spleen/liver, gut/liver, pancreas/liver and hypothalamus/pituitary. For serial arranged resistance the formula is:

  Eq. 8-5: $R_{total} = R_1 + R_2 + ...R_n$.

- *The law of Laplace*. For a thin-walled organ with two main radii, the relationship between transmural pressure (D*P*) and tension (*T*) is determined by the radii:

  Eq. 8-6: $DP = T/(r_1 + r_2)$. – See Fig. 8-9.

- *The Starling equation*

  Starling described the transvascular fluid flow (J_f, volume per min in 100 g of tissue), determined by the combined effect of the Starling forces, in 1896 in the equation:

  Eq. 8-7: $J_f = Cap_f \times [(P_c - P_t) - s(\pi_c - \pi_t)]$.

Cap_f is the capillary filtration coefficient (ml of fluid per min per kPa in 100 g of tissue). The Starling forces are the pressure differences in brackets. P_c is the capillary hydrostatic pressure, P_t is the tissue hydrostatic pressure (zero), π_c is the capillary colloid osmotic pressure (3.6 kPa or 27 mmHg), π_t is the tissue colloid osmotic pressure (0.5 kPa), and s is the capillary protein reflection coefficient.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Solutes are exchanged in capillaries and small venules, because of the large surface area and the thin endothelial vessel walls with many pores.

B. Oxygen diffuses from the blood to the interstitial fluid mainly across the total surface of the endothelial cell walls.

C. Systemic oedema is caused by a small increase in mean arterial pressure.

D. The bloodflow through the capillaries is regulated by arteriolar tone.

E. Oxygen is a water-soluble gas.

II. Each of the following five statements have True/False options:

A. Oedema is always caused by a hydrostatic pressure gradient, which is too great.

B. Macrocytic anaemia without megaloblasts in the bone marrow is found in pregnancy, in newborn babies, in hepatic disorders, in hypothyroidism and in aplastic anaemia.

C. The erythrocyte sedimentation rate is normally only a few mm per first hour, 15-20 with a common cold and 50-100 during pregnancy.

D. The reticulocyte count is normally less than 2.5% of the red cell count, but following haemorrhage or haemolysis the relative number of reticulocytes increases reflecting increased erythropoiesis.

E. The three-leaflet mitral valve prevents the leakage of blood backward from the left ventricle to the left atrium.

III. Each of the following five statements have True/False options:

A. The lack of vitamin B_{12} in the liver and the red bone marrow inhibits the methyl-malonyl Co-A mutase and spoils the purine-pyrimidine-DNA-synthesis. The inhibition of these two processes leads to the neurological and the haematological disorders in pernicious anaemia.

B. Mean corpuscular volume expresses the mean volume of each red cell, and mean corpuscular haemoglobin concentration provides the mean concentration in each red cell.

C. Pores of 0.4-0.6 mm permeate the venous end of the capillaries.

D. Fenestrations are round windows found in the capillaries of organs that transport lots of water (the bowels, glomerular capillaries of the kidneys, pancreas and salivary glands). The protein lattice in the fenestrae is so tight, that it keeps plasma proteins back.

E. Newtonian fluids are defined as fluids with a viscosity that is dependent of the shear rate.

8. Case History A

A grey-haired male with blue eyes, 52 years old, is complaining of precordial pain, Dyspnoea upon stair climbing, and nausea. He is depressed and suffers from frequent coughs.

The doctor observes icteric skin and eyes, ataxic walking, dysdiadochokinesis, and positive Babinski. Massive subcutaneous bleeding was found at the left hip.

Laboratory tests revealed the following abnormal results: Lack of HCl in the gastric fluid during fasting and following a pentagastrin test. Haematology tests revealed large erythrocytes - many with nuclei. The red cell count was $1.4*10^{12}$ per l. The haematocrit was 0.21, and the blood [haemoglobin] was 4 mM. The bleeding time was 90 min and the platelet count was $50*10^{9}$ per l. The concentration of vitamin B_{12} in serum was 90 ng per l. The total [bilirubin] in serum was 18 mg per l, and the rise mainly due to non-conjugated bilirubin. A test with radioactive B_{12} was specific for lack of intrinsic factor production from the patient's parietal cells.

1. *What was the cause of this severe pancytopenia (lack of all blood cell types)?*
2. *Calculate the oxygen capacity for haemoglobin.*
3. *Why did the patient develop leucopenia and thrombocytopenia? Was the lack of leucocytes and platelets of any consequences to the patient?*

4. *Does a severe, chronic anaemia trigger physiologic adaptations?*

8. Case History B

In a healthy 20-year old male, with a mean cardiac output of 7 l per min and a haematocrit of 45%, 20 l of fluid are filtered per day in the capillaries. The concentration of protein in the fluid is 5 g per l.

A daily volume of 3 l of fluid passes into the lymphatic vessels and is returned to the blood as lymphatic fluid. The capillaries absorb the rest of the filtered fluid, supposedly together with a small amount of protein (10 g).

*The total amount of plasma reaching the capillary system every day must be 55% of all the whole blood. Each day has 1440 min, so the plasma flow is: (7*1440*0.55) = 5544 l per day.*

1.Calculate the mean protein concentration in the lymphatic fluid.

2.Compare this concentration to that of liver lymph.

Try to solve the problems before looking up the answers

Highlights

- *Erythrocyte sedimentation rate (ERS) is abnormally increased (above 20 mm) in anaemia, infections, immunology reactions, ischaemia, malignancy or traumata. Normally, the level is only a few mm.*
- *Haemopoiesis is the formation of blood cells. All blood cells are derived from stem cells. Stem cells produce erythroid cells, granulocytes, lymphoid cells, megacaryocytes and monocytes by a number of differentiation steps. Stem cells maintain normal cell populations in a healthy bone marrow controlled by haemopoietic growth factors, and stem cells have the capacity for self-renewal.*
- *The erythropoiesis is controlled by the hormone erythropoietin. Erythropoietin is liberated to the circulating blood as a response to hypoxia of any cause (eg, cardiac-pulmonary-renal disease).*
- *The successive change in affinity during binding of the 4 oxygen molecules to each haemoglobin, explains the sigmoid shape of the oxygen dissociation curve.*
- *When blood is saturated under the normal ambient oxygen partial pressure (20 kPa or 150 mmHg), the oxygen capacity of haemoglobin is 1.34 and not the theoretical maximum 1.39 ml STPD g^{-1}.*
- *In many small vessels bloodflow is non-Newtonian and Poiseuilles law is not applicable.*
- *Bloodflow tends to become turbulent, when the flow velocity is high, the viscosity is low, and the vessels are irregular.*
- *Starling's paracapillary circulation plays a dominating role in the transport of crystalloids through the capillary wall.*
- *Of the daily capillary filtration, 9/10 is reabsorbed in the venous end of the capillaries, and 1/10 forms the lymph.*

- *Lymphatic fluids from liver and kidney have a protein concentration equal to plasma's, whereas those from skin and skeletal muscles only contain 2% protein, and brain lymph no protein at all.*

- *Severe anaemia increases respiration, metabolic rate, and temperature due to the large cardiac work.*

- *Aging and arteriosclerosis increase the stiffness of elastic arteries, causing the arterial compliance to fall from 1 to about 0.5 ml of blood per mmHg.*

- *Pernicious anaemia is the most common cause of vitamin B_{12} (cobalamin) deficiency. This is a disorder with an atrophic gastric mucosa. Parietal cell antibodies are present in the plasma of 90% of all patients with pernicious anaemia. The parietal cells of the gastric glands fail to secrete HCl and intrinsic factor.*

Further Reading

Pries, A.R., T.W. Secomb, and P. Gaetgens: Design principles of vascular beds. *Circ. Res*. 77: 1017, 1995.

Rostgaard J and K Qvortrup: Electron microscopic demonstrationsof filamentous molecular sieve plugs in capillary fenestrae. *Microvascular Research* 53: 1-13, 1997.

Return to content

Return to chapter 8

Chapter 9

Systemic Resistance And Hypertension

Study Objectives

- To *define* arterioles, autoregulation, mean arterial pressure, metabolic vascular control, preferential channels, systemic hypertension, systemic resistance, and inflammatory hyperaemia.
- To *describe* arteriolar calibre, vascular resistance, the measurement of arterial blood pressure and the pressures of the pulmonary system, the role of myoglobin, respiratory arrhythmia, alterations of mean arterial pressure and pressure amplitude.
- To *calculate* one variable when relevant variables are given.
- To *explain* the control of the arterial pressure, hypertension, and reactive hyperaemia.
- To *use* these concepts in problem solving and in case histories.

Principles

- *The haemodynamic features of the cardiovascular system are determined by Newtonian and non-Newtonian relations between the driving blood pressure, the bloodflow and the vascular resistance.*
- *These features are related to the ability of each tissue to control its own bloodflow in accordance with its needs. The control of local vascular resistance is a combination of neural and metabolic factors affecting a basal smooth muscle tone.*

Definitions

- **Arterioles** are vessels that range from 150 to 10 mm in diameter. They control the distribution of blood to different tissues.
- **Arteriolar calibre** is the internal diameter of the arteriole, and the size is determined by the contractile activity of its smooth muscle cells and by the transmural arteriolar pressure.
- **Autoregulation** is an automatic control phenomenon that aims at maintaining a constant bloodflow when the driving pressure is changed.
- **Capillary intermittence**: Krogh presumed that a tissue capillary shifts between a closed and an open state. The capillary diameter varies with the oxygen tension. In well-perfused tissue, high O_2 tension causes vasoconstriction and thus tends to reduce its perfusion.
- **Inflammatory hyperaemia** refers to increased bloodflow with accumulation of leucocytes. This reaction is mainly caused by leukotrienes released by the leucocytes.
- **Malignant hypertension** (*accelerated*) refers to a rapid and serious rise of the arterial

blood pressure. The condition can start as paralysis, unconsciousness, or blindness.

- **Metabolic control** is the sum of all metabolic factors that match the oxygen supply to the energy requirement.
- **Myoglobin** is a red, iron-containing, oxygen-binding globin similar to haemoglobin.
- **Reactive hyperaemia** is the increase in bloodflow following temporal vascular interruption of surgical or experimental character.
- **Standard affinity of myoglobin** towards O_2 is the reaction rate at 50% binding. This standard affinity is much higher than that of haemoglobin towards oxygen.
- **Secondary hyperaldosteronism** is recognised by high serum concentrations of renin and aldosterone. This occurs in malignant hypertension or following prolonged use of diuretics. The patients develop cerebral oedema and haemorrhage, cardiac failure and hypertensive nephropathy with proteinuria and microscopic haematuria.
- **Systemic hypertension** - according to WHO - is defined as an arterial blood pressure exceeding 160/95 mmHg (21.3/12.6 kPa) for several months. The pressure increase is either systolic, diastolic or a combination.
- **Systemic resistance** is the total peripheral vascular resistance (*TPVR*), mainly consisting of the arteriolar resistance (in particular that of the essential arterioles in the large striated muscles).
- **Thoroughfare channels** or *preferential channels* shunt the blood directly into the venules bypassing the true capillary bed.
- **Vasomotion** is the rhythmic changes in the arteriolar diameter that causes bloodflow to fluctuate. Vasomotion is brought about by active changes in the tension of vascular smooth muscles. The arteriole can relax completely and then close completely.
- **Vasopressin** is another name for *anti-diuretic hormone (ADH)* from the hypophyseal posterior lobe. ADH controls renal water retention and acts as a moderate vasoconstrictor.
- **VIP** is *Vasoactive Intestinal Polypeptide* from the intestine, the salivary glands and the penile cavernous bodies. VIP is a neurotransmitter and a potent vasodilatator, which is used in the treatment of impotence.

Essentials

This paragraph deals with *1. Autoregulation, 2. Autonomic nervous control, 3. The baroreceptors and other regulators, 4. Oxygen release to the mitochondria, 5. Measurement of blood pressure, 6. Age and MAP.*

1. Autoregulation

The resistance vessels of the coronary system tend to *diminish* any change in the bloodflow in the coronary vessels that are triggered by changes in the driving pressure within a certain range. Increases or reductions in the driving pressure are immediately followed by similar alterations of coronary bloodflow. However, the resistance of the vessels is then changed - metabolically and mechanically - so that the *final* coronary bloodflow is maintained at control levels at all times (changes along the arrows in Fig. 9-1).

Autoregulation has been explained by at least two theories:

1) ***The myogenic theory*** considers autoregulation as a *myogenic* response - an intrinsic property of vascular smooth muscle. Increased stretch of the smooth muscle elicits contraction, whereas diminished stretch elicits vasodilatation. This is illustrated in Fig. 9-1, where an abrupt rise in perfusion pressure from 115 mmHg passively stretches the wall (increases the transmural pressure) and produce an initial increase in bloodflow. Then the vascular smooth muscles contract and the bloodflow falls along the arrows, so that the coronary bloodflow is maintained at 200-250 ml min^{-1}. Similarly, an abrupt fall in perfusion pressure from 115 mmHg has the opposite effect, so the normal bloodflow is re-established.

2) *The* ***metabolic control theory.***

Fig. 9-1: Autoregulation: Changes in bloodflow triggered by changes of the driving pressure has a tendency to be diminished. The example here is the coronary bloodflow, which is described further in relation to Fig. 10-7.

Metabolic control is the sum of all metabolic factors that match the oxygen supply to the energy requirement.

There is a remarkable proportionality between changes of myocardial oxygen consumption and coronary bloodflow. If the oxygen supply is insufficient compared to the myocardial demand, a vasodilatator is releases from the myocytes to the interstitial fluid, so the coronary resistance vessels dilatate.

Adenosine is continuously produced by breakdown of ATP. *Adenosine* is a likely candidate for the role of *metabolic mediator*, because it is such a potent vasodilatator and because it diffuses readily across the cell membranes. Adenosine may work via *presynaptic inhibition* of sympathetic nerve fibres to the smooth muscles of the coronary resistance vessels. Falling perfusion pressure leads to diminished rate of adenosine washout and thus to local vasodilatation. Adenosine dilatates the vessels and causes increased coronary bloodflow. Increased perfusion pressure washes out adenosine, which leads to vasoconstriction and local decrease in bloodflow until it is re-established. - The myogenic and the metabolic control frequently co-operate during autoregulation.

Reactive hyperaemia (ie, increased limb bloodflow following experimental vascular interruption) is probably explained by the metabolic vascular control theory.

Autoregulation protects not only the coronary bloodflow, but also the cerebral, intestinal and renal bloodflow to mention the most important organs.

2. Autonomic nervous control

The sympathetic and the parasympathetic division of the autonomic nervous system control the tone of the resistance vessels by opposing actions. Almost all blood vessels receive efferent nerve fibres from the *sympathetic* nerve system to their smooth muscles.

True capillaries do not contain smooth muscles and do not receive autonomic nerve supply. Metarterioles and capillary sphincters do not receive nerve fibres at all.

The sympathetic vasoconstrictor fibres and circulating catecholamines control both arteriolar, venous and venule tone. The vessels are innervated by postganglionic neurons from the paravertebral sympathetic trunk. The noradrenergic control releases noradrenaline and ATP. The transmitter transport is axonal. Noradrenaline binds to a-adrenergic constricting receptors. Adrenaline binds to both a-adrenergic *constricting* receptors and to b-adrenergic *dilatating* receptors. Consequently, adrenaline elicits vasoconstriction in arterioles where a-receptors predominate, and vasodilatation where b-adrenergic receptors predominate. Adenosine

dilatates vessels, because it inhibits release of noradrenaline possibly via presynaptic purine receptors. In the synapse, the neurotransmitter is eliminated by re-uptake, by enzymatic breakdown and by diffusion. The arterioles of the skeletal muscles, the skin, the kidneys and the splancnic region are densely innervated.

Hunting predators are claimed to have sympathetic vasodilatator fibres to the skeletal muscle vessels, which is consequential during hunting, but such fibres have not been found in humans (Uvnaes).

The cholinergic system is almost exclusively *parasympathetic*. The vessels of the head, neck and thoraco-abdominal organs receive parasympathetic nerve fibres (the 3rd, 7th, 9th and 10th cranial nerves). The large intestine, bladder and genital organs receive parasympathetic fibres from the sacral segments 3-5. The nerve fibres to the external genitals are

active during sexual excitation. Acetylcholine is the vasodilatating transmitter for muscarinic and nicotinic cholinergic receptors. Purinergic receptors use vasodilatating transmitters as ATP, AMP and the potent adenosine.

Cholinergic sympathetic fibres innervate sweat glands and release acetylcholine as stimulus.

3. The baroreceptors

Rapid regulators of the arterial blood pressure are the *arterial baroreceptors* originating from the carotid sinuses and the aortic arch. These classical arterial pressor-receptors are well established and work within seconds following dynamic changes in blood pressure. The arterial baroreceptors probably do not regulate chronic blood pressure changes with constant tone.

The baroreceptor reflex is triggered by stretch of the wall, and the receptors are also called stretch receptors or pressor-receptors. The baroreceptors are mainly located in the walls of the internal carotid arteries (known as the carotid sinuses) and in the aortic arch. Signals are transferred from each carotid sinus via afferent nerve fibres forming the sinus nerve to the glossopharyngeal nerve, and conducted to the nucleus of the solitary tract of the brain stem.

The impulse frequency in the nerve afferents increases with the arterial pressure maintained over a period (Fig. 9-2). The curve is S-shaped with a steep rise in the normal range of arterial pressures, indicating an optimal sensitivity in this area. There is no activity below 60 mmHg.

An increasing rate of pressure change (dP/dt; a sudden rise in pulse pressure amplitude) also increases the firing rate in a single nerve fibre (Fig. 9-2, right). Thus, baroreceptors act as *differential-sensors*. The frequency during the rising systolic pressure is distinctly greater than that in the diastole.

Fig. 9-2: Activity in the carotid sinus nerve at maintained arterial pressure (left) and during a single cardiac cycle with low, normal and high blood pressure.

Baroreceptors convey information about mean arterial pressure (MAP), pulse pressure, and the rate of pressure change (dP/dt). Arterial baroreceptor nerve fibres are *buffer nerves* concerned with *short-term buffering* of the blood pressure.

The afferent signals are conducted to the nucleus of the solitary tract in the medulla. This nucleus is the site confluence for both baroreceptor and chemoreceptor signals. Stimulation here *inhibits* sympathetic structures and *enhances* parasympathetic structures. Thus, a rise in arterial pressure causes vasodilatation and a fall in heart rate, both of which contribute to a lowering of blood pressure. A primary fall in arterial pressure elicits vasoconstriction and a

rise in heart rate, both of which contribute to a rising blood pressure.

Change of body posture from lying to erect reduces the arterial pressure in the carotid sinuses, which elicits an immediate reaction with strong sympathetic tone and diminished vagal tone. This minimises the fall in brain blood pressure, and prevents loss of consciousness. – Hypotensive drugs, exposure to weightlessness, and immobilisation interfere with the baroreceptor reflex, which normally protects us during standing. Such individuals may develop *orthostatic hypotension*, when they stand up and they may faint.

Behavioural and emotional control of blood pressure and heart rate is exhibited by the *hypothalamus*. This autonomic control centre also includes a temperature centre from where contraction of skin vessels is instituted in cold environments.

In *hypertension* the baroreceptor system adapts to the rising pressure within days by moving up the set point. Patients with hypertension have stiff arterial walls as a result of the high arterial pressure, so their baroreceptors are less sensitive than in healthy persons. The increased arterial stiffness is not the main phenomenon in hypertension. Most hypertensive patients are dominated by increases peripheral vascular resistance, which mainly affects the diastolic arterial pressure.

Patients with *hypersensitive* baroreceptors in the carotid sinuses to external pressures are in danger of hypotension with fainting and death from external pressure over the neck at the site of the carotid sinus (so-called *carotid collar syncope or collar death*). Tight collars or other types of external pressures elicit fainting due to marked vasodilatation and hypotension. - Another cause is *emotional fainting* (vasovagal syncope) with a strong emotional activation of the vagus tone via hypothalamus.

Three types of regulators are involved in the adjustment of blood pressure. They are classified as short-term, intermediate-term and long-term regulators.

1. *The arterial baroreceptor reflexes* described above operate rapidly.

2. *Transcapillary volume shifts* in response to changes in capillary blood pressure, begin their function within minutes. When veins are stressed by increased pressure, they slowly expand so that the blood pressure decreases. Conversely, when the intravascular volume decreases, the opposite occurs.

3. *Renal regulation of the body fluid volume*.

When arterial pressure rises, more urine is excreted. Hereby, the plasma and interstitial volume is reduced. The *diminished* plasma volume decreases venous return to the heart, reducing cardiac output, so that elevated arterial blood pressure is brought back towards normal (Fig. 9-6).

A *decrease* in arterial pressure elicits the opposite reaction: The *renin-angiotensin-aldosterone-cascade* is triggered (Chapter 24). Aldosterone from the adrenal cortex promotes Na^+-reabsorption and K^+-secretion from the renal tubules. The reabsorbed Na^+ augments water retention (Fig. 9-6), as does also increased vasopressin (ADH) secretion from the posterior pituitary. A falling arterial pressure also diminishes the release of atrial natriuretic peptide (ANP), and its Na^+ - and water- excreting actions are reduced (Fig. 9-6).

4. Oxygen release to the mitochondria

The factors that ease O_2-diffusion and delivery are:

1. *Myoglobin* in muscle cells releases O_2 during muscular contraction, when the blood supply is blocked. Myoglobin is important as a dynamic O_2 store in muscle cells, although myoglobin is not totally saturated with O_2. During muscular contraction the bloodflow is blocked, and the O_2 tissue tension falls drastically. Myoglobin then gives off O_2 to the cell. The P_{50} for oxymyoglobin is only 5 mmHg (compare to 27 mmHg for oxyhaemoglobin). Bloodflow is re-established during muscular relaxation. Thus, myoglobin is rapidly *reloaded*, even when there is only a small rise in O_2 tension.

2. *Heat energy* releases O_2 during work, since increasing heat energy equals increasing movement of O_2 molecules.

3. *Carbon dioxide:* With rising P_{CO2}, oxygen binding to haemoglobin decreases (Bohr effect, Fig. 8-3).

4. *Binding of 2,3 - DPG* (diphosphoglycerate) to haemoglobin eases the release of O_2 at low tensions (see Chapter 8, paragraph 3).

5. *Mitochondria* located close to capillaries have reduced diffusion pathway.

6. *Short distance* capillary networks, as following capillary recruitment, improve the oxygen delivery.

Oxygen is lipophilic. Since almost the entire capillary surface is identical to the lipid containing plasma membrane of the endothelial cells, oxygen is able to use the total capillary surface for diffusion. The transport of lipophilic molecules is *perfusion limited.*

Oxygen diffuses so easily over the capillary endothelium, that there is *tension equilibrium* between blood and tissues already at the ***arterial*** capillary end.

With rising perfusion the tension equilibrium point is shifted towards the venous part.

Due to the oxyhaemoglobin, the O_2 tension can be maintained through the entire capillary. The oxygen tension varies in the tissues. There is a longitudinal *tension drop* towards the venous end of the capillary, and radial tensions drop in the tissue itself. In brain tissue, the O_2 tension can vary from an arterial level in certain small areas (P_{aO2} of 13.3 kPa or 100 mmHg) towards zero, when bloodflow is insufficient.

Brain and heart tissues are extremely sensitive to a fall in P_{O2}.

Brain tissue is found in the nerve cells of the retina. These nerve cells are deprived of oxygen in 4.5 s (occurrence of *black out*). This can be verified by pressure on the upper eyelid. Consciousness is lost (*grey out*) a few seconds after cardiac arrest. After 90 s, the brain interstitial fluid $[K^+]$ increases drastically from 3 to 60-70 mM, and both action potentials and synapse transmissions are eliminated. There is ion equilibrium over the cell membranes. Intracellular $[Na^+]$ also increases drastically and intracellular brain oedema develops. A high extracellular $[K^+]$ is life threatening.

The EEG of an anoxic brain is recognisable as a *straight* EEG trace (no electrical activity) indicating brain death. Because $[Ca^{2+}]$ rises in the nerve cell, this increases the K^+ conductance, so that more K^+ leaks out into the interstitial fluid.

The kidneys only use 15 ml O_2 each min but they receive 25% of the cardiac output at rest (1200 ml per min containing 200 ml O_2 per l). The kidneys have the lowest arteriovenous O_2 content difference of all the larger organs in our body. The large *safety margin* is important for this vital organ during bleeding or when the renal bloodflow is reduced (more in Chapter 25).

5. Measurement of blood pressure

The arterial blood pressure is measured indirectly in the brachial artery with Korotkoff´s *auscultatory* method. WHO has proposed standardisation of this method. Continuous intra-arterial recordings can obtain exact arterial blood pressure measurements. Comparison with intra-arterial recordings have shown that Korotkoff´s method estimates the systolic pressure too low (about 10 mmHg), and the diastolic pressure differs a few mmHg.

The blood pressure increases in some patients due to the presence of a doctor (ie, *white coat hypertension*). This is revealed by repeated measurements – preferably performed before, during and following exercise.

Ejection of blood from the left ventricle triggers a pulse wave in the wall of the arterial tree, and the volume-pressure variations here distribute with a large velocity along the arterial tree. In young persons the velocity is 5-10 m per s; with age, atherosclerosis and hypertension the arterial tree becomes stiffer and the velocity increases (see Ch. 8 about compliance and also Fig. 8-8).

Fig. 9-3: Changes in pressure in the arterial tree of a supine healthy person.

The systolic pressure increases progressively along the arterial tree, whereas the diastolic and the MAP decrease (Fig. 9-3). The pulse amplitude, which is the difference between systolic and diastolic pressure therefore, increases clearly (Fig. 9-3). The end of systole is marked by a brief sharp fall in pressure (dicrotic notch), caused by the relaxation of the ventricle with backflow of blood as the aortic valves close. This backflow pressure moves with the blood all along the arterial tree (Fig. 9-3).

The blood pressure has to be measured repeatedly, with the patient sitting comfortably in a relaxed environment, and measured at more than three consultations in order to avoid false alarm with white coat hypertension. A diastolic pressure above 95 mmHg (12.6 kPa) expresses an increased MAP and the age of the patient influences the strategy of the treatment.

Fig. 9-4: Normal pressures in the circulation of a supine healthy person.

Essential *diurnal* variations are present, but repeated blood pressure measurements over three consultations seem to define a reasonable diurnal mean level. Continuous recording of the arterial pressure is sometimes necessary.

Patients below 40 years of age, with a diastolic pressure above 100 mmHg must be followed and examined further. Patients above 40 years of age, with a diastolic pressure above 120 mmHg, must be examined further.

Normal values for blood pressures measured in different locations of the circulation are given in Fig. 9-4.

6. Age and MAP

Populations living under *natural conditions* - including Indian troops in Brazil and healthy living persons in the Western Hemisphere - maintain their mean arterial pressure (MAP) throughout life. Their distribution curve for MAP is close to the normal distribution.

The MAP and the systolic pressure measured as an average for the total population, increases with increasing age in the rich part of the World.

As an order of thumb, the systolic blood pressure in mmHg is equal to 100 plus age in years, because these values are close to typical statistical mean values from examination of large population groups. This is because general diseases, with consequences for the systolic blood pressure and MAP, are accumulated with age in the Western Hemisphere. Quite a few of the accumulated disorders (such as atherosclerosis – see Chapter 10) probably occur as a consequence of our life style - operating in a heterogeneous genetic pool.

Previously, systemic hypertension was therefore characterised by a MAP larger than normal for the age. Practically difficult comparisons had to be made with a statistical, so-called normal material. Today, most doctors use the *WHO definition* (see below).

The MAP is a good estimate of the driving pressure, and the cardiac output is the stroke volume multiplied by the cardiac frequency. MAP and cardiac output are easy to determine, so the *TPVR* can be calculated.

With pressure expressed in mmHg and cardiac output expressed in ml per s, the unit for *TPVR* is 1 $mmHg*s*ml^{-1}$. This unit is complicated in writing and the abbreviation is 1 PRU (Pressure Resistance Unit). The normal value for *TPVR* in the systemic circulation at rest is one PRU, and during exercise it is only 0.3 PRU.

Pathophysiology

This paragraph deals with *1. Natural history of hypertension, 2. Symptoms and signs of hypertension, 3. Risk factors (Western lifestyle), 4. Types of systemic hypertension, 5. Therapeutic principles, 6. Future strategy.*

1. Natural history of hypertension

Primary hypertension always has a diastolic element reflecting involvement of the resistance vessels (eg, muscular arterioles etc). *Secondary hypertension,* caused by atherosclerosis or other types of stiff arterial walls, is often purely systolic.

In the early stages of hypertension, the arterial blood pressure is oscillating between hypertensive episodes and normal periods. The hypertensive episodes are typically dominated by sympathetic overactivity with increased cardiac output and almost unchanged *total peripheral vascular resistance (TPVR).* Eventually, the pressure changes the distensibility of the arteriolar walls and thus leads to sustain structural changes of the resistance system. As the hypertension develops the *TPVR* is increased.

Any rise in blood pressure is a strong stimulus to the high-pressure baroreceptors, but these essential sensors do not always work appropriate in hypertension. The expected bradycardia from the high arterial pressure acting normally on the arterial baroreceptors is not seen in hypertensive patients.

The initial sympathetic tone is also depicted in the high resting heart rate, in contrast to the bradycardia found normally, when the blood pressure rises. The abnormal baroreceptor reflex is probably an adaptive consequence of the variable but lasting initial hypertension.

Permanent structural changes of the resistance vessels, with strongly reduced specific compliance (reduced distensibility) and reduced lumen of arterioles and small muscular arteries, eventually leads to permanent hypertension.

The rising *TPVR* implies a rising workload for the left ventricle and thus creates left

ventricular hypertrophy.

2. Symptoms and signs of hypertension

The typical patient with hypertension is *asymptomatic*. This is what makes the development of this disorder dangerous. The first sign of systemic hypertension is sometimes acute myocardial infarction with sudden death. Of all acute cases of myocardial infarction up to 25% only experience a sudden pain, there is cardiac arrest, and the cases are recorded as sudden death from myocardial infarction at section.

Hypertonic patients with coronary artery disease experience angina at exhaustion or from myocardial hypertrophy.

Malignant or accelerated hypertension refers to a rapid and serious rise of the arterial blood pressure. The condition can start as paralysis, unconsciousness, or blindness.

Secondary hyperaldosteronism is recognised by high serum concentrations of renin and aldosterone. This occurs in *malignant hypertension* or following prolonged use of diuretics. The patients develop cerebral oedema and haemorrhage, cardiac failure and hypertensive nephropathy (with proteinuria and microscopic haematuria). Patients with malignant hypertension develop *dissecting aortic aneurysms* and *retinal damage with papiloedema*, so they die rapidly without specific therapy.

Ophtalmoscopy for hypertonic changes of the retina also provides the diagnosis hypertension. These changes include haemorrhages in the retinal nerve fibre layer, exudates as yellow-white spots called *cotton wool spots*, irregular arteriolar diameter, microaneurysms, and papillary stasis (Fig. 9-5).

Fig. 9-5: Hypertonic changes of the retina seen by ophtalmoscopy. The patient has malignant hypertension. - A normal retinal fundus is found in Fig. 6-5.

A necrotic arteriolitis is often found by ophtalmoscopy in malignant hypertension.

3. Risk factors (Western lifestyle)

Causative or *risk factors* for essential hypertension include *genes*, because there is a clear racial and familial accumulation of hypertension. A risk factor is a factor showing statistical covariance with the disease - see also Chapter 10.

Africans have higher arterial blood pressure than Caucasians, and some families accumulate cases of hypertension. Specific genes have not been identified.

The environmental factors are numerous, but *Western Hemispher lifestyle* is the key word, since the occurrence of increasing systemic blood pressure with increasing age is obviously related to accumulation of disease. However, accumulation of hypertension with age is not a law of nature.

Western lifestyle is sedentary, with psychological stress in career and family life. Existential procedures have to be performed rapidly including buying and eating fast food. Persons with a stressful everyday life, with smoking, alcohol and large meals following long work hours, practice little exercise if any, and become obese with hyperlipidaemia, hyperglycaemia and hyperuricaemia.

The hunting human has become a stressed user of automatic tools (cars, mobile telephones, household utilities, TV, PC etc). This lifestyle pattern frequently implies a serious *sympathetic overactivity* with a typical rise in resting cardiac rate and thus in cardiac output.

One essential and measurable variable in the life style pattern is the lack of exercise (eg, physical inactivity). A low maximum oxygen capacity or *fitness number* is measurable with the submaximal exercise test of Åstrand (Fig. 18-3), and reproducible in each individual. The fitness number is expressed as the *maximum oxygen uptake* in $ml*min^{-1}*kg^{-1}$.

A maximum oxygen uptake *below 34* ($ml*min^{-1}*kg^{-1}$) is related to *risk factor accumulation* and early death from hypertensive or other related complications (Fig. 18-14). Such a low maximum uptake is a clear indication of physical inactivity, where dilatation of muscular arterioles is seldom or almost never achieved.

The unknown cause of essential hypertension in the Western Hemisphere may well prove to be *physical inactivity* and the related life style patterns described above.

In some cases of hypertension there is a clear relation to the renin-angiotensin-aldosterone cascade (Chapter 24)**.** The series of events starts with a rise in *TPVR* due to increased vascular tone. Over months and years, the walls of arteries and arterioles thicken and atherosclerosis is spread in the arterial tree. Such changes reduce the driving pressure in the renal arteries, which leads to a fall in glomerular filtration rate (GFR) and increased NaCl/water retention (Chapter 25). The falling pressure in the renal artery triggers b-receptors on the JG-cells of the juxtaglomerular apparatus (Fig. 25-17). Renin is released from these cells located in the afferent glomerular arteriole. Renin separates the decapeptide, angiotensin I, from the liver globulin, angiotensinogen. When angiotensin I passes the lungs or the kidneys, a dipeptide is cut off from the decapeptide by an angiotensin-converting enzyme (ACE). Hereby angiotensin II (octapeptide) is produced. Angiotensin II stimulates the aldosterone secretion from the adrenal cortex, and thus stimulates the Na^+ reabsorption and the K^+ secretion in the distal, renal tubules. The renin-angiotensin-aldosterone cascade further contributes to the salt and water retention. Angiotensin II is also a circulating vasoconstrictor just as adrenaline and vasopressin found in high plasma concentrations in many hypertonics. ACE inhibitors (see later) are rational choices for hypertonics with high angiotensin II, but also for other categories for reasons unknown (diabetics etc). The *cascade* is further described in Chapter 24 – paragraph 6.

4. Types of Systemic Hypertension

There are two forms of hypertension, I) primary or essential, and II) secondary hypertension.

I) Primary hypertension is a multifactorial syndrome without known cause. Approximately 90% of all cases are classified as primary or essential hypertension, because the causative factors are not clarified in detail. Increased peripheral resistance is responsible for most cases of primary hypertension.

Fig. 9-6: Factors contributing to systemic hypertension. Abbreviations: ECV = Extracellular fluid Volume; TPVR = Total Peripheral Vascular Resistance.

II) Secondary hypertension

In about 10% of all cases the cause of the hypertension is clarified, and these patients are classified as secondary hypertension. This condition must always be suspected in young hypertonics.

Renal, endocrine or cardiovascular diseases cause secondary hypertension or it relates to pregnancy or to drugs. Endocrine disorders are treated systematically in Chapters 26>30.

1. *Renal disorders* (Chapter 25) account for more than 80% of all cases of secondary hypertension. The disorders are chronic cases of glomerulonephritis, pyelonephritis and

other permanent damage of the kidneys, where salt and water retention dominates. Hyperparathyroidism and Ca^{2+} overload can lead to *renal failure* and severe hypertension. A renal artery stenosis sufficient to reduce the glomerular pressure leads to renin release from the juxtaglomerular apparatus, aldosterone release and thus to *increased salt-water retention* (see the *renin-angiotensin-aldosterone cascade*, Chapter 24, paragraph 6). Renal artery stenosis (atherosclerosis or fibromuscular hyperplasia), chronic renal inflammation (glomerulonephritis or pyelonephritis), and congenital polycystic kidneys can lead to secondary, systemic hypertension. Renal function is examined with endogenous creatinine clearance and the renal vessels by scanning or arteriography. The plasma renin concentration is measured.

2. *Hyperaldosteronism* has a primary and a secondary form. Conn´s syndrome is *primary* hyperaldosteronism. This condition is characterised by an isolated rise in serum aldosterone, since the cause is hyperfunction of the zona glomerulosa of the adrenal cortex - not the renin release. Secondary hyperaldosteronism is a condition with abnormally high stimulation of the adrenal zona glomerulosa. The serum concentrations of the whole renin-angiotensin-aldosterone cascade are increased.

3. *Cushing's syndrome* describes clinical conditions with increased glucocorticoid concentration in the blood plasma. The classical *Cushings disease* is caused by *excess* liberation of ACTH from the adenohypophysis, but ACTH excess is also known to originate from ectopic ACTH producing tumours or from excess administration of ACTH. - Non-ACTH related adrenal adenomas or carcinomas, glucocorticoid excess administration, and alcohol abuse (so-called *Pseudo-Cushing*) cause *Cushing's syndrome*. - The *dexamethasone suppression test* is described in Chapter 30.

4. A pituitary tumour producing an excess of growth hormone (Ch.28) causes *acromegaly*. The patient sometimes has a *diabetic* glucose tolerance test (Ch.27). These patients die from heart failure, IHD or hypertension.

5. *Phaeochromocytoma.* This is a tumour of the sympathetic nervous system (Ch. 28) releasing both noradrenaline and adrenaline. The signs are intermittent or constant systemic hypertension, tachycardia with other arrhythmias, orthostatic hypertension and flushing.

6. In the last three months of pregnancy some females develop hypertension, oedema and proteinuria (*pre-eclampsia* or *toxaemia of pregnancy*). If this condition develops into severe hypertension with fits and lung oedema, it is called *eclampsia*. This is a life threatening condition, which must be treated immediately with intravenous hydralazine or minoxidil, and if necessary termination of pregnancy. Hydralazine is orally active vasodilatators, which work by direct relaxation of smooth muscles.

7. *Drugs* such as steroids or oral contraceptives with high oestrogen, sympatomimetics, aldosterone, and vasopressin all cause severe systemic hypertension. Monoamineoxidase-inhibitors combined with tyramine (cheese) or wine sometimes cause hypertension. A careful medical history is helpful.

8. *Cardiovascular disorder* - as coarctation of the aorta - is the cause of hypertension in a few young patients. The coarctation produces a late systolic murmur. These hypertonics have a low pressure distal to the coarctation.

9. *Atherosclerosis* (see Chapter 10) is characterised by a special systolic hypertension frequently found in the elderly without any diastolic hypertension. These patients do not have any arteriolar disease.

5. Therapeutic principles

Systemic hypertension is a health threat to the person as a whole, since the untreated disease shortens life expectancy with approximately 20 years. Target organs for damage are the heart, aorta, brain, eyes and the kidneys.

The positive effect on life expectancy of a moderate reduction of an abnormally high systemic arterial blood pressure is well documented.

The simple *resistance model* presented in Eq. 9-1 is applied for the therapy of systemic hypertension. The *driving pressure* in the systemic circulation is equal to the *cardiac output* multiplied with the *Total Peripheral Vascular Resistance (TPVR).*

The cardiac output is equal to the *cardiac frequency* multiplied with the *stroke volume*, and the stroke volume depends of the *total blood volume*. *TPVR* depends of the degree of contraction of the resistance vessels and of the distensibility (eg, specific compliance) of the arterial system.

Principally, systemic hypertension is therefore treatable through one or more of the following strategies:

1. Reduction of the *total blood volume* (and thus the stroke volume) with diuretics results in reduction of the driving pressure,

2. Reduction of the *cardiac frequency* reduces cardiac output and thus the driving pressure,

3. Reduction of *TPVR* with vasodilatators reduces the driving pressure.

Two strategies of therapy and their combination are available: *Change of life style* with or without drug therapy. *Drug therapy* must usually be continued for the lifetime of the patient.

Life style modifications (relaxed duration exercise and healthy habits):

In healthy individuals, the opening of resistance vessels during exercise typically reduces the *TPVR* to 30% of the value at rest. This vasodilatation expresses an enormous capacity, which is only present in the resistance vessels of the striated muscular system at large. The only natural way to break the vicious circle described above is to maintain the dilatation capacity throughout life by frequent use of the locomotor system. The exercise must include large muscle groups for some time. The exercise must be relaxed and comfortable in order to become a life style. Other beneficial effects of *relaxed duration exercise* (such as walking, golf, jogging, swimming, badminton, tennis etc) is *improved glucose tolerance, weight loss, improved heart function, improved lipid profile, normal gastrointestinal functions and psychological benefits such as improved mood and a healthy sleep pattern.* Healthy food and drinking habits are important, and smoking has to be given up.

Hypotensive drugs can be divided into 5 categories:

5.1. Diuretics

Hypertensive patients seem to handle Na^+ just as healthy persons (see Chapter 25).

Initial administration of diuretics produce a pronounced renal salt and water excretion, which lead to a reduction in ECV, and a fall in systemic blood pressure. The urinary salt and water excretion returns to normal after several days, but the blood pressure remains at the reduced level. This is difficult to explain. Perhaps some diuretics have a direct relaxing effect on vascular smooth muscle in the arterioles or other vessels.

The different groups of diuretics are treated in Chapter 25.

5.2. β-adrenergic receptor blockers

β-blockers antagonise competitively the effects of adrenaline and nor-adrenaline on b-adrenergic *vasodilatating* receptors. The typical non-selective b-adrenergic receptor blocker is propranolol, which is a potent reversible antagonist at both b_1-and b_2-adrenergic receptors. Propranolol acts on the heart and reduces the chronotropic (reduced heart rate) and inotropic effect (reduced force and cardiac output); the reduced cardiac function is most pronounced during high sympatho-adrenergic activity, such as during exercise or stress, so the drug can release acute cardiac failure. The anti-arrhythmic effect of propranolol is probably due to its local anaesthetic action on cardiac cells including pacemaker cells. The effect of propranolol on hypertension is not clarified, since it seems to increase peripheral vascular resistance slightly. Simultaneously, propranolol reduces the release of renin from the juxtamedullary apparatus. This inhibits aldosterone secretion, and thus reduces the potassium secretion of the distal tubular system. The result is potassium retention, which is further aggravated by *β-blockade* of receptors on cell membranes, whereby the adrenaline-stimulated Na^+-K^+ pump is inhibited. Following meals containing carbohydrate and potassium, there is a release of insulin, which stimulates the Na^+ - K^+ pump, and thus the K^+ uptake in cells. Adrenaline also stimulates the Na^+- K^+pump through activation of b_2 - receptors, whereby the plasma-[K^+] is reduced. The normal effect of insulin is hypoglycaemia, which is compensated by lipolysis and glycogenolysis (with FFA and glucose liberation), by increased sympathoadrenergic activity. Propranolol inhibits lipolysis from adipocytes and glycogenolysis from hepatocytes, myocardial and skeletal muscle cells. This is a problem with diabetics or for patients with reduced glucose tolerance. b-blockade may lead to life threatening hypoglycaemia or a serious rise in blood pressure, if adrenaline release dominates. Propranolol is thus contraindicated in persons with diabetes, sinus bradycardia, partial heart block and congestive heart failure. Propranolol increases airway resistance, which is a hazard to patients with COLD or asthma, because of bronchoconstriction.

Many b-blockers act selectively, but all compounds have effects as described below:

Selective b_1-blockers acts on the cardiac b_1-receptors and reduces the force of cardiac contraction and thus lowers the blood pressure.

Blockade of b_1-adrenergic receptors located on the renin-secreting juxtaglomerular cells reduces the renin release and the blood pressure in persons with renin-dependent hypertension (eg, patients with a high renin level in the plasma from renovascular disease).

Many b-blockers reach the brain tissue through the blood-brain barrier, and others reach the brain cells through the large fenestrae of the circumventricular organs. The CNS-effect is an inhibition of the sympatho-adrenergic output, and beneficial effects on paroxysms of panic and anxiety. The hypotonic CNS-effect is probably dominating, and explains the maintained lowering of blood pressure, although the initial reduction in cardiac output is often only temporary.

5.3. α_1-adrenergic antagonists

inhibit the effect mediated through noradrenaline released from sympathetic presynaptic fibres to the postsynaptic a_1-receptors and produce vasodilatation. Also a central effect of these compounds (doxazosin, prazosin) may be involved. The hypotensive efficiency of these drugs give rise to the main complication, which is a serious fall in blood pressure following the first

dose.

5.4. Angiotensin Converting Enzyme (ACE) Inhibitors

Angiotensin converting enzyme is found to have the highest activity in the endothelium of the long pulmonary capillaries. Converting enzyme is a ***kininase II***, which convert the decapeptide, angiotensin I, to the vasoconstrictive octapeptide, angiotensin II. ACE inhibitors (captopril, enalapril, and lisinopril) reversibly inhibit converting enzyme and thus act as a vasodilatator of both resistance and capacitance vessels. Angiotensin II is a potent vasoconstrictor, in particular when its concentration in plasma is high. Patients with 100 pg l^{-1} or more of angiotensin II react beneficial on ACE inhibitors. Also other hypertonics such as diabetic patients reduce their risk of vascular insults by the use of ACE inhibitors for reasons unknown.

5.5. Calcium-channel blocking agents

Ca^{2+}-antagonists (amlodipine, nifedipine, diltiazem, and verapamil) acts as effective vasodilatators, because they relax the smooth muscles of the arterioles. They also inhibit the cardiac contractile force. Ca^{2+} -antagonists inhibits the Ca^{2+} -entry into the cells, because they bind to the proteins of Ca^{2+} -channels in the membrane. The overall effect is beneficial incongestive heart failure, because the vasodilation dimishes *TPVR* and thus reduces afterload. Hereby, the cardiac output is improved despite cardiac contractile depression.

6. Future strategy

- *Systemic hypertension is the most frequently diagnosed and treated risk factor for the development of atherosclerosis (including ischaemic heart disease).*
- *A risk factor is a factor showing covariance with atherosclerosis. The remaining risk factors for atherosclerosis are physical inactivity, hypercholesterolaemia, hypertriglyceridaemia, increased LDL concentration, smoking, diabetes, and familiar factors (genes, social inheritance or life style patterns).*
- *A rational strategy is to control the risk factors for the patients. A successful lowering of arterial blood pressure with a hypotensive drug must not be accompanied by an unrecognised consequential rise in other risk factors.*
- *Relaxed exercise is an alternative therapeutic strategy to antihypertensive drugs in many cases of essential hypertension.*
- *Mild and relaxed exercise has other beneficial effects, namely a consequential reduction of most of the known risk factors for atherosclerosis.*
- *Healthy food, exercise and drinking habits are important to hypertonics, and smoking has to be given up.*

Equations

The driving pressure (D*P*) in the systemic circulation is equal to the cardiac output (Q°) multiplied with the *TPVR* according to Poiseuille´s law:

Eq. 9-1: $DP = Q^{\circ} * TPVR$.

This is a simple resistance model for circulating fluid, and the model is applied for therapeutic strategies.

Fick's first law of diffusion: The flux (*J*) of O is equal to the diffusion coefficient of oxygen

(D is 10^{-9} m^2 s^{-1}) multiplied with the concentration gradient (dC) per distance unit (dx) through a given area (A). Fick's first law is written:

Eq. 9-2: $J = (- D \times dC) \times A/dx$ (mol per time unit) with a diffusion gradient (dC) through the area A. Notice that D/dx is a permeability coefficient (m per s).

The first law can also be written: $J = (D \times DP \times A)/dx$.

Self-Assessment

Multiple Choice Questions

I. The following five statements have True/False options:

A. Nitrovasodilatators has side effects such as hypotension, reflex tachycardia and headache.

B. The blood-brain barrier is impermeable to all b-blockers.

C. The ACE-inhibitor, captopril, dilatates both arterioles and capacitance vessels.

D. Depolarisation of the vascular smooth muscle cell membrane opens voltage-gated Ca^{2+} -channels, whereby Ca^{2+}-ions enter the cell, combine with calmodulin and activate myosin light-chain kinase.

E. Thiazides have serious side effects such as hyperglycaemia (glucose intolerance), hypercholesterolaemia, hypokalaemia and hyperuricaemia.

II. The following five statements have True/False options:

A. MAP and Q∘ are easy to determine, so the *TPVR* can be calculated.

B. Angiotensin converting enzyme is a kininase II, which convert the decapeptide, angiotensin II, to the vasoconstrictive octapeptide, angiotensin I.

C. The driving pressure (D*P*) in the systemic circulation is equal to the cardiac output *(Q∘*) *divided* by the *TPVR* according to Poiseuille´s law.

D. Noradrenaline binds to a-adrenergic constricting receptors. Adrenaline binds to both a-adrenergic constricting receptors and to b-adrenergic dilatating receptors. Consequently, adrenaline elicits vasoconstriction in arterioles where a-receptors predominate and vasodilatation where b-adrenergic receptors predominate.

E. Ca^{2+}-antagonists (amlodipine, nifedipine, diltiazem, and verapamil) act as effective vasodilatators, because they relax the smooth muscles of the arterioles.

Case History A

A male, age 50 years, visits an ophthalmologist in order to have measured new lenses for myopia and astigmatism. Ophtalmoscopy reveals irregular vessel diameter, bleeding, yellow-white spots, and papillary stasis. The patient is advised to see his general practitioner, which finds a constant arterial blood pressure of 200/110 mmHg (26.66/14.66 kPa). The heart frequency is 85 beats per min and the cardiac output at rest is normal.

The patient is an office clerk, and also has a sedentary off- duty life. The patient is a heavy smoker using 40 cigarettes per day. His father had high blood pressure and died from

cerebral infarction at the age of 62 years.

An X-ray of thorax reveals clear lung fields and left ventricular hypertrophy.

1. *What is the diagnosis?*
2. *What is the treatment of choice?*
3. *What is the main risk for this patient?*
4. *What happens in the lungs and the left ventricle of this patient?*
5. *Compare the left ventricular pressure-volume work rate of this patient to that of a healthy individual. Assume that they are both at rest with a cardiac output of 5 l per min. Assume that the healthy person has a mean arterial pressure of 90 mmHg (12 kPa).*
6. *Convert the work rate units used into watts, and explain the development of ventricular hypertrophy.*

Conversion factors are found in Symbols or here:

1 litre = 10^{-3} m^3. 1 mmHg = 133.3 Pa (N/m^2). 1 watt = 1 Nm/s = 1 J/s.

Case History B

A 59-year old office worker is known to have systemic hypertension. From the initial arterial pressure of 195/115 mmHg, he was brought down to a stable level of 160/95 mmHg by antihypertensive drugs. During work the patient suddenly collapses, and he is brought to hospital in an unconscious state with an arterial blood pressure of 75/45 mmHg. There are no signs of hemiplegia. Assume that the brain is hypoxic, and that the brain is producing lactic acid out of 30% of all glucose molecules combusted here. Among other values the blood glucose concentration is determined to 5 mM, and the arteriovenous glucose concentration difference increases to 300% of normal (0.5 mM). The cerebral bloodflow (CBF) is reduced to 50% of the normal value (650 ml min^{-1}). The total production by oxidative phosphorylation is 36 ATP per glucose molecule, and by anaerobic metabolism 2 ATP per glucose molecule.

1. *What is the most likely diagnosis?*
2. *Calculate the anaerobic and aerobic contribution to brain metabolism.*
3. *Calculate the net glucose flux and the ATP production in a normal brain, and compare the results to those of the patient.*

Case History C

A female, 66 years of age, complains of frontal headache. She has been treated for migraine for the last 40 years. The new headache is different from migraine. The doctor measures her arterial blood pressure to 195/115 mmHg (25.9/15.3 kPa). By ultrasound screening the length of her left kidney is measured to be half the length of the right. Renal arteriography reveals a stenosis of the left renal artery. The stenosis is relieved by balloon dilatation, where a catheter with a balloon at its tip is inflated at the right site. The success of the treatment is confirmed over the following weeks, where her blood pressure reach a level of 145/95 mmHg (19.3/12.6 kPa).

1. *What is the cause of her hypertension?*
2. *Explain the pathophysiological mechanism.*

3. What is the most likely cause of her renal artery stenosis?

Case History D

A female, age 22 years, is sitting on a bicycle ergometer with her calf muscles 0.9 m below heart level. She is at rest and the venous pressure is 10 mmHg (1.3 kPa) at the level of the heart. The oxygen uptake(V_{O2}) is 0.247 l STPD per min and the muscle bloodflow is 3 ml per min per 100 g of tissue (3 Flow Units, FU). The total weight of all her skeletal muscles is 30 kg.

Following 5 min of rest, she starts cycling, hereby increasing her oxygen uptake to 4.5 l STPD per min, and her muscular arterioles dilatate to reach a three-fold increase in inner radius. During exercise the arterio-venous O_2 content difference is 170 ml STPD per l, and the oxygen uptake in the skeletal muscles increases from 1 to 100 ml STPD per min per kg.

1. *Calculate the venous pressure in the calf muscles at rest.*
2. *Calculate the relative alteration of the muscular vascular resistance during exercise.*
3. *Calculate the driving blood pressure over the working muscles during exercise, where the arterial blood pressure is 170/70 mmHg (22.7/9.3 kPa) and the venous pressure in the calf muscles is reduced to 20 mmHg (2.6 kPa).*
4. *Calculate the rise in muscle bloodflow during exercise.*

Try to solve the problems before looking up the answers.

Highlights

- *The cardiac output(Q) is the stroke volume multiplied by the cardiac frequency. The heart must pump harder to provide a given Q° with increasing age, because the arteries become increasingly stiff with age.*
- *The distensibility or compliance of the arterial system diminishes with age due to atherosclerosis.*
- *When the pressure wave travels through the arterial tree, the arterial compliance is always less in the distal part of the system.*
- *The mean arterial pressure (MAP) is a good estimate of the driving pressure (ΔP), whereas the pulse pressure varies almost directly with the stroke volume.*
- *MAP and Q are easy to determine, allowing a calculation of total peripheral vascular resistance (TPVR). The systemic TPVR is one PRU at rest and 0.3 PRU during exercise.*
- *The velocity of the systemic arterial pressure wave varies inversely with the arterial compliance, whereby the velocity increases with age and with increasing degrees of hypertension.*
- *The high frequency components of the systemic arterial pressure wave are damped in the periphery, and the systolic peak components are elevated.*
- *The MAP varies directly with the Q and the TPVR.. The resistance model for circulating blood: ΔP = Q° * TPVR is applicable for therapeutic strategies in hypertension.*

- *The EEG of an anoxic brain is recognisable as a straight EEG trace (no electrical activity) indicating brain death. Because [Ca^{2+}] rises in the nerve cell, this increases the K^+ conductance, so that more K^+ leaks out into the ISF.*
- *The arterial blood pressure is measured indirectly in the brachial artery with Korotkoff´s auscultatory method (standardised by WHO). The systolic pressure is recorded by the occurrence of a tapping sound, and the diastolic pressure is manifested by the disappearance of the sound.*
- *Continuous intra-arterial recordings can obtain exact arterial blood pressure measurements. Comparison with intra-arterial recordings have shown that Korotkoff´s method estimates the systolic pressure too low (about 10 mmHg), and the diastolic pressure differs a few mmHg.*
- *A risk factor is a factor showing covariance with atherosclerosis. The remaining risk factors for atherosclerosis are physical inactivity, hypercholesterolaemia, hypertriglyceridaemia, increased LDL concentration, smoking, diabetes, and familiar factors (eg, genes, social inheritance or unhealthy life style).*
- *Populations living under natural conditions - including Indian troops in Brazil and healthy living persons in the Western Hemisphere - maintain their mean arterial pressure (MAP) throughout life.*
- *In the rich part of the World, the MAP and the systolic pressure, measured as an average for the total population, increases with increasing age.*
- *Systemic hypertension - according to WHO - is defined as an arterial blood pressure exceeding 160/95 mmHg (21.3/12.6 kPa) for several months. The pressure increase is either systolic, diastolic or a combination.*
- *Systemic hypertension is the most frequently diagnosed and treated risk factor for the development of atherosclerosis including ischaemic heart disease.*
- *The cause of essential hypertension in the western world may well prove to be physical inactivity and related life style patterns.*
- *Relaxed exercise is an alternative therapeutic strategy to antihypertensive drugs. Mild and relaxed motion (such as walking, bicycling, golf, jogging, swimming, badminton, tennis etc) is utilised, whenever possible, in the treatment of essential hypertension.*
- *Mild and relaxed exercise has other beneficial effects, namely a consequential reduction of other known risk factors for atherosclerosis: Improved glucose tolerance, weight loss, improved heart function, and improved lipid profile, normal gastrointestinal functions and psychological benefits such as improved mood and a healthy sleeping pattern.*
- *Healthy food and drinking habits are important to hypertonics, and smoking has to be given up.*
- *A rational strategy for the future is to control the risk factors for the patients. A successful lowering of arterial blood pressure must be accompanied by improvement of other risk factors.*

Further Reading

Hypertension. Monthly journal published by the *Am. Heart Association*, 7272 Greenville Av., Dallas TX 75231-4596, USA.

Julian DG, Camm AJ, Fox KS, Hall RJC & Poole-Wilson PA (1995) *Diseases of the Heart*, 2nd Edn. London: Bailliere Tindall.

Katzung BG. *Basic & Clinical Pharmacology*. Appleton & Lange, Stanford, Connecticut, 1998.

Return to content

Return to chapter 9

Chapter 10

Cardiac Performance And Disorders

Study Objectives

- To *define* concepts such as cardiac insufficiency, central venous pressure, compliance, contractility, venous pump, venous return and ventricular stroke work.
- To *describe* four capacities of the cardiac pump called bathmotropic, chronotropic, inotropic, and dromotropic states. To describe the pressures of the left and the right half of the heart. To *describe* atherosclerosis, arteriosclerosis and risk factors involved.
- To *draw* pressure-volume curves and contractility lines for the heart at different conditions, and pressure variations in the left side of the heart and the aorta.
- To *calculate* the external work on the blood by the right and left ventricle, and the kinetic energy.
- To *explain* the autonomic innervation of the heart, venous return, cardiac contraction, Starling´s law of the heart, cardiac performance and oxygen demand, and cardiac filling pressure. To explain ischaemic heart disease, rheumatic heart disease, Buerger´s disease, and Raynaud´s disease.
- To *use* the above variables and concepts in problems and case histories.

Principles

- *Starling´s law of the heart: With an increased ventricular filling during diastole (venous return), the ventricular fibre length increases, so the ventricular contraction and stroke volume increases. - This is an intrinsic adaptation of the pumping capacity to the venous return. - Starling's law of the heart is also called the Frank-Starling relationship, because it was described independently by Otto Frank and Starling.*
- *The Fick principle (cardiac output) states that the volume of oxygen taken up by the blood in the lungs, divided by the arteriovenous oxygen content difference, is equal to the cardiac output. This example utilises the law of conservation of matter.*
- *The law of Laplace - Eq. 8-6*
- *Poiseuille´s law - Eq. 8-3.*

Definitions

- **Angina pectoris** (chest pains) is pain felt beneath the sternum or in the precordial area. The hypoxic pains are typically provoked by exercise or cold and submitted by subendocardially situated nerve fibres. The pains are relieved rapidly by nitro-glycerine and rest.
- **Arteriosclerosis** refers to atherosclerosis (and further changes) of the peripheral arteries.
- **Atherosclerosis** is a process of progressive lipid accumulation (*atheromatosis*) and calcification of the inner arterial walls in the abdominal aorta, lower extremities and the arteries of the heart, brain and kidneys.
- **Bathmotropic state** refers to the irritability of the myocardium.
- **Cardiac insufficiency** is a disorder, where the heart cannot pump enough blood to satisfy

the nutritive needs of the body.

- **Central venous pressure** (CVP) is the pressure measured in the caval veins at the level of the heart or in the right atrium.
- **Chronotropic state** refers to the cardiac frequency.
- **Compliance** of the resting cardiac chambers refers to d*V*/d*P* (chamber compliance) - the reverse of the elastance (d*P*/d*V***)** of relaxed tissue.
- **Cardiac contractility** is the d*P*/d*V* of the contracting ventricle. The contractility is depicted on the pressure-volume loop of the cardiac ventricle. Contractility refers to the change in slope of the pressure and volume increase from isovolumetric rest to contraction. Contractility is a function of contraction by crossbridge cycling.
- **Ejection fraction** refers to the stroke volume of blood as a fraction of the end-diastolic ventricular volume. This is a useful index of contractility.
- **Inotropic state** is another term for the myocardial contractility.
- **Intermittent claudication** refers to chronic ischaemia of the legs with hypoxic pains while walking.
- **Dromotropic state** refers to the conduction velocity of the myocardial conduction system.
- **Maximum oxygen uptake** is the oxygen uptake during maximum exercise. This is a measure of *endurance capacity*, and when expressed per kg of body weight it is also called the *fitness number* (ml min^{-1} kg^{-1}).
- **Mean circulatory equilibration pressure** (MCEP) is the filling pressure everywhere in the circulatory system following cardiac arrest.
- **Venous pump** refers to all local external forces acting on valvular veins and facilitating venous return.
- **Venous return** is the bloodflow reaching the right atrium (in steady state a similar bloodflow reaches the left atrium).
- **Ventricular stroke work** is the work applied to the blood at each ejection from the ventricle.

Essentials

This paragraph deals with *1. Cardiac electro-mechanics, 2. Pressure-volume work, 3. Venous return, 4. The venous pump, 5. The lipoprotein metabolism, and 6. Risk factors.*

1. Cardiac electro-mechanics

The heart is a four-chambered double pump. Every 24 hours the heart is ejecting more than 10^4 l of blood, and contracting more than 10^5 times. The total amount of work performed over the lifetime of a person is enormous. The order of size is calculated at the end of this chapter.

The cardiac cycle describes volume, pressure, and electric phenomena in the left ventricle as a function of time, and one heartbeat is shown below (Fig. 10-1). The red clock-shaped curve is the *intraventricular pressure.* The left ventricle is closed to the aorta in diastole and blood flows from the atrium to the left ventricle (Fig. 10-1). The ECG is explained in Chapter 11. Contraction or *ventricular systole* results in closure of the mitral valve (Fig. 10-1).

Fig. 10-1: Electro-mechanical events in the cardiac cycle. The aortic pressure is shown with a green curve. The blue atrial pressure curve has a, c, and v waves. For the ECG see Chapter 11.

The systole is *isovolumetric* (ie, the volume of blood in the ventricle is unchanged) until the

intraventricular pressure exceeds the aortic pressure. Then the aortic valve opens and ventricular ejection occurs (see thick upward arrow in Fig. 10-1). *Bulging* of the cuspidal mitral valve into the left atrium during isovolumetric contraction causes a rise in left atrial pressure (see the *c-wave* in the thin *a-c-v-curve* of Fig. 10-1). The intraventricular pressure reaches a plateau around 15-16 kPa and then begins to decrease. The aortic valve closes when the intraventricular pressure falls below aortic pressure (see small arrow indicating *retrograde* flow in aorta in Fig. 10-1). This is the end of the ejection phase or the *left ventricular ejection time* (LVET).

The *electrocardiogram* (ECG) does not reflect the mechanical performance of the heart, but *closure* of the aortic valve corresponds in time with the *end of the T-wave* in ECG (Chapter 11).

The *a-wave* occurs during the contraction of the right *atrium* by which it squeezes out extra blood just before ventricular systole (early atrial contraction). As the atrium relaxes, the CVP is reduced (blue atrial curve in Fig. 10-1). The next atrial wave is the *c-wave*, and this wave is produced by closure of the *cuspidal* valves and by the right ventricular contraction, because the increased ventricular pressure is transmitted backwards to the right atrium and large veins.

Filling of the atrial chambers with blood is aided by the large *atrial compliance*. This is why the atrial pressure only rises modestly. The third atrial wave is the *v-wave* for *venous return*. Throughout ventricular systole and isovolumetric relaxation, *venous blood* returns to the heart, but the tricuspid valve is closed, so that the central veins and right atrium are distended. The coinciding pressure build-up is relieved, when the tricuspid valve opens at the start of diastole, and the pressure is reduced (Fig. 10-1). The increase in *right atrial pressure* is transmitted backwards to the large veins near the heart. Prominent waves can often be seen in the neck veins when supine.

The periods just described for the left heart can be shown to be the same in the right heart, except for the fact that the systolic pressures are considerably lower in the right ventricle and pulmonary artery. The stroke volumes of the two ventricles are the same. Contraction of the right ventricle begins just after that of the left side and lasts for a shorter time (the peak pressures obtained are much less).

2. Pressure- volume work

At the end of the ventricular diastole the left atrial pressure increases, because the atrial systole begins just before the ventricular systole.

The intraventricular *pressure-volume loop* is a time-independent representation of the cardiac cycle, where the instantaneous intraventricular pressure and volume is plotted (Fig. 10-2). In diastole from B to C, the ventricle receives blood from the left atrium. The small increase in ventricular pressure reflects passive expansion and elastance of the myocardial wall. Pressure and volume increase with a slope that is related to *contractility*. The green line represents minimal contractility. The line starts from the ventricular *dead volume***,** that is a virtual minimal volume of blood, which can never be ejected (50 ml in Fig. 10-2). A steep rise in pressure occurs from C to D, with no change in ventricular volume (the isovolumetric contraction). At D the aortic orifice opens, because the end-diastolic pressure in the aorta is passed. During the rapid ejection phase, the fall in ventricular blood volume, is accompanied by a continuous increase in pressure. During ejection the volume falls by a size equal to stroke volume, pressure rises and falls until the *residual* ventricular volume is attained (about 80 ml in Fig. 10-2). The last ejection phase is slow, because the pressure decreases towards A, where the aortic orifice closes. The last event from A to B is the *isovolumetric relaxation* with a sharp drop in pressure at constant volume (Fig. 10-2). The steep slope of the curve refers to optimal contractility at a given condition (Fig. 10-2). Actually, the precise contractility concept is the *change* in slope from isovolumetric rest to the highest slope of the curve.

Fig. 10-2: The left ventricular pressure-volume loop from a healthy person at rest. – The pressure-volume loop is a widely applicable pathophysiological tool.

The *area* of the loop represents the pressure-volume work on the stroke blood volume performed by the ventricular contractile elements during ejection.

The pressure in C is the *end-diastolic* intraventricular pressure or the so-called *preload*. The force against which the ventricle contract is termed *afterload*. A good index of the afterload is the *peak aortic* (or intraventricular) pressure during systole, equal to the highest pressure shown in Fig. 10-2. The afterload is almost equal to the *peak* systolic pressure in the arterial tree. When the afterload is increased at constant end-diastolic pressure and volume, a greater ventricular pressure develops in order to expel blood (the dashed curve from D in Fig. 10-2). The result is a smaller stroke volume (and hence a greater residual ventricular volume), because of the high aortic and intraventricular pressure (Fig. 10-2).

The Frank-Starling relationship states that increasing left ventricular end-diastolic volume increases the stroke volume of the next heart beat. During isovolumetric contraction, the *end-diastolic* intraventricular pressure (EDIP) or *preload* increases. Increasing end-diastolic volume increases the ventricular filling pressure and thus stroke volume and stroke work. An increase in afterload occurs when the aortic pressure increases. Such an increase causes a decrease in stroke volume. Hereby, the EDIP and the end-diastolic volume increase, so the cardiac *stroke work* increases concomitantly (Fig. 10-2). The *ventricular* stroke work is the sum of the *pressure-volume work* and the *kinetic work* and calculated according to Eq. 10-1. *Preload* is the end-diastolic filling pressure of the ventricle just before contraction (C and E in Fig. 10-3).

Fig. 10-3: Left ventricular pressure-volume loops at rest (red area) and following an increase in end-diastolic volume by increased diastolic filling (blue loop with E-F).

When the left ventricle expands from C to E, by receiving more blood than before, the end-diastolic volume is increased. The greater diastolic filling results in a larger stroke volume according to *Starling's law* of the heart. The *Frank-Starling relationship* can be formulated as follows: Any increase of preload (ventricular filling) invokes a progressive increase in the blood volume ejected by the ventricles beat-by-beat until the cardiac output equals the input.The ventricular pressure increases with the rise in systolic aortic pressure (increased *afterload* in Fig. 10-3). The stroke work of the heart on the blood is the area A, B, E, F, A. Accordingly, the stroke work is increased. A sympathetically mediated increase in contractility without a change in end-diastolic pressure (*preload*) results in an increased intraventricular pressure (Fig. 10-4). The slope of the line through G illustrates the *increased contractility*. The larger stroke volume, smaller residual volume, and larger stroke work on the blood (area C, D, G, H) is also shown in Fig. 10-4.

Fig. 10-4: Left ventricular pressure-volume loops from a healthy person at rest and following an increase in contractility during exercise (G-H).

The transmural pressure rises during ventricular contraction even at constant fibre tension. This is because the short ventricular radius is reduced to the same extent (see the law of Laplace, Eq. 8-5).

Echocardiography shows that the short ventricular radius is reduced by 5-6 mm during systole in healthy people at rest. The Laplace law is acceptable in such a situation.

If a cardiac chamber of a heart patient increase its diameter to double, and a spherical chamber is assumed, it implies a *two-fold* greater fibre tension (Fig. 8-9B). To the patient, this means an enormous energy requirement in the myocardium to maintain the necessary pressure. This disorder is called *cardiac failure* or insufficiency (see later).

3. Venous return

Veins are highly distensible or *compliant* vessels (ie, they have a large volume/pressure ratio, d*V*/d*P*) that have *one-way* valves. The venous system (veins, venules, and venous sinuses) controls the amount of blood that is translocated from the venous to the arterial side of the circulation.

In this paragraph the circulatory system is simplified to a venous system connected by a heart pump to an arterial system (Fig. 10-5A). The *central venous pressure* (CVP) is the pressure measured in the caval veins at the level of the heart or in the right atrium.

When the heart pump stops, the pressure is the same in all compartments of the circulatory system, (ie, the mean circulatory equilibration pressure, MCEP). MCEP depends upon the blood volume and the compliance of the vessels (Fig. 10-5A). As the heart pump starts and moves blood from the venous system, the pressure here will fall and the arterial pressure will rise. A further rise in pumping activity (cardiac output) reduces the central venous pressure (CVP), and finally the CVP is negative, so the central vessels collapse. This impedes the bloodflow into the atrium (the venous return), so that the cardiac output cannot rise any longer (Fig. 10-5A).

Fig. 10-5: The central venous pressure (CVP) as a function of cardiac output (A). – The cardiac output as a function of CVP (B). – Combined venous return and cardiac output curves as a function of CVP (C).

The cardiac output must be a function of CVP at a given steady state. Increasing the venous return will increase CVP from –1 kPa towards zero, and increase cardiac output as well (Fig. 10-5B). As CVP becomes increasingly positive, it exerts a backpressure on the venous system to impede venous return. The rise in cardiac output levels off, and there is no further rise at values around MCEP.

The third step is to combine the curves in Fig. 10-5 A and B. At the normal CVP (or right atrial pressure) the venous return curve crosses the cardiac output curve, and both flows are 5 l per min. If the atrial pressure is suddenly increased to the *mean circulatory equilibration pressure*, then all flow of blood is stopped (Fig. 10-5C). The low pressure during arrested circulation is mainly due to the very distensible venous system. The right atrial pressure has only increased slightly, but enough to decrease the venous return to zero and thus the cardiac output is zero (Fig. 10-5C). The more the atrial pressure falls below the venous pressure, the more the venous return will rise up to a certain level at an atrial pressure of -0.2 kPa (almost zero). Negative atrial pressures have the same venous return. This is because negative transmural pressures in the central, thoracic veins imply collapse. The venous return is therefore constant, regardless of a further fall in right atrial pressure (Fig. 10-5C).

The cardiac output must equal venous return in the steady state. Thus the cardiac output- curve for the left ventricle must cross the venous return curve in one point (Fig. 10-5).

The steep part of the cardiac output curve shows that the cardiac output can double following a small rise in pressure.

The *driving pressure* for the systemic circulation is the mean aortic pressure (MAP) minus CVP. The relationship to cardiac output and total peripheral vascular resistance (*TPVR*) is given by Eq. 10-2. A small fraction of *TPVR* is found in the venous system.

Eq 10-3 expresses the venous return. Small variations in CVP alter the volume of blood considerably in the venous system. A normal value for venule pressure is 1.3 kPa (10 mmHg) and for CVP about zero. Since venous return must equal cardiac output in steady state, the venous resistance is only about 10 % of *TPVR*.

4. The venous pump

The venous pump is defined as all local external forces that facilitate venous return to the heart. Two important pump mechanisms are involved:

1. *The skeletal muscle- pump*. The deep veins of the arms and legs are affected by pressure exerted by exercising skeletal muscles. The veins are compressed by muscle contraction, and the one-way valves prevent the blood from flowing backward, and thus secure the transfer of blood toward the heart. Even the superficial veins are compressed during contraction.

   As soon as a venous segment is emptied of blood, its *transmural* pressure is so low that

the *filling* pressure from more peripheral veins can fill the empty segment with blood.

The skeletal muscular venous pump is also called the *peripheral venous heart.* In the erect position the peripheral venous heart must overcome the force of gravity and prevent overpressure in the dependent limb. During muscle rest there is an added hydrostatic pressure load of 13.3 kPa (100 mmHg) in the dependent limb. With a MAP of also 13.3 kPa the total pressure in a foot artery is 26.6 kPa, and in the dependent vein just above 13.3 kPa. During muscle contraction the venous pressure rises driving blood into the heart and just after muscle contraction the venous pressure falls again.

2. *The thoraco-abdominal pump.* The large veins are also affected by the positive intra-abdominal pressure and by the negative pressure in the thoracic cavity. The inferior caval vein returns blood from lower regions to the heart. During *inspiration* the intrathoracic pressure becomes more negative, and blood is *sucked* into the caval veins facilitating venous return to the heart. The inspiratory contraction of the diaphragm increases abdominal pressure favouring venous return. During *expiration* the intra-abdominal pressure decreases and intrathoracic pressure increases but remains negative, so that the venous return is maintained. *Intrinsic* cardiac mechanisms, including the *length-tension relation* (Chapter 2), allow the heart to increase stroke volume *beat-by-beat,* when the venous return increases. – Straining against a closed glottis is called Valsalva´s manoeuvre, and it is part of our every day life during coughing, defaecation, urination and lifting of heavy weights. The intrathoracic pressure increases abruptly, whereby the venous return is inhibited and the cardiac output is reduced (Starling´s law). Normally, fainting is prevented by a strong arteriolar and venous constriction released by the baroreceptor reflex.

5. The lipoprotein metabolism

This paragraph is inserted here in order to help the reader understand the pathophysiology of the most common cardiovascular disorders.

Lipoprotein particles are build up by a non-polar core containing triglycerides (TG) and cholesterol esters (E in Fig. 10-6). The polar shell of each particle consists of phospholipids, apoproteins and non-esterified cholesterol. These substances provide the particle with a negative electrical charge, which allow it to remain soluble in plasma (Fig. 10-6 right).

Hepatic synthesis of cholesterol varies inversely with the dietary intake.

Fig. 10-6: Lipoprotein metabolism (left). - The lipoprotein particles are found both in the plasma of healthy persons and of patients with atherosclerosis (right).

Four different lipoprotein particles (VLDL, IDL, LDL and HDL) control the transport of cholesterol to the cells.

a) *VLDL, IDL and LDLVery low-density lipoproteins* (VLDLs), which contain mainly the liver-produced TG, are synthesised in the liver and carries a characteristic surface protein called apoproteins B-100 (Fig. 10-6 right). VLDL is liberated from the liver in the postabsorptive phase

Chylomicrons are formed in the enterocyte from dietary fat after a meal and absorbed from the intestine into the blood (Fig. 22-13). Cylomicrons contain all the dietary lipids including lipophilic vitamins and have a half-life of 5 min.

VLDL from the liver, and chylomicrons absorbed from the intestine, are hydrolysed by the enzyme lipoprotein lipase (LPL) on the endothelial surfaces in the capillaries into glycerol, chylomicron remnants and free fatty acids (FFAs). From the FFAs the TG molecules are resynthesised and used or stored in adipocytes, heart and striated muscle cells (Fig. 10-6).

As more and more TG is removed from VLDL, the density of the particles becomes greater, and they are now termed *intermediate* density lipoproteins (IDLs). Normally, the liver cells take up half of the IDL particles, because they have receptors for the apoprotein B-100 on the IDL surface. The hepatic lipase removes TG from the IDLs to produce *low-density*

lipoproteins (LDLs) still maintaining their apoprotein B-100. This apoprotein is recognised by the LDL receptors of all cells. LDL is the largest cholesterol fraction in blood plasma, and has a half-life of 24 hours. LDL delivers cholesterol and other lipids from the liver to the cells for metabolic and structural purposes (forward transport). An increasing concentration of cholesterol inside the cell automatically down-regulates LDL receptors and thus regulates the receptor-mediated endocytosis of additional LDL. The circulating LDL concentration is controlled by the number of hepatic LDL receptors and by enzyme activity in the cholesterol synthetic pathway.

Genetic LDL receptor *deficiency* elevates the ratio of LDL to HDL in blood plasma, and a ratio greater than 4 is a serious risk factor for cardiovascular disease.

b) *HDL* The remains of the chylomicrons co-operate with IDL and *high-density* lipoproteins (HDLs) to form cholesterol esters. Cholesterol esters are then exchanged for TG in VLDL and chylomicrons by the *cholesterol ester-transfer protein*, whereby HDL_3 changes to the less dense HDL_2.

HDL is the substrate for *lecithin-cholesterol acyltransferase* (LCAT). This enzyme catalyses the conversion of free cholesterol to cholesterol ester. LCAT is reduced in severe liver disease.

HDLs in plasma are disk formed particles, mainly produced in the liver. They contain an entirely different apoprotein called apoprotein Apo-I or Apo-II, and also Apo-E (Fig. 10-6 right). In the fasting state, the HDL concentration in the blood plasma is generally increased in females, by oestrogens, by exercise, and by moderate alcohol intake.

Similarly, fasting HDL concentrations are reduced (and LDL increased) in males, by androgens, by smoking, by obesity, and by an inactive sedentary life-style.

The cell membranes contain specific HDL receptors. HDL absorbs cholesterol in peripheral tissues and thereby matures from the nascent state (Fig. 10-6 right). LCAT is activated by the apoprotein A on the HDL surface. Mature HDL facilitates the transport of cholesterol back to the liver ((backward transport or reverse cholesterol uptake), where it binds to Apo-E.

Normally, HDL particles carry 30% of the total quantity of cholesterol in the blood. HDL protects against development of atherosclerosis, and a high HDL/LDL ratio reduces the risk of cardiovascular disease.

Population groups at risk are advised to eat a low fat diet with unsaturated lipids and a low cholesterol content, in order to prevent or delay the development of atherosclerosis.

Persons with extremely low total cholesterol in their plasma demonstrate a higher mortality than persons with values around 5 mM. The reason for the increased mortality (mainly death of cancer and gastrointestinal diseases), is probably insufficient immune defence and genetic factors.

6. Risk factors

A *risk factor* for a disease is a factor showing a statistical co-variance with the disease. A risk factor is not identical with a definite *disease factor*, where all causal steps are clearly understood. Nevertheless, risk factors may obtain an increasing degree of causal relationship to the disease. Two or more risk factors present frequently potentiates the risk.

Major risk factors that predispose to atherosclerosis and ischaemic heart disease (IHD is atherosclerosis if the coronary arteries) are consequences of the Western World lifestyle. These consequences are often notified as age changes. The fact remains that the western life style is characterised by unhealthy fast-food, a high dietary fat fraction, obesity, years with lack of exercise, low fitness, smoking, hypertension, hyperlipidaemia, diabetes, gout, oral contraceptives (synthetic steroids), drugs and doping (testosterone or other steroids in excess).

Consideration of risk factors must be supplied with other relevant information in order to provide a whole patient status, including genetic and immunological factors as mentioned above.

The inactive lifestyle of the Western World is documented by measurement of a low maximal oxygen uptake (endurance capacity) as an average for large population groups. Values below 34 ml per kg and per min are unhealthy for any age. Such a low endurance capacity is related to high mortality (Chapter 18), especially from IHD (males and postmenopausal females). Regular exercise seems to protect against IHD.

The following risk factors for IHD are dealt with below: Obesity, male sex, smoking, hyperlipidaemia, familial hypercholesterolaemia, diabetes mellitus, gout and asymptomatic hyperuricaemia, oral contraceptives or synthetic steroids for doping.

Obesity is clearly associated with IHD, but probably not linked independently to IHD. The obese patient is characterised by an inactive lifestyle, low fitness, and preferring a fatty diet.

Sex. Males are more frequently affected by coronary artery disease (IHD) than fertile females. The smaller incidence in females is obviously related to the presence of natural female sex hormones (oestradiol in natural dosage). After the menopause, the female incidence approaches that of males. Female sex hormones in natural dosage may be protective, and male sex hormones *atherogenic*. Testosterone stimulates hepatic cholesterol synthesis.

Smoking a certain number of cigarettes per day is directly related to the incidence if IHD, and following 10 years of abstention, the risk declines towards the normal level.

Hypertension is associated with an increased risk of IHD (Chapter 9). Most forms of hypotensive therapy only reduce the risk of cardiac events to some extent, but clearly reduce the risk of *stroke*. Some hypotensive drugs reduce the arterial blood pressure but still have unwanted side effects. Light exercise reduces the blood pressure and implies other beneficial effects.

Hyperlipidaemia. Total cholesterol, HDL with calculation of LDL, and triglyceride

concentrations should be measured in all patients. High total serum cholesterol combined with a low HDL, and also high triglycerides are associated with an increased risk of IHD (Fig. 10-6).

Heterozygous familial hypercholesterolaemia is a relatively common genetic defect caused by mutations in the gene coding for the LDL receptor. With defective genes there is malproduction of LDL receptors in the liver. Some patients are without physical signs – others have cholesterol deposition around the eyes (xanthelasma) or in the tendons (xanthomas). These patients require diet with fibres and reduction of the cholesterol intake. Alcohol consumption must be reduced. The body weight must be kept close to ideal with exercise. Usually the patients require treatment with lipid-lowering drugs.

Homozygous familial hypercholesterolaemia is extremely rare. These patients are without LDL receptors in the liver, so they accumulate cholesterol and other lipids in the aorta, arteries, organs and skin. The HDL/LDL ratio in blood plasma is greatly reduced (below 1:4), and the fasting total cholesterol increases towards 30 mM. Drugs are needed, but the patients usually die young from ischaemic heart disease.

Diabetes mellitus. Increased blood glucose after fasting and an abnormal glucose tolerance test is associated with increased risk of atherosclerosis and a high LDL (Chapter 27).

Gout and asymptomatic *hyperuricaemia* (Chapter 20) is associated with an increased risk of ischaemic heart disease and atherosclerosis.

Intake of several types of oral contraceptives or synthetic steroids for doping (Chapter 18) increases the risk of atherosclerosis and thrombo-embolic phenomena.

Pathophysiology

This paragraph deals with heart disorders and atherosclerosis, which cause most people of the Western Hemisphere to die. *Coronary or Ischaemic Heart Disease* (CHD or IHD) is the most widespread. The remaining cases are caused by arteriosclerotic damage of the brain (strokes)

and other organs (liver, kidney etc).

The two typical cardiovascular disorders are I. Atherosclerosis and II. Rheumatic heart disease. - Atherosclerosis is involved in several widespread disorders (ischaemic heart disease, peripheral arterial disease and hypertension). These diseases are related to the typical life style of the Western Hemisphere, whereas rheumatic heart disease is more frequent in poor countries with high frequency of infections and malnutrition.

I. Atherosclerosis

Atherosclerosis is a process of progressive lipid accumulation (*atheromatosis*) and calcification of the inner arterial walls in the abdominal aorta, lower extremities and the arteries of the heart, brain and kidneys.

Atherosclerotic plaques contain cholesterol, and the most important single factor for their development is a high plasma concentration of total cholesterol, in particular a high concentration of LDL.

Atheromas are yellow streaks or lesions found in arteries at autopsy. They are formed in the intima (lamina intima) by lipid accumulation in macrophages and monocyte adhesion. As more and more cholesterol crystals are deposited, the atheromas grow and the surrounding fibrous and smooth muscle tissue is involved. Finally – as the subendothelial distortion leads to platelet aggregation - large arteriosclerotic plaques are formed. They consist of cholesterol and other lipids, dead cells, collagenous fibres, and there is excessive proliferation of the smooth muscle cells. The fibrosis or sclerosis makes the arterial wall stiff, which lead to systolic hypertension. Later Ca^{2+} salts precipitate and a factual calcification of the arterial wall may occur.

Typical for atherosclerosis patients are a high total cholesterol concentration in the blood plasma (total cholesterol above 6.2 mM), a dangerously high LDL and a low HDL fraction in fasting plasma (below 20% of the total). Often the atherosclerotic patient also has a high total triglyceride concentration (above 2 mM). In a fasted patient the triglyceride concentration depicts the precursor concentration of dangerous cholesterol: Very Low Density Lipoprotein (VLDL).

Large atherosclerotic plaques narrow the arterial lumen and produce arterial stenosis with reduced bloodflow. Insufficient oxygen delivery to the tissue is called ischaemic hypoxia, and hypoxic pains develop as in angina pectoris and intermittent claudication. Total occlusion of the arterial lumen is caused by a thrombus or an embolus in the lumen, or by wall bleeding. Disruption of the endothelium results in accumulation of thrombocytes and fibrin with thrombus formation and a complex atheromatous lesion is produced.

Arteriosclerosis (atherosclerosis) manifests itself in the coronary arteries as Ia. ischaemic heart disease and in peripheral arteries as Ib. peripheral arteriosclerosis.

Ia. Ischaemic heart disease (IHD)

Atherosclerotic coronary artery disease remains a leading cause of death, and is manifested as focal narrowing in the epicardial coronary arteries. The gradually narrowed vessel segment can be abruptly occluded by clot formation (thrombus) or by vasoconstriction at the atherosclerotic lesion. When a thrombus flows along the arterial tree with the blood and occludes the vessel, it is called an embolus.

Ischaemic heart disease is caused by reduced bloodflow to a region of the myocardium. Myocardial ischaemia diminish delivery of oxygen and nutrients, and accumulate potentially toxic substances such as lactic acid and K^+ around the cardiac cells, whereby necrosis may result. The causes are *atherosclerosis* with atheromas, thrombosis, emboli, or spasms in the coronary arteries.

Fig. 10-7: Bloodflow through the left coronary artery at rest and during exercise in a healthy person (upper) and in a patient with coronary obstruction (lower) and angina

pectoris.

Coronary bloodflow is restricted in the systole by strong myocardial contractions and in diastole by the high heart rate of exercise. Normally, the coronary bloodflow in healthy persons is small during systole and increases during diastole (Fig. 10-7A). The high heart rate at exercise implies a short diastolic duration, but the rise in pressure secures a great *diastolic* bloodflow (Fig. 10-7B).

Four clinical manifestations of IHD are treated below:

1) Angina pectoris (chest pains) is pain felt beneath the sternum or in the precordial area - often referred to the left arm-shoulder-neck-jaw etc. Exercise and cold bring on hypoxia pains in the substernal or precordial area. Hypoxic pains are transmitted by subendocardially situated nerve fibres. The coronary resistance vessels contain a-adrenergic constrictor receptors and badrenergic dilatator receptors. The vasodilatatory capacity of the coronary resistance vessels can be maximally mobilised already at rest (Fig. 10-7C). Exercise shortens the diastolic duration and restricts the rise in diastolic bloodflow further (Fig. 10-7D). The aggravated myocardial ischaemia results in a lactate acidosis.

Following sublingual administration of nitro-glycerine, peak concentrations are achieved in the plasma within 1 min. Organic nitrates dilatate constricted coronary vessels, improve the bloodflow to the subendocardial (pain sensitive) part of the myocardium, and dilatate resistance & capacitance vessels. This dilatation reduces the venous return to the heart and the arterial pressure (reduced preload and afterload).

The beneficial effect of drugs such as glycerol- trinitrate in angina have been known for more than a century. Recently it was realised that the drugs act by releasing nitric oxide (NO) in the vascular wall (see Chapter 5).

Ca^{2+}-channel blockers block the Ca^{2+} flux into the smooth muscle cells of the coronary arteries, so they relax. The Ca^{2+} -channel blockers also reduce the force of contraction and thus the oxygen demand of the myocardium.

Coronary angioplasty is a method by which atheromatous obstructions are dilatated by an inflated balloon.

The arterial oxygen concentration is also reduced in anaemia, CO poisoning and in shock. Patients with hyperthyroidism or hypertension may have increased coronary oxygen demand and all these paitients may experience *chest pains* caused by myocardial hypoxia.

Another manifestation of ischaemic heart disease is myocardial infarction.

2) Acute myocardial infarction (AMI) is due to a sudden coronary thrombosis from an atheromatous plaque causing cellular death (*infarct*) of a myocardial area. Distal to the coronary occlusion the blood pressure is low. The thin-walled subendocardial vessels are squeezed most and receive the smallest bloodflow, often leading to *subendocardial* infarcts. The myocardial infarcts are sometimes *silent* (which means *without pains*; the pain relief is due to destruction of subendocardial nerve fibres). The typical infarct causes severe and long lasting pain.

Acute myocardial infarction renders the heart incapable of pumping the minimal blood volume required to transfer sufficient oxygen to the mitochondria.

The patient experiences a sudden chest pain and the pain is lasting for hours in contrast to *angina*. The patient may develop signs of shock. Necrotic myocardial cells liberate cellular enzymes such as *creatine kinase*, which peaks in the blood plasma within 24 hours. The total enzyme release depicts the size of the infarction. Lactic dehydrogenase (LDH) isoenzymes peaks a few days later, and *LDH 1* is rather specific for myocardial necrosis.

Read Chapter 11 before this paragraph:

Non-Q wave infarction: When only part of the wall is necrotic there are deeply inverted,

symmetrical T-waves (*coronary T- waves*) and mostly ST depression in the ECG. These signs of ischaemia are often transient - only during the acute attack - and found in all precordial leads located above the infarcted area. Such a subendocardial infarct does not show deep Q waves, and epicardial involvement implies ST segment elevation.

Fig. 10-8: Myocardial infarction of the left ventricular wall with lack of movement of infarcted tissue during systole. The ECG changes are typical for the anterolateral location of the infarct (see Ch. 11).

Q-wave infarction: A wide and deep Q wave in the ECG is a lesion wave, and the sign of transmural myocardial infarction with necrosis through the whole of the myocardial wall. The deep Q wave is maintained for years after the event. A typical *anterior wall infarction* shows changes in lead I and in V_2-V_6 dependent upon the localisation (Fig. 10-8). During systole the infarcted area does not contract or move due to cell necrosis (paradoxical movement). There is always a danger of rupture of necrotic tissue.

A typical *posterior wall infarction* is diagnosed by a *mirror image* with changes in V_1-V_2 (*reciprocal* changes) and ST- depression in lead I. ST-depression or ST-elevation is indicative of myocardial ischaemia.

Chapter 11 is a prerequisite for the understanding of the above paragraph!

Fig. 10-9: Left ventricular pressure-volume loops in a healthy person (red) and for persons with acute (blue area) or chronic, congestive (green area) cardiac failure.

The AMI patient is extremely tired. Even the work of breathing is a heavy task. The condition is often a case of general ischaemic hypoxia (Fig. 10-9) and can develop into cardiogenic shock (see below).

Interaction of platelets with collagen in the vessel wall is the first step in platelet aggregation leading to thrombosis. The activated platelets release thromboxanes A2 (TxA2) from arachidonic acid in the phospholipids of the platelet membrane. Platelet aggregation is inhibited by cAMP and by acetyl-salicylic acid, which inhibits platelet cyclo-oxygenase.

Eicosapentaenoic acid (EPA) in the diet reduces the frequency of thrombosis by reduction of the TxA2 production.

3) Cardiac failure or *cardiac insufficiency* is a disorder, where the heart cannot pump enough blood to satisfy the nutritive needs of the body. Cardiac insufficiency is manifest by a consequential decrease in cardiac output (*lower output failure*) or by an increase in cardiac output (*higher output failure*). The cardiac failure can be acute or chronic.

Damming of blood in the vessels behind the insufficient heart pump is typical.

Acute cardiac failure is caused by AMI, acute intoxications, anaesthesia etc. Occlusion of the coronary artery to the left ventricle impairs contractility, and left ventricular failure develops due to the diminished cardiac output from the left ventricle. Initially, the right ventricular output is maintained, whereby the left atrial pressure (and pulmonary venous pressure) is increased beat-by-beat. As a consequence, the left ventricular output will increase until the cardiac outputs of the two ventricles are equal. The increased pulmonary venous pressure leads to reduced lung compliance (dV/dP) and increased respiratory elastic resistance with increased respiratory work and distress. Eventually, plasma fluid flows into the alveoli and *pulmonary oedema* is developed (Fig. 10-10).

Chronic or congestive cardiac failure occurs in conditions such as IHD and following severe hypertension. In *chronic cardiac failure* blood is accumulated and expands the venous system and the left ventricle (Fig. 10-9).

Cardiac oedema develops during congestive cardiac failure, because the kidneys retain NaCl and water. The accumulated fluid increases venous return, which in turn elevates the right atrial pressure. The rising atrial pressure elevates the venous and the capillary pressure. This causes loss of fluid into the interstitial fluid volume. Accumulation of abnormal volumes of

interstitial fluid is the definition of *oedema*.

The low cardiac output and blood pressure causes an increased sympathetic tone with constriction of the afferent renal arterioles to the glomeruli. As a consequence, the renal bloodflow (RBF), and the glomerular filtration rate (GFR) decrease (Fig. 10-10). Also the NaCl concentration decreases in the renal macula densa (Fig. 25-17). The *renin-angiotensin-aldosterone cascade* is activated, which enhances salt-and water-retention. Angiotensin II is a strong vasoconstrictor, which further decreases the renal bloodflow, and aldosterone promotes the reabsorption of NaCl and water from the distal renal system. A certain salt-water retention is beneficial in the early stages of cardiac failure, because of improved venous return and thus improved cardiac output according to Starlings law of the heart. However, prolonged activation of the renin-angiotensin-aldosterone cascade and the *sympathetic nervous system*, damage the heart muscle further and reduce its contractility. This is because circulating vasoconstrictors, such as catecholamines, vasopressin and angiotensin II, imply an extra workload on the damaged myocardium.

When the salt and water retention results in even a small rise in the osmolarity of plasma, there is a stimulus of osmoreceptors, located close to the neurosecretory cells in the hypothalamus. The osmoreceptors stimulate both production and secretion of vasopressin (antidiuretic hormone, ADH) in the neurosecretory cells. ADH eases the renal reabsorption of water in the outer cortical collecting ducts leading to a low urine flow (antidiuresis).

Vasopressin is also an universal vasoconstrictor.

Fig. 10-10: Formation of pulmonary oedema in left ventricular failure (mitral stenosis) and in congestive cardiac failure with ankle oedema.

Increased venous pressure with stasis of blood dilatates the central vessels and the heart chambers. The distended atrial wall liberates *atrial natriuretic peptide* (ANP), which increases Na^+-excretion and dilatates peripheral vessels (Chapter 24). This is a partial compensation of the increased *preload* (the water loss by high urine flow reduces venous return) and *afterload* (the vasodilatation reduces outflow resistance).

Venoconstriction shifts significant quantities of blood from the peripheral to the central circulation. Since *central venous pressure* (CVP) varies inversely with *TPVR*, it is possible to maintain cardiac output in resting patients with congestive heart failure (insufficient contractile force) at the expense of increased CVP, by reduction of *TPVR* (Eq. 10-2).

When cardiac output decreases more and more during development of congestive heart failure, the compensation fails, and both CVP and end-diastolic ventricular pressure (preload) and volume rises further (Fig. 10-9). The superficial meck veins are expanded, when CVP is abnormally elevated. Eventually, large volumes of plasma water flow from the liver into the peritoneal cavity due to the elevated CVP. Fluid accumulation in the abdomen is called *ascites*.

Patients with a cardiac output much *higher* than normal can develop cardiac failure. The venous return is much too high, and after some time with an overexpansion of the heart, the cardiac pump fails to eject the same blood volume as it receives, and an increasing blood volume is accumulated behind the insufficient ventricle. The rise in left atrial pressure leads to pulmonary oedema and eventually the right ventricle fails so peripheral oedema develops.

Examples of this condition are cardiovascular disorders with a drastic reduction of the *TPVR*. The low opening pressure in the left ventricle being equal to a low end-diastolic aortic pressure (Fig. 10-11) illustrates the low *TPVR*.

Fig. 10-11: Left ventricular pressure-volume loops from a healthy person, and from a person with high metabolic rate (hyperthyroidism) - or a person with arteriovenous shunts.

In *hyperthyroidism* (a disease with an abnormally high metabolic rate described in Chapter 28), all the vessels in the systemic circulation dilatate, and the venous return overloads the

heart. On the other hand, short-term administration of L-thyroxin to patients with chronic heart failure improves cardiac and exercise performance.

Any major *arteriovenous shunt* leads a large fraction of the arterial blood directly into the veins. This greatly increases venous return and overloads the heart (Fig. 10-11).

4) Cardiogenic shock - terminal pump failure - is such a severe reduction of cardiac output that the peripheral tissues suffer seriously from lack of oxygen, the cells deteriorate and within hours or days the patient die. The pulmonary capillary wedge pressure is normal or elevated in contrast to other types of shock (blood loss or vasodilatation).

During insufficient pumping capacity and cardiac arrest, the cardiac pump do not get rid of the blood volume received and a large blood volume is therefore accumulated in the distensible venous system, pulmonary system and the thin-walled chambers of the heart. This is why the lower part of a newly diseased body is filled with blood in distensible vessels (*livores*), and the upper part of the body is pale.

I b. Peripheral arteriosclerosis

Peripheral arteriosclerosis refers to atherosclerosis and other changes of the large and medium large peripheral arteries. The muscular lamina media grows and becomes fibrotic often with atheromas in small arteries. The walls of the elastic arteries become thick of hyaline and the lumen narrows, which causes systolic hypertension. The narrowing also leads to ischaemia, which further promotes systolic hypertension.

The most frequent of these disorders is chronic ischaemia of the legs, also called *intermittent claudication* from its prominent symptom. Claudication is a cramp-like pain, which occurs during exercise and subsides at rest. Occlusive atheromatous lesions between the common iliac and the common femoral artery lead to claudication of the thigh and calf. Lower femoral artery disease usually causes claudication of the calf, and occlusive lesions of the popliteal artery causes claudication of the calf or the foot.

If possible, regular relaxed exercise should be undertaken in an attempt to develop anastomoses. In severe ischaemia there is pain at rest. Balloon dilatation is often useful, and amputation may become necessary.

II. Rheumatic Heart Disease

Rheumatic fever is caused by repeated pharyngeal infections with group A streptococcus. An autoimmune reaction is triggered by the streptococci and the patient develops fever, joint pains, diastolic mitral murmur caused by mitral valve inflammation, cardiac enlargement with pericardial effusion and pericarditis (raised ST-segment in ECG – see Chapter 11) or myocarditis (inverted T-waves). More than 50% of those who suffer from *acute rheumatic fever* with carditis develop rheumatic heart disease many years later.

The rheumatic valvular disease mainly affects the mitral and the aortic valves.

Mitral stenosis and regurgitation

Practically all cases of mitral stenosis are caused by rheumatic heart disease. Severe mitral valve stenosis is present, when the mitral valve orifice is reduced to 10^{-4} m^2 as compared to the normal area of ($5 * 10^{-4}$ m^2). The left atrium dilatates and its walls hypertrophy in order to maintain sufficient bloodflow to the left ventricle. Obviously, the pressures also increase in the entire pulmonary vascular bed: veins, capillaries, arteries and right ventricle.

The stenotic mitral valve and the resistance of the pulmonary arteries determine the pressure in the pulmonary capillary bed. Up to a certain point, pulmonary arteriolar vasoconstriction protects the patient from pulmonary oedema.

Patients with severe mitral stenosis have cyanotic cheeks and ears (mitral faces) caused by stasis of the blood. Auscultation at the apex, lying on the left side just following exercise reveals a split second heart sound with a *mitral snap* (second component) as the mitral valves open, then a mid-diastolic rumbling murmur - like a sack of potatoes falling on a floor -

caused by turbulence through the narrowed orifice. The murmur ends in a loud first heart sound, because the cusps are kept open until the start of the ventricular systole (Fig. 10-12). As the left atrium grows it becomes activated later than the right, and the P-wave is bifid (P-mitral). The large left atrium favours the development of atrial fibrillation with thrombo-embolism and emboli to the brain, kidneys and gastrointestinal tract.

Mitral stenosis is often combined with regurgitation. Regurgitation is recognised by a systolic murmur without any 1.heart sound. The 1.heart sound is caused by the closure of the cusps, and in mitral stenosis they do not close. Usually the condition is asymptomatic for decades following rheumatic fever. Light pulmonary oedema presents itself as coughing and exertion dyspnoea.

Replacement of the mitral valve is performed with artificial valves, which may work for decades with adequate anticoagulant therapy.

Fig. 10-12: Mitral stenosis, aortic stenosis and aortic regurgitation.

Aortic stenosis and regurgitation

Disorders of the normal tricuspid aortic valve are mainly caused by rheumatic fever or by atherosclerosis. Almost half of all cases of rheumatic heart disease include the aortic orifice, usually associated with the mitral orifice.

Symptoms and signs are characteristic: Exercise-induced syncope and angina. This occurs when the disease is severe and the area of the aortic orifice is reduced to 1/3 of normal. The left ventricular pressure rises and the left ventricle hypertrophies. The increased oxygen demand of the myocardium leads to ischaemia with angina pectoris, arrhythmias and left ventricular failure. Healthy persons can increase their cardiac output by a factor of 5 or more, but this is not possible for patients with severe aortic stenosis. The arterial blood pressure falls, the patient is pale (*aortic face*), chest pains worsen, and the patient may lose consciousness. There is a strong systolic murmur in the aortic area and over the carotid arteries. The echocardiogram demonstrates thick aortic valve cusps and left ventricular hypertrophy. A ventricular-aortic pressure gradient above 6.7 kPa (50 mmHg) measured by cardiac catheterisation, is indicative of surgery. Without surgical intervention death frequently ensues within a few years from the occurrence of the first serious signs.

Tricuspid stenosis and regurgitation

This is an uncommon complication, which is related to rheumatic heart disease or is congenital. Regurgitation is more frequent than stenosis, but often the two conditions are combined.

Isolated tricuspid stenosis dilatates the right atrium and the caval veins, and the liver swell just like the condition with constriction of the heart. Atrial fibrillation is frequent from the dilatated chamber.

Tricuspid regurgitation is a condition where the right ventricle delivers blood to both the pulmonary artery and the right atrium at each systole. The right heart is dilatated, whereas the lung vessels are not. Pulsation of the neck veins and a large tender liver are typical signs. There is often a blowing systolic murmur over the sternum.

Replacement of the tricuspid valve is performed with artificial valves.

Left ventricular hypertrophy

Left ventricular hypertrophy is an abnormal increase of the left ventricular mass caused by increased demands of cardiac work. If due to pressure overload it is called *concentric* hypertrophy in which new contractile elements are lined up in parallel, and if due to volume overload it is called *eccentric* hypertrophy, in which new contractile elements elongate the myocardial cell. - This disease is described in Chapter 11.

Other cardiovascular disorders

Thrombo-angiitis obliterans (Buergers disease). *Buerger´s disease* occurs in small arteries of

the limbs of young smokers – typically males. The vessel wall is inflamed, but many lesions look like atherosclerosis. The patient is invalid by intermittent claudication, and the only choice of treatment is to stop smoking.

Raynaud´s disease is a condition with cold precipitated attacks of spasms of the small arteries and arterioles, supplying the fingers and toes. The disease is usually bilateral and affects predominantly young girls and female smokers.

First the skin becomes pale and white from vasoconstriction, due to slow bloodflow, and finally red because of hyperaemia. The vasoconstriction occurs in the digital arteries, arterioles and skin capillaries. A few minutes later the capillary smooth muscle spasm is released due to local vasodilatators, and the capillaries are filled with oxygen poor blood (the skin becomes blue and is still cold). Finally, the arterial and arteriolar constriction is released and the classical physiological reactive hyperaemia occurs, with red, warm fingers and *paresthesia* (numbness). Centrally controlled vasoconstrictor tone, sensitive to cold signals, is probably implicated in cases with symmetrical spasms. A centrally increased vasoconstrictor tone involving the coronary bloodflow is consistent with the fact that some of the Raynaud patients also suffer from *chest pains* (angina pectoris) and *migraine*.

Raynaud´s disease occurs in a primary and a secondary form. Primary Raynaud´s disease is a condition where the cause is unknown (ie essential). There is a benign familiar occurrence of so-called *dead fingers* (ie, digiti mortui familiaris), and a malignant form with symmetrical gangrene (ie, symmetric gangrene).

Secondary Raynaud´s disease (*Raynaud´s phenomenon*) occurs together with connective tissue disorders (dematomyositis, polymyositis, systemic lupus erythematosus and systemic sclerosis). Raynaud´s phenomenon is also a side effect to treatment with b-blockers, in which case they must be withdrawn. The patient has to wear warm clothes in order to protect both the shell and the core temperature. *Nifedipine* is sometimes beneficial.

Varicose veins have incomplete valves. Normally, the muscular venous pump maintains the venous bloodflow towards the heart. Patients with defective valves can develop venous pooling or stasis and ankle oedema. This is because the contracting leg muscles squeeze the blood in the retrograde as well as in the anterograde direction.

Equations

Ventricular stroke work rate is the sum of the pressure-volume work and the kinetic work:

$$\textit{Eq. 10-1: Stroke work rate} = [(P \times V) + \tfrac{1}{2} m * v^2].$$

Both the pressure-volume work and the kinetic work are work per stroke duration or time unit that is comparable to work-rate or effect in Watts.

The driving pressure for the systemic circulation is the mean aortic pressure (MAP) minus CVP. The relationship to cardiac output and total peripheral vascular resistance (TPVR) is given by:

$$\textit{Eq. 10-2: Cardiac output} = (MAP - CVP)/TPVR.$$

A small fraction of TPVR is found in the venous system. The venous return is expressed by the approximative equation:

Eq. 10-3: Venous return = (venule pressure - CVP)/ venous resistance.

Self-Assessment

Multiple Choice Questions

I. The following five statements have True/False options:

A. Smoking has no effect on Buerger´s disease.

B. When the brain is deprived of new blood, gray out occurs 4.5 seconds following blockage of its bloodflow.

C. There is a mid-systolic rumbling murmur at the apex in severe mitral stenosis.

D. The central venous pressure varied inversely with the total peripheral vascular resistance.

E. Organic nitrates dilatate constricted coronary arteries, improve the bloodflow to the subendocardial part of the myocardium especially, and dilatate resistance and capacitance vessels.

II. The following five statements have True/False options:

A. Exercise-induced syncope and angina is characteristic of aortic stenosis and regurgitation.

B. Concentric ventricular hypertrophy is due to volume overload, whereas eccentric hypertrophy is due to pressure overload of the ventricles.

C. IDL particles are further hydrolysed by LPL, resulting in slightly higher density of the so-called low-density lipoprotein still maintaining its *apoprotein* B-100.

D. Echocardiography shows that the short ventricular radius is reduced by 15-20 mm during systole in healthy persons at rest.

E. The residual ventricular volume is approximately 150 ml of blood in healthy persons at rest.

Case History A

Three years following heart-lung transplantation, a patient is examined at the hospital for the cause of frequent exertion syncope. Heart catheterisation reveals that the systolic/diastolic pressure in the left ventricle is 26.7/0 kPa (200/0 mmHg), and in the aorta 10.7/6.7 kPa (80/50 mmHg).

1. *What is the cause of exertion syncope?*
2. *What is a likely diagnosis?*
3. *Argue for the size of the left ventricular cavity and wall thickness.*

Case History B

A 24-year old sporty male consults the doctor because of syncope while playing handball.

The examination reveals a systolic murmur to the right in the second intercostal space aortic site). The systolic murmur is audible also over the neck. The arterial blood pressure is 145/85 mmHg (19.3/11.3 kPa). There is a history of rheumatic fever at the age of 11.

ECG shows a deep S-wave in V_1 and a high R-wave in V_6 (the sum is 5 mV), and there are asymmetrically negative T-waves in V_4-V_6.

1. *What is the most likely diagnosis?*
2. *What is the ECG diagnosis*
3. *Describe the prognosis and the therapy.*

Case History C

A 55-year old female complains of headache and an arterial blood pressure of 190/100 mm Hg (25.3/13.3 kPa) is found. Her ECG shows deep negative S-waves in V_1 and V_2, and high R-waves in V_5 and V6. The sum of one S- and one R-wave is above 4 mV.

1. *Calculate the mean arterial pressure and compare the result to a normal value.*
2. *What is the diagnosis?*
3. *Is the patient suffering from any cardiac disease?*
4. *Why is the arterial pressure amplitude (eg, systolic minus diastolic pressure) much larger than normal?*

Case History D

A girl, 17 years of age, with frequent episodes of acute tonsillitis and articular pain during her childhood, developed a cardiac disease. The diagnosis was made as the doctor heard a diastolic murmur over the precordial area. The main complaint of the girl was dyspnoea at exertion. Cardiac catheterisation revealed a mean pressure in the pulmonary artery of 58 mmHg (7.7 kPa), and in the left atrium 28 mmHg (3.7 kPa), whereas the pressure in the left ventricle was only 2 mmHg (0.27 kPa) in early diastole. The pressure in the aorta was 105/80 mmHg (14/10.6 kPa). The cardiac output was measured at rest with the Fick principle to be 2.5 l min^{-1}. Her circulating blood volume is somewhat less than her total blood volume, and exactly 5000 ml. The blood volume of the systemic capillaries is 3% of 5000 ml.

1. *What is the name of the condition with frequent episodes of tonsillitis?*
2. *Calculate the pulmonary vascular resistance.*
3. *In what orifice is the cardiac disorder located?*
4. *Calculate the mean passage time in the systemic capillaries.*

Case History E

A female, 33 years old, complains of attacks of severe pain in the fingers of both hands, when she is outdoors in cold weather. She is a heavy smoker since the age of 16. The patient also suffers from another pain disorder. She has half-sided headache with visual disturbances at least twice a month in the cold season. The patient explains the pain attacks in the fingers as follows. First the skin of both hands and the fingers (not the thumb) becomes pale and white, and they feel like dead. A few minutes later the skin is blue, and the pain is severe. After 5-10 minutes the skin suddenly becomes red and it is very painful.

1. *What is the diagnosis of the finger disease?*
2. *Are the finger disease and the headache related?*
3. *Describe the pathophysiology of the finger disease.*

Try to solve the problems before looking up the answers.

Highlights

- *The ventricular stroke work is the sum of the pressure-volume work and the kinetic work.*
- *A good index of the afterload is the peak aortic pressure during systole.*
- *A ventricular-aortic pressure gradient above 6.7 kPa (50 mmHg) is indicative of surgery in aortic stenosis.*
- *The venous pump is defined as all local external forces that facilitate venous return to the heart.*
- *The skeletal muscular venous pump is also called the peripheral venous heart, because its force must be equal to or larger than that of the heart in order to return the blood in the upright position.*
- *During inspiration, the intrathoracic pressure becomes more negative, and blood is sucked into the large thoracic veins facilitating venous return to the heart.*
- *Atherosclerosis is a process of progressive lipid accumulation (atheromatosis) and calcification of the inner arterial walls in the abdominal aorta, lower extremities and the arteries of the heart, brain and kidneys.*
- *High-density lipoproteins (HDLs) in plasma are disk formed particles, mainly produced in the liver for cholesterol transport. HDLs protects against the development of atherosclerosis as they transport the cholesterol back to the liver, where the elimination*

begins.

- *Acute myocardial infarction renders the heart incapable of pumping the minimal blood volume required to transfer sufficient oxygen to the mitochondria at rest.*
- *Cardiogenic shock - or terminal pump failure - is such a severe reduction of cardiac output that the peripheral tissues suffer seriously from lack of oxygen, the cells deteriorate and within hours or days the patient die. The pulmonary capillary wedge pressure is normal or elevated in contrast to other types of shock.*
- *Left ventricular hypertrophy is an abnormal increase of the left ventricular mass caused by increased demands of cardiac work. If due to pressure overload it is called concentric hypertrophy in which new contractile elements are lined up in parallel, and if due to volume overload it is called eccentric hypertrophy, in which new contractile elements elongate the myocardial cell.*
- *Cardiac oedema develops during chronic cardiac failure, because the kidneys retain fluid.*
- *Buerger´s disease occurs in small arteries of the lower limbs of young male smokers. The vessel wall is inflamed, but many lesions look like atherosclerosis. The patient is invalid by intermittent claudication, and the only choice of treatment is to stop smoking.*
- *Raynaud´s disease is a condition with cold precipitated attacks of spasms of the small arteries and arterioles, supplying the fingers and toes. The disease is usually bilateral and affects predominantly young girls and female smokers.*
- *Rheumatic fever is caused by repeated pharyngeal infections with group A streptococcus. The streptococci trigger an autoimmune reaction and the patient develops fever, joint pains, diastolic mitral murmur caused by mitral valvulitis, cardiac enlargement with pericardial effusion and pericarditis or myocarditis.*
- *More than 50% of those who suffer from acute rheumatic fever with carditis develop rheumatic heart disease many years later.*
- *The rheumatic valvular disease mainly affects the mitral and the aortic valves.*
- *Practically all cases of mitral stenosis are caused by rheumatic heart disease. Severe mitral valve stenosis is present, when the mitral valve orifice is reduced to 10^{-4} m^2 as compared to the normal area of ($5 * 10^{-4}$ m^2).*
- *Disorders of the normal tricuspid aortic valve are mainly caused by rheumatic fever or by atherosclerosis. Almost half of all cases of rheumatic heart disease include the aortic orifice, usually associated with the mitral orifice.*

Further Reading

von Spiegel T, G Wietash and A Hoeft (1998). Basics of myocardial pump function. *Thorac Cardiovasc Surgery* 46, Suppl 2, 237-41.

Return to content

Return to chapter 10

Chapter 11

Cardiac Action Potentials and Arrhythmias

Study Objectives

- To *define* asystolia, atrial fibrillation, atrial flutter, axis-deviation, bradycardia, heart block, tachycardia, sinus rhythm, ventricular fibrillation, ventricular tachycardia and vulnerable period.
- To *describe* cells with pacemaker-function, the normal impulse conduction in the heart, the recording of the electrocardiogram (ECG), and different types of heart block.
- To *draw* membrane potentials of sinus cells, atrial cells, and myocardial cells during a cardiac cycle, and draw a normal ECG.
- To *estimate* the electrical QRS-axis from the R-waves of the standard limb leads.
- To *explain* the effect of the autonomic nervous system on the membrane potential of the sinus node. To *explain* the genesis of the ECG in health and disease. To *explain* the Adam-Stokes syndrome.
- To *use* these concepts in problem solving and case histories.

Principles

- *The body conducts electrical signals uniformly in all directions (an almost uniform volume conductor increasing the potential electrical field around the heart generator).*
- *Einthoven's law states that any two of the three bipolar standard limb leads determine the third one with mathematical precision.*

Definitions

- **An electrocardiogram (ECG)** is a curve showing the potential variations against time in the whole body stemming from the heart, which is an electrochemical generator suspended in a conductive medium.
- **Asystolia** refers to cardiac arrest.
- **Atrial fibrillation** is a continuous atrial activation with 400 or more contractions per min. Contractions spread through the atrial tissue almost without mechanical effect and only few electrical signals are conducted to the ventricles.

Atrial flutter is an atrial contraction rate around 300 per min, often with every second contraction conducted to the ventricles. Sawtooth-like flutter waves characterise the ECG.

- **Bradycardia** is an unduly slow heart rate. ***Sinus bradycardia*** is a sinus rhythm at rest below 60 beats per min during the day or less than 50 at night.
- **Calcium concentration** in plasma (total): Normal range is 2-2.5 mM.

Heart block is a blockage somewhere along the pathway for impulse conduction in the heart.

-
- **Mean QRS- axis** of the ventricles or the *mean cardiac vector* is the net force in the frontal plane during ventricular depolarisation and repolarisation. Many electrical potentials are propagated in different directions and most of these cancel each other out. The main direction of the mean cardiac vector is from the base of the ventricles towards the apex.
- **Left-sided axis deviation** is characterised by a positive R_I and a negative R_{III} (Fig. 11-7). Cardiologists use the net area of the QRS-complex for precise diagnosis.
- **Pacemaker cells** are small pale cells located in the sinus node of the heart. The sinus node is the primary determinator of the cardiac rhythm, because its cells have the highest spontaneous frequency.
- **Potassium concentration** in plasma: Normal range is 3.5-5 mM.
- **Right-sided axis deviation** is characterised by a negative R_I and a positive R_{III} (Fig. 11-7). Cardiologists use the net area of the QRS-complex for precise diagnosis.
- **Sinus rhythm** refers to the normal cardiac pacemaker rhythm from the sinus node. The spontaneous discharge at rest is usually 100 beats per min, but the parasympathetic inhibitory tone predominates in healthy individuals resulting in a resting heart rate around 75 beats per min.
- **Tachycardia** refers to a cardiac rate above 100 beats per min. *Sinus tachycardia* is a sinus rhythm above 100, which can be caused by anaemia, cardiac failure, catecholamines, emotion, exercise, fever, pregnancy, pulmonary embolism or thyrotoxicosis.
- **Ventricular tachycardia** is defined as three or more ventricular beats occurring at a rate of 120 beats per min or more.
- **Ventricular fibrillation** is an extremely rapid ventricular activation without pumping effect. Electrical defibrillation is the only effective therapy.
- **Vulnerable period** is a dangerous period in cardiac cycle just at the end of the contraction (simultaneous with the T-wave in the ECG). Electrical conversion (an electrical shock) given during this period may in itself initiate ventricular fibrillation. Refractory areas of cardiac muscle are spread among non-refractory areas.

Essentials

This paragraph deals with *1. Neural regulation of the heart function, 2. The action potential, 3. The role of calcium, 4. Spontaneous depolarisation, 5. Development of electrocardiography, and 6. The normal ECG.*

1. Neural regulation of the heart function

A large number of sympathetic (S) and vagal (X) motor nerve fibres end close to the sinoatrial node (SA in Fig. 11-1).

Fig. 11-1: The neural control of the heart.

Sympathetic stimulation speeds up the sinus node (*sinus tachycardia*) and vagal activity slows the

node (*sinus bradycardia*). Increased concentration of the sympathetic transmitter noradrenaline, and of adrenaline from the adrenal glands, cause *positive inotropic state* (increased contractility), *positive chronotropic state* (increased frequency), *positive dromotropic state* (increased conduction velocity), and *positive bathmotropic state* (increased irritability) on the heart. Noradrenaline activates a-adrenergic constrictor receptors in the coronary vessels, whereas adrenaline activates b-adrenergic vasodilatator receptors.

The neurotransmitter acetylcholine, activating muscarinic receptors, and vagal stimulation causes *reduced contractility* (negative inotropy), *reduced frequency* (negative chronotropy), *reduced conduction velocity* (negative dromotropy), and *reduced irritability* (negative bathmotropy).

2. The action potential

Across the ventricular cell membrane there is a steady potential difference of almost the same size as the equilibrium potential for K^+ (-94 mV), that is -*90* mV (Fig. 11-2). This negative potential is referred to as the *resting membrane potential* (RMP), because it represents the potential difference across the cell membrane (inside negative) at rest between successive action potentials.

Fig. 11-2: Recordings of ECG (above), intracellular membrane potential (red curve) and contraction (blue curve) of one heart cycle in a ventricular fibre.

Any process that reduce the absolute size of the RMP (ie, depolarise the membrane) tends to activate (open) *fast* Na^+-channels. These channels contain fast opening and fast closing gates (inactivation gates). Electrochemical forces favour the abrupt influx of Na^+ from neighbouring regions. Hereby, the potential is further diminished and more and more Na^+-channels are activated or opened. The threshold potential for release of an action potential is a rise of 25 mV from -90 mV. The cardiac action potential is an *all-or-none response*, which can be divided into *five phases:*

The fast depolarization (phase 0) is shown by the abrupt upstroke, which is related to the rapid entry of Na^+ into the cell through the *fast* Na^+-channels, suddenly allowing the electrostatic and chemical forces to work (Fig. 11-2). The *fast* Na^+-influx causes phase 0 of atrial, ventricular and Purkinje action potentials (Fig. 11-2). The fast Na^+-channels are both *voltage*- and *time*-dependent. Phase 0 stops at about +30 mV, because the fast Na^+-channels become voltage-inactivated by closure of inactivation gates. The potential difference approaches the equilibrium potential for Na^+ (+ 60 mV), but only reaches +30 mV. The conduction velocity along the fast response fibre increases with the AP-amplitude and especially with the slope of phase 0.

Phase 1 is the early repolarization from the upstroke. This is related to K^+-outflux.

Phase 2 is the plateau of the action potential, where the *slow* Ca^{2+}-Na^+-channels remain open for a long period - up to 300 ms. The net influx of Ca^{2+} and Na^+ is almost balanced by a net efflux of K^+, so the balance is forming the *plateau* (Fig. 11-2). Ca^{2+} activates the muscle contractile process. When the slow Ca^{2+}-Na^+-channels close at the end of the plateau, the voltage-gated K^+-channels are activated, and the permeability for K^+ increases rapidly.

Phase 3 is the terminal repolarization. With *all* the K^+-channels open, large amounts of K^+ diffuse out of the ventricular fibre. The equilibrium potential for K^+ (-94 mV) and the RMP is rapidly approached.

Phase 4 is recognized by the RMP of - 90 mV (Fig. 11-2). The Na^+-K^+ pump restores ionic

concentrations by exchanging Na^+ for K^+ in a ratio of 3:2.

Phase 5 covers the *relative* refractory period (RR), and the T-wave in the ECG. The long *absolute* refractory period (AR) of the ventricular cells covers the whole shortening phase of the contraction (Fig. 11-2 blue curve). In the absolute refractory period all fast Na^+-channels are voltage-inactivated and closed, which prevents sustained tetanus. As a consequence, no stimulus is sufficient to trigger contraction regardless of size.

In the relative refractory period, enough of the fast Na^+-channels are recovered, so that a sufficiently large stimulus can break through and produce an action potential although smaller than normal.

The long absolute refractory period protects the cardiac pump, as it is not possible to bring ventricles into smooth tetanus.

As described above the cardiac muscle fibre has a horizontal plateau, because of the *slow* Ca^{2+}-Na^+-channels (phase 2). - Skeletal muscle fibres have no plateau, because they do not open slow Ca^{2+}-Na^+-channels for such a long time.

3. The role of Ca^{2+}

Cardiac fibres contain many mitochondria and intercalated desks with gap junctions, transverse tubules of invaginated sarcolemma, and sarcoplasmic reticulum (Fig. 11-3). The fibres require a continuous supply of oxygen and they are provided with a rich capillary supply, about one capillary for each cardiac fibre.

The spontaneous firing from the pacemaker spreads over the entire heart as a propagating wave.

The wave of excitation is conducted rapidly along the long axis of a cardiac fibre, and spreads along the myocardial sarcolemma from cell to cell via electrically conducting *gap junctions.*

The excitation also reaches the interior of the cell through the large T-tubules filled with mucopolysaccharides. During the phase 2 plateau of the action potential, Ca^{2+} permeability increases, and Ca^{2+} flows down its electrochemical gradient into the cell through the *slow* Ca^{2+}-channels.

The *channel proteins* are phosphorylated by a *cAMP-dependent protein kinase A*. The small Ca^{2+} influx is called *trigger*-Ca^{2+}, because it releases large amounts of Ca^{2+} from the sarcoplasmic reticulum (Fig. 11-3). Hence the cytoplasmic [Ca^{2+}] increases from the resting level of 10^{-7} molar by a factor of 10-100 during excitation. The free Ca^{2+} binds to troponin C, just as in striated muscle cells, and the complex interacts with tropomyosin to activate sites between the actin and the myosin filaments. This process starts *crossbridge cycling* and thus contraction of the myofibrils.

Fig. 11-3: Excitation-contraction coupling in a cardiac fibre.

When the Ca^{2+}-influx ceases at the end of systole, the Ca^{2+} -movement is reversed. Now, Ca^{2+} is pumped into the sarcoplasmic reticulum by a Ca^{2+} -pump. The binding of Ca^{2+} to *troponin C* is inhibited by phosphorylation of troponin I, and the binding sites between actin and myosin are blocked resulting in diastole. In diastole the Ca^{2+} surplus is removed by a 3 Na^+- 1 Ca^{2+}-exchanger, and by an electrogenic Ca^{2+}-pump (Fig. 11-3).

Catecholamines and increasing extracellular [Ca^{2+}], raise the cytoplasmic [Ca^{2+}] and thus the

developed force of contraction. This is accomplished in the following way. Adrenaline and noradrenaline activate adenylcyclase, whereby cAMP is formed and the dependent *proteinkinase A* phosphorylates and activates Ca^{2+}-channel proteins (Fig. 11-3). Hereby, the cytoplasmic Ca^{2+} is increased, and thus the force of contraction.

The excitability of cardiac fibres, striated muscle cells and neurons is reduced by a *reduction* of the cellular Ca^{2+}-gradient (hypocalcaemia), a *rise* in the Na^{+}-gradient across the cell membrane, or the *administration* of Ca^{2+}-blockers that prevent Ca^{2+} from entering the cell.

Cardiac glycosides, such as digoxin, are beneficial for patients with chronic cardiac failure. Digoxin inhibits and reduces the number of Na^{+}-K^{+}-pumps in the membranes of cardiac cells, hereby producing high intracellular levels of Na^{+}. Some of the intracellular Na^{+} is exchanged for extracellular Ca^{2+}, and cytoplasmic Ca^{2+} enhances the force of cardiac contraction (positive inotropic effect). Exactly the necessary minimum dose of digoxin must be used for each patient with cardiac failure.

In myocardial cells, as in nerve and skeletal muscle cells, K^{+} plays a major role in determining the RMP**.** At physiologic concentrations of K^{+} outside the myocardial cell ($[K^{+}]^{o}$ about 4 mM), the RMP is determined by a dynamic balance between the membrane conductance to K^{+} and to Na^{+}. As $[K^{+}]^{o}$ is increased (hyperkalaemia), the membrane depolarises. Depolarisation inactivates voltage-dependent K^{+} -channels and activates Na^{+} -channels, allowing Na^{+} to make a proportionally larger contribution to the RMP. Increased $[K^{+}]$ in the extracellular fluid reduces the force of contraction, so the heart becomes dilatated and flaccid.

Because the equilibrium potential for Na^{+} is positive (+ 60 mV), it tends to depolarise the RMP. Since RMP is -90 mV at normal $[K^{+}]$, and the equilibrium potential of K^{+} is always more negative (-94 mV), there is a small outflux of K^{+} from the cell at equilibrium. The K^{+} and Na^{+} gradients are maintained by the efficient Na^{+}-K^{+}-pump.

4. Spontaneous depolarisation

The rhythm of the heart is initiated by a complex flow of electrical signals, which are called *action potentials*. The action potentials generated in the sinus node (SN) display *automaticity* (ie, they undergo spontaneous and rhythmic depolarization without external stimuli). The sinus node is the *primary pacemaker,* because it has the highest frequency. The automaticity (Fig. 11-4) is associated with small *pale round cells* in the SN.

The SN and the AV-node also contain elongated cells that react with a special AP, a so-called *slow response*. The AP is smaller than the fast response, phase 0 is long and caused by slow Ca^{2+}-influx, phase 1 is absent, phase 2 is slow repolarisation, and the real repolarisation in phase 3 is due to inactivation of the slow Ca^{2+}-channels and increased K^{+}–outflux. Phase 4 is horizontal, and the relative refractory period is long, extending well into phase 4. The slow response propagates slowly and tends to be blocked, which causes cardiac rhythm disturbances.

The *membrane potential* of pacemaker cells (about -55 to -60 mV) is *never* constant (Fig. 11-4). The diastolic depolarization of pacemaker cells (producing the very special slope in phase 4 of these cells) is ascribed to an influx of Na^{+} through special channels in their cell membrane. Towards the end of phase 4 there is also a certain influx of Ca^{2+} through voltage-activated Ca^{+2} -channels. Hypocalcaemia diminishes the amplitude of the action potential and the slope of the

pacemaker potential. The diastolic pacemaker depolarization is opposed by the K^+-outflux.

Fig. 11-4: Pacemaker potentials from a sinus nodal fibre: a: Sympathetic stimulation; b. Normal heart rate; c. Vagal stimulation. There is a constant Na^+-influx between two heartbeats. Reduction in the slope of the pacemaker potential and increase of the threshold both diminish the cardiac frequency.

The K^+-channels remain open during the repolarization, so the membrane potential approaches -60 mV at the end of the action potential. The rising pacemaker potential between two heartbeats, is caused by the inherent leaks of the pacemaker membrane to Na^+. The *constant* Na^+-influx causes the membrane potential to approach the threshold potential. When the membrane potential reaches the threshold potential (about -40 mV), the slow Ca^{2+}-Na^+-channels open and the cycle starts again. This process continues for a lifetime (Fig. 11-4).

The heartbeat is self-initiating, and normally the propagating wave (impulse) originates in the sinus node (see above). The impulse propagates from the sinus node via *three* bundles of internodal syncytial cells, through the left and right atrial wall to the *atrioventricular* (AV) *node*. This point is typically reached within 40 ms. After passing through the AV node (with a so-called AV- *delay* of 100 ms), the propagating wave (excitation) reaches the *bundle of His*. Here, specialised conduction fibres activate almost synchronously all the ventricular tissue and thus impart maximal thrust to the blood. The propagation velocity in these large *Purkinje fibres* is 1-4 m per s, which is the fastest velocity possible in the heart. The Purkinje system terminates just under the endocardial surface on *gap junctions* in the myocardial cells. The AV-delay provides time for the atrial systole to pass extra blood to the ventricles before the ventricular systole occurs.

Adrenergic transmitters and sympathetic stimulation increase the slope of the diastolic pacemaker depolarization. Acetylcholine and vagal stimulation increase the K^+-efflux, so the slope is reduced and thus the cardiac frequency is reduced.

Some cardiac fibres show a slow response comparable to that of the pacemaker cells, but with a constant phase 4. Ischaemia may activate *ectopic* pacemaker cells.

5. Development of electrocardiography

In 1903 Einthoven began a systematic study, with a string galvanometer of the potential differences between electrodes placed on the skin surface during heart beats. The string galvanometer was developed to become the first electrocardiograph. The fluids of the body conduct electricity quite easily. The body conducts electrical signals uniformly in all directions.

An electrocardiogram (ECG) is a curve showing the potential variations in time in the body stemming from the heart, which is an *electrochemical generator* suspended in a conductive medium acting as a volume conductor. The myocardial cell membranes have separate charges, and small ion gradients produce electrochemical gradients. Small ion fluxes across the cell membrane only occur during depolarization and repolarization, where potential differences are produced between polarised and depolarised tissue regions in the heart. Each cardiac fibre behaves as a dipole, the magnitude and direction of which is symbolised by an arrow or a vector. The dipole vector points from minus to plus by definition.

Originally Einthoven assumed that the sum of all electrical activity in the heart resulted in an electromotive force - a *main cardiac vector* originating in the middle of the heart.

Einthoven used three standard bipolar limb leads forming a triangle in the frontal plane (Einthoven's triangle in Fig. 11-5 and 11-7). Lead I records the potential difference between the right and left arms, lead II between the right arm and left leg; and lead III between the left arm and left leg. The right leg is used to ground the patient (Fig. 11-7; observe the positive and negative

signs). The connections were arbitrarily chosen so most healthy individuals had dominating upright (positive) QRS-complexes and T-waves in their leads.

The actual recording sites are at the junctions between limbs and trunk, because the arms and legs act as *extended electrodes*. Einthoven actually placed the limbs of the patient in bathing tubes and used these as the first electrodes. Einthoven arranged the ECG equipment so that the *direction* of the mean QRS-axis towards the *positive pole* of a bipolar lead produces an *upright* deflection (*positive* R-waves in the 3 leads shown in Fig. 11-5). When directed towards the *negative pole* a *downward* deflection (a *negative* RIII-wave in left-sided axis deviation, Fig. 11-7) was recorded.

Einthoven's law states that any two of the three bipolar limb leads determine the third one with mathematical precision. The potential differences recorded over time in an ECG can be estimated using the rules of *vectorial* projection with force parallelograms in the frontal plane (Fig. 11-5).

Fig. 11-5: Einthoven's triangle. To the left is shown the main direction of the conduction system in the frontal plane.

The R-waves (and QRS deflections) in two of the three standard leads are drawn graphically in Einthoven's triangle and their resultant is the *mean QRS-axis* of the heart or more exact the mean ventricular axis in the frontal plane (Fig. 11-5). This is a mathematical concept representing the integral cardiac vector operating in Einthovens triangle.

The *mean QRS-axis* of the heart is usually located around 60 degrees in Einthoven's triangle.

Right- sided axis deviation is found in children or high thin individuals with a vertical located heart, and in persons with right ventricular hypertrophy. The axis is now located to the right and the condition is diagnosed by a positive R_{III} combined with a negative R_I (between +90 to +180 and even +90 to –90 degrees in Fig. 11-7).

Left-sided axis deviation is found with left ventricular hypertrophy, in fat individuals and in late pregnancy, when the heart is pressed upwards to the left. The condition is diagnosed by a negative R_{III} combined with a positive R_I (Fig. 11-7).- Actually cardiologists operate with the *net areas* of the QRS-complexes in order to diagnose significant deviations.

6. The normal ECG

The ECG is recorded from the surface of the body, and used to demonstrate the presence of AV-blocks, ectopic foci, premature beats, sinoatrial arhythmias, atrial fibrillation and ventricular fibrillation etc.

Each heart cycle has a fixed pattern of ECG waves (P,Q,R,S and T). The sinus node is a minimal muscle mass, and there is no potential difference (wave in the ECG) before the atria depolarise with a P-wave. When the propagating impulse wave is directed towards the *positive* electrode (as in lead II) the atrial depolarization will produce a *positive* P-wave (Fig. 11-6). The P-waves correspond to the impulse distribution in the atria. Retrograde excitation of the atria creates a *negative* P-wave. When atrial excitation coincides with the QRS, the P-wave is often hidden. The PR-interval (normally 0.12-0.2 s) measures the impulse speed through the supraventricular tissue from the sinus node to the bifurcation of the Hiss bundle. Prolonged PR-interval is caused by disturbances of AV conduction. The QRS-complex measures the passage through the left and right bundle branch followed by depolarization of the strong ventricular myocardium (Fig. 11-6). The QRS-complex is prolonged (> 0.12 s), when the left or right bundle branch is blocked by disease. When excitation originates in the ventricles, it spreads slowly and the QRS-complex is severely deformed.

Since the activation of the septum is mainly from left towards right, the propagating wave moves away from the exploring electrode, and the *Q-wave* becomes *negative*, whereas the dominating

ventricular transfer of the propagating wave towards the apical electrodes provides the large, *positive* R-wave in almost all leads. The small propagating wave moving away from the electrode at the apex and to the right to reach the thin-walled right ventricle, is responsible for the small, *negative* S-wave. The T-wave represents ventricular repolarization and has the same direction as the QRS-complex in most normal leads. When the QRS-complex is positive and the T-wave is negative, it indicates that the repolarization proceeds in a wrong direction. Abnormal T-waves are due to cardiac hypertrophy, myocardial damage or electrical disturbances.

The isoelectric ST-segment represents the period, where the entire ventricular myocardium is depolarised (therefore isoelectric). Any deviation (up or down from the isoelectric level) indicates *anoxic damage* of the myocardium.

The QRS- plus ST-intervals correspond to the duration of the ventricular systole.

Fig. 11-6: Normal ECG (II. lead). The action potentials from an atrial (green curve) and a ventricular fibre (blue curve) are shown above. – To the right is shown the direction of propagating waves in the frontal plane and their relation to the ECG waves.

The *T-wave* is positive in most leads, and due to the apical directed repolarization of the ventricular action potential (Fig. 11-6).

Unipolar, precordial leads (Wilson) have *one* exploring electrode as the actual recording electrode, detecting changes in the local potential relative to zero. The exploring electrode is defined relative to the three standard extremity leads, which are connected to form one indifferent *reference electrode*. Conventionally, the *6 Wilson leads* are recorded from specific locations on the precordium (V_1 to V_6).

The exploring electrodes of leads V_1 to V_3 are located to the right and are looking at the right side of the heart. The QRS complexes of the normal heart are mainly negative in these leads. The exploring electrodes of leads V_4 to V_6 are located to the left and are looking at the left side of the heart, where the QRS complexes are typically positive.

Fig. 11-7: Standard limb leads (Einthoven's triangle) and precordial ECG leads.

In unipolar recordings, upright deflection indicates movement of the propagating wave towards the positive, exploring electrode, and downward deflection (a negative ECG wave) indicates that the propagation wave is moving away (Fig. 11-7).

Pathophysiology

This paragraph deals with *cardiac arrhythmias.*

Arrhythmia is any cardiac rhythm different from the normal sinus rhythm. Sinus rhythm is the automaticity elicited from the normal cardiac pacemaker, the sinus node. The sinus node depolarises spontaneously with a frequency between 60 and 100 beats per min (bpm) at rest. Normally the parasympathetic tone predominates, whereby the heart rate is reduced from 100 to 60-70 bpm at rest. Any reduction in parasympathetic tone or increase in sympathetic tone results in tachycardia (ie, a ventricular pumping frequency above 100 bpm). Reduction in sympathetic tone or increase of the parasympathetic tone leads to bradycardia (ie, a ventricular pumping rate below 60 bpm).

Cardiac arrhythmias are divided into two groups: I. Pacemaker abnormalities, and II. Conduction abnormalities (cardiac block).

I. Pacemaker abnormalities

arise in the sinus node (ie, sinus tachycardia, sinus bradycardia or sinus arrhythmia) or outside the node (ie, ectopic beats, tachycardia, and fibrillation and shifting pacemaker). These arrhythmias

are disorders of rhythmogenesis.

Sinus node disease is a mixture of conditions with insufficient discharge of signals from the sinus node. The *sick sinus syndrome* is caused by damage of the nodal tissue. Sinus pauses lead to sinus bradycardia, tachycardia, tachy-brady-cardia syndrome, ectopic beats or atrial arrest (sinus arrest).

Sinus tachycardia (heart rate above 100 bpm) is caused by psychological (panic, anxiety) or physical stress (anaemia, hypoxia, exercise, fever, intoxication, shock etc). Any condition with increased sympathetic tone results in tachycardia. Adrenergic b-blockers are effective in slowing down the heart rate.

Sinus bradycardia (heart rate below 60 bpm) is a normal phenomenon in well-trained people and in elderly persons. Athletes at rest have a high stroke volume and a low pulse, because they are dominated by vagal tone in this condition. All persons react with sinus bradycardia, during hypothermia, hypothyroidism, jaundice, increased intracranial pressure, treatment with digoxin or b-blockers, and in sinus node disease.

Sinus arrhythmia is a normal phenomenon. There is a rise in heart rate during inspiration followed by a fall during expiration. During inspiration, the low intrathoracic pressure improves the venous return, which – with a delay through the pulmonary circulation - increases the stroke volume of the left ventricle. The resulting increase in aortic intravascular pressure combined with the low intrathoracic pressure around the aorta leads to an increased transmural pressure over the aortic wall and thus to a strong stimulation of the aortic baroreceptors. This is why the vagal tone falls and the sinus node discharge quickens. The reflex and circulation delay explains why the fall in heart rate manifests itself during expiration and not during inspiration.

Ectopic beats originate in pacemaker cells outside the sinus node. The abnormal pacemaker tissue is triggered by ischaemia, mechanical or chemical stimuli. Cardiac catheterisation can trigger ectopic beats. Ectopic beats are either of atrial or ventricular origin.

1) *Atrial ectopic beats* appear as early (premature extrasystoles) and abnormal P-waves in the ECG; they are usually followed by normal QRS-complexes (Fig. 11-8). Following the premature beat there is often a compensatory interval. A premature beat in the left ventricle is weak because of inadequate venous return, but after the long compensatory interval, the post-extrasystolic contraction (following a long venous return period) is strong due the Starling´s law of the heart. - Adrenergic b-blockers are sometimes necessary.

2) *Ventricular ectopic beats* (extrasystoles) are recognized in the ECG by their wide QRS-complex (above 0.12 s), since they originate in the ventricular tissue and slowly spread throughout the two ventricles without passing the Purkinje system. The ventricular ectopic beat is recognized by a double R-wave (Fig. 11-8). The classical tradition of simultaneous cardiac auscultation and radial artery pulse palpation eases the diagnosis. Now and then a pulsation is not felt, and an early frustraneous beat is heard together with a prolonged interval. A beat initiated in the vulnerable period may release lethal ventricular tachycardia, since the tissue is no longer refractory.

Fig. 11-8: Atrial (left) and ventricular (right) ectopic beats.

After a period with high cardiac frequency from activity of an ectopic focus (ie, overdrive), there often follows a period with a remarkable fall in frequency (ie, overdrive suppression). During the overdrive, the Na^+-K^+-pump is extremely active in order to extrude Na^+ from the myocardial cells in the short phase 4 periods, and the Na-influx exceeds the K-influx (Na:K= 3:2). Hereby, the cell becomes hyperpolarised in the end, which may stop the high frequency and turn it into suppression.

Tachycardia occurs in *paroxysms* and is either of atrial or ventricular origin.

1) *Atrial tachycardia* is elicited in the atrial tissue outside the SN as an atrial frequency around 200 bpm. Often only every second impulse passes the AV-node to the ventricles, so a 2.1 AV-block is found in the ECG (Fig. 11-9).

2) *Ventricular tachycardia* is elicited from one focus in the ventricular tissue with a frequency around 200 bpm (more than 120 bpm) and abnormal intraventricular impulse conduction (disturbed QRS complexes). Of course, there are no P-waves in the ECG, and the QRS-complexes are broad and irregular (Fig. 11-9).

Fig. 11-9: Left: Atrial tachycardia with a QRS-frequency of 100 bpm. - Right: Ventricular tachycardia with a QRS-frequency of 200 bpm following 2 sinus beats.

Fibrillation is either atrial or ventricular in origin.

1) *Atrial fibrillation* is a condition in which the sinus node no longer controls the rhythm and the atrial muscle fibres undergo a tumultuous rapid twitching. A total irregularity of ventricular contractions characterise the fibrillation. An excitation wave with 400-600 cycles per min, courses continuously through the atrial wall over a circular pathway about the origin of the great veins (the *circus motion theory*). There is a continuous activation with more than 400 P-waves per min, where regular atrial contraction is impossible. It is difficult to see and count the P-waves of the ECG. Because of the refractoriness of the AV-bundle, only some of the excitation waves result in ventricular beats. The pulse of the patient is therefore irregular as the occurrence of QRS-complexes in the ECG. The many P-waves (also called f-waves for fluctuations) are characteristic for atrial fibrillation. Untreated atrial fibrillation has a QRS-frequency of 150-180 bpm (Fig. 11-10). Old patients with chronic heart disease often show the so-called *slow atrial fibrillation* with a QRS-frequency below 60 bpm.

Most cardiac disorders can lead to atrial fibrillation or flutter.

Atrial flutter is related to atrial fibrillation, but the atrial frequency - counted from the P-waves - is much lower than 400 bpm - usually around 300 bpm and the AV-conduction is more regular. The consequences to the patient depend upon the number of impulses conducted from the atria through the AV-node to the ventricles (recorded as QRS-complexes). Often every second impulse reaches the ventricles, so the ratio of AV-blocks is 2:1, but the ratio can also be 3:1, 4:1 etc. Atrial flutter is recognized in the ECG as sawtooth-like P-waves (Fig. 11- 10).

2) *Ventricular fibrillation* is a tumultuous twitching of ventricular muscle fibres, which are ineffectual in expelling blood. The condition is lethal without effective resuscitation. The irregular ventricular rate is 200-600 twitches/min. Without contractile co-ordination the force is used frustraneous. Actually, the heart does not pump blood, so within 5 s unconsciousness occurs, because of lack of blood to the brain. In patients with coronary artery disease, ventricular fibrillation is a cause of sudden death. The trigger is *anoxia* (with an ineffective Na^+-K^+-pump) and the impulses arise from several foci in the ventricular tissue. There is no regular pattern in the ECG. Ventricular fibrillation is initiated when a premature signal arrives during the downslope of the T-wave (*vulnerable period*). Electrical shock (electrocution) also triggers ventricular fibrillation.

Fig. 11-10: Ventricular defibrillation of a patient with cardiac arrest.

Ventricular fibrillation is the most serious cardiac arrhythmia. It must be converted to sinus rhythm at once by the application of a large *electrical shock* to the heart (ventricular defibrillation) or the patient will die. Alternating current is applied for 100 ms or 1000 volts direct current is applied for a few milliseconds. The *vulnerable period* (VP in Fig. 11-7 is actually phase 3 and represented in the ECG as the T-wave) is dangerous, because an electrical shock, when given during this period, will cause in itself *ventricular fibrillation*.

Here is shown sinus rhythm and one ectopic beat followed by ventricular fibrillation. The only effective treatment is rapid institution of electrical defibrillation.

Fig. 11-11: Left: Atrial fibrillation with a QRS-frequency of 180 bpm. - Right: Ventricular fibrillation following 2 beats of sinus rhythm and one ectopic beat.

Shifting pacemaker is a condition where the impulse originates in shifting locations inside the SN, or the pacemaker shifts from the SN to the AV-node. In the first case the P-wave change size from beat to beat, and in the second case the P-wave is found either in front of the QRS-complex or behind.

II. Conduction abnormalities

are 1) sino-atrial block, 2) atrioventricular block, 3) bundle branch block, 4) WPW- syndrome, and 5) the long QT-syndrome.

1) Sino-atrial block (see SN disease) is characterized by long intervals between consecutive P-waves, and caused by blockage of the formation or conduction of the stimulus from the SN to the atrial tissue (ischaemia or infarction of the SN).

2) Atrioventricular block is blockage of the conduction from the atria to the AV-node.

The first-degree AV block is a prolongation of the PQ (PR)-interval (above 0.2 s) implying a delay of the conduction - not a real block. All beats are conducted, so there is a QRS-complex following each P-wave, although with delay (Fig. 11-12).

The second-degree AV block occurs when some signals are not conducted to the AV-node, so some of the P-waves are not followed by QRS-complexes. The ventricles actually drop some beats. A typical example is Mobitz type I block or Wenchebach block, which is a predictive loss of a QRS-complex. The PQ-interval is increased progressively until a P-wave is not followed by a QRS-complex. Mobitz type II block occurs without warning. Suddenly, a QRS- complex falls out (Fig. 11-12).

The third degree AV block (complete AV-block) is a total block of the conduction between the SN and the ventricles. Also blocked Hiss bundle conduction results in an AV-block (Fig. 11-12). An AV- or ventricular pacemaker maintains life with a spontaneous escape rhythm around 40-50 bpm, or cardiac arrest occurs with the fainting paroxysms of Adam-Stokes syndrome.

The Adam-Stokes syndrome is a clinical disorder caused by a partial AV-block, with a long P-Q interval and a wide QRS complex in the ECG, suddenly becoming a total bundle block. The condition results in unconsciousness and cramps caused by brain hypoxia and sometimes resulting in universal cramps (*grand mal*) due to violent activity in the motor cortex. The keyhole is the AV node and the bundle of His. Disease processes here elicit the Adam-Stokes syndrome. Therapy is to provoke sinus rhythm by a few forceful strokes in the precardial area of the thorax, accompanied by mouth-to-nose-resuscitation and external heart massage. A sympathomimetic drug can be injected if necessary even in the heart directly. The patient must be immediately brought to hospital for special intensive care. Permanent pacemaker treatment may become necessary.

Fig. 11-12: Four types of atrio-ventricular (AV)-block. From above downwards: First-degree AV-block, Second-degree Mobitz I block (Wenchebach), Second-degree Mobitz II block, and Complete AV-block.

3) Bundle branch block is a block of the right or the left bundle branches. The signal is conducted first through the healthy branch and then it is distributed to the damaged side. This distribution takes more time than usual, so the QRS-complex is wider than normal (more than 0.12 s in Fig. 11-13).

In *right* bundle branch block, the right ventricle is activated late, which is shown by a tall double R-wave in V1 (ie, the second late R-wave is from the right side), and a deep wide S-wave in leads

I and V6 .

The *left* bundle branch block is characterized by a late activation of the left ventricle from apex towards basis. This results in a solid R-wave in the left precardial leads (V5 andV6), whereas there is a deep broad S-wave in V1 and III (Fig. 11-13).

Fig. 11-13: Right and left bundle branch block.

4) WPW-syndrome or *Wolf-Parkinson-White block* is not a direct block of the conduction through the Hiss bundle and branches, but is caused by a short cut through an extra conduction pathway from the atria to the ventricles. This abnormal conduction pathway is congenital and called the *bundle of Kent* (Fig. 11-14). Due to this short-cut, the slow conduction through the AV-node is bypassed and the ventricles are depolarised faster than normal. The WPW-syndrome is recognized in most ECG leads as a short PQ (PR)-interval followed by a wide QRS-complex with a delta wave (Fig. 11-14). The patients often have paroxysmal tachycardia or they may develop atrial fibrillation.

Some patients are treated with ablation of the *bundle of Kent*. Other patients are asymptomatic and in good physical condition.

Fig. 11-14: The WPW-syndrome and the long QT-syndrome.

5) The long QT-syndrome. This is frequently a genetic condition, where fast repolarised cells are restimulated by cells that have not repolarised. When acquired the condition is caused by myocardial ischaemia, by drugs or by a low serum $[Ca^{2+}]$ - below 2 mM. Normally, the QT-interval is less than 50% of the preceding RR-interval (Fig. 11-14). The long QT-interval symbolises a long ventricular systole. Actually, the ST-interval is simultaneous with the phase 2 plateau of the ventricular membrane action potential. Here, the slow Ca^{2+} -Na^{+} - channels remain open for more than 300 ms as normally. The net influx of Ca^{2+} and Na^{+} is almost balanced by a net outflux of K^{+}. Hereby, a long phase 2 plateau or isoelectric segment is formed.

Cardiac pacemakers

Implanted cardiac pacemakers are successful in keeping heart patients alive. This is often a beneficial treatment of Adam Stokes syndrome or ventricular tachycardia. An electrical pacemaker is a small stimulator with battery planted underneath the skin. The electrodes are connected to the right ventricular muscle tissue, whose contraction rate is controlled by the stimulator.

Cardiopulmonary resuscitation

Cardiac arrest is cessation of all spontaneous cardiac rhythmicity. Cardiac arrest is most often caused by anoxia. The cause of anoxia is inadequate respiration due to terminal lung disease, thoracic trauma, and shock or deep anaesthesia.

Cardiopulmonary resuscitation is important in keeping the heart alive until electrical defibrillation can be performed with a large electrical shock. Alternating current is applied for 100 ms, or 1000 mV direct current is applied for a few ms.

Ventricular hypertrophy

The consequences of the rise in cardiac mass are the same for the locally recorded ventricular fibre action potentials. The action potentials increase in magnitude, with a parallel increase of QRS voltage observed in the ECG. The heart partially adapts to the increase in workload by an increase in muscle mass (hypertrophy). The degree of hypertrophy is roughly proportional to the increase in load. As the electrically active surface area is increased, there is an increase in ventricular fibre action potential, and thus in the amplitude of the R wave in the left precordial leads.

The *ECG criterion* of left ventricular hypertrophy is that the sum amplitude of S in V_1 and R in V_6 is larger than 3.5 mV (3.5 cm).

With a delay in conduction through the large left ventricle - or with left bundle branch block - there is a wide QRS complex.

The normal positive T-wave is due to the apical directed repolarization of the ventricular AP. Therefore, asymmetrical T-inversion or bi-phasic T-waves and downward-sloping ST-segment signal abnormal repolarization with the propagating wave moving away from the apical electrode in most leads (so-called strain pattern in Fig. 11-15).

Left axis deviation is often found in the standard leads of left ventricular hypertrophy patients (augmented R-wave in I and S-wave in III).

Fig. 11-15: ECG from a patient with left ventricular hypertrophy.

The disorders causing *left ventricular hypertrophy* are:

Heart diseases: Myocardial disorders, pericarditis, valvular disorders, congenital heart disease.

Vascular disorders: Atherosclerosis, systemic hypertension (Chapter 12) , aortic stenosis, renal disorders, arteriovenous shunts, and aneurysms.

Thoracic diseases: Diseases of the lungs & pleura, and kyphoscoliosis.

Pumping of increased volume load: Acromegaly, anaemia, obesity and excessive alcohol intake, thyrotoxicosis, severe manual work and sports.

Left ventricular hypertrophy is often demonstrated by echocardiography or found on the ECG.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

Stem statement: The ventricular action potential is

A. initiated by rapid entry of Na^+.

B. characterised by slow Ca^{2+} -Na^+- channels.

C. characterised by closed K^+- channels in phase 3.

D. dependent upon Ca^{2+}-influx.

E. independent of the Na^+-K^+ -pump in phase 4.

II. Each of the following five statements have True/False options:

A. In myocardial cells, as in nerve and skeletal muscle cells, K^+ plays a minor role in determining the resting membrane potential.

B. The impulse propagates from the sinus node via five bundles of internodal syncytial cells through the left and right atrial wall to the atrioventricular node.

C. The long absolute refractory period of the ventricular cells, covers the whole shortening phase of the contraction, where all the fast Na^+-channels are voltage-inactivated. As a consequence, no stimulus is sufficient regardless of size.

D. The fast Na^+-influx causes phase 0 of atrial- , ventricular- , and Purkinje- action potentials. The fast Na^+-channels are both voltage- and time-dependent.

E. Noradrenaline activates a-adrenergic constrictor receptors in the coronary vessels, whereas adrenaline activates b-adrenergic vasodilatator receptors.

III. The following five statements have True/False options.

A. WPW-syndrome or Wolf-Parkinson-White block is caused by a short cut through an extra conduction pathway from the atria to the ventricles.

B. Atrial fibrillation is more malignant than ventricular fibrillation.

C. All pacemaker abnormalities arise in the sinus node.

D. Premature beats are also called atrial ectopic beats.

E. Only few cardiac arrhythmias can lead to atrial fibrillation and flutter.

Try to solve the problems before looking up the answers.

Highlights

- *Heart rate is controlled mainly by the autonomic nervous system. Sympathetic stimulation speeds up the sinus node (sinus tachycardia) and vagal activity slows the node (sinus bradycardia).*
- *The autonomic nervous system controls myocardial contraction by varying the Ca^{2+} - permeability of the sarcolemma via hormones and the adenylcyclase system.*
- *Increased concentration of the sympathetic transmitter noradrenaline, and of adrenaline from the adrenal glands, cause increased contractility, increased frequency, increased conduction velocity, and increased irritability of the heart.*
- *Catecholamines and increasing extracellular $[Ca^{2+}]$, raise the cytoplasmic $[Ca^{2+}]$ and thus the developed force of contraction.*
- *Cardiac digitalis glycosides (digoxin) block the Na^+-K^+-pump. This blockage increases the internal $[Na^+]$ to the extent that less Ca^{2+} is removed from the cell. This - and any - form of elevated cytoplasmic $[Ca^{2+}]$ enhances contractile force. Increased $[K^+]$ in the extracellular fluid reduces the force of contraction, so the heart becomes dilatated and flaccid.*
- *Noradrenaline activates α-adrenergic constrictor receptors in the coronary vessels, whereas adrenaline activates β-adrenergic vasodilatator receptors.*
- *The neurotransmitter acetylcholine, activating muscarinic receptors, and vagal stimulation*

cause reduced contractility (negative inotropic state), reduced frequency (negative chronotropic state), reduced conduction velocity (negative dromotropic state), and reduced irritability (negative bathmotropic state).

- *Changes in blood concentrations of gasses and protons affect the cardiac function directly and indirectly via chemoreceptors.*
- *The long absolute refractory period of the ventricular cells, covers the whole shortening phase of the contraction, where all the fast Na^+-channels are voltage-inactivated. As a consequence, no stimulus is sufficient regardless of size.*
- *The electrocardiogram (ECG) is a surface recording of the electrical field generated in the entire body by the heart.*
- *The sinus node is a minimal muscle mass, and there is no potential difference (wave in the ECG) before the atria depolarise with a P-wave. When the propagating wave is directed towards the electrode (as in lead II) the atrial depolarization will produce a positive P-wave.*
- *The P-waves correspond to the impulse distribution in the atria, and the QRS-complex origin from depolarisation of the strong ventricular myocardium.*
- *The QRS deflections in two of the three standard leads can be drawn graphically in a triangle and their resultant is the mean QRS-axis of the heart.*
- *The T-wave is caused by the spread of repolarization over the ventricles.*
- *The small propagating wave moving away from the electrode at the apex and to the right to reach the right ventricle, is responsible for the small, negative S-wave.*
- *The Adam-Stokes syndrome is a clinical disorder caused by a partial AV-block, with a long P-Q interval and a wide QRS complex in the ECG, suddenly becoming a total bundle block. The condition results in unconsciousness and cramps caused by brain hypoxia and sometimes resulting in universal cramps (grand mal) due to violent activity in the motor cortex. Disease processes in the AV node and the bundle of His elicit the Adam-Stokes syndrome.*
- *Ventricular tachycardia is defined as three or more ventricular beats occurring at a rate of 120 beats per min or more.*
- *Ventricular fibrillation is an extremely rapid ventricular activation without pumping effect. Electrical defibrillation is the only effective therapy.*
- *Vulnerable period is a dangerous period in cardiac cycle represented in the ECG as the downslope of the T-wave. Electrical conversion (an electrical shock) given during this period causes in itself ventricular fibrillation.*

Further Reading

Cardiovascular Reviews & Reports. Monthly journal published by Le Jacq Communications Inc, 777 West Putnam Av., Greenwich CT, 06830, USA.

Kastor, JA (1994) *Arrhythmias*. Philadelphia: W.B.Saunders Co.

Surawicz, B (1995) *Electrophysiological Basis of ECG and Cardiac Arrhythmias*. Baltimore: Williams and Wilkins.

Return to content

Return to chapter 11

Chapter 12

Blood Flow, Distribution And Shock

Study Objectives

- To *define* concepts such as anaphylactic shock, bloodflow, hydrostatic indifference point, hypotension, mean circulatory equilibrium pressure, mean transit time, and shock.
- To *describe* the principle of mass balance (Fick principle) for cardiac output determination, the dilution principle (tracer or indicator bolus), the isotope-wash-out-method, the venous occlusion plethysmography, and the mean transit time method.
- To *draw* an indicator dilution curve, an isotope-wash-out-curve and a plethysmography curve with flow determinations.
- To *calculate* one variable when relevant information is given.
- To *explain* the control of bloodflow to the brain the myocardium, the kidneys, the muscles, the gastrointestinal channel, the skin, and the foetus. To explain the compensatory reactions to shock. To explain the cerebral ischaemic response.
- To *use* the concepts in problem solving and case histories.

Principles

- **The law of conservation of matter** (see Chapter 8). This principle is used to measure bloodflow.
- **Fick´s principle** for determination of cardiac output (see Chapter 10).
- **Poiseuille´s law**. When the radius of a vascular bed is doubled, its bloodflow may increase by as much as 16 times. This is because of Poiseuille´s law (see Eq. 8-3).

Definitions

- **Anaphylactic shock** *(anaphylaxis)* refers to a severe allergic disorder in which the cardiac output and the mean arterial pressure fall rapidly and drastically (see Chapter 32).
- **Hydrostatic indifference point** is the point in the cardiovascular system, in which the pressure does not change with change of body position.
- **Hypotension.** Severe hypotension refers to a condition with a systolic blood pressure below 75 mmHg (10 kPa).
- **Mean circulatory equilibrium pressure** is a pressure of 1 kPa measurable in all divisions of the circulatory system just after cardiac arrest.
- **Mean transit time** for indicator particles in a system is equal to the sum of all transit times for all single particles divided with their number.

- **Shock** is a clinical condition characterized by a gradual fall in arterial blood pressure and rapid heart rate. Respiration is also rapid and the skin is moist, pale or bluish-grey.

- **Vasovagal syncope** or ***emotional fainting*** is a condition, where the fainting is caused by a strong emotional activation of the parasympathetic nervous system via the hypothalamus with bradycardia, vasodilatation and decreasing arterial pressure.

Essentials

This paragraph deals with *1. The coronary bloodflow, 2. The regulation of coronary bloodflow, 3. Brain bloodflow, 4. Skin & fat bloodflow, 5. The splancnic circulation, 6. The foetal circulation, 7. Fick's principle, 8. The dilution principle, 9. Clearance, 10. The isotope-wash-out-method, 11. The mean transit time, 12. Vascular pressure reference (supine to standing).*

1. The coronary bloodflow

The myocardial metabolism is an exclusively *aerobic* process under normal conditions. It depends on oxidative phosphorylation in order to re-synthesise ATP. The O_2 needs of the myocardium are therefore great, even at rest. Exercise can increase the needs *six-fold;* however, the myocardium cannot extract a greater fraction of the O_2 delivered, since the myocardial O_2 extraction is already close to maximum at rest. Thus the coronary bloodflow must rise importantly during exercise in order to deliver the O_2 needed.

Two main coronary arteries arise from the aorta. The *left main coronary artery* (Fig. 12-1) divides into two major branches: The *left anterior descending artery*, which courses down the interventricular groove towards the apex of the heart, and the *left circumflex artery*, which courses leftward and posteriorly in the atrioventricular groove to the postero-lateral wall of the left ventricle.

The *right coronary artery* (Fig. 12-1) arises from the right aortic sinus and courses rightward and posteriorly in the atrioventricular groove to reach the right atrium, and via the *posterior descending artery* to the posterior wall of the left ventricle and the lower part of the interventricular septum. Later the right coronary artery also gives off branches to the *posterolateral wall* of the left ventricle.

This arrangement of coronary vessels exists in *half* of the population in western countries. In 30% of the population the *posterior descending artery* arises from the *right coronary artery*, and the posterior left ventricular branch arises from the *left circumflex artery*. In another 20% of the population the *right coronary artery* is small and supplies only the right atrium and the right ventricle with blood, and all the blood supply to the left ventricle comes from the *left main coronary artery*.

The main arteries run along the epicardial surface and divide several times on the surface of the heart before they send off small penetrating vessels forming a network of intramural arteries, arterioles and capillaries in their way to the endocardium. The myocardial capillaries feed into a net of intramural venules. They drain eventually into the epicardial collecting veins. *Right ventricular* venous blood drains into the right atrium. *Left ventricular* venous blood drains into the *coronary sinus* that empties in the right atrium, except for a small blood volume, which drains into the left ventricle. The epicardial coronary vessels contain a preponderance of constrictor receptors called *adrenergic α-receptors*, whereas the intramuscular and endocardial coronary receptors have a preponderance of dilatator receptors

called *adrenergic β-receptors.*

Fig. 12-1: Coronary bloodflow and receptors.

Due to the contraction of the myocardium in systole, the myocardial bloodflow is blocked and the heart receives its nutrition in the diastolic period (Fig. 10-7). The coronary bloodflow is *phasic*.

2. The regulation of coronary bloodflow

The coronary bloodflow is described before in Chapter 9 (paragraph 3) and in relation to Fig. 10-7.

The coronary bloodflow is mainly controlled by *local metabolic autoregulation,* and sympathetic stimulation does not always cause significant vasoconstriction. Accordingly, a moderate decrease in arterial blood pressure down to 9.3 kPa (70 mmHg) does not significantly reduce the bloodflow through the myocardium.

Unlike skeletal muscle tissue, the myocardium cannot function anaerobically for extended periods by building up an *oxygen debt.* Thus, oxidative ATP synthesis must continuously match ATP utilisation in the heart. At rest the heart produces 70% of its ATP from oxidation of fatty acids and 30% from oxidation of carbohydrates.

During exercise with lactate production by skeletal muscles, this lactate becomes an important substrate for the myocardial metabolism, entering the tricarboxylic acid cycle after conversion to pyruvate.

Catheterisation of the venous sinus of the heart in healthy subjects at rest reveals a venous haemoglobin saturation fraction of 0.30. Hence, 0.7 parts of the haemoglobin concentration of the venous blood is desaturated. Thus, arterial blood with a normal oxygen concentration (C_{ao2} of 200 ml per l) liberates (200 × 0.7) or 140 ml of oxygen per l to the myocardium. Variations in the arteriovenous O_2 content difference at the *steep* part of the O_2-haemoglobin dissociation curve can only change the myocardial O_2 tension modestly. The extremely high O_2 content difference of the heart at rest implies that a *rise in coronary bloodflow* must be the main source of extra O_2 to the heart during exercise.

Most of the blood entering the coronary circulation is delivered during the diastolic phase. This is because the myocardial tissue pressure increases during systole, and the contraction squeezes the blood/myocardium - in particular in the subendocardial layer. Therefore, the systolic bloodflow through the inner layer of the left ventricular wall approach zero. The duration of each diasole is reduced with increasing heart rate, so the increased oxygen demands during exercise calls for a higher coronary artery pressure in diastole in order to secure the necessary bloodflow.

3. Brain bloodflow

The blood reaches the brain through the internal carotid and the vertebral arteries. The dominant control of cerebral bloodflow (CBF) is metabolic autoregulation but also a pressure dependent myogenic autoregulation is present. The smooth muscle walls of the small cerebral arteries respond immediately to changes in the transmural pressure gradient. Hereby, the CBF is maintained constant despite changes in systolic blood pressure between 80 and 160 mmHg (10.7-21.4 kPa). The small brain vessels are metabolically regulated. Increased P_{aCO_2} and reduced P_{aO_2} dilatates brain vessels and increase CBF. CO_2 (not H^+) passes the blood-brain

barrier easily. The mean arterial pressure can double without any appreciable rise in CBF. A neuropeptide released in response to transient hypotension *(calcitonin gene-related peptide)* is probably involved in the autoregulation.

The vertebral arteries join to form the basilar artery, which forms the circle of Willis together with blood from the internal carotids. When brain arterioles dilatate the CBF increases, and since the brain tissue within the cranium is relatively incompressible, the venous outflow must balance.

CBF is normally 55 FU in humans at rest. One FU is one ml of blood per min per 100 g of brain tissue. With a normal brain weight of 1300 g this value corresponds to a total of (55×13) = *715 ml* of blood per min. This resting CBF and the oxygen uptake of the brain can double during *cerebral activity* and triple in active brain regions during an *epileptic attack* (Chapter 7).

The sympathetic nervous system plays a secondary role for the CBF. Some brain vessels contract by *sympathetic stimulation.* This neurogenic control only concerns the *larger* cerebral arterioles.

Some degree of autoregulation is found in many other organs including the *skeletal muscle mass, the splancnic area, and the kidneys.*

4. Skin & fat bloodflow

Blood flows through the skin and subcutaneous tissues in order to *nourish* the cells, and to *regulate* shell temperature. Blood flows much faster through the arteriovenous anastomoses in the skin of the face, the fingers and toes in a cold environment. The sympathetic activity constricts the metarterioles that lead to the skin, so the blood bypasses the cutaneous circulation. Hereby, the skin bloodflow can fall from about 5 FU´s and approach zero. The heat content of the blood returns to the body core, which helps to maintain the core temperature.

In a warm environment, the sympathetic tone is minimal and the arterioles dilatate, so that the skin perfusion can rise to perhaps 70 FU, and much energy is given off to the atmosphere. *Psychological influence* can cause one to blush or to have a white face, by *changing* α*-adrenergic* constrictor tone and through the effect of local, vasoactive substances normally found in the skin. When a large fat combustion occurs (during hunger and distance running), the fat bloodflow can increase from 3 to 20 FU´s. Cold and warm environments alter the fat perfusion just like the skin perfusion (Chapter 21). The sympathetic regulation of the arteriolar tone in fat and skin tissue is also similar. The sympathetic change in tone is not related to the classical baroreceptors.

5. The splancnic circulation

The splancnic area is drained (1.5 l per min) via the hepatic veins at rest, so *all blood* passes through the liver. The liver receives more than one litre of blood from the portal vein and less than 0.5 l from the hepatic artery each minute. A special characteristic for the splancnic circulation is that two large capillary beds are partially in series with one another forming a *portal system*. The splancnic perfusion increases after meals, and decreases during fasting and duration exercise. The sympathetic nervous system has a tonic activity on splancnic vessels via a-adrenergic nerve fibres. Vagal fibres dilatate the splancnic vessels. Haemorrhagic shock can elicit a fatal splancnic hypoxia.

6. The foetal circulation

The foetus depends completely on the mother and her placenta. The *placental barrier* can be passed by low molecular substances (nutrients, gasses and waste) by diffusion.

The *foetal* haemoglobin (F) has a sigmoid dissociation curve, which is shifted to the left relative to haemoglobin A for adult, because haemoglobin-F has a greater affinity for oxygen than does adult haemoglobin (Chapter 15). Haemoglobin F is not affected by 2,3-DPG. Foetal blood has a high haemoglobin concentration (200 g l^{-1}), so the foetus takes up large amounts of O_2 in placenta. This occurs even at low P_{O_2}, and the maximal value in the placental blood is only 6.7 kPa or 50 mmHg. Often the foetus achieves an arterial oxygen concentration similar to that of the mother.

Steroid hormones, maternal thyroid hormones, and catecholamines cross the placental barrier to the foetus. Peptide hormones cannot traverse the placental barrier, except small peptides such as thyrotropin-releasing hormones (TRH) and antidiuretic hormone (ADH).

Foetal insulin contributes to anabolism and lipid storage. Human chorionic somatomammotropin, prolactin, and IGF-2 are the most important growth factors during foetal life. Foetal parathyroid hormone stimulates the transport of Ca^{2+} to the foetus.

The *maternal blood* rich in O_2 and nutrients is injected into the *intervillous spaces* of the placenta via spiral arteries, and returns with CO_2 and waste to the mother via veins draining to the uterine veins. The *chorionic villi* dip into the internal sinuses of the placenta. The exchange of nutrients, metabolic waste and gasses across the placental barrier occurs by diffusion. Foetal blood rich in O_2 and nutrients returns to the foetus from the placenta in the umbilical veins.

The placenta has a diameter of 16-20 cm and the placental barrier has an area of about 10 m^2. The blood flowing in the umbilical veins continues in the *ductus venosus* or the blood enters the foetal liver and then all blood is gathered in the inferior caval vein.

The 80%-saturated blood from the umbilical veins flows into the foetus. The saturation is reduced to 2/3 (67%) when passing through the oval communication between the right and left atria (*foramen ovale*) to reach the left ventricle (Fig. 12-2). The blood in the left ventricle is mixed with desaturated blood from the lungs, whereby the oxygen saturation is reduced to about 65% in the blood passing to the upper-body foetal organs (brain and heart).

The blood in the right ventricle (a mixture of blood mostly from the superior caval vein, but also the coronary sinus and the inferior caval vein to some extent) is only half-saturated. The foetal pulmonary vascular resistance is high due to the compressed inactive lungs (Fig. 12-2). Thus, the major part of the right ventricular output to the pulmonary artery bypasses the pulmonary circulation and flows through a foetal channel (*ductus arteriosus*) between the pulmonary trunk and into the descending aorta (Fig. 12-2).

When the major part of the blood passes the ductus arteriosus and joins the left ventricular output, the resulting oxygen saturation is around 60% in the blood of the descending aorta reaching the lower part of the body and back to the placenta in the umbilical arteries. This shunt delivers well-saturated nutritive blood to upper-body foetal organs (brain and heart). Venous drainage from these essential organs returns to the foetal heart in the superior vena Cava. The right ventricle ejects the venous return.

Much of the blood flowing in the descending aorta of the foetus is directed toward the placenta, so venous drainage from all organs is shunted toward the placenta, where wastes are eliminated from foetal blood, whereas O_2 and nutrients are acquired. Maternal hypoxia and

reduced venous return or pressure on the umbilical vessels during birth reduces the oxygen supply to the foetus, which is reflected as bradycardia.

At birth, P_{CO_2} increases, and the first breath reduces the intrathoracic pressure in the new-born, so placental blood is sucked into the baby (placental transfusion). When the bloodflow through the umbilical vein ceases, the muscular sphincter of the ductus venosus contracts. Massive sensory stimuli of the baby caused by labour and delivery, cutaneous cooling after delivery, the falling P_{AO2} and rising P_{ACO2} without the placenta, and withdrawal of a placenta-produced respiratory inhibitor all adds up in activation of the respiratory centre and maintaining breathing in the newborn. The newly established air-liquid interface reduces pulmonary surface tension, which eases the lung expansion.

Distension of the lungs with air also distends pulmonary vessels, so pulmonary vascular resistance (*PVR*) decreases drastically. Hereby, pulmonary bloodflow increases. As a consequence, bloodflow via ductus arteriosus slows, and pulmonary venous return to the heart increases (Fig. 12-2).

Fig. 12-2: The foetal circulation.

The left atrial pressure increases above that in the right atrium and the inferior vena cava by the newly established decrease in pulmonary vascular resistance, because of the large pulmonary bloodflow to the left atrium.

Occlusion of the umbilical vein reduces the bloodflow to the right atrium, and occlusion of the umbilical arteries increases the resistance to the left ventricular output of blood. The resulting elimination of the pressure gradient across the atria, abruptly closes the valve over the *foramen ovale*, and the septal leaflets fuse within a couple of days.

The low pulmonary vascular resistance reduces the pressure in the pulmonary artery, whereas the aortic pressure rises. This reverses the flow of oxygenated blood through the *ductus arteriosus*. Within minutes after lung expansion, the muscular wall of the ductus arteriosus constricts, and its closure is complete within 10 days.

Failure of the foramen ovale or the ductus arteriosus to close gives rise to two congenital cardiac abnormalities (see later in this Chapter).

Normal arterial blood gas tensions are established by 30 minutes of age. Left atrial pressure increases and foramen ovale closes soon after birth. This *reverses* the blood pressure gradient across the foramen ovale, so now the left atrial pressure exceeds the right. When the umbilical cord is closed, and the placental circulation is thus eliminated, the *TPVR* of the newborn increases. The decrease in pulmonary vascular resistance (*PVR*) and increase in *TPVR* means a great difference in the size of the blood pressures in the aorta and in the pulmonary artery.

In conclusion, the *parallel* foetal circulation is transformed into a *series circulation* in the baby. The foramen ovale and ductus venosus close within 3 days of birth (Fig. 12-2). The *sharp increase* of O_2 content (C_{aO_2}) in the baby's blood is a potent and universal vasodilatator.

The dramatic changes in gas exchange affect cardiopulmonary and vascular regulation, probably via local mediators such as *arachidonic acid* and *prostacyclin*.

7. Fick's principle

According to the *law of conservation of matter*, mass or energy can neither be created nor destroyed (the principle of mass balance). *Adolph Fick* applied natural occurring indicators like

O_2 and CO_2 when measuring cardiac output. Using O_2 as an indicator and the law of mass balance, he claims that the O_2 flux, taken up by the lung blood, plus the venous O_2 flux to the lung, must be equal to the O_2 flux, which leaves the lung in the oxygenated blood. Thereby, Fick proposed that the cardiac output could be calculated according to Eq. 12-1.

A classical example of the usefulness of Fick's principle is to consider the data of a young healthy male at rest. The typical data for such a person are an O_2 uptake of *250* ml STPD per min and an arteriovenous O_2 content difference of *50* ml STPD per l of blood. According to Fick's principle, this male can only satisfy his O_2 demands, if 5 l of blood is oxygenated in his lungs every minute. Thus, a cardiac bloodflow of 5 l per min is his cardiac output.

The oxygen concentration of mixed venous blood is usually obtained through a venous catheter inserted up the median cubital vein, through the subclavian vein, and finally into the right ventricle or pulmonary artery, where the blood is well mixed. Arterial blood is easily obtained from the radial artery (C_{aO2}). The disappearance rate of oxygen from the respired air can be recorded in a metabolic ratemeter as the oxygen uptake.

The principle of mass balance is valid only for a system in *steady state*. Steady state is a state where the indicator is administered at a constant rate, and is neither stored, mobilised, synthesised nor used by the system, and where no shunts are present.

This method has been used to measure a large increase in cardiac output in different patient groups. For example, patients with anaemia have been found to have higher cardiac output at rest.

8. The dilution principle

When an *indicator* bolus (mass or dose of *tracer* in weight or molar units) is instantaneously injected in the right side of the heart, the indicator and blood will mix. The mixture leaves the right ventricle through a *well-mixed outlet,* passes the pulmonary circulation and then returns to the left side of the heart. The indicator concentration during the first passage of any peripheral artery is recorded continuously or by multiple sampling. The resulting curve is shown in a semilog scale (Fig. 12-3). The indicator concentration (in mol/ml of blood) reaches a peak and then decreases in a few seconds, before it again rises due to indicator recirculation with the blood (Fig. 12-3). The first decrease in concentration is assumed to be mono-exponential. Hence it is easy to extrapolate to the concentration zero, and read the so-called *first passage time,* T_1. In this case the T_1 is 9 s (Fig. 12-3). The mean concentration (c mol/ml) of indicator in the period T_1 seconds is determined by planimetry.

The average amount of indicator (in moles) leaving the left ventricle per second in one ml of blood is c, hence c is given in mol/ml of blood. The volume of blood (V) in which the indicator dose is distributed is dose/c. Since blood carries only c mol of tracer in each ml, the heart needs at least a bloodflow of V ml (dose in mol/ c) in order to carry the entire dose through the aortic orifice in T_1 (9) seconds. Accordingly, the cardiac output per second is dose/($c*T_1$). The product ($c*T_1$) is the area under the curve (Fig. 24-1). Thus the *dose/area ratio* must be equal to the bloodflow (ml/second) leaving the left ventricle.

Put simply, the bloodflow (cardiac output in ml of blood per s or more convenient per min) can be measured by dividing the *dose* of indicator injected upstream by the *area* under the downstream concentration curve.

Fig. 12-3: The indicator dilution principle.

The bloodflow equation is also called the *dose/area equation.*

An attractive choice of indicator is to use *cold saline*, of known temperature and volume. A flexible catheter, with a thermistor located at its tip, and an opening through which cold saline can pass, is used. The catheter tip is advanced to the pulmonary artery, while the opening supplies the right atrium with saline. The thermistor records the downstream alterations in temperature as the saline bolus passes. This is the *thermodilution technique*. This technique can be frequently repeated without having harmful effects. Moreover, there is negligible recirculation, and the method spares the patient the ordeal of an arterial puncture.

This method is widely used. For example, interesting indicator dilution studies have shown the pump effect of external cardiac massage to be modest.

9. Clearance

Clearance is a theoretical tool for estimating bloodflow in the kidney and other organs. *Clearance* is the volume of blood plasma, which is *totally cleared* each minute of a given indicator by a specific organ (eg, renal clearance). The extraction (E) is the fraction of substance, which is extracted from the total amount transported to the organ per minute (Eq. 12-3).

Clearance for para-amino hippuric acid (PAH) at low plasma concentrations is a measure of the *renal plasma flow* (RPF – see Ch. 25). The high hepatic extraction of bromsulphalein or of indocyanine is used to estimate the *splancnic perfusion.*

10. The isotope-wash-out-method

A lipid-soluble indicator, such as ^{133}Xenon dissolved in saline, is injected in the *tibialis anterior muscle* (* in Fig. 12-4 and Eq. 12-4). At steady state, the tracer concentration in the venous blood (C_v) is assumed to be the *average* blood concentration, and C_{tis} the *mean* tissue concentration.

Fig. 12-4: Isotope (^{133}Xe) wash out from the gastrocnemius muscle before, during and after walk on a treadmill. Upper curve is when the femoral artery is occluded - lower curve is from the healthy leg.

The *fractional fall* in the mean tissue concentration of Xenon (C_{tis}) per time unit (dt) is *constant* during the whole elimination period (a rate constant = $\ln 2/T_{½}$). The flow/W_{tis} (weight of tissue) is a perfusion coefficient in FU (ml of blood per min and per 100 g tissue). The fall in mean tissue concentration per time unit (ratio dC_{tis}/C_{tis}) is measurable as $T_{½}$ on the skin surface at the Xenon deposit in muscle tissue with a scintillation detector (Fig. 12-4 and Eq. 12-4).

The method (see Eq. 12-4) is used clinically to detect peripheral vascular diseases. An example is *intermittent claudication* that refers to constricting pain arising during activity of any muscle group but most commonly in the calf muscles. The hypoxic pain and cramp appear after having walked a certain distance and is promptly relieved by rest. The cause is *femoral occlusion* due to arteriosclerosis with insufficient local bloodflow and *ischaemic hypoxia* (Fig. 12-4).

11. The mean transit time

The *mean transit time* (t_{mean}) for indicator particles in a system with the volume, V, is equal to the sum of all transit times for all single particles divided with their number.

This concept is used in a wide variety of indicator methods (Eq. 12-5).

By means of intravascular catheters it is possible to measure the *partial* circulation time through most parts of the circulation. For a healthy adult at rest the normal ranges include the following: arm-ear 8-12 s, arm-lung 5-7s, and lung-ear 3-5 s.

12. Vascular pressure reference

The heart is not always the correct reference point for blood pressure measurements. The elastic properties of the vascular tree differ throughout the body.

Actually, the point in which the pressure does not change with change of body position is approximately 5 cm beneath the diaphragm during expiratory relaxation (Fig. 12-5). This is called the *hydrostatic indifference point* (HIP). Above this horizontal level, all vascular pressures are lower in the erect than in the recumbent position. The subatmospheric intrathoracic pressure counteracts venous collapse, so the intrathoracic veins remain open and the atrial pressures are zero in the erect position. The veins of the neck and face are collapsed. The venous sinuses of the brain are kept open by attachment to the surrounding tissues, and their pressures are around -1.3 kPa (-10 mmHg) in the erect position.

Fig. 12-5: The hydrostatic indifference point (HIP) in an adult male. The subject changes position from recumbent to erect.

HIP must not be mixed up with the *mean circulatory equilibration pressure*, which is a pressure of 1 kPa (6 mmHg) measurable in all divisions of the circulatory system just after cardiac arrest. This is also called the *mean circulatory filling pressure*, because it is a determinant of the venous return.

When a supine person arise, his TPVR increases, the systolic blood pressure falls andthe diastolic blood pressure rises. Thus, the pressure amplitude falls, but the mean arterial pressure is unchanged. The stroke volume is reduced more than the heart rate rises, so the cardiac output will decrease when attending the standing position.

An elegant way of studying circulatory consequences of standing is by use of *lower-body-negative-pressure* (LBNP). LBNP applied to a recumbent subject simulates the circulatory effects of standing.

The venous return to the heart is dependent upon the body position, and upon the total blood volume. The venous return is also dependent upon the venous compliance and upon the sympathetic tone in the venous system and in the arteriolar system.

When a person is located on a tilt table in a horizontal position, his blood pressures in a superficial vein on the feet is approximately 1.6 kPa (12 mmHg) and in the femoral veins 0.8 kPa (6 mmHg). When the person is turned upright towards the vertical plane, the venous pressure increases by the hydrostatic column up to the *hydrostatic indifference point* (HIP) just below the heart as long as he is not standing and using his skeletal muscle pump. If the tilt table is turned, so the head of the person is downward (*Trendelenburg position*), then the venous pressure increases in neck and head. The Trendelenburg position is rational during neck and head surgery. With the head upward the patient risks *air* embolism, if blood vessels are cut during neck and head surgery due to the subatmospheric pressure in the vessels.

Pathophysiology

This paragraph deals with *1. Shock* and *2. Congenital heart disease*.

1. Shock

Shock is defined as a clinical condition characterised by a gradual fall in arterial pressure and a rapid heart rate. Respiration is also rapid and the skin is pale, moist and grey. The general circulatory insufficiency causes the bloodflow to vital tissues to be inadequate, so delivery of oxygen and other nutrients as well as elimination of waste products is insufficient.

In principle the circulatory insufficiency can be caused by disorders in the heart (cardiac insufficiency with imminent or manifest *cardiogenic shock*) or in the vessels (vascular insufficiency developing into vascular shock).

The *cardiogenic shock* can be caused by restricted ventricular filling (bi- or tricuspidal stenosis, pericardial fibrosis, or cardiac *tamponade*); the cause can also be myocardial disorders (infarctions, myocarditis etc) or restricted/ineffective ventricular ejection in in cases wwith semilunar stenosis/insufficiency or shunts.

The *vascularly generated shock* is caused by loss of blood or other fluids (absolute hypovolaemia) or by vasodilatation (relative hypovolaemia). *Absolute hypovolaemia* is caused by blood loss, plasma loss (burns or other denuding conditions, ascites, hydrothorax etc) or dehydration (water deprivation, severe diarhoea or vomiting, excessive sweating, intestinal obstruction with luminal fluid accumulation, urinary loss of proteins/salt/water, excessive use of diuretics, hypoaldosteronism etc). *Relative hypovolaemia*, sometimes with universal vasodilatation, is released by endotoxins (septic shock from viral or bacterial infections), anaphylactic shock (see Ch. 32) or by a neurogenic vasodilatation (neurogenic shock by severe pains or stress, anaesthetics or brain stem lesions close to the vasoconstrictor centre).

The *reduced delivery* of oxygen and nutrients to virtually all cells of the body, is consequential: The mitochondria synthesise less ATP, the Na^+-K^+-pump operates insufficiently, the metabolic processing of nutrients is depressed which profoundly depresses muscular contractions, and finally digestive enzymes destruct the damaged cells. Glucose transport across the cell membranes in the liver and in the skeletal muscles is depressed including a severe inhibition of the actions of insulin and other hormones (Chapters 26 and 30). During progressive shock the metabolism is reduced and thus the heat energy, so the *body temperature* tends to decrease, if the patient is not kept warm.

Compensatory mechanisms in shock are called *negative feedback* mechanisms, because they operate to counteract the fall in blood pressure. Baroreceptor responses and many hormonal control systems, that tend to raise the falling blood pressure, are examples of negative feedback (Fig. 12-6). The gain of a feedback system is defined as the ratio of the response to the stimulus itself.

Decompensatory mechanisms exaggerate the primary fall in blood pressure. This is called *positive feedback*. A positive feedback mechanism can lead to a *vicious cycle* and death, if its gain is above one. Two examples with ischaemic brainstem depression and cardiac depression are shown in Fig. 12-6.

Shock is divided into 3 stages by severity:

1A. *Mild shock* is a condition, where compensatory reactions can cure the patient without external help. A latent shock is produced when a healthy blood donor delivers more than the usual 500 ml of blood for transfusion, but the volume is often replaced within an hour. A

number of negative feedback mechanisms oppose the induced changes of shock. The fall in MAP and pulse pressure reduces the stimulation of the high-pressure baroreceptors in the carotid sinus and the aortic arch. The negative stimulation of the cardiovascular control centres in the brainstem enhances the sympathetic tone (and reduces the vagal tone) leading to increased heart rate and contractility as well as to arteriolar and venous constriction mainly in the skin, skeletal muscles and the splancnic area. The bloodflow favours the brain and the heart as long as possible.

An array of other compensatory reactions are given in Fig. 12-6: Increased vascular permeability, reduced capillary pressure with *autotransfusion* from the interstitial fluid, thirst and drinking followed by absorption of fluid from the gastrointestinal tract, and release of powerful vasoconstrictors such as adrenaline, angiotensin II, vasopressin etc.

Catecholamines and enkephalins are released from chromaffine granules in the adrenal medulla. Catecholamines increase the heart rate and the cardiac output by stimulation of the adrenergic b_1-receptors in the myocardium. Catecholamines constrict vessels all over the body by stimulating a_1-receptors located on the surface of vascular smooth muscles.

ADH (*vasopressin*) is secreted from the posterior pituitary gland in response to shock, because the *sinoaortic* baroreceptors are under-stimulated. Vasopressin is a modest vasoconstrictor and a strong antidiuretic hormone. The increased ADH secretion causes increased fluid reabsorption by the kidneys and restores blood pressure and volume.

Renin is secreted from the juxtaglomerular apparatus, when blood pressure and renal perfusion falls drastically. Renin acts on the plasma protein, angiotensinogen, to form inactive angiotensin I, which is transformed to the powerful vasoconstrictor, angiotensin II by angiotensin converting enzyme, ACE. The most likely trigger of the *renin-angiotensin-aldosterone cascade* is described in Chapter 24. - The rise in normal plasma-$[K^+]$ due to the ischaemia of shock also releases aldosterone.

ACTH and β-endorphins are released into the blood from the anterior pituitary gland Chapter 26) in response to haemorrhage or other forms of stress. ACTH and endorphins both exaggerates and restricts the development of shock. These opioids depress the brainstem control centres that normally mediate autonomic responses to stress. Hence, naloxone (an opioid antagonist) improves the circulation and increases the rate of survival from life-threatening shock. - On the other hand, ACTH has a small aldosterone and a strong cortisol stimulating effect.

Initially, the bleeding patient suffers from *hypercoagulability*. Thromboxane A_2 (TxA_2) aggregates thrombocytes, and the aggregate releases more TxA_2. This positive feedback prolongs the clotting tendency. In this phase anticoagulants (heparin) reduce the mortality from shock.

Fig. 12-6: Development of shock conditions. Effects, effectors and reactions are shown.

1B. *Serious shock* leads to *myocardial damage*, because the arterial pressure is too low to secure a coronary bloodflow adequate for nutrition. Myocardial contractility is depressed, and the ventricular function curve shifts to the right (Fig. 10-5B).

Loss of more than 35% of the total blood volume of a healthy person is a threat, if the loss is unaided by blood transfusion. An arterial blood pressure below 8 kPa, where there is no additional baroreceptor response can stimulate the chemoreceptors of the carotid body and

increase ventilation. The sucking effect of the low inspiratory pressure improves the venous return to the heart.

Cerebral ischaemia is consequential at arterial pressures below 5 kPa. The cerebral hypoxia elicts a generalised and powerful sympathetic stimulation with a pronounced arteriolar-and venous-constriction. Further hypoxia in the brainstem activates the vagal centres resulting in bradycardia.

***1C**. Irreversible shock* is a terminal condition, where all therapy is frustraneous. Nothing can save the patient. The progressive deterioration becomes irreversible at a blood loss of more than 50% of the total blood volume. The drastic fall in arterial blood pressure reduces the renal glomerular filtration pressure below the critical level, so filtration is diminished or abolished, leading to abolish urine output (anuria). The low cardiac output and bloodflow result in stagnant hypoxia of all mitochondria. Hypoxia increases lactic acid liberation. Renal failure with tubular necrosis prevents excretion of excess H^+. The high H^+-concentration further depresses the myocardium, reduces blood pressure and thus the tissue bloodflow. This aggravates the metabolic acidosis - a classical *vicious* cycle (Fig. 12-6).

In the later stage of haemorrhagic shock, there is fibrinolysis and prolonged coagulation time (*hypocoagulability*). Hence, heparin therapy can be lethal.

The phagocytic activity of the reticulo-endothelial system (RES) is depressed during shock. Endotoxins constantly enter the blood from the bacterial, intestinal flora of a healthy person. The macrophages of the RES (Chapter 32) normally inactivate these endotoxins and release mediators such as hydrolases, proteases, oxygen free radicals, coagulation factors, prostaglandins, thromboxanes and leucotrienes. Some of these mediators modulate the temperature control and hormone secretion.

Following loss of half the total blood volume the shock patient must have lost about 50% of his circulating macrophages, and control substances modulating the phagocytic activity of RES. The depressed defence mechanisms in RES result in an endotoxic shock which aggravates the haemodynamic shock - a vicious cycle.

The patient loose consciousness and falls into a state of *stupor* or of *coma*.

A severe shock becomes irreversible, when the high-energy phosphate stores of the liver and heart are depleted. All of the creatine phosphate is degraded and almost all ATP has been degraded to ADP, AMP and eventually to the even more efficient vasodilatator, *adenosine*. Adenosine diffuses out of the cells and into the circulation, where it is converted into hypoxanthine and uric acid, a substance that cannot re-enter the cells. Cellular depletion of high-energy phosphate is probably causing the final state of irreversibility.

The cerebral bloodflow is now so low that the function of the cardiovascular brainstem centres is depressed. The loss of sympathetic tone leads to cardiac depression with terminal bradyardia and vasodilatation with falling peripheral resistance. The fall in arterial pressure intensifies the damage, and a vicious cycle is established.

Two types of shock deserve special consideration:

Anaphylactic shock (anaphylaxis with relative hypovolaemia) is a severe allergic disorder in which the cardiac output and the mean arterial pressure fall rapidly and drastically due to relative hypovolaemia. As soon as an antigen to which the patent is sensitive, has entered the blood the *antigen-antibody reaction* (Chapter 32) triggers release of histamine from basophilic cells in the blood and mast cells in the tissues. Histamine dilatates arterioles and most

peripheral vessels. This results in falling arterial pressure and increased capillary permeability with rapid loss of plasma water into the interstitial fluid.

Septic shock (relative hypovolaemia). Septic shock or blood poisoning is a widespread bloodborn bacterial infection - often life threatening. Examples are gas gangrene bacilli spreading from a gangrenous limb, colon bacilli with endotoxin spreading into the blood from infected kidneys (pyelonephritis), and *fulminant peritonitis* due to acute abdominal disease. Frequent causes of fulminant peritonitis are rupture of the infected gut or the uterus, and rupture of the uterine tube due to extrauterine pregnancy. Septic shock is characterized by tremendously high fever, high cardiac output, marked vasodilatation, red cell agglutination, disseminated intravascular coagulation withmicroclots spread all over the circulatory system. When the clotting factors are used up, internal haemorrhages occur. Endotoxins produce vasodilatation, induce synthesis of nitric oxide (NO) synthase in the vascular smooth musculature. Overproduction of NO may contribute to the vasodilatation and the depressed myocardial contractility found in septic shock. The high cardiac output is due to a high stroke volume and a high heart rate. The diastolicpressure is low and the systolic pressure is high until endotoxins begin to inhibit myocardial contractility seriously. Now the condition is in a vicious cycle, which is often fatal.

***Therapy keypoints are**:*

First of all the *cause* has to be established in order to give the appropriate therapy.

1. *Head-down position* (placing the patient's head below the level of the heart) is the immediate therapy of haemorrhagic and neurogenic shock.

2. *Haemorrhagic arrest* (closing the abdominal aorta with the pressure of a fist, or blocking the bleeding from an artery with a finger) is often life saving when applied without unnecessary delay.

3. *Replacement transfusions*. The best possible therapy of haemorrhagic shock is whole blood transfusion. The best treatment of shock caused by plasma loss is plasma transfusion, and the best therapy of dehydration shock is transfusion with the appropriate solution of electrolytes.

4. *Oxygen breathing* is always helpful in shock with insufficient delivery of oxygen.

5. *Sympathomimetic drugs* (noradrenaline, adrenaline etc) are often beneficial in neurogenic and anaphylactic shock. They are seldom useful in haemorrhagic shock, where the sympathetic nervous system is already activated to its maximum.

2. Congenital heart disease

On a global scale approximately 1% of live births result in *congenital heart disease*. The mother may have suffered from rubella infection, abuse of drugs or alcohol, exposure to influential radiation or other factors causing genetic or chromosomal abnormalities.

Two pathophysiological phenomena occur: *Right-to-left shunting of blood* results in cyanosis, clubbing of fingers, reduced growth in children, *exertion syncope*, and paradoxical emboli from veins to systemic arteries. Children with Steno-Fallot´s tetralogy (Fig. 12-7) use *squatting*. The advantage of squatting for the children is improved cerebral oxygenation as the position reduces the right-to-left shunt.

Left-to-right shunts in the heart result in *pulmonary hypertension*, because of persistently increased pulmonary bloodflow and vascular resistance (the so-called *Eisenmenger* response - frequently caused by ventricular septal defect).

Three classical congenital heart diseases are Steno-Fallot´s tetralogy, Coarctation of the aorta, and persistent ductus arteriosus.

Steno-Fallots tetralogy

The four elements suggested by the name tetralogy are: 1. Ventricular septal defect, 2. Overriding aorta (ie, the aortic orifice is located above the ventricular septal defect and therefore receives blood from both ventricles), 3. Pulmonary stenosis (right ventricular outflow obstruction), and 4. Right ventricular hypertrophy.

The pulmonary stenosis causes a high right ventricular pressure, and as this pressure supersedes the left ventricular pressure there is a right-to-left shunting of blood through the ventricular septal defect. The mixed blood passes through the overriding aorta, and the patient is cyanotic with all the hypoxic consequences described above (eg, exertion syncope, squatting, small stature, finger clubbing and polycythaemia). The children are tired and dyspnoeic, and growth is retarded although they often demonstrate a surprising appetite, because of the enormous cardiac work. Complete surgical correction is possible.

Fig. 12-7: Three common congenital heart disorders.

Coarctation of the aorta

Coarctation is a narrowing of the aorta distal to the insertion of ductus arteriosus (Fig. 12-7), and often associated with stenosis of the aortic orifice (a bicuspid aortic valve). The obstructed aortic bloodflow forces blood through collateral arteries such as the intercostal and periscapular arteries. The high blood pressure may cause nose bleeds and headaches, and the low distal bloodflow may cause claudication and cold legs.

Turbulent bloodflow through the coarctation is often recognized as a forceful systolic murmur even on the back.

Surgical excision of the coarctation with end-to-end anastomosis must be performed in childhood, because a low renal bloodflow, maintained over years, frequently results in irreversible systemic hypertension

Persistent ductus arteriosus (PDA)

In the foetus the ductus arteriosus leads blood into the systemic circulation instead of through the unexpanded lungs. Hereby, the foetal blood is oxygenated during its passage of the placenta. At birth, the expansion of the lungs with atmospheric air triggers contraction and closure of ductus arteriosus by constriction of its muscular wall. Premature babies and children borne by mothers, who suffered from rubella in the first trimester, are often born with PDA.

A *PDA* shunts blood from the aorta to the pulmonary artery throughout the cardiac cycle (Fig. 12-7). This is because the aortic pressure is much higher (150/80 mmHg) than that of the pulmonary artery (40/20 mmHg or 5.3/2.7 kPa).

The condition may be symptom-less for years, but a large shunt increases the work of the left

ventricle and causes left heart failure, which increases the risk of pulmonary congestion and oedema. The person is *easily fatigued* even from moderately strenuous exercise.

The frequency of *infective endocarditis* is increased in PDA. Infective endocarditis commonly occurs on congenitally or rheumatically damaged valves. The endocardium also suffers from jet lesions located on the endocardial surface opposite to a shunt with a high driving pressure. This is common for all types of congenital heart disease apart from atrial septal defect, where the driving shunt pressure is too small to damage the endocardium.

The continuous shunting of turbulent blood causes a continuous machinery murmur best heard below the left clavicle.

Surgical ligation of the duct should be performed as early as possible.

The general trend today is an increasing survival rate of congenital heart disease. Adults living with congenital heart disease for years may present themselves as cardiac arrhythmias resistant to standard therapy or at autopsy following sudden cardiac death.

Terminal heart failure is now managed by heart-lung transplantation.

Equations

- The Fick cardiac output equation states that the cardiac output is calculated from the ratio between alveolar oxygen uptake and arteriovenous oxygen content difference:

  Eq. 12-1: $\dot{Q} = \dot{V}_{O2}/(C_{aO2} - C_{v\,O2})$; [ml STPD*min^{-1}/ml blood*min^{-1}]. The last oxygen concentration is in mixed venous blood.

- The law of mass balance is applied to both bloodflow and oxygen flux in Eq. 14-7 and 14-8. The flow and flux relations implies the following shunt equation:

  Eq. 12-2: $\dot{Q}_{shunt}/\dot{Q}_{total} = (C_{aO2} - C_{c'02})/(C_{v\,O2} - C_{c'02})$. The last oxygen concentration is in pulmonary end-capillary blood.

- Clearance is the volume of blood plasma, which is totally cleared each minute of a given indicator by a specific organ (renal clearance). The extraction (E) is the fraction of substance that is extracted from the total amount transported to the organ per minute.

  Eq. 12-3: $E = \dot{Q}(C_a - C_v)/(\dot{Q} \times C_a)$; $E = (C_a - C_v)/C_a$

- *The isotope-wash-out-method*. A homogenous muscle tissue of the weight, W_{tis}, is presumed. A lipid-soluble indicator such as ^{133}Xenon dissolved in saline, is injected in the tibialis anterior muscle (* in Fig. 12-4). At steady state, the tracer concentration in the venous blood (C_v) is assumed to be the average blood concentration, and C_{tis} the mean tissue concentration. A distribution coefficient is introduced: $C_{tis}/C_v = l$, which is known

for Xenon. The decrease of the mass of indicator in the tissue per time unit (dt) must be equal to the mass supplied (which is zero) minus the mass of indicator leaving the tissue in the venous blood. The principle of mass balance provides the following equation:

$$(W_{tis} \times dC_{tis}) = \text{[mass supplied minus mass eliminated]}.$$

$$(W_{tis} \times dC_{tis}) = (- C_v \times Flow \times dt) \text{ or } W_{tis} \times dC_{tis}/C_{tis} = (- C_v/C_{tis} \times Flow \times dt).$$

$$\text{Eq. 12-4: } dC_{tis}/C_{tis} = - Flow \times dt/(W_{tis} \times 1).$$

The fractional fall in the mean tissue concentration of Xenon (C_{tis}) per time unit (dt) is constant during the whole elimination period (a rate constant = $\ln 2/T_{½}$). $Flow/W_{tis}$ is a perfusion coefficient in FU (ml of blood per min and per 100 g tissue). The ratio dC_{tis}/C_{tis} is measurable as $T_{½}$ on the skin surface above the Xenon deposit in fat or muscle tissue with a scintillation detector.

- The volume equation for a cylindrical system implies that the flow per second ($\dot{Q}_s$) and its volume (V) is related by t_{mean}:

$$\text{Eq. 12-5: } t_{mean} = V/\dot{Q}_s.$$

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. The capillaries have the greatest cross-sectional area of the systemic circulation.

B. The systemic arterioles offer the greatest vascular resistance to bloodflow.

C. The sympathetic regulation of the arteriolar tone in fat and skin tissue depends upon the classical baroreceptors.

D. Increased compliance in the venous system means decreased venous return.

E. The lactate produced by skeletal muscles during exercise is not an important substrate for the myocardial metabolism.

II. Each of the following five statements have True/False options:

A. Vasopressin is a strong antidiuretic hormone and a modest vasoconstrictor.

B. ACTH and b-endorphins are released into the blood from the posterior pituitary gland in response to haemorrhage or other forms of stress.

C. Fulminant peritonitis is a frequent form of acute abdominal disease.

D. With the head of the patient downward, the patient risks air embolism if blood vessels are cut during neck and head surgery.

E. Hypotension and ischaemic hypoxia depresses myocardial contractility.

Case History A

20-year old soldiers, body weight 70 kg, is brought to the field hospital with a profusely bleeding gun wound in the left shoulder. The skin is cold and wet, the arterial pressure is 65/ 40 mmHg, and the heart frequency is 144 beats per min (bpm). There is no diuresis.

The bleeding is stopped by ligation of the bleeding arteries, and blood transfusions are given instantly (2 litres of whole blood with a normal packed cell volume, PCV, of 0.45). Hereby, the arterial pressure increases to a steady level of 105/70 mmHg, and the heart rate is reduced to 100 bpm. The condition of the soldier is clearly improved and his cardiac output is measured to 3.5 l per min.

24 hours later the soldier had a relapse, and his PCV was measured to 0.35. The PCV is corrected for trapped plasma, and assumed representative for the body as a whole (ie whole body haematocrit).

1. Describe the initial cardiovascular events leading to the shock condition.

2. *Describe the effect of the blood transfusions.*

3. *Calculate the order of size of the blood loss.*

4. *Was the blood transfusion therapy sufficient?*

Case History B

A male, 21 years old, is located in the supine body position. He has a cardiac output of 5.4 l per min at rest, a circulating blood volume in the pulmonary circulation of 650 ml and in the systemic circulation of 4750 ml. His total mass of skeletal muscle is 35 kg. The muscular perfusion is 3 ml of blood per min per 100 g of muscle tissue (3 FU), and the mean passage time in the muscular capillaries is 5 s.

1. *Calculate the mean transit time for all red blood cells in the total circulatory system.*

2. *Calculate the mean transit time for the pulmonary circulation only.*

3. *Calculate the total perfusion of the skeletal muscle mass at rest (ml/min) and calculate the bloodflow per second.*

4. *Calculate the functioning capillary volume in the muscular capillaries.*

Case History C

A 64-year old male normally has a body weight of 74 kg, a total blood volume of 5 l and a blood [haemoglobin] of 10 mmol per l (mM). One day he suddenly vomits large quantities of fresh blood. For two days his stools have been tarry. The last weeks have been stressful at work. The patient calls his doctor and the emergency ward at the hospital is alerted. Due to an incompetent local ambulance service, the patient is brought to hospital without delay by taxi.

Here, the mean arterial blood pressure is below 10 kPa (75 mmHg) and falling. The heart rate is above 150 beats per min and rising. The blood [haemoglobin] is 5 mM measured one hour after the first massive blood loss.

The emergency team immediately institute transfusion of blood. The following 8 days the patient receives three transfusions of blood and at least 10 l of physiological saline. On the second day his [haemoglobin] has increased to 7.2 mM, but on the third day it falls again to 5 mM. On the 4th day at the hospital the patient develops high fever (maximum 40.6 ^{o}C), and a broad-spectrum antibiotic program is started without delay. On the 8th day at hospital the patient has normal temperature, but he develops watery swellings of legs and lower abdomen, in spite of pronounced urination. The body weight is now 80 kg.

***1**. What is the most likely cause of the haemorrhage?*

***2**. Estimate the size of his blood loss.*

***3**. Why did the patient develop high fever?*

***4**. Why did the patient accumulate water?*

Case History D

A small girl, borne with ventricular septal defect, pulmonary stenosis, overriding aorta and a right to left shunt through the septal defect is examined with cardiac catheterisation. The oxygen concentrations in her blood are 138, 195, and 220 ml per l in the pulmonary artery, brachial artery and pulmonary veins, respectively. Her oxygen uptake at rest is 164 ml per min, and she has polycythaemia with a blood haemoglobin of 164.2 g per l. The girl has a 30% right-to-left shunt through the septal defect and directly to the overriding aorta.

1.What is the diagnosis?

2. Calculate the oxygen concentration in the pulmonary veins and estimate the saturation degree.

3. Calculate the bloodflow through the lungs.

4. Calculate the total bloodflow through the aorta assuming that the lung bloodflow ($Q°_{lung}$) and the shunt bloodflow ($Q°_{shunt}$) equals the total cardiac output, $Q°_{total}$.

5. Estimate the size of the venous return.

6. Is it likely that this patient develops cyanosis? Calculate the concentration of reduced haemoglobin in mean capillary blood.

Case History E

A male, 18 years old, suspect of congenital heart disease, is examined at the hospital. His body weight is only 60 kg and he has always abstained from exercise. The patient is tall and slim, although he is always hungry and is actually eating more than normal. Cardiac catheterisation is performed with the patient resting in the supine position, and reveals the following: Mixed venous blood from the right ventricle and from the pulmonary artery (C_{vO2}):

160, arterial blood from the aorta (C_{aO2}) 195, and blood from the right atrium 130 ml STPD l^{-1}. The oxygen uptake ($V°_{O2}$) at rest is 310 ml STPD per min.

1. *Calculate the cardiac output (Q°) from the right ventricle.*
2. *Calculate the Q° from the left ventricle.*
3. *Provide arguments for a certain cardiac abnormality, which explains the findings.*
4. *Why has the patient always avoided exercise?*
5. *Why is the patient slim although he is eating a lot and not performing exercise?*

Try to solve the problems before looking up the answers.

Highlights

- *The coronary bloodflow is mainly controlled by local metabolic autoregulation. Accordingly, a moderate decrease in arterial blood pressure down to 9.3 kPa (70 mmHg) does not significantly reduce the bloodflow through the myocardium.*
- *The blood reaches the brain through the internal carotid and the vertebral arteries. The dominant control of cerebral bloodflow (CBF) is autoregulation. The sympathetic nervous system plays a secondary role.*
- *Some degree of autoregulation is found in many other organs including the skeletal muscle mass, the splancnic area, and the kidneys.*
- *CBF is normally 55 FU in humans at rest. One Flow Unit (FU) is one ml of blood per min per 100 g of brain tissue. This resting CBF and the oxygen uptake of the brain can double during cerebral activity and triple in active brain regions during an epileptic attack. Brain vessels are metabolically regulated. Increased P_{aCO_2}, and reduced P_{aO_2} dilatate brain vessels and increase CBF.*
- *Blood flows through the skin and subcutaneous tissues in order to nourish the cells, and to regulate shell temperature. Blood flows much faster through the arteriovenous anastomoses in the skin of the face, the fingers and toes in a cold environment. The sympathetic activity constricts the metarterioles that lead to the skin, so the blood bypasses the cutaneous circulation.*
- *Psychological influence can cause one to blush or to have a white face, by changing α-adrenergic constrictor tone and via the effect of local, vasoactive substances normally found in the skin.*
- *When a large fat combustion is occurring (eg, hunger and distance running), the fat bloodflow can increase from 3 FU to 20 FU. Cold and warm environments alter the fat perfusion just like the skin perfusion.*
- *The placental barrier has an area of 10 m^2, and can be passed by low molecular substances (nutrients, gasses, and waste) by diffusion. The foetal haemoglobin (F) has*

a dissociation curve that is shifted to the left relative to adult haemoglobin, and F is not affected by 2,3-DPG..

- *Foetal blood has a high haemoglobin concentration, so the foetus takes up large amounts of O_2 in placenta.*

- *Much of the blood flowing in the descending aorta of the foetus is directed toward the placenta, so venous drainage from all organs is shunted toward the placenta, where wastes are eliminated from foetal blood, whereas O_2 and nutrients are acquired.*

- *The parallel foetal circulation is transformed into a series circulation in the baby. The foramen ovale and ductus venosus close within 3 days of birth, and the ductus arteriosus closes within 10 days. The sharp increase of O_2 content (C_{aO_2}) in the baby's blood is a potent and universal vasodilatator. The dramatic changes in gas exchange affect cardiopulmonary and vascular regulation, probably via local mediators such as arachidonic acid and prostacyclin.*

- *When a supine person arise, his TPVR increases, the systolic blood pressure falls and the diastolic blood pressure rises. Thus the pulse pressure amplitude falls, the MAP is unchanged. The stroke volume is reduced more than the heart rate rises, so the cardiac output will decrease in the standing position.*

- *The bloodflow (cardiac output in ml of blood per s or more convenient per min) can be measured by dividing the dose of indicator injected upstream by the area under the downstream concentration curve.*

- *Clearance is the volume of blood plasma. Which is totally cleared each minute of a given indicator by a specific organ.*

- *Catecholamines constrict vessels all over the body by stimulating α_1-receptors located on the surface of vascular smooth muscles.*

- *ADH (vasopressin) is secreted from the posterior pituitary gland in response to shock, because the sinoaortic baroreceptors are under-stimulated. Vasopressin is a modest vasoconstrictor and a strong antidiuretic hormone.*

- *The enzyme renin is secreted from the juxtaglomerular apparatus, when blood pressure and renal perfusion falls drastically. Renin acts on the plasma protein, angiotensinogen, to form inactive angiotensin I, which is converted to the powerful vasoconstrictor, angiotensin II by ACE in the lungs.*

- *Angiotensin II is a powerful stimulator of the aldosterone secretion from the renal cortex. Aldosterone promotes the reabsorption of Na^+ and increases the secretion of K^+ and H^+ in the distal tubular system of the kidneys. Water follows by osmosis, so the extracellular volume is increased.*

- *ACTH and β-endorphins are released into the blood from the anterior pituitary gland in response to haemorrhage or other forms of stress.*

- *Septic shock or blood poisoning is a widespread bloodborn bacterial infection - often life-threatening. Examples are gas gangrene bacilli spreading from a gangrenous limb, colon bacilli with endotoxin spreading into the blood from infected kidneys, and fulminant peritonitis due to acute abdominal disease.*

- *Haemorrhagic arrest (haemostasis) is often life saving in shock, when applied without unnecessary delay.*

- *Replacement transfusions. The best possible therapy of haemorrhagic shock is whole blood transfusion, of shock caused by plasma loss plasma transfusion, and of dehydration shock transfusion with the appropriate solution of electrolytes.*

- *Oxygen breathing is helpful in shock with insufficient delivery of oxygen.*

- *Sympathomimetic drugs (noradrenaline, adrenaline etc) are often beneficial in neurogenic and anaphylactic shock. They are seldom useful in haemorrhagic shock.*

Further Reading

Calver, A., J. Collier and P. Vallance. "Nitric oxide and cardiovascular control." *Experimental Physiology* 78: 303-326, 1993.

Return to chapter 12

Return to content

Chapter 13.

Mechanics Of Breathing And Lung Disorders

Study Objectives

- To *define and apply* the law of ideal gasses, gas partial pressures and fractions, solubility coefficients, Poiseuille´s and Laplace´s laws.
- To *describe* flow-limitation in the airways, dynamic airway compression, respiratory work, and surface tension (and how it is affected by pulmonary surfactant). To describe obstructive and restrictive lung disorders respiratory failure, cystic fibrosis, and the respiratory distress syndrome in new-borns.
- To *explain* the function of the lung from its structure, respiratory volumes, normal and abnormal airway resistance, static and dynamic pressure-volume relations including compliance, with their measurement and normal values.
- To *use* these concepts in diagnosis, problem solving and apply them to case histories.

Principles

- *The law of conservation of matter (see Chapter 8). This principle is used to measure physiological volumes and volume rates.*
- *The law of ideal gasses is defined in Eq. 13-1 (late in this Chapter).*
- *Poiseuille´s law is used both in the circulatory and the respiratory system (Eq. 13-3).*
- *The law of Laplace (Eq. 13-4).*
- *Bernouille´s principle (Eq. 13-6).*

Definitions

- **Alveolar ventilation-perfusion ratio** ($\dot{V}_A/\dot{Q}$-ratio) is an estimate of the gas exchange capacity. This key point is the cardinal variable of cardiopulmonary function - see Chapter 14.
- **Apnoea** is a temporary stop in breathing.
- **Asthma** is an inflammatory lung disease characterised functionally by broncho-constriction, hypersecretion and oedema of the bronchial wall - all contributing to obstruction of the airflow.
- **Chronic bronchitis** refers to an inflammatory process in the wall of the bronchioles with excessive production of mucus and sputum from hypertrophic glands. The small airways are narrow, and there is morning cough more than 3 months per year (WHO).

- **Compliance** is an index of expandability of elastic organs and defined as the change in volume per unit change in pressure.
- Emphysema refers to destruction of lung tissue distal to the terminal bronchioles (a lung unit termed a primary lobulus or acinus). There is degenerative loss of radial traction of the bronchial walls.
- Expiratory reserve volume (ERV) is the volume of air, which can be expired following a normal expiration at rest.
- Forced expiratory volume in one second (FEV_1) is the 1-s-volume exhaled with forceful pressure from maximal inspiration. FEV_1 is often expressed in relation to the total forced expiratory volume (FEV).
- Functional residual capacity (FRC = ERV + RV) is the lung volume in the normal end-tidal expiratory position (airways open and relaxed respiratory muscles).
- Obstructive lung disorders are characterised by an abnormally low airflow.
- Restrictive lung disorders are characterised by small lung volumes (low total lung capacity).
- Residual volume (RV) represents the volume of air left in the lungs after a maximal expiration - normally1.2 litre (l) in adult persons.
- Solubility (a or *Bunsen's solubility coefficient*) is the volume of a particular gas (ml at Standard Temperature Pressure Dry, STPD) dissolved per ml of body warm blood at a partial pressure of one atmosphere (101.3 kPa or 760 mmHg). The solubility coefficients for nitrogen, carbon monoxide, oxygen, and carbon dioxide are listed in Box 13-1.
- **Spontaneous pneumothorax** means that the pleura surface suddenly ruptures without known cause. The rupture typically occurs apically, where the mechanical tension is largest due to expansion. The sudden pain is pleuritic (accentuated by inspiration or coughing), there is dyspnoea, drum sounds by percussion of the affected area and no breath sounds by auscultation.
- **Total lung capacity** (TLC) is the total volume of air in the lungs, when they are maximally inflated (RV + VC) - approx. 6 l of air.
- **Vital capacity** (VC) is the largest volume of air that can be exhaled after a maximal inspiration. VC is measured with or without forced expiration. The size is typically around 4 litres, but it depends on age, sex, and height in the healthy individual.

Box 13-1: Solubilities or Bunsen's solubility coefficients (α) for gasses in body-warm blood. The unit for the solubility is ml STPD*(ml of fluid)$^{-1}$*(101.3 kPa)$^{-1}$

Carbon dioxide:	0.52
Carbon monoxide:	0.018
Nitrogen: 0.012	(Water: 0.013; Fat: 0.065)
Oxygen:	0.022

Essentials

This paragraph deals with *1. Air, 2. Lung volumes, 3. Pneumotachography, 4. The lungs, rib cage and compliance, 5. Dynamic airway compression, 6. Dynamic flow-volume loops, 7. Elastic recoil, 8. Airway resistance, 9. Surface tension, 10. Pulmonary defence mechanisms.*

1. Air

Air passes through the nose and mouth further into the airways, where it is warmed, humidified and filtered. From the trachea to the alveoli, there are 23 branching generations of airways. The first 16 (as an average) constitute the *conducting zone*, which is an anatomic dead space, because no gas exchange takes place. The 17-23 generations form the *respiratory zone*. Each generation of branching increases the total cross-sectional area of the airways, but reduces the radius of each airway and the velocity of air flowing through that airway. The exchange effective *respiratory zone* comprises of the respiratory bronchioles, alveolar ducts and alveolar sacs. At this dead end of the airways, there are approximately 300 million alveoli. Each alveolus has a diameter of 100 -300 mm and is surrounded by approximately 1000 capillaries. Each capillary is in contact with several alveoli, so the capillaries present a sheet of blood to the alveolar air for *gas exchange*. The total area between pulmonary capillary blood and alveolar air ranges from 70-140 m^2 in adult humans (increased during exercise) through recruitment of new capillaries in particular in the apical parts of the lungs.

Turbulent flow is the agitated random movement of molecules, which accounts for the sounds heard over the chest during breathing. This flow develops at the branch points of the upper airways even in quiet breathing. Turbulence also develops when constriction, mucus, infection, tumours, or foreign bodies decrease the radius of the airways. Vagal stimulation (by smoke, dust, cold air, and irritants) leads to airway constriction, whereas sympathetic stimulation dilatates the airways.

2. Lung volumes

Lung volumes are measured by *spirometry* (Fig. 13-1). A spirometer consists of a counterbalanced bell, which is connected to a pen writing on a rotating drum. The air-filled bell is inverted over a chamber of water, so an airtight chamber is formed. The bell is counterbalanced so it moves up and down with respiration with minimal resistance (Fig. 13-1). Volume changes can be recorded on volume and time calibrated paper. - If the spirometer includes a CO_2 absorber, the device is called a metabolic ratemeter (construction Benedict-Krogh).

The *residual volume* (RV) represents the volume of air left in the lungs after a maximal expiration (1.2 l in Fig. 13-1). The vital capacity (VC) is the maximum volume of air that can be exhaled after a maximal inspiration (4.8 l in Fig. 13-1). VC has three components. The first is the inspiratory reserve volume (IRV), which is the quantity of air that can be inhaled from a normal end inspiratory position. The second component of VC is the tidal volume (V_T), which is the volume of air inspired and expired with each breath (about 0.5 l at rest).

Fig. 13-1: A healthy person, connected to a spirometer, is performing a vital capacity manoeuvre from maximal inspiration to maximal expiration (RV).

The third component of VC is the expiratory reserve volume (ERV), which is the amount of air that can be exhaled from the lungs from a normal end-tidal expiratory position that is characterised by a relaxed expiratory pause (Fig. 13-1). This is the easiest position to reproduce, and the lung volume in this position is called functional residual capacity (FRC = ERV + RV). The total lung capacity (TLC) is the total volume of air in the lungs, when they are maximally inflated (RV + VC) - approx. 6 l of air.

When the person in Fig. 13-1 exhales with maximal effort, the *forced expiratory volume* in one-second (FEV_1) may be recorded by spirometry.

The forced vital capacity manoeuvre is performed with all the expiratory and accessory

muscles. When we contract our strong expiratory accessory muscles, we generate high airflows at lung volumes near total lung capacity (Fig. 13-6). Just following peak-expiratory flow (PEF) the airflow velocity decreases linearly with volume no matter how hard the subject tries. This is the *effort-independent airflow* (Fig. 13-6) caused by dynamic airway compression (see later).

3. Pneumotachography

A pneumotachograph is a device for measuring airflow. It consists of a respiratory tube with a small resistance (typically a fine network) to airflow (Fig. 13-2). The two chambers separated by the resistance, connects to the differential transducer chambers by thin tubes. – A transducer consists of 2 chambers separated by a membrane, whose position in space reflects the pressure difference. During respiration through the pneumotachograph tube, a small pressure difference (D*P*) is generated across the resistance, and this pressure difference is directly proportional to the laminar airflow across the resistance according to Poiseuille's law: *Resistance* = D*P*/*V*°$_E$.

Fig. 13-2: Pneumotachograms from a healthy person at rest: The upper curve is the classical flow curve and below is the flow integrated to a volume curve with a tidal volume of 500 ml.

4.The lungs, rib cage and compliance

The lung-thoracic wall system consists of two elastic components that work together: The lungs, which behave like a balloon trying to collapse and the thoracic cage trying to expand. The following *three important pressures* influence the elastic properties of these two components:

1. The *barometric pressure* (P_B). The barometric or atmospheric pressure is one atmosphere and is used as reference pressure.

2. The *alveolar pressure* (P_{alv}). - The pressure in the alveoli is equal to the static mouth pressure when there is no airflow (during apnoea with the glottis open). The static mouth pressure (= P_{alv}) depends upon the lung volume (Fig. 13-3).

3. The *intrathoracic pressure* (P_{it}). - The intrathoracic pressure is the pressure in the fluid-filled *pleural space* between the parietal and visceral layers of pleura (Fig. 13-3). The intrapleural fluid reduces the friction between the two layers. The P_{it} can be measured as the pressure in a sensitive balloon, which is passed into the oesophagus. Pressure changes in the intrapleural space equal the oesophageal balloon pressure changes, because the oesophagus traverses the intrapleural space. The P_{it} is subatmospheric due to the opposing directions of the elastic recoil of lungs and thoracic cage (Fig. 14-5).

Fig. 13-3: Transmural, static pressures and lung volumes in a healthy person. – The red compliance curve is for the total system (see later). The blue and green compliance curves are for the lungs and the chest wall, respectively.

Compliance is an index of distensibility of elastic organs and defined as the change in volume per unit change in pressure (d*V*/d*P*). The following is a description of the measurement of the combined lung plus rib cage compliance, the lung compliance and the rib cage compliance alone.

The combined lung plus rib cage compliance

The subject expires completely to residual capacity (RV in Fig. 13-3), and then inspires a measured volume of air from a spirometer. The subject then relaxes the respiratory muscles during apnoea while the glottis is open at each volume. The alveolar pressure gradient to the ambient air (P_{alv} - P_B) can then be measured as the *mouth pressure* by a manometer (Fig. 13-3). The procedure is repeated by varying the inspired lung volume between residual volume

(RV) and total lung capacity (TLC), and the red relaxation curve of Fig. 13-3 is recorded. This relaxation curve shows the *specific standard compliance* for the combined system (defined at static conditions as the slope of the curve at functional residual capacity, FRC). In healthy persons, the specific standard compliance for the combined system is 1 ml per Pascal (0.1 l BTPS per cm of water) at FRC. The lungs and chest wall move together and support each other. This is what makes the *standard compliance* of the combined system less than that of the lungs or rib cage alone (Fig. 13-3).

The lung compliance

The volume of the lungs before each apnoea is varied in the same way as when measuring the lung compliance for the combined system. The pressure difference between the mouth and intrathoracic (oesophageal) pressure (P_{alv} - P_{it}) is the blue curve marked transmural lung pressure in Fig. 13-3. In healthy persons, the *specific lung compliance* is 2 ml per Pascal (0.2 l BTPS per cm of water) at FRC, where the blue lung curve is almost linear (dV/dP at FRC).

Normal lungs are very distensible at functional residual capacity (FRC), but stiffen progressively towards total lung capacity (TLC). The falling compliance is caused by an increase in the air-liquid surface tension, because the liquid contains tension-reducing molecules (*surfactant*, see below) that are spread further and further apart. Thereby the compliance of the lung is reduced. Compliance also decreases with age; there are corresponding decreases in lung volumes.

The rib cage compliance

This is a calculated variable. The static transmural wall pressure (P_{it} - P_B) is indirectly obtained from the two static transmural pressures measured above. With the two elastic systems in static equilibrium (P_{it} - P_B) must be equal to the difference between the static transmural pressure of the combined system and that of the lungs: (P_{alv} - P_B) - (P_{alv} - P_{it}). According to this equation a minimal increase in lung volume (dV) implies the following relationship: d(P_{it} - P_B)/dV = d(P_{alv} - P_B)/dV - d(P_{alv} - P_{it})/dV. Each of the entities is an *elastance* (dP/dV). The elastance of the thoracic cage is equal to the combined elastance minus the lung elastance.

The *specific compliance* of the thoracic or rib cage (dV/dP) is the specific reciprocal elastance, and equal to 2 ml per Pascal (0.2 l BTPS per cm of water). The chest wall compliance curve is constructed by using the above equation (green curve in Fig. 13-3).

5. Dynamic airway compression

The driving pressure for air to move is (P_{alv} - P_B). The driving pressure and airway resistances are studied when air moves into and out of the lungs, and the condition is therefore called dynamic. The driving pressure for inspiration is a negative alveolar pressure (P_{alv}) relative to P_B (Fig. 13-4).

Respiratory volume is recorded graphically with a (x, y)-recorder. The tidal volume is plotted against the driving pressure, which is equal to the dynamic alveolar pressure (Fig. 13-4).

The resistance to airflow, and the viscous resistance of lung tissue, causes the dynamic pressure-volume curve (Fig. 13-4) to deviate from the static (Fig. 13-3). The slanting straight line (diagonal) is sometimes called *dynamic compliance* for the combined system (Fig. 13-4).

Integrating pressure with respect to volume gives the two green areas corresponding to the *elastic work* of *one inspiration* (Fig. 13-4). This is the work needed to overcome the elastic resistance against inspiration. The red area to the right of the diagonal is the extra work of inspiration called the *flow-resistive work* or alternatively *non-elastic work* (Fig. 13-4).

During expiration, the flow-resistive work is equal to the light green area (Fig. 13-4). The inspiratory and expiratory curve forms a so-called *hysteresis loop*. The lack of coincidence of

the curves for inspiration and expiration is known as *elastic hysteresis*. With deeper and more rapid breathing the hysteresis loop becomes larger, and the non-elastic work relatively greater.

Fig. 13-4: Tidal volume (V_T) and the dynamic transmural pressure (P_{alv} - P_B) in a healthy person during one respiratory cycle. – The expiratory curve from a patient with obstructive lung disease is shown to the left (see pathophysiology).

During a forceful expiration, the intrathoracic or pleural pressure (P_{it}) rises and causes the alveolar pressure (P_{alv}) to exceed the downstream pressure at the airway openings (P_B). As flow resistance dissipates the driving energy along the bronchial tree, the driving pressure of the cartilaginous bronchi falls towards zero at the mouth (Fig. 13-5). At a certain point the forces that expand the airway equal the forces that tend to collapse. This is the *equal pressure point*. Beyond the equal pressure point the driving pressure falls below the external pressure, and the bronchi are compressed (Fig. 13-5). At this point the person cannot voluntarily increase the rate of expiratory airflow, because increased effort also increases the external pressure. This phenomenon is called *dynamic airway compression* with airway collapse.

The maximum expiratory airflow is *effort-independent* according to Bernoulli's law (Fig. 13-6). Bernoulli's law states that the driving energy equals the sum of the kinetic energy, the constant positional energy and the laterally directed energy (ie, the lateral pressure directed towards the walls). Thus, during expiration the lateral pressure is lowest where the cross sectional area is smallest (the trachea), and the last part of the trachea collapses (Fig. 13-5).

Coughing causes momentary collapse of the tracheal wall. The airway closure occurs when the equal pressure point has moved to a part of the airway that is not supported by cartilage and has the smallest cross sectional area (highest kinetic energy).

Fig. 13-5: Alveolar sac, bronchiole and a cartilaginous bronchus collapsing during expiration.

The *peak airway resistance*, where flow limitation takes place, is found in the medium-sized segmental bronchi around the 4^{th}-7^{th} generation moving peripherally as lung volume decreases. In healthy people the least resistance to airflow is found in the numerous terminal bronchioles. At low lung volumes the elastic pull in the bronchioles becomes smaller (the structures relax with falling volume) and the airways tend to collapse more easily (Fig. 13-5).

6. Dynamic flow-volume-loops

Fig. 13-6 shows a dynamic flow-volume loop generated by plotting airflow velocity measured at the mouth with a pneumotachograph against lung volume (integrated airflow velocity). The computer makes a mark after one s of forced expiration.

The large loop is from a *healthy person* performing a forced expiration from full lung inflation (Total Lung Capacity, TLC) to full lung deflation (Residual Volume, RV) that is the vital capacity, VC. This is a so-called *forced vital capacity manoeuvre* (Fig. 13-6). The inspiratory airflow velocity increases rapidly when inspiring from maximal expiration, and reaches a plateau dependent on muscle force until maximal inspiration, at which point the velocity falls rapidly to zero (Fig. 13-6). Inspiration is limited by the force-velocity relationship of the inspiratory muscles - not flow-limited, as is forceful expiration.

Fig. 13-6: Dynamic flow-volume loops for forced vital capacity from a healthy person (yellow area/Residual Volume = 1.2 l). Patients with restrictive (red area/RV 0.6 l) and obstructive lung disease (blue area/RV 2.4 l) are also shown – see pathophysiology.

7. Elastic recoil

Two forces oppose lung expansion:

A. *The overall elastic recoil* (dP/dV) is the sum of the pulmonary elastic recoil and its surface tension. These forces relate to the *elastic work* of Fig. 13-4. - Traditionally,

the reciprocal elastance or the compliance (dV/dP) is preferred as an index of the distensibility of the lungs and the rib cage. There is a thin fluid layer on the inner surface of the alveolus. Because of the alveolar fluid - air interface a surface tension is created that tends to collapse the alveolus (just like the elastic recoil). The contribution of surface tension to the overall lung elasticity is more than 50%.

B. *The airflow resistance forces* relate to the *non-elastic work* of Fig. 13-4. The total airflow resistance is the sum of all the resistances of the nose and mouth (a substantial portion of the total) and of the 23 generations of the tracheobronchial tree. The friction between gas molecules and between gas molecules and the walls also contributes to airway resistance. The airway resistance is important and makes the sliding of lung tissue over each other (viscous tissue resistance) a minor issue.

In the lungs, the term *static* compliance is used because the volume and pressure measurements are made when there is no airflow. Increased lung compliance is caused by reduced lung elasticity, and means that lungs with elastic tissue degeneration are easier to inflate. Reduced lung compliance is caused by increased lung elasticity and means that stiff, fibrotic lungs are harder to inflate – see pathophysiology.

8. Airway resistance (R_{aw})

The driving pressure (DP) for laminar airflow ($V\dot{}_E$) through the airway resistance is the intra-alveolar pressure (P_{alv}) minus the ambient or barometric pressure, P_B.

The *airway resistance* (R_{aw}) is defined by Poiseuille's law - see Eq. 13-3.

R_{aw} is directly related to the air viscosity (h) and to the length (L) of the tube, and inversely related to its radius in the 4th power: $R_{aw} = 8\ h\ L/r^4$. Doubling the length of the airways only doubles the airway resistance, but halving the radius increases the resistance sixteen-fold. Such a rise in resistance takes place in the small airways during bronchiolitis. The walls of the bronchioles are inflamed causing oedema (swelling), constriction, sloughing of epithelium, and excessive secretion. A similar reversible bronchoconstriction takes place in hyperirritable airway disease (asthma – see pathophysiology).

In the clinic the *airway resistance* is sometimes measured in a body plethysmograph, which is technically demanding to operate, so in everyday clinical practice the *Forced Expiratory Volume* in 1 second (FEV_1) from total lung capacity is used as an indirect measure. This indirect method requires only simple, reliable and accurate spirometers. The patient is asked to expire as fast as possible from TLC by creating a high driving pressure. The driving pressure is considered an arbitrary unit, because it is a reproducible, fast muscle force in each patient. As long as a patient applies an expiratory pressure above the threshold pressure needed to create dynamic airway compression, its absolute size is immaterial. The airway resistance is obtained by dividing the arbitrary expiratory pressure unit by the airflow velocity, FEV_1.

With increasing lung volume the expanding lung tissue pulls the airways open and thereby decreases the airway resistance. There is a continuum from the top to the bottom of the upright lung, with respect to the degree of airway and alveolar distension. The greatest relative lung distension - at any lung volume - is found at the top due to a more negative P_{it}. Consequently, the distended top of the lung has the lowest relative ventilation. Thus airway calibre is larger at the top than at the bottom of the upright lung, causing airway resistance to increase progressively from the top to the base of the lung (Fig. 14-5).

9. The surface tension

When the pressure-volume curves in air (Fig. 13-4) are compared to those from saline-inflated excised lungs, it appears that the air-filled lungs are less compliant and show a larger hysteresis loop than when they are inflated with saline. Saline-filled lungs have no air-liquid interface,

and thus no surface tension. More than half the total elastic recoil force of the lungs is caused by *surface tension*.

The lungs are suspended in a gravity field tending to separate the parietal and the visceral pleura. The gravity causes the tendency to be greater at the lung apex than at its base. Thus, the intrathoracic pressure (P_{it}) is more subatmospheric at the apex of the lung than at its base (-900 compared to -200 Pa or -9 to -2 cm of water at FRC). Since the alveolar pressure (P_{alv}) is more or less the same throughout the alveoli, it follows that the transmural pressure gradient (P_{alv} - P_{it}) over the lung tissue is larger at the apex than at the base of the lung. Therefore, the alveoli at the apex always expand more than the alveoli at the base of the lung. However, the expanded apical alveoli will distend less during inhalation than the small, compliant alveoli located in the middle and basal regions.

Pulmonary surfactant lowers the surface tension in the alveoli, which increases the lung compliance. Surfactant is secreted into the thin air-filled interface of the alveolar lining. Surfactant is a complex phospholipid that is a combination of dipalmitoyl phosphatidyl-choline (DPPC) and other lipids and proteins. DPPC orients perpendicular to the air-water interface, such that the charged choline base is dissolved in water (hydrophilic) and the nonpolar, hydrophobic fatty acids projects toward the alveolar air. The type 2 alveolar epithelial cells secrete surfactant.

Surfactant prevents alveolar collapse. According to the law of Laplace, the transmural distending pressure in spherical alveoli is equal to $T/(2r)$, where T is the total wall tension (elastic recoil plus surface tension; N m^{-1}) and r is the radius (Fig. 8-9). Because the distending pressure is essentially the same in communicating alveoli, the total wall tension changes with diameter. During expiration the diameter decreases, surfactant molecules are packed tightly together, separating the water molecules and reducing the total wall tension. During inspiration the diameter increases, the surfactant molecules scatter, and the water molecules move closer to each other so the total alveolar wall tension increases progressively; the lung becomes stiffer.

10. Pulmonary defence mechanisms

During normal breathing most of the particles of more than 10 mm in diameter - such as pollen - are deposited and removed in the nose and nasopharynx. Particles below 1 mm are deposited in the alveoli. Particles between 1 and 10 mm are deposited in the bronchi - the smaller the particles are the lower they reach.

Although sneezing and coughing with expectoration can eliminate many inhaled particles, the mucociliary escalator assisted by bronchus-associated lymphoid tissue (BALT) and alveolar macrophages perform the main clearance of the airways.

Mucociliary escalator. The airways are protected by humidification all the way to the alveoli with a mucous layer, which prevents dehydration of the epithelium and surrounds the epithelial cilia (Fig. 13-7).

Fig. 13-7: Bronchial wall during an attack of asthma. The protective layer in the lumen is abundant, and consists of a gel phase and a liquid phase surrounding the cilia of the epithelial cells. The lamina propria swells, and the smooth muscle layer is hypertrophic.

The airway mucous consists of polysaccharides from goblet cells and from mucous glands in the bronchial wall. Serous and seromucous glands are also active. The mucous forms a gelatinous blanket on top of the liquid layer (Fig. 13-7). The cilia continuously move the gelatinous blanket with inhaled particles on the top upward towards the pharynx, where they are swallowed.

Clearance of the respiratory bronchioles may take days, whereas clearance of the main bronchi

is typically accomplished within an hour. Smoking reduces mucociliary transport, and indirectly impairs gas exchange. Smoking also reduces surfactant production and thus increases the work of breathing.

The lung secretions contain bactericidal lysozyme and lactoferrin from granulocytes. The a1-antitrypsin normally neutralises chymotrypsin, trypsin, elastase, and proteases secreted by granulocytes during inflammation, and thus prevents destruction of lung tissue.

BALT.

Bronchus-associated lymphoid tissue (BALT) in the walls of the main bronchi is part of the mononuclear phagocytotic system or *Reticulo-Endothelial-System*. These tissue aggregates contain macrophages originating from monocytes and lymphocytes. The lymphocytes are also present in the lamina propria (Fig. 13-7).

Following sensitisation of B-lymphocytes to specific antigens, the cells produce specific antibodies or immunoglobulins (IgA, IgG and IgE) in response to new contact with the antigen (Fig. 33-4). IgA inhibits the attachment of poliovirus, bacteria and toxins in the respiratory tract. IgE is related to the pathogenesis of allergic disorders - see Ch. 33.

Lungs do have endocrine functions. Alveolar macrophages are amoebic cells that swallow particles and bacteria in the alveoli. While they execute microbes in their phagolysosomes, the cells migrate to the mucociliary escalator, or they are removed by the blood or by the lymphatic system. Smoking impairs the normal macrophage activity.

The inactive polypeptide, angiotensin I, is converted into the potent vasoconstrictor, angiotensin II, by the angiotensin converting enzyme (ACE), located on the pulmonary endothelial cells. Angiotensin is important for the regulation of the arterial blood pressure – also during chock.

Adrenaline, dopamine, histamine, prostaglandins A1 & A2, prostacyclin (PGI2) and vasopressin (ADH) pass unaffected through the lungs. Bradykinin, leucotrienes, prostaglandins E2 & F2a, and serotonin are almost completely cleared during passage through the lungs by enzymatic activity.

Adrenergic sympathetic activity (and sympathomimetic drugs) relax bronchial smooth muscle via adrenergic b_2-receptors, whereas parasympathetic cholinergic activity (and parasympathomimetics) constrict bronchial smooth muscles via muscarinic receptors.

Smoke, dust and other irritants (perhaps also adenosine, histamine and substance P) constrict the airway smooth muscles via a reflex triggered by the rapidly adapting irritant-receptors (Chapter 16). Decreased P_{ACO2}, thromboxane and leucotrienes (see Chapter 32) also act as bronchoconstrictors.

Vasoactive intestinal peptide (VIP) can dilatate airways and reduces airflow resistance. Substances that dilatate airways include increased P_{ACO2}, adrenergic alpha-blockers, catecholamines and atropine.

Pathophysiology

Lung volumes, as measured with a spirometer, are needed in order to differentiate between two major functional types of lung-airway disorders, and in quantifying the degree of abnormality. The two types are called A. obstructive and B. restrictive disorders (Box 13-2).

Box 13-2. Classical respiratory disorders	
A.	*Obstructive disorders (increased flow-resistance)*
A1	Asthma: Acute attacks chronic inflammation of bronchial wall Expiratory flow-limitation - Hyper-reactive bronchial wall –

	Eosinophils-Hypersecretion – Bronchoconstriction. Criterion: A low FEV_1 improves more than 15% following inhalation of broncho-dilatators	
A2	Chronic obstructive bronchitis & emphysema	
A3	Respiratory failure	
A4	Cor pulmonale	
A5	Sleep apnoea	
A6	Cystic fibrosis	
B.	*Restrictive disorders (small lung volumes - especially VC)*	
B1	Restrictive disorders *in the lung parenchyma*	
		Granulomatosis (sarcoidosis – often also obstructive) Systemic connective tissue diseases (Rheumatoid arthritis, lupus, sclerosis) External allergic alveolitis (organic dust) Diffuse progressive pulmonary fibrosis Collapsed alveoli and alveolar oedema
B2	Restrictive disorders in the chest wall	
		Rib fractures - Kyphoscoliosis - Ankylosing spondylitis Pneumothorax - Pleural disorders and effusions (transudates and exudates)
B3	Restrictive disorders in the newborn	

A. Obstructive lung disorders

The most common disorders are *A1. asthma* and *A2. chronic obstructice bronchitis and emphysema*. A special condition called cystic fibrosis is also dealt with.

These disorders are all characterised by low expiratory airflow as measured by low Forced Expiratory Volume in one s (low FEV_1). The low FEV_1 is due to narrowing of the airways with increased airflow resistance.

The patient with *obstructive* lung disease has a smaller flow-volume-loop than that of a normal subject - performed as a forced vital capacity manoeuvre (Fig. 13-6). The RV of the patient is 2.4 l or twice as high as that of the healthy individual, because of *air trapping* (a large volume of trapped air). Sometimes the so-called *saw-tooth* phenomenon is observed (Fig. 13-6). This is an unspecific sign of intrathoracic airflow limitation neither related to obstructive sleep apnoea nor to body mass index (Chapter 20).

Although the flow-volume curve in obstructive lung disease is consistently reduced in the flow direction, it is not always reduced in the total volume direction (Fig. 13-6).

A1. Asthma

Attacks of asthma occur acute and episodic, but the underlying cause is a chronic inflammation of the lung airways. Asthma is characterised by expiratory airflow limitation due to hyperactive bronchi with eosinophilic inflammation, mucous hypersecretion and bronchoconstriction. Asthma is diagnosed when there is an improvement of FEV_1 greater than 15% following inhalation of broncho-dilatators. If necessary, airway hyperreactivity can be demonstrated by histamine or metacholine provocation of bronchoconstriction. Large numbers of eosinophils are present in the sputum, and often in the blood. Skin-prick tests often identify extrinsic causes, which the patient must avoid.

Stimulation of the vagal nerve or metacholine provocation causes a forceful reflex bronchoconstriction in asthmatics. This probably explains the hypersensitivity to non-specific stimuli (eg, exercise, cold air or water, pollution, dust, vapours and fumes).

Fig. 13-8: Asthma is an acute obstructive lung disease, with reduced lumen due to broncho-constriction, hypersecretion and oedema of the bronchial wall. Emphysema is a chronic obstructive lung disease with degenerative loss of radial traction of the bronchial walls. Diffuse lung fibrosis is characterised by the thick and stiff alveolar interstitium all over the lung. The Pickwick Syndrome is a restrictive lung disorder due to obesity.

Asthma occurs in the form of extrinsic asthma, which is caused by a specific allergen (*extrinsic*) in an atopic person. A person suffering from *atopy* has a personal history of symptoms from nose, lungs and skin with hay fever (allergic rhinitis), eczema or urticaria often from childhood (childhood asthma). Common allergens are the house-dust mite or its faeces, pollen grains, moulds and domestic pets.

When a middle-aged non-atopic person develops asthma, an allergen is seldom identified. The condition is sometimes called *intrinsic asthma*.

Within minutes from inhalation of an allergen there is an immediate, anaphylactic reaction (Fig. 33-5). Eosinophils recognise the allergen and release allergen specific IgE antibodies. The allergen-IgE complex is bound to IgE-receptors on the surface of granule containing mast cells, eosinophils and basophils. Hereby, mediators of anaphylactic reactions are released from the mast cell granules: Leukotrienes (= SRS-A) are strong bronchoconstrictors and also cause mucosal inflammation with oedema and hypersecretion. Prostaglandin D2 is also contributing with bronchoconstriction and vasodilatation with increased capillary permeability. Eosinophils release leukotrienes C4, PAF, major basic protein and eosinophilic cation protein, all of which are toxic to epithelial cells.

Histamine is a powerful vasodilatator and may play some role for the hypersecretion and bronchoconstriction in asthma, but antihistamins have no effect.

During an attack of asthma the bronchial wall is suffering (Fig. 13-8). The hypertrophic smooth muscles contract, capillary leak of plasma water results in oedema of the lamina propria, and hypersecretion of the mucous glands produces thick mucous, which the cilia can hardly move. All of this causes universal narrowing of the airways or occlusion by mucous.

The airflow limitation results in wheezing respiration, and the patient feels dyspnoa. The attacks usually occur during the night often with coughs. Stethoscopy of the lungs reveals wheezing. A severe asthmatic attack may continue for hours and days, in which case the condition is called *status asthmaticus*. There is tachycardia and sometimes *pulsus paradoxus* (the pulse disappears due to a marked fall in both systolic and pressure amplitude during inspiration).

Some asthma patients develop a relative insensitivity of the adrenergic b_2-receptors of the bronchial smooth muscles (down-regulation of the b_2-receptors). All b-receptors act through activation of adenylcyclase and cAMP. When noradrenaline binds to b_2- receptors it causes bronchodilatation, but not always sufficient in asthma.

Therapy keypoints

b_2-adrenergic agonists interact with the bronchodilatating b_2-receptors, but they also cause tachycardia by stimulating the b_1-receptors of the myocardium. The b_2-receptor agonists (salbutamol or terbutaline aerosols) are effective in mild asthma. - Cardioselective adrenergic b_1-blockers must be administered with care to patients with a combination of asthma and cardiac disease.

Anticholinergic bronchodilatators bind to muscarinic M_1 and M_3 receptors of the airways, and selective muscarinic antagonists such as ipratropium or oxitropium by aerosol are used as supplement to salbutamol. Previously, atropine - a non-selective muscarinic antagonist - was used. These drugs are rather ineffective in asthma.

Anti-inflammatory drugs (Na^+ -cromoglycate, nedocromil- Na^+) blocks a chloride channel in the inflammatory cells, thus preventing Ca^{2+} influx, and thereby liberation of mediators. They are given to cases of mild asthma (children) before stimulation such as exercise.

Corticosteroids (beclomethasone dipropionate or fluticasone propionate) are administered by inhalation, and they are effective in severe cases. The activation of inflammatory cells is rapidly decreased by local corticosteroids, which have minor systemic side effects only.

A2. Chronic bronchitis & emphysema

Chronic obstructive bronchitis is an inflammation of the bronchioles characterised by excessive production of mucous from hypertrophic mucous glands. There is an increase in the number of gel secreting goblet cells (Fig 13-7). The lumen contains large amounts of mucus and pus. The ciliated columnar epithelium is ulcerated and sometimes replaced by squamous cells without cilia (*metaplasia*).

The small airways are narrowed, and their walls thickened by inflammation and oedema. There is morning cough with sputum for at least 3 months of the year for at least two years (WHO). This is a clinical diagnosis based on the patient history.

Emphysema is an patho-anatomical diagnosis characterised by enlargement of the air spaces and destruction of the lung tissue distal to the terminal bronchiole (a tissue unit called an acinus or a primary lobule). The emphysematous lung has increased compliance (increased d*V*/d*P*) because the elasticity is decreased. Destruction of the alveolar walls includes the capillary bed with increased pulmonary vascular resistance causing pulmonary hypertension. Emphysema is established at autopsy. The most common type of emphysema is centri-lobular emphysema, where the damage is limited to the central part of the lobule or acinus, whereas the peripheral alveolar ducts and alveoli are preserved. The rare type of emphysema is called *pan-lobular* or *pan-acinar*, because the entire lobule is destroyed. - Bullous emphysema is when the entire lung consists of large cysts or bullae with hardly any normal tissue left.

Chronic bronchitis & emphysema is synonymous with many alternative terms: Chronic Obstructive Airway/Lung/Pulmonary Disease or abbreviated COAD/COLD/COPD.

Both bronchitis and emphysema co-exist in many patients. Some of these patients are dominated by the first, others by the second, and some have elements of asthma too.

These disorders are almost exclusively confined to smokers (cigarette smokers in particular), and the severity of the disease is proportional to the amount of tobacco (number of cigarettes) smoked per day. The disorders are progressive during years of smoking and cause impaired exercise capacity. A patient smoking more than 25 cigarettes per day have a mortality that is 20 times higher than that of a non-smoker. Other airway irritants such as atmospheric pollution contributes to the death rate.

The maintained irritation of the epithelium from smoke causes the hypertrophy and the

hypersecretion of the mucous glands in the larger airways (Fig. 13-7). Surfactant normally lowers the surface tension of the alveolar fluid layer, but smoke has an adverse effect on surfactant. A high frequency of acute airway infections increases the pulmonary damage. The small airways of smokers are infiltrated with neutrophils, which are also present in their lumen. Neutrophils release elastase and proteases. These enzymes destroy lung tissue and produce emphysema, when not balanced by antienzymes such as antiproteases. A typical *antiprotease* in normal serum is hepatic a_1-antitrypsin, which is inactivated by smoking. The main phenotypes of the a_1-antitrypsin gene are MM= normal, MZ= heterozygous deficiency, and ZZ= homozygous deficiency. Hereditary deficiency only accounts for a minor part of emphysema. In the population with the susceptible phenotypes (MZ and ZZ), smokers develop emphysema 20 years sooner than non-smokers. Smokers with the phenotype PiZ are at a high risk of developing emphysema.

Symptoms and signs

The patient complains of smoker's cough with large morning expectoration (ie, large quantities of sputum - purulent during exacerbation). Fog and pollution worsen the condition. As the lung function deteriorates, breathlessness (dyspnoea) becomes so severe that even dressing or tooth brushing feels like heavy exercise.

The patient expires for a long time and with a snapping inspiration. The intercostal space is drawn in under inspiration and the accessory respiratory muscles are active. The lungs are hyperinflated and extend between the heart and the chest wall.

Some patients are thin "pink puffers" or "type A emphysematous fighters". Despite their severe dyspnoea, their arterial blood gasses are close to normal (*point* i in Fig. 14-7). They are often emphysematous with over-expanded lungs but with little bronchitis. As the term implies these patients are not cyanosed.

Other patients are overweight "blue bloaters" or "type B bronchitis non-fighters", because they seem to have given up the tiring respiratory effort at the expense of hypoxia and cyanosis. They often have symptoms of bronchitis and suffer from coughing, hypoxaemia, secondary polycythaemia, *cor pulmonale* (see below), and carbon dioxide retention with respiratory acidosis (Ch 17). These patients are typically cyanosed, and they suffer predominantly from chronic bronchitis - maybe with little centriacinar emphysema. The patients show little respiratory effort, which explains why they become cyanotic. Cor pulmonale, with body fluid retention (increased extracellular fluid volume), is demonstrated by a raised jugular venous pressure and ankle oedema. The hypercapnia (see below) results in peripheral vasodilatation, tremor, confusion, coma, and papiloedema.

The two clinical pictures are seldom clean and often overlap.

Many COLD patients are thin – typically the fighters - perhaps due to the high oxygen cost of breathing (often 30% of their oxygen uptake at rest), even food intake is a demanding task, and they tend to develop wasting. Obstruction of the airflow causes uneven ventilation; destruction of lung tissue reduces the capillary blood volume. The maldistribution of bloodflow upsets the normal matching of ventilation and perfusion (Ch. 14). This is the major cause of hypoxia and later hypercapnia. Hypercapnia signals alveolar hyperventilation.

Half of all patients with severe breathlessness die within 5 years from hypoxia.

Therapy of COLD

Bronchodilatators, such as the b-adrenergic agonist salbutamol, may reduce the dyspnoea of many patients.

Inhaled corticosteroids are marginally beneficial, but with only a few systemic side effects.

Long term treatment with antibiotics is controversial, and most double blind controlled clinical trials show no effect.

Mucolytics are ineffective for patients with COLD.
Influenza vaccines are recommended for these patients as in other individuals at increased risk.

A3. Respiratory failure

The terminal stage is respiratory failure, which is characterised by hypoxia (ie, P_{aO2} of less than 7.3 kPa or 55 mmHg) and by hypercapnia (P_{aCO2} of more than 6.4 kPa or 48 mmHg) and severe disability. The hypercapnic patient has sacrificed a normal (Box 14-1) P_{aCO2}, because it is much too costly for him in terms of energy expenditure for breathing. The patient is a *blue bloater* in a state of deep hypoxia with extremely low P_{aO2} and malnutrition.

As the hypoxia persists the pulmonary arterioles constrict, which lead to pulmonary hypertension and cor pulmonale.

A4. Cor pulmonale

Cor pulmonale is a heart condition secondary to pulmonary disorder with loss of pulmonary vessels and increased pulmonary resistance (Chapter 14). The lung disease causes pulmonary hypertension, increased work load of the right heart, right ventricular hypertrophy, and finally right heart failure.

Fig. 13-9: Upper airways of a healthy person and of a patient with obesity and obstructive sleep apnoea. This patient also suffers from the Pickwick syndrome. - The airway pressures are given in kPa.

A5. Sleep apnoea

Patients with chronic bronchitis and emphysema – among others - may complain of sleep apnoea. The normal short periods of apnoea occurring during REM sleep are prolonged, the patient wakes up in fear of suffocation, and snoring is frequent. The upper airways are obstructed in the supine position (Fig. 13-9) and the patient may develop severe hypoxia. Obstruction of the pharynx causes an abnormally high negative pressure in the airways (Fig. 13-9, below). During daytime the patient complains of headache and somnolence.

Some of these patients are extremely overweight and develop a restrictive lung disorder due to accumulation of fat in the mediastinum and abdomen (Fig. 13-8). They suffer from frequent periods of somnolence and cyanosis during daytime due to hypoventilation, and they exhibit secondary polycythaemia and low P_{AO2} (the *Pickwick syndrome*).

A6. Cystic fibrosis

Cystic fibrosis is an autosomal recessive inherited disease of all exocrine glands, caused by a gene mutation on chromosome 7.

A specific error results in a defective transmembrane regulator protein (the cystic fibrosis transmembrane conductance regulator). This is an a-adrenergic gated *chloride channel*. Normally, an elevated cAMP in the epithelial cell will open the chloride channel. In cystic fibrosis this does not happen, and with less excretion of NaCl to the airways, sweat ducts and pancreatic ducts, there is less excretion of water and thus increased viscosity of the secretions. Viscous secretions plug the airways and duct systems.

New-borns with recurrent bronchopulmonary infections must be suspected of cystic fibrosis until disproved. It is too late to await airflow limitations. The diagnosis is confirmed by three sweat tests with a NaCl concentration above 60 mM.

Patients with cystic fibrosis not only suffer from pulmonary infection with obstructive lung disease, bronchiolitis and bronchiectasis, but also from *pancreatic insufficiency*.

Therapy is improved by administration of ATP, which stimulates nucleide receptors independent of cAMP. Hereby, the Cl^- excretion is stimulated.

B. Restrictive lung disorders

Low lung volumes, measured as reduced total lung capacity (TLC) and vital capacity (VC), characterise restrictive lung diseases. *Restrictive* lung disease with small lung volumes (VC) is shown in Fig. 13-6. Also the TLC is small, and all volumes are often proportionally decreased. The airflow velocity and relative forced expiratory volume in 1 s is typically normal. One way or the other, the normal expansion of the lungs is restricted or the pulmonary compliance is decreased.

The restrictive disorders are divided into those localised in the *B1. lung parenchyma*, and *B2. chest wall* including its neuromuscular apparatus and the pleura. In newborns there is a third restrictive disorder: *B3. The respiratory distress syndrome.*

B1. Restrictive disorders of the lung

These disorders comprise granulomatous disorders, systemic connective tissue disorders, extrinsic allergic alveolitis, diffuse pulmonary fibrosis, collapsed alveoli (atelectasis) and alveolar oedema.

Granulomatous disorders are characterised by granulomas, which are nodules of inflammatory cells (macrophages, histiocytes, T-lymphocytes, and multinucleated giant cells) that are reacting to an irritant. The irritant is derived from micro-organisms (tubercle bacilli, fungi etc.), helminths, neoplasm or hypersensitive cells. There is a decrease in the number of circulating T-lymphocytes, because they are concealed in the lung tissue and the hilar lymph nodes. Restriction is often combined with obstruction.

Sarcoidosis is a restrictive, granulomatous disorder characterised by bilateral hilar lympadenopathy with pulmonary infiltration and fibrosis on chest X-ray. In some cases there are skin sarcoidosis in the form of erythema nodosum (painful red nodules on the shins), inflammation of the uvea (uveitis), parotitis, acute arthritis or fatal involvement of the CNS.

Systemic connective tissue disorders such as rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and systemic scleroses (generalized scleroderma) are sometimes complicated by interstitial lung fibrosis with restrictive disease. In RA there is pulmonary effusion with adhesions and restriction of thoracic expansion. In SLE, with plasma antibodies against nuclear components, there is recurrent pleurisy with pleural exudates and restricted breathing. In systemic sclerosis there is interstitial fibrosis occurring as widespread cysts called the honeycomb lung. The honeycomb lung is a term used for a typical radiological picture of small, thick-walled lung cysts (dilatated lung lobules).

Extrinsic allergic alveolitis is caused by inhalation of many types of organic dusts containing antigens. Heavy occupational exposure results in a type 3-hypersensitivity *pneumonitis* (Chapter 33). Three examples are described here.

Farmer's lung is due to inhalation of the spores of thermophilic actinomycetes and micropolyspora faeni.

Maltworkers lung is due to inhalation to the spores of aspergillus clavatus, when turning germinating barley.

Bird fanciers lung is due to antigens present in feathers and in avian excreta. - Many different drugs and also radiation may cause pulmonary tissue reactions, with or without allergy, with interstitial fibrosis, alveolitis and restrictive disease.

Diffuse interstitial pulmonary fibrosis is the terminal stage of many lung disorders, which may end up as a honeycomb lung (see above). The characteristic pathology is the thick alveolar interstitium all over the lung (Fig. 13-8). The serious condition is emphasised by the fact that the arterial blood gas tensions are low. The patient is dyspnoeic at rest, and the degree of hypoxaemia is explained by the ventilation-perfusion-inequality. The pulmonary diffusing capacity for oxygen is often reduced by 75% compared to normal (Ch 15).

Pulmonary compliance is reduced, when *alveoli collapse* (atelectasis) or when left heart failure causes *alveolar oedema* (Fig. 10-10).

B.2. Restrictive disorders of the chest wall

These disorders comprise rib fractures, damage of the trachea or major bronchi, kyphoscoliosis (hunchback), ankylosing spondylitis, pneumothorax, pleural effusion, and neuromuscular disorders (poliomyelitis, myasthenia gravis, Guillain-Barre syndrome etc).

Rib fractures are mainly caused by trauma, but they can also occur as the consequence of metastases or osteoporosis. Rib fractures are painful and the pain prevents sufficient ventilation and coughing. Traffic accidents may produce so many rib fractures that a flail segment moves inward during inspiration (*flail chest*). Such paradoxical movements cannot ventilate the underlying lung tissue sufficiently.

Kyphoscoliosis should be corrected as early as possible, because later the restrictive ventilatory condition develops into hypoxaemia due to ventilation/perfusion inequality. Finally hypercapnia ensues, and the patient die from respiratory failure or pneumonia.

Ankylosing spondylitis present itself as back pain with morning stiffness. The disease is genetic with a high frequency among B27-positive subjects (HLA-B27). Thoracic movements are reduced due to fixation of the vertebral joints and the ribs. Low lung volumes and chest wall compliance is combined with normal airway resistance.

Pneumothorax means that air has entered the pleural space from the lung, or pneumothorax is due to a chest wall trauma. Normally, the pressure in the pleural space is subatmospheric, because the elastic recoil forces of the lung and chest wall drag in opposite directions. Communication with the atmosphere eliminates the pressure difference, the lung collapses and the thorax expands.

Spontaneous pneumothorax means that pleura is suddenly ruptured without known cause. The rupture typically occurs apically, where the mechanical tension is largest due to expansion. The sudden pain is pleuritic (accentuated by inspiration or coughing), there is dyspnoea, no breath sounds by auscultation, and drum sounds by percussion of the affected area. Chest X-ray (Fig. 13-10) confirms the diagnosis. Once the leak has closed, air will be reabsorbed, because the total sum of partial pressures of the gases in the passing blood is lower than the atmospheric pressure prevailing in the pleural space or cavity (Krogh´s *sliding equilibrium*).

Fig. 13-10: Drawing of the characteristic X-ray findings in a large spontaneous (left) and

a case of tension (right) pneumothorax.

Tension pneumothorax is, fortunately, quite rare - but an emergency. When air can be sucked into the pleural cavity through the rupture during inspiration, and cannot escape during expiration, the pneumothorax enlarges from breath to breath. The condition arises when the rupture functions as a one-way-valve. Signs of mediastinal shift are tracheal deviation and movement of the heart towards the other side on the chest X-ray. The large veins and heart becomes progressively compressed and the venous return to the heart suffers (Fig. 13-10). The incomplete diastolic filling leads to imminent cardiac failure and death may occur suddenly. Intervention must be rapid and primarily consists of establishing an open pneumothorax until tube drainage with underwater seal can be established.

Pleural effusion is defined as accumulation of liquid (fluid except air) in the pleural space. Chest movements are reduced, breath sounds are absent, and percussion is dull. The X-ray confirms the diagnosis. Analysis of pleural fluid allows distinction between an exudate, transudate or pus.

Transudates have a low content of proteins including lactic dehydrogenase. Transudates are pleural fluids that complicate hypoproteinaemia of any cause (hunger-oedema, nephrotic syndrome and heart failure).

Exudates are pathological pleural fluids caused by malignant tumours (bronchial carcinoma, mesothelioma) or by malignant inflammations (tuberculosis, pulmonary infarction), and exudates are often bloodstained. Haemo- and chylo-thorax describe the accumulation of blood and lymph in the pleural space, respectively.

Poliomyelitis, postinfective polyneuropathy (Guillain-Barré), myasthenia gravis and other neuropathies can affect the *nerve supply* to the respiratory muscles (the most important is the diaphragm) and cause restrictive lung disease with sleep apnoea and respiratory failure.

B3. The Respiratory Distress Syndrome of newborns

Respiratory distress in premature infants is caused by inadequate synthesis of surfactant by the type 2 cells. Such infants have lungs with enormous surface tension forces and low compliance, causing collapse of alveoli (atelectasis) and oedema (Fig. 10-10). Positive-pressure ventilation opposes these changes and may improve gas exchange. Administration of aerosolised surfactant is effective.

Equations

- ***The law of ideal gasses*.** The ideal gas equation reads:

$$\text{Eq. 13-1:}\quad P \times V = n \times R \times T.$$

One mol of gas occupies a volume of 22.4 l at Standard Temperature (273 K), Pressure (one atmosphere = 101.3 kPa = 760 mmHg), Dry air (STPD). - The fact that the product of the pressure and volume of a fixed mass of gas is constant at constant temperature was discovered by Robert Boyle in 1660 (Boyles law).

If the temperatures are the same in two states of one mass of gas, then:

$$\text{Eq. 13-2: (Boyles law):}\quad P_1 \times V_1 = P_2 \times V_2 = \text{constant}.$$

Boyle's law (Boyle-Mariottes law) is not a fundamental law like Newton's laws or the law of conservation of energy, but a practical approximation for real gas.

- ***Poiseuille's law*.** The driving pressure (DP) for laminar airflow ($V\dot{}_E$) through the airway resistance is the intrapulmonary pressure (P_{alv}) minus the external barometric pressure, P_B. Poiseuille's law for laminar air flow is an analogy to Ohm's law:

$$\text{Eq. 13-3: } R_{aw} = DP/\, V\dot{}_E.$$

Airway resistance (R_{aw}) is directly related to the air viscosity (h) and to the length (L) of the tube, and inversely related to its radius in the 4th power: $R_{aw} = 8\ h\ L/r^4$. Doubling the length of the airways only doubles the airway resistance, but halving the radius increases the resistance sixteen-fold.

- ***The law of Laplace*:** For a thin-walled organ with two main radii, Laplace assumed that the transmural (internal) pressure at equilibrium was identical with the fibre tension in the wall (T) divided by the two main radii. For a spherical organ (bubble) such as the alveoli, the two radii are the same, which simplifies the equation

$$\text{Eq. 13-4: Transmural pressure} = T/(2r),$$

where T is the total wall tension (elastic recoil plus surface tension; N per m) and r is the radius.

- ***Rohrers equation*.** Laminar or streamline flow (the streamlines move parallel to the sides of the tubes) is limited to airways with low airflow velocities and smooth walls. Such conditions are normally present in small airways. Laminar flow is silent. However, all airways branch and there is transitional flow at each bifurcation. This transitional (laminar-turbulent) flow depends on the following driving pressure:

$$\text{Eq. 13-5: } DP = [V\dot{}_E \times R_1 + V\dot{}_E^2 \times R_2].$$

The R-symbols denote constants. The second (turbulent) component is small during quiet breathing.

- ***Bernoulli's equation*** states that the total driving energy, applied to a continuously flowing ideal fluid volume (dV flowing frictionlessly and laminarly), equals the sum of 3 types of energy the kinetic energy ($1/2\ \rho\ v^2$ - fluid density multiplied by the squared velocity) , the potential energy at the height (h) and the gravity (G), and the laterally directed energy (the lateral pressure, P, directed towards the walls).

$$\text{Eq. 13-6: Total energy} = dV\,(1/2\ \rho * v^2 + h * \rho * G + P).$$

The lateral pressure is highest where the velocity is lowest. The equation of continuity states that the velocity varies inversely with the cross-sectional area of the tube. Consequently, the lateral pressure is highest where the cross sectional area of the tube is largest. This is surprising as it may seem.

- ***Reynolds equation*** for turbulent flow (eg, energy demanding whirls or eddy currents in a fluid) states that the critical, volumetric mean velocity (v) is directly proportional to the viscosity of the fluid (h), Reynolds-number (Rey), and inversely proportional to the

density (ρ) and the radius, r:

Eq. 13-7: v = h*Rey/(r*ρ).

Turbulence is audible. The viscosity (h) of body-warm air (in the airways) and of blood is $1.88*10^{-2}$ and 4 mPa*s, respectively. The density (ρ) of body-warm air (saturated with water vapour) and of blood is 1.15 and 10^3 kg m^{-3}, respectively. The critical Reynolds number is 1200 for most fluids including body-warm air and blood.

- ***The three lung volume conditions*** are derived from the law of ideal gasses.

Consider a cylinder with a movable piston containing n moles of a gas at volume, V, at a certain pressure and temperature. At the above described standard conditions (STPD) the following equation applies: $V_{STPD}\times760/273 = n\times R$.

Consider the same mass of gas at room temperature (t ^{o}C or 273+t K), saturated with water vapour, and at actual barometric pressure (P_B). These conditions are known as ATPS (Ambient Temperature, Pressure, and Saturated with water vapour at tension P_{water}). The air volume rises with temperature: $V_{ATPS}\times(P_B-P_{water})/(273+t) = n\times R$.

Now consider the same amount of gas at the conditions present in the alveoli; the air is saturated with water vapour, which exerts a partial pressure of 6.3 kPa (47 mmHg) at 37 ^{o}C, at ambient pressure. These conditions are known as BTPS (Body Temperature, ambient Pressure, and Saturated with water vapour). At BTPS conditions the air volume only varies with barometric pressure: $V_{BTPS}\times(P_B-47)/(273+37) = n\times R$.

Since we have considered one mass of gas the products of volume and pressure divided by temperature, must in all three states equal $n\times R$. This is expressed in the following three equations, the constants being those applicable with pressures in kPa:

Eq. 13-8: $V_{STPD}\times101.3/273 = V_{ATPS}\times(P_B-P_{water})/(273+t) = V_{BTPS}\times(P_B- 6.3)/(273+37) = n\times R$.

A *real gas* has a finite size of the molecules, which reduces the effective volume of the space. The real (actual) gas pressure is smaller than the ideal, because of attractive intermolecular forces. Real gasses do not obey the ideal gas equation (Eq. 13-1). However, the deviations from the ideal gas law are acceptable at the pressure (P) and temperatures (T) of life on earth.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. The lateral pressure is highest where the cross sectional area of a tube is smallest.

B. Emphysema is destruction of lung tissue distal to the terminal bronchioles.

C. Clearance of the respiratory bronchioles is typically accomplished within an hour.

D. Asthma is an acute obstructive lung disease, with reduced lumen due solely to bronchoconstriction.

E. Reduced airway resistance characterises obstructive lung disease.

II. Each of the following five statements have True/False options:

A. Coughing is the criterion of bronchitis.
B. Pulmonary surfactant is a combination of dipalmitoyl phosphatidyl-choline and other lipids and proteins.
C. Adrenergic sympathetic activity contracts bronchial smooth muscle via b_2-receptors.
D. Silicon dioxide intoxicates the alveolar macrophages and triggers the fibrinogenic mechanism, so the picture is that of progressive massive fibrosis
E. A lung unit is termed a primary lobulus or acinus.

Case History A

A 24-year-old male, with an oxygen uptake of 333 ml STPD/min, is breathing from a metabolic ratemeter containing 50 l of atmospheric air with an oxygen fraction of 0.2093. The room temperature is 293 K, the water vapour tension is 18 and P_B is 760 mmHg.

1. *Calculate the original STPD volume of the metabolic ratemeter.*

2. *Calculate the time period in which it was safe for the person to breathe in this device. Use a safety margin from 21% to 14% in the breathing medium.*

Case History B

A male with a FRC of 2.5 l shows an intrathoracic pressure change of 3 cm of water during normal tidal breathing of 0.5 l. His chest wall compliance is 0.15 l BTPS per cm of water. He has a total alveolar wall tension force (T) of 0.07 N/m tending to collapse two alveoli with radius 0.00004 and 0.00008 m, respectively. The total alveolar wall tension consists of the surface tension plus the elastic recoil forces.

1. Calculate the ΔP, which can prevent collapse of the two alveoli.

2. Is this result consequential for the stability of his alveolar design?

3. Is there a natural solution to this problem?

4. Calculate the specific lung compliance of this patient and compare it to the normal value.

5. What pressure must be applied to supply this person with one l (BTPS) of air per breath under positive-pressure ventilation.

Try to solve the problems before looking up the answers.

Highlights

- *Air passes into the airways through the nose and mouth, where it is warmed, humidified and filtered. From the trachea to the alveoli, there are 23 generations of airways. The first 16 (as an average) constitute the conducting zone, which is an anatomic dead space, because no gas exchange takes place. The 17-23 generations are the respiratory zone.*

- *The lung-thoracic wall system consists of two elastic components that work together: the lungs, which behave like a balloon trying to collapse, and the thoracic cage trying to expand.*

- *The barometric pressure (P_B) is one atmosphere and is frequently defined as a zero reference point. All pressures measured are given in reference to the barometric pressure, which is the pressure at the mouth or at the surface of the thoracic cage.*
- *The lateral pressure is highest where the cross sectional area of the tube is largest.*
- *Dynamic airway compression: The lateral pressure is lowest where the cross sectional area of the tube is smallest and the velocity largest (Bernoulli's law). The external pressure exceeds the lateral pressure and the airway is compressed.*
- *Coughing causes momentary collapse of the tracheal wall.*
- *Turbulent flow is audible.*
- *Normal lungs are very distensible at functional residual capacity (FRC), but stiffen progressively towards total lung capacity (TLC). Lung distensibility is called compliance.*
- *Compliance is an index of expandability of elastic organs and defined as the change in volume per unit change in pressure (dV/dP). The falling compliance during inflation near TLC is caused by an increase in the air-liquid surface tension, because the liquid contains tension-reducing molecules (surfactant) that are spread further and further.*
- *The lungs and chest wall move together and support each other. This is what makes the total standard compliance of the respiratory system less than that of the lungs or rib cage alone.*
- *Two forces are opposing lung expansion: The overall- elastic recoil of the lung, and the non-elastic or airflow resistance.*
- *Surfactant is a complex phospholipid that is a combination of dipalmitoyl phosphatidylcholine (DPPC) and other lipids and proteins. DPPC orients perpendicular to the air-water interface, such that the charged choline base is dissolved in water (hydrophilic) and the nonpolar, hydrophobic fatty acids projects toward the alveolar air. The type 2 alveolar epithelial cells secrete surfactant.*
- *Surfactant lowers the surface tension importantly in the alveoli, thereby increasing the lung compliance.*
- *β_2-adrenergic agonists interact with the bronchodilatating β_2-receptors, but they also cause tachycardia by stimulating the β_1-receptors of the myocardium.*
- *Respiratory distress syndrome in premature infants is caused by inadequate synthesis of surfactant by the type 2 cells. Such infants have lungs with enormous surface tension forces (low compliance), causing atelectasis (collapse of alveoli) and oedema.*
- *Chronic bronchitis is characterised pathologically by hypertrophy of the mucous glands of the bronchi, in combination with an increase in the number of gel secreting goblet cells. The bronchial wall is oedematous and inflamed.*
- *The most common type of emphysema is centri-lobular emphysema, where the damage is limited to the central part of the lobule or acinus, whereas the peripheral alveolar ducts and alveoli are preserved.*

Further Reading

The glossary Committee of the International Union of Physiological Sciences (IUPS): "Glossary on respiration and gas exchange". *J. Appl. Physiol.* 35: 941-961, 1973.

Rahn, H., A.B. Otis, L.E. Chadwick, W.O. Fenn. The pressure-volume diagram of the thorax and lung. *Amer. J. Physiol*. 146: 161-166, 1946.

Katzung BG. *Basic & Clinical Pharmacology*. Appleton & Lange, Stanford, Connecticut, 1998.

Return to chapter 13

Return to content

Chapter 14.

Gas Exchange And Disorders

Study Objectives

- To define cardiac output, diffusion, diffusion- and perfusion-limited gas exchange, hypercapnia, hypocapnia, hypoxia, respiratory quotient (RQ), ventilatory exchange ratio(R) , and the ventilation-perfusion ratio ($\dot{V}_A / \dot{Q}$ -ratio) .

- To describe Henry's and Dalton's laws, factors of importance for the lung diffusion capacity and pulmonary perfusion, and its measurements, describe the P_{O2}-P_{CO2}-diagram, hypo-and hyperventilation, pulmonary water balance, and mixed venous blood composition.

- To calculate the pulmonary perfusion, use the alveolar gas equation, the alveolar ventilation equation, the final $\dot{V}_A / \dot{Q}$ -equation, and the law of mass balance in calculations.

- To explain the alveolar oxygen uptake and carbon dioxide output, pulmonary vascular resistance and pressures. To explain the alveolar dead space, veno-arterial shunts, uneven regional ventilation-perfusion ratio in health and disease, and peripheral gas exchange.

- To use the concepts in problem solving and case histories.

Principles

- ***Bernoulli's principle or equations (see Eq. 13-8).***

Definitions

- **Alveolar oxygen uptake** per min is the uptake of oxygen molecules into the passing pulmonary blood - into the cardiac output.

- **Cardiac output** is the volume of blood leaving the left (or the right) ventricle each min.

- **Diffusion** is a transport of atoms or molecules caused by their random thermal motion.

- **Diffusion capacity for the lung** (D_L) is defined as the volume of gas diffusing through the lung barrier per min and per unit of pressure gradient ($D_L = \dot{V}_{O2} / DP$).

- **Diffusion-limitation** of *gas exchange* is a condition where equilibration does not occur between the gas tension in the pulmonary capillaries and the alveolar lumen. When the distance between the capillary blood and the cells is large, diffusion becomes a limiting factor even at high bloodflow.

- **Hypercapnia** refers to a resting condition with *hypoventilation*, where P_{aCO2} is higher than 6.4 kPa (48 mmHg).

- **Hypocapnia** is a *hyperventilation* disorder with abnormally reduced P at rest (below

aCO2

33 mmHg or 4.4 kPa).

- **Hypoxia** denotes *oxygen deficiency in tissues* due to insufficient delivery of oxygen or inability to utilize oxygen. Hypoxia may be present both with low P_{aO2} and with normal P_{aO2}.

- **Hypotonic** or *hypobaric* **hypoxia** is characterised by a P_{aO2} less than 7.3 kPa (55 mmHg). This is the threshold, below which the ventilation starts to increase by carotid body stimulation. As the altitude increases, the barometric pressure decreases and the partial pressure of oxygen in the alveolar air falls.

- **Hypoxic pulmonary vasoconstriction** is a compensatory mechanism in alveoli with low ventilation and low oxygen partial pressure. The mechanism is triggered directly by smooth muscle contraction in the vessel walls at a P_{aO2} less than 7.3 kPa (55 mmHg).

- **Multiple inert gas technique** is a procedure, where multiple inert gases of different air—to-blood solubility ratios are infused intravenously until steady state of pulmonary gas elimination is reached. The partial pressure of each gas is measured inthe infused fluid and in the expired air. The $V\dot{}_A / Q\dot{}$ -equation (Eq. 14-5) and the law of mass balance is used to compute the most likely regional $V\dot{}_A / Q\dot{}$ -distribution. - Clinically, the alveolar-arterial oxygen tension gradient is measured instead of this complicated reseach procedure.

- **Perfusion-limited** or flow-limited *gas exchange* is limited by the bloodflow. The only limitation to net movement of small molecules across the capillary wall is the rate at which bloodflow transports the molecules to the capillaries.

- **Peripheral Resistance Unit (PRU)** is measured as driving pressure per bloodflow unit (eg, $mmHg*s*ml^{-1}$).

- **Pulmonary hypertension** is a condition with a mean pulmonary artery pressure above normal - a pressure above 2 kPa or 15 mmHg.

- **Pulmonary vascular resistance (*PVR*)** is the ratio between the pressure gradient and the bloodflow. The basic equation is: *PVR* (PRU) = D*P* / bloodflow (PRU in $mmHg*s*ml^{-1}$).

- **Pulmonary oedema** is an emergency caused by filtration of fluid out of the pulmonary capillaries into the interstitial space (interstitial oedema), and eventually in the alveolar spaces (alveolar oedema).

- **Respiratory Quotient (RQ)** is a metabolic ratio between the carbon dioxide output and the oxygen uptake of all cells of the body.

- **Standard affinity** is the binding force between two molecules, when half of the binding sites are occupied (at 50% saturation). In the case of oxyhaemoglobin the P_{50} is used. Here, standard affinity is equal to $1/P_{50}$.

- **Ventilatory exchange ratio (*R*)** is the ratio between the carbon dioxide output and the oxygen uptake measurable with gas exchange equipment at the mouth.

Essentials

This paragraph deals with

1. Gas exchange, 2. A key to lung disorders, 3. Uneven distribution of tidal volume and perfusion, 4. Blood gasses, 5. The P_{O2} - P_{CO2} diagram, 6. The $V°_A$ / $Q°$ - curve, 7. Blood-R-curves, 8. Dead space, 9. Anatomic venous-to-arterial shunt, 10. Ficks law of diffusion, 11. Single -breath diffusing capacity, 12. Compensation of $V°_A/Q°$ - mismatch, 13. Pulmonary bloodflow, and 14. Regional ventilation.

1. Gas exchange

Gasses are exchanged between the atmosphere and the alveolar air, and gasses diffuse between the alveolar air and the blood flowing through the pulmonary capillaries.

Oxygen is transported from the atmosphere, via the alveolar ventilation and then carried by the pulmonary bloodflow (equal to the cardiac output), into the cells and their mitochondria for metabolic purposes. Carbon dioxide, the final end-product of metabolism, migrates from the cells to the atmosphere.

A healthy normal person at rest, ventilates his lungs with 5 litres (l) min^{-1} of fresh air ($V°_A$). The *Respiratory Quotient* (RQ) is a metabolic ratio between the carbon dioxide output

($V°_{CO2}$) and the oxygen uptake ($V°_{O2}$) defined for all body cells as a whole. In respiratory steady state, RQ can be measured as the *ventilatory exchange ratio (R)* (Fig. 14-1).

On a diet dominated by carbohydrate the metabolic RQ for all cells of the body is approaching 1, and in a *respiratory steady state*, identical to the *ventilatory exchange ratio, R*, which is measured in the expired air (Fig. 14-1).

Fig. 14-1: The respiratory quotient (RQ) is compared to the measurable ventilatory exchange ratio (*R*).

The normal resting carbon dioxide output is 10 mmol or 224 ml STPD per min from an adult person, and the cardiac output is typically 5 l min^{-1}. The blood volume of 5 l carries each min about 10 mmol (or 224 ml STPD) of oxygen towards the mitochondria. Following passage of the capillary system, the same amount of CO_2 is carried towards the lungs in the venous blood as long as RQ and *R* is 1.

Blood passing the pulmonary capillaries of a healthy person is rapidly equilibrating with the alveolar air. Oxygen from the air diffuses into the blood and binds reversibly with haemoglobin. The *normal oxygen capacity* is 200 ml STPD per l of blood (150 g haemoglobin per l carrying 1.34 ml STPD per g).

The six zones of the *alveolar-capillary barrier* are: 1) a fluid layer containing *surfactant,* 2) the *alveolar epithelium*; 3) a *fluid-filled interstitial space* ; 4) the *capillary endothelium* with *basement membrane*; 5) the *blood plasma;* and 6) the *erythrocyte membrane*. The six zones form an almost ideal gas exchanger for oxygen and carbon dioxide diffusion.

There are 300 million tiny blind end sacs (alveoli) in both lungs together. Fortunately, the alveoli are diluted continuously with fresh air as we breathe.

2. A key to lung disorders

The alveolar ventilation-perfusion ratio is presented as a straight line in Fig. 14-2.

Alveolar ventilation ($V°_A$) and pulmonary bloodflow (equal to the cardiac output, $Q°$) is considered in three extreme situations:

1. The *normal condition* in which $V°_A$ and $Q°$ are matched (ideal $V°_A / Q°$ -ratio = 5/5 = 1), is shown with the typical normal arterial gas tensions (Fig. 14-2).

2. *Pulmonary embolism* creates an *alveolar dead space*. The $V°_A$ is maintained, but there is no bloodflow ($Q°$ regional), so the $V°_A / Q°$ -ratio of the lung region approaches infinity. In the alveolar dead space, alveolar gas pressures approach the levels in inspired air.

3. *Occlusion of the airway* represents an extreme mismatch of venous to arterial shunting of blood, namely perfusion with no ventilation at all (ie, the total ratio approaches zero). The arterial blood gas tensions approach those of venous blood (Fig. 14-2).

The straight line (or $V°_A / Q°$ -axis) of Fig. 14-2 represents an infinite row of ventilation-perfusion-values. Each value refers to an alveolus with equilibrated blood flowing by.

Two well-known equations are relevant here: the *Fick cardiac output equation* (Eq. 14-1) and the *alveolar gas equation* (Eq. 14-3).

The hyperbolic relationship between $V°_A$ and F_{ACO2} is described in the *alveolar ventilation equation* (Eq 14-4).

These three equations can be combined to one equation, which can be expressed in several ways. The calculations are not shown here. The *final equation* reads as Eq. 14-5:

$$V°_A / Q° = R(C_{aO2} - C_{v\text{-}CO2}) / F_{ACO2}$$

Solutions of this equation provide values from zero to infinity for the ventilation-perfusion-ratio. These solutions can be plotted in a P_{O2} - P_{CO2} diagram (Fig. 14-6), where complicated calculations are performed and solved graphically at a glance by looking at the red $V°_A / Q°$ --curve. In the venous point the $V°_A / Q°$ -ratio is zero, and in the I-point on the abscissae the $V°_A / Q°$ -ratio is infinite.

The regional $V°_A / Q°$ -ratio is the *all-important variable*. In any cardio-pulmonary disease, the normal variation of the ratio for the entire system (Fig. 14-2) is exagerated.

Fig. 14-2: Three pulmonary regions or alveoli representing 3 $V°_A / Q°$ ratios from zero to infinity. Normally, the ventilation/perfusion-ratio is 0.8-1.2 for the entire system. Blood gas tensions are given in kPa (133.3 Pa equals 1 mmHg).

3. Uneven distribution of tidal volume and perfusion

can eventuate from *uneven resistance to airflow* within the lung (bronchoconstriction, collapse and compression of airways). Uneven distribution can also be caused by *uneven regional lung compliance* (insufficient surfactant, loss of elastic recoil as in destruction of alveolar tissue, and increase of elastic recoil as in connective tissue scarring or fibrosis with stiff lungs). *Hypoperfusion* can be caused by *compression* of pulmonary vessels, *obliteration* of vessels by fibrosis, or *blockage* by emboli or thrombosis.

Functional shunts arise with any consolidation of alveolar regions that continue to have bloodflow (pneumonia, oedema, haemorrhage, cell necrosis, lack of surfactant).

4. Blood gasses

Blood gases from an arterial blood sample of a healthy person typically show the values of Box 14-1:

Box 14-1: Blood gas values (ranges) from healthy persons at rest. - Normal mean tensions for mixed venous blood and for alveolar air are shown below.

PaO2 :	10-13 kPa (75-95 mmHg)
PaCO2:	4.8-6 kPa (36-45 mmHg),
Base Excess:	Zero (Chapter 17),
pHa :	7.35-7.45 (ie,[H+] = 35-44 nM)
PAO2 :	10 - 13.3 kPa (75-100 mmHg)
PACO2 :	4.8-6 kPa (36-45 mmHg). Mean: 5.3 kPa (40 mmHg),
Mean PvO2 :	6 kPa (45 mmHg)
Mean PvCO2 :	6.1 kPa (46 mmHg)

The blood gasses are essential in the management of severely ill persons with respiratory or circulatory diseases. Any patient - as any healthy person - has some degree of *ventilation-perfusion mismatch*. The P_{aO2} in itself is a good detector of consequential mismatch, but skilled management necessitates interpretation of P_{O2} -P_{CO2} combinations. This is quite easy with the use of Fenn & Rahn´s P_{O2} - P_{CO2} diagram.

Let us develop this excellent clinical tool from simple mathematics and geometry.

5. The P_{O2} - P_{CO2} diagram (Fenn-Rahn)

The general P_{O2} - P_{CO2} - diagram is actually a rectangular triangle with the following corners: (0,0), (P_{O2} , 0) and (0, P_{CO2}) in Fig. 14-3.

The total tension of all three dry gasses is equal to (P_{O2} + P_{CO2} +P_{N2}) or the barometric pressure (P_B) minus the tension of water vapour in the alveolar air at 37^o C.

The total dry tension at P_B=760 mmHg is thus (760 - 47) = 713 mmHg or (101.3 - 6.3) = 95 kPa. Accordingly, the 713 mmHg on the abscissa refers to pure oxygen (O_2), the 713 on the ordinate refers to pure carbon dioxide (CO_2), and (0,0) represents pure nitrogen (N_2).

Let us assume that analysis of an alveolar air sample results in values shown in point A of the diagram. The diagonal is the hypotenuse of the triangle with a slope of –1. The vertical or the horizontal distance to the diagonal gives the size of the P_{N2}, so in each point all three tensions can be read.

Maintained breathing of pure oxygen leaves all possible expiratory values on the diagonal of Fig. 14-3, namely -1, which is an *R*-value of 1.

In any event, *R* is always equal to one, when one mole of CO_2 is given off to the alveolar air for each mole of O_2 uptake by the blood (Fig. 14-3 and 14-4).

Fig. 14-3: The general P_{O2} - P_{CO2} diagram as developed by Fenn and Rahn. The two axes

show the total partial pressure sum of the three dry gasses at one atmosphere. A is a representative alveolar point when breathing air. - Typical alveolar points in the Andes and on the top of Mt. Everest are also shown. – The total dry tension is 713 mmHg or 95 kPa.

Maintained breathing of atmospheric air at sea level implies a P_{IO2} around 150 mmHg (Fig. 14-3). A line from this point to the "CO_2 corner " of Fig. 14-3 represents the situation, where the venous blood delivers CO_2 to the alveolar air without uptake of O_2. Accordingly,

$V°_{O2}/ V°_{CO2}$ approach infinity and thus *R* approaches infinity.

Fig. 14-4: The ventilatory exchange ratio, *R*, in different lung regions when breathing air at one atmosphere of pressure (modified from Fenn-Rahn).

When only O_2 is given off to the blood and no CO_2 is removed, the *R*- line falls on the abscissa (Fig. 14-4). All possible *R-values* are easily constructed graphically from any P_B on the P_{O2} - P_{CO2} diagram. The gas-*R* lines fan from the inspiratory or *I* -point (Fig. 14-4).

6. The $V°_A$ / Q°- curve

All alveolar ventilation-perfusion ratios ($V°_A$ / *Q° -ratio*) from zero to infinity were represented by the *straight line* of Fig. 14-2. This line covers all possible combinations of regional ventilation-perfusion-units in the lungs in health and disease. This *line* can be changed to a *curve* by transfer to the P_{O2} - P_{CO2} diagram (Fig. 14-5). Such a curve connects the points for the regional $V°_A$ */Q° -ratio* equal to zero (v^- with all perfusion-no ventilation) and for the regional $V°_A$ */Q° -ratio* equal to infinity (the inspired point I with no perfusion-only ventilation).

Alveolar tensions around **A** refer to the healthy upright lung, where the regional $V°_A$ */Q° -ratio* is slightly less than *one* in most alveolar units.

Normally, alveolar ventilation ($V°_A$) and perfusion (*Q°°)* are matched and the total $V°_A$ */Q° ° -ratio* is between *0.8 to 1.2* with normal alveolar and blood gas tensions. In the normal upright lung the regional $V°_A$ */Q°°* -ratio is approximately 0.6 at the lower and about 3 at the upper lung region.

Fig. 14-5: Three alveolar regions in the upright lung of a healthy person at rest. The upper alveolus and airway is distended and its bloodflow is minimal. The lower alveolus is compressed by gravity and its bloodflow is high.

The *pulmonary bloodflow decreases* from the lower to the upper parts of the lung of a resting person (Fig. 14-5). Likewise, the relative ventilation of the lung also decreases linearly from the base to the apex, but at a slower rate. Thus, the regional ventilation-perfusion ratio varies from *zero* in the lower region, where there is only bloodflow and no ventilation to *infinity* in the upper region, where there is only ventilation and no bloodflow. At the lower lung region, regional *V°* approaches zero and at the top of the lung regional perfusion approaches zero. In a P_{O2} - P_{CO2} diagram each point on the curve represents partial pressures at which alveolar air and blood can equilibrate at a certain $V°_A$ / *Q°* -ratio. Thus for any practically obtainable

point, a single value exists for blood gas concentrations (see later in Fig. 14-6). - Lung regions at the base with low $V°_A/Q°°$ has low P_{AO2} and high P_{ACO2}, relative to normal mean values. Upper lung regions with *high* $V°_A/Q°°$ have relatively *high* P_{AO2} and *low* P_{ACO2}.

7. Blood-R-Curves

For a person in respiratory steady state, the *R*- value of the *blood* is equal to the *R* -value of the *alveolar* gas (the gas-*R*). As respiratory gasses are exchanged with a certain *R*- value (eg, *R*=1), the passing blood must do the same. Accordingly, the *blood-R* is equal to the *gas-R*. The blood -*R* curves fan out from the venous point (Fig. 14-6). One green *blood-R* curve is shown. The green curve intersects with its blue *gas- R*-line on the $V°_A/Q°$ - curve.

The shape of the blood-*R* curves is dictated by the oxyhaemoglobin dissociation and the carbon dioxide binding curves, which in turn are affected by the Bohr- and the Haldane-shifts.

Fig. 14-6: Gas-*R* and the related blood-*R* curve (0.8) drawn together with the $V°_A/Q°$ - curve. An ideal alveolar point is shown (i) together with normal values for arterial (a), alveolar (A), and expired (E) gas tensions. - The symbol t for tissue (co-ordinates 1,47 mmHg) denotes minimal tensions in the peripheral tissues of a healthy person.

The regional $V°_A/Q°$ -ratios (Fig. 14-6) show the *lower* lung regions to be *relatively underventilated* (ratio below one), the *middle* lung regions to be *well matched* (ideal regional ratio of 1), and the *upper* lung regions to be relatively *overventilated* (ratio above 1 and approaching infinity). Also *R* approach infinity when we approach the inspired point **I**, Fig. 14-6).

Points **A** and **E** refer to the alveolar and expired air tensions, respectively. Every regional deviation from the average total $V°_A / Q°$ -ratio of 1 in healthy subjects, will result in alveolo-arterial gas tension differences (see later in Fig. 14-6).

Underventilated and overperfused alveoli have increased P_{AN2} and thus increased P_{aN2}, whereas overventilated alveoli reduce their P_{ACO2} almost as much as they increase P_{AO2}. Hereby, P_{aN2} becomes greater than P_{AN2}, and a precisely measured difference is used as a measure of mismatch.

8. Dead space and shunt

Ideal lungs have a matched ventilation-perfusion ratio resulting in an ideal composition of the alveolar air throughout the lung. With all alveoli having identical $V°_A / Q°$ -ratio s, the alveolar point **A** would be located in the *ideal* point **i** for all alveoli (see Fig. 14-7). Doubling of alveolar ventilation will move the point **A** halfway down the blue diagonal towards **I,** and as alveolar ventilation approaches infinity the gas concentrations of the alveolar air approaches those of the inspired air (**I**).

Normally, the expired air values are always represented by a point **E** on the diagonal between **A** and **I** (Fig. 14-6).

All the displacement from ideal point **i** to real life point **A** is caused by *alveolar dead space*, and all the displacement from **i** to E is caused by *alveolar* plus *anatomic* dead space. This sum is also termed the *physiological* dead space (Fig. 14-7). The physiological dead space of a healthy adult at rest is approximately *150 ml* out of a tidal volume of 500 ml (30%). During

exercise the physiological dead space will rise to perhaps *200 ml* simultaneously with a rise in tidal volume to 2000 ml as an example. This is a relative physiological dead space of only 10%, which is an advantage to the individual during work.

Fig. 14-7: Alveolar (A), expired (E) and arterial (a) gas tensions from a patient with chronic obstructive lung disease. Both a large alveolar dead space and a serious shunt are present. The ideal point (i) is also shown. The different locations of the symbol's t illustrate the tensions in peripheral tissues of a patient and of a healthy person.

In healthy persons, the alveolar gas tensions vary during a respiratory cycle around a mean value, although the oscillations are close to the *ideal point*. These variations are called alveolar gas tension *oscillations* (see Chapter 16, Fig. 16-8).

Patients with lung disorders often have $V°_A/Q°°$ -mismatch by a combination of areas of *veno-arterial shunting* often in the lower lung regions, and areas of *increased* alveolar dead space often in the upper lung regions (Fig. 14-7). The location of the arterial point **a** (50,40 mmHg) on the green curve indicates that the **i-a** *distance* is larger than 50% of the total **i-v** *distance*, which must be caused by more than 50% veno-arterial shunting. This **i-a** *distance* is an essential clinical concept, called the *ideal alveolar-arterial* P_{O2} *gradient* (Fig.14-7). The closer point **a** is to point $\mathbf{v}^-$**,** the larger is the shunt. The veno-arterial shunt is *total* (100%), when the point **a** is moved to the point $\mathbf{v}^-$

9. Anatomic venous-to-arterial shunts

Normally, up to 5% of the venous return passes directly into the systemic arterial circulation. This shunt-blood includes nutrient bloodflow coming from the upper airways and collected by the bronchial veins. Also the coronary venous blood that drains directly into the left ventricle through the Thebesian veins is shunt-blood.

The classical way to determine the relative size of a shunt is by the law of conservation of matter. Adolph Fick used the naturally occurring indicator oxygen as substance (Fig. 14 -8). The law of mass balance is applied to both bloodflow and oxygen flux in Eq. 14 -7 and 14-8, where $C_{c'02}$ is the oxygen concentration in the *pulmonary end capillary blood* of an ideally functioning alveolus (Fig.14-8).

The flow and flux relations lead to Eq. 14-9, which shows that the classical method, necessitates cardiac catheterisation to get mixed venous blood (C_{v^-O2}) for the determination of mixed venous gas tensions.

Fig. 14-8: The classical method of determining the size of a shunt implies cardiac catheterisation and measurements of blood gas concentrations.

The location of point **E**, more than half way down the diagonal to **I**, suggests a *large physiological dead space* - more than 50% of the tidal volume (Fig. 14-7).

As the disease progresses, the venous point (v^-) moves to the left and upward, so that peripheral tissues with the smallest P_{O2} gradient become increasingly hypoxic.

The broken curve shows the tensions in tissues from the mixed venous driving tension to tissue tensions (**t**) of only one mm Hg (Fig. 14-7). The slopes of these *tissue tension curves* are

about 1/20, reflecting that CO_2 diffuses 20 times faster than oxygen (23.2 * 0.85= 20). In the final phase of lung disorders also hypercapnia becomes prominent (see the venous point with high P_{CO2} in Fig. 14-7).

10. Fick's law of diffusion

states that the flux of gas transferred across the alveolar-capillary barrier is related to the *solubility* of the gas, the diffusion area (A), the length of the diffusion pathway from the alveoli to the blood (L), and the driving pressure (P_1 - P_2). These factors are all included in the simplified version of Ficks law marked Eq. 14-2. The *solubility* is also called the Bunsen solubility quotient, α**.**

Although the diffusion area at rest is close to the size of half a tennis court, and the diffusion distance (L) is 0.5 - 1 micrometer, it is difficult to predict their variations between individuals. Therefore, Marie Krogh developed the individual *lung diffusion capacity* (D_L) defined as the flux of gas transferred per pressure unit through the lung barrier of a certain person. Since the counter pressure of CO in the blood is virtually zero, a simple measure of P_{ACO} provides us with the pressure gradient in Eq. 14-2. The *standard affinity* of the haemoglobin-CO reaction is very large and *250* times greater than that of O_2 (Eq. 14-10). The standard affinity is measured as the reciprocal value of P_{50}. The P_{50} for haemoglobin-CO is just a fraction of one mmHg, and the haemoglobin-CO dissociation curve is too close to the ordinate of Fig. 14-9 to show - so an enlargement is drawn to the left. - D_L consists of a *barrier-factor* (consequential in lung oedema and in lung fibrosis) and a *haemoglobin-factor* (which reflects the binding rate of oxygen to haemoglobin). The presence of haemoglobin permits blood to absorb 65-fold as much O_2 as the content in plasma at normal P_{aO2}.

Fig. 14-9: Dissociation curves for Oxy- and CO-haemoglobin.

CO competes with O_2 for binding sites on haemoglobin, and thus exposure to CO *reduces* the O_2 binding to haemoglobin. Persons breathing traces of CO occupy a large fraction of all binding sites by CO. The CO binding causes a leftward shift of the oxy-haemoglobin dissociation curve. All the binding sites that are bound to CO, do not respond to falling P_{aO2}. The remaining O_2 molecules on the CO-haemoglobin molecule are much more avidly bound and unload slower than normal.

Diffusion is rapid over short distances. In normal lungs there are *trans-barrier pressure gradients* for diffusion of both O_2 and CO_2. D_{LCO} is measured by measuring the carbon monoxide uptake and the driving pressure (see *single-breath diffusing capacity* below).

11. Single-breath diffusing capacity

The subject takes a deep breath of 0.3% carbon monoxide (CO) and holds the breath for 10 s before exhaling and alveolar sampling. During the breath holding, CO is taken up by the haemoglobin of the passing blood in proportion to its alveolar tension (P_{ACO}).

A simple assumption is that the CO uptake is directly proportional to the mean alveolar P_{CO} (symbolised with P_{ACO}). The *diffusing capacity of the lung* (D_L) is also called the *transfer factor*, because D_L measures not only diffusion, but the barrier thickness and ventilation-

perfusion mismatch as well. Patients with lung disease often have abnormal size and thickness of the alveolar barrier or ventilation-perfusion mismatch. In such cases measurement of the *CO transfer* need not be a true measure of the total diffusing capacity.

Box 14-2: Diffusing capacities of the lungs for different gasses in healthy persons

Units	**ml STPD s^{-1} kPa^{-1}**		**ml STPD min^{-1} $mmHg^{-1}$**	
	Rest	**Exercise**	**Rest**	**Exercise**
DLCO	3	7.5	25	62.5
DL02	3.6	9	29	73
DLCO2	70	175	565	1412

The *single-breath CO diffusing capacity* is normally *3 ml* STPD s^{-1} kPa^{-1} at rest. The values during rest and exercise - and in two units - are shown in Box 14-2.

The *transfer factor* is reduced by diseases affecting the lung parenchyma, such as emphysema, pneumonectomy and fibrotic diseases (the alveolar barrier is too small in area or too thick or both).

12. Compensation of $V°_A$ /Q°- mismatch

Low P_{AO2} in poorly ventilated alveoli, causes arteriolar *constriction,* which redistributes bloodflow to well-ventilated alveoli.

Low P_{ACO2} exists in alveolar regions with a *high* ventilation-perfusion-ratio. Low values constrict the small airways leading to these alveoli. Their reduced ventilation results in redistribution of gas to alveoli with better bloodflow.

13. Pulmonary bloodflow

Pulmonary vascular *resistance* (*PVR*) is minimal compared to that of the systemic circulation. The pulmonary vascular system is basically a low-pressure, low-resistance, highly compliant vessel system with a bloodflow sensitive to gravity and to P_{AO2}.

The system is meant to accommodate the entire cardiac output - and not to meet special metabolic demands as in the case of the systemic circulation.

Box 14-3: Blood pressures in the pulmonary system of a healthy supine person at rest

Units	**mmHg**	**kPa**
Right ventricle	25/-1	3.3/-0.133
Pulmonary artery	25/8	3.3/1
Mean pulmonary Artery	13	1.7
Pulmonary capillaries	8	1
Left atrium	5	0.7
Driving pressure	8	1

The pressure in the right ventricle is *3.3 kPa* systolic and - *0.133 kPa* diastolic in a healthy, supine person at rest. The pressure in the pulmonary artery is about *3.3 kPa* systolic and *1 kPa* diastolic, with a mean of 1.7 kPa (Box 14-3). The blood flow of the pulmonary capillaries pulsates and its mean pressure is *below 1 kPa*. The pressure in the left atrium is 0.7 kPa. This value implies a pressure drop across the pulmonary circulation of (1.7 - 0.7) = 1 kPa. This *driving pressure* is *less than 1/10* of the systemic driving pressure.

The walls of the pulmonary vessels are thin, hence their pressure must fall at each inspiration, because the intrapulmonic pressure falls.

Change of posture from supine to erect position, will reduce the pressure toward zero in the apical vessels, whereas it increases the pressure in the basal vessels due to gravity.

When the *driving pressure* in the apical blood vessels approaches zero, the blood flow will also approach zero. Apart from its implication for gas exchange, this phenomenon limits the supply of nutrients. Lung disorders often occur in the apical regions.

The *pulmonary vascular resistance (PVR)* is the ratio between the pressure gradient and the bloodflow. A *peripheral resistance unit (PRU)* is measured as driving pressure per bloodflow unit. The basic equation is: *PVR* (PRU) = D*P*/bloodflow (mmHg*s*ml^{-1}).

At rest, the pulmonary driving pressure is 8 mmHg (Box 14-3), and the bloodflow is 5 l per min (83 ml per s). The ratio is 8/80 = 1/10 PRU (normal *PVR* is only 10% of the systemic resistance at rest: *TPVR* = 1 PRU). Calculated in kPa the *PVR* is 1/80 kPa s ml^{-1}. Such low values for *PVR* are only found in the lungs of healthy, non-smokers.

The *PVR* remains low in healthy persons, even when cardiac output increases to 30 l per min, because of distensibility and recruitment of pulmonary vessels. Stretch receptors, found in the left atrium and in the walls of the inlet veins, are believed to be stimulated by distension. Such a distension blocks liberation of *vasopressin* (antidiuretic hormone, ADH) from the *posterior pituitary* and releases atrial natriuretic factor (ANF) from the atrial tissue. Hereby, the urine volume increases and the extracellular volume decreases.

Changes in *pulmonary vascular resistance* are achieved mainly by *passive* factors, but also by *active modification.*

Passive factors: The larger arteries and veins are located outside the alveoli (extra-alveolar); they are tethered to the elastic lung parenchyma, and are exposed to the pleural pressure. The pulmonary capillaries lie between the alveoli and are exposed to the alveolar pressure.

Alveolar capillary volume. The intra-alveolar vessels are wide open at low alveolar volumes, so that their *PVR* must be minimal. With increasing alveolar distension these vessels are compressed. This increases the *intra-alveolar PVR*. However, at low alveolar (lung) volumes, the extra-alveolar vessels are small because of the small transmural vascular pressure gradient, and their *PVR* is high.

With increasing lung distension, the intrathoracic pressure becomes more subatmospheric. This elevates the transmural vascular gradient and is coupled with the radial traction on these vessels by the surrounding lung parenchyma as it expands. Thus, the *extra-alveolar PVR decreases.* The greatest cross-sectional area exists in the many *intra-alveolar* vessels, hence increasing *PVR* in these vessels offsets decreased *extra-alveolar PVR.*

Thus, total pulmonary vascular resistance is increased at higher alveolar volumes when intra-alveolar *PVR* is high. *PVR* is minimal at FRC, where there is air enough to open the extra-alveolar vessels with minimal closure of the intra-alveolar vessels.

Pulmonary artery pressure. A healthy person at rest (FRC) has approximately *half* of the pulmonary capillaries open, but with increasing arterial pressure, the previously closed capillaries open (recruitment). As the arterial pressure continues to rise, the capillaries become distended. The net effect is a rise in the total cross-sectional area of the lung capillaries, leading to decreased *PVR.*

Left atrial pressure. Patients with *high* left atrial pressure have distended capillaries due to the venous backpressure. As a result of the reduced driving pressure their *PVR* is decreased

further.

Gravity. The pulmonary bloodflow per unit lung volume is greatest at the lower and decreases towards the upper lung regions. Gravity creates a gradient of vascular pressures from the top to the bottom of the lungs. The intravascular pressure is much lower at the upper than at the lower lung regions, unlike the *alveolar* pressure, which is essentially constant throughout the lung. At the top of the lung all vascular pressures can approach zero (with the alveolar pressure as reference). Under these conditions there is no bloodflow through the upper region, and if it is still ventilated, it is an *alveolar dead space*.

Active modification is essential: Both sympathetic and parasympathetic fibres sparsely innervate the pulmonary blood vessels. *Sympathetic stimulation* constricts the pulmonary vessels, whereas *parasympathetic stimulation* dilatates them. *Vasoconstrictive agents* include: Arachidonic acid, catecholamines, leucotrienes, thromboxane A, prostaglandin F, angiotensin-II, and serotonin. The *vasodilatators* are acetylcholine, bradykinin, nitric oxide (NO) and prostacyclin.

A *decrease in* P_{AO2} in an occluded region of the lung produces hypoxic vasoconstriction of the vessels in that region as mentioned above. The reduced P_{AO2} causes constriction of the precapillary muscular arteries leading to the hypoxic region. The hypoxic effect is *not nerve-mediated*. This reaction shifts blood away from poorly ventilated alveoli to better-ventilated ones. NO seem to dilatate the vessels of the well-ventilated segments of the lung. Perfusion is hereby matched with ventilation.

14. Regional ventilation

Milic-Emili has developed the elegant *onion skin diagram* of the regional ventilation (Fig. 14-10). The first 25% of the lower abscissa is the residual volume or RV, and this axis shows the total lung capacity (TLC) up to 100% TLC (maximal inspiration). The upper abscissa shows the vital capacity, VC, from zero to 100%. The ordinate is the *regional* ventilation volume in % of the *maximal* regional total lung capacity (TLC). The maximal regional TLC is any given lung region totally filled with air by a maximal inspiration (Fig. 14-10).

The slope of the *onion skin-lines* are constant above FRC, thus the fraction of the tidal volume reaching each lung region, must be constant during the whole inspiration from FRC (Fig. 14-10). The slope is larger in the lower than in the upper lung region, because the lower alveoli are the ones most compressed by the gravity-sensitive pleural pressure. Accordingly, they can distend most during inspiration. The upper alveoli are always more expanded than the lower due to the *pull* of gravity. The upper alveoli follow the *first in - last out* principle. During expiration to residual volume (RV) the upper alveoli are the last to empty (Fig. 14-10). - During inspiration from RV, the lower alveoli are closed up to FRC (*closing volume* and *closing capacity* - see the horizontal blue curve in Fig. 14-10). Around FRC the lower alveoli open.

At the start of the inspiration from FRC the lower alveoli are the smallest, so any inspiration will always distend the lower alveoli most.

Fig. 14-10: The relative, regional ventilation (ordinate) depending upon total ventilation from RV to TLC (modified from Milic-Emili).

The upper alveoli are always expanded by gravity. At TLC all alveoli are assumed to be *maximally* distended (Fig. 14-10). The alveoli and small airways are increasingly distended from the lower to the upper lung regions. As a consequence, their compliance must decrease progressively, and the pleural pressure also decreases towards the top of the lung (Fig. 14-5).

Conclusion:

The multiple inert gas technique has confirmed that the major problems in pulmonary disorders are not true shunts, diffusion barriers, and lamination of alveolar gasses, but dominantly ventilation/perfusion inequality *with functional veno-arterial shunts and alveolar deadspace.*

Pathophysiology

This paragraph deals with *1. Hypocapnia*, *2. Acute hypercapnia* and *3. Vascular lung disorders*. - *Hypoxia* is described in Chapter 15.

1. Hypocapnia

Hypocapnia or *hyperventilation* is a disorder with abnormally reduced P_{aCO2}. The hyperventilation reduces P_{aCO2} and produces an *acute respiratory alkalosis*, characterised by increased pH, and normal or unchanged Base Excess (BE = Zero). Changes in *Base Excess* are effected by renal mechanisms, which take hours to develop.

2. Acute hypercapnia (CO_2-poisoning)

Hypercapnia is a condition, where P_{aCO2} is higher than 6.4 kPa (48 mmHg). Patients with a large dead space and $V°_A$ /$Q°°$ -mismatch develop hypercapnia, due to hypoventilation. Reduced alveolar ventilation increases P_{CO2} and lowers P_{O2}. Since the CO_2 stores are much larger than the O_2 stores, the initial rise of P_{CO2} is lower than the drop in P_{O2}. Thus, the *R*-value must fall, as seen typically in anaesthetic depression of the respiratory centre. The arterial tensions follow the alveolar. The changes in mixed venous tensions are small, because $Q°$ is maintained and the slope of the oxyhaemoglobin dissociation curve is steep at a mixed venous P_{vO2} around 45 mmHg.

The patient with acute hypercapnia is flushing, nervous, horrified of death, and has increasing dyspnoea. The death-horror and hallucinations are followed by loss of consciousness and respiratory arrest. The blood gasses show increased P_{aCO2} and reduced pH (acute respiratory acidosis, Chapter 17) with a base excess of zero.

For patients with chronic pulmonary disease, the hypoxia increases the 2,3-DPG concentration in the red cells, which - together with the hypercapnia and fever - displaces the oxyhaemoglobin curve to the right. This is beneficial for tissue oxygenation, because it increases the tissue tension gradient during oxygen unloading .

Fig. 14-11: The oxyhaemoglobin dissociation curve.

Abnormal blood gas values are indicators of the severity of the disorder. The first phase is characterised by *normal* blood gasses at *rest*. The second phase is *respiratory insufficiency* with abnormal blood gasses at rest (*hypoxia*: P_{AO2} less than 7.3 kPa or 55 mmHg, and *hypercapnia:* P_{aCO2} higher than 6.4 kPa or 48 mmHg). The term *terminal* respiratory insufficiency refers to the grave prognosis.

Hypoxia is dangerous because its effects are irreversible, while hypercapnia is reversible. The oxygen treatment increases P_{aO2}, which is vital, so oxygen therapy should be administered instantly to patients with hypoxia – irrespective of hypercapnia. A few patients may have adverse effects with respiratory arrest, when the hypoxic drive for the peripheral chemoreceptors is eliminated. The ventilation will fall, which elicits a substantial rise in P_{aCO2} with anaesthetic effect on the respiratory centre.

The advantage of *oxygen enriched air* can be shown by an example. A patient with asthma is hospitalised with a P_{aO2} of 5.5 kPa (41 mmHg) and a S_{aO2} of 0.75 (Fig. 14-11). Oxygen enriched air is valuable to such a patient. Oxygen enriched air is administered with a nasal catheter or accurately with a simple plastic mask using the Venturi or Bernoulli principle (Chapter 13).

A small increase in the oxygen concentration of atmospheric air from 21% to 24% leads to a rise in P_{IO2} (3% of 95 kPa is 2.9 kPa; 3% of 713 mmHg is 21.4 mmHg). The major part of this rise reaches the arterial blood (2.6 kPa or 20 mmHg) and this rise in P_{aO2} from 41 to 61 mmHg is often enough to save the patient, because S_{aO2} increases to *0.94* (Fig.14-11). The oxygen flux to the tissues depends upon a normal haemoglobin concentration and a normal cardiac output.

3. Vascular lung disorders

Diseases of the pulmonary vascular tree are diagnosed as pulmonary oedema, pulmonary embolism, and pulmonary hypertension.

3a. Pulmonary oedema is an emergency caused by filtration of fluid out of the pulmonary capillaries into the interstitial space (*interstitial oedema*), and eventually in the alveolar spaces (*alveolar oedema*) – see Fig. 10-10.

The amount of fluid filtered out of the pulmonary capillaries is determined by the Starling equation (Eq. 8-7). The capillary hydrostatic pressure is the main outward force, and this pressure is larger at the base than at the apex of the upright lung. The main inward force is the colloid osmotic pressure of the proteins of the blood. Normally, the alveoli are kept free of fluid, because a net outflux of fluid from the vasculature is balanced by a small lymph flow to the hilar lymph nodes.

Pulmonary oedema has at least 3 causes:

1. *Increased pressure.* Patients with *left cardiac failure* (acute myocardial infarct, chronic myocardial failure, mitral stenosis, aortic stenosis, and hypertension) can drown in their own plasma transudates. The increased venous backpressure distends all pulmonary vessels (lung congestion), and as soon as the pulmonary capillary pressure is higher than the *colloid osmotic pressure* (normally 3.3 kPa or 25 mmHg), there is a filtration of plasma water into the pulmonary interstitial tissues and into the alveoli. The pulmonary vascular pressure rises in the supine position causing attacks of lung oedema to occur at night.

2. *Increased capillary permeability.* Pulmonary oedema can be caused by capillary damage with war gas, toxins, pneumonia etc.

3. *Reduced concentration of plasma proteins* increases net filtration at the arteriolar end of the lung capillary and reduces net reabsorption of filtered fluid at the venular end.

Oedema is particularly serious in the lungs, because it widens the diffusion distance between the alveolar air and the erythrocytes. There is not enough time for oxygen to travel from the air to the individual erythrocyte. Thus, the blood leaving the lungs is only partially oxygenated. Both the VC and the compliance are reduced.

Increased pulmonary capillary pressure is caused by any type of left ventricular failure (acute myocardial infarction or chronic heart failure) and by mitral valve stenosis. A pressure above

2.6 kPa (20 mmHg) causes interstitial oedema, and as the pressure rises above 4 kPa, alveolar oedema develops. Interstitial oedema may not be recognised, but *alveolar oedema* is dramatic.

The patient is severely dyspnoeic, with tachypnoea, tachycardia, and coughing up a frothy pink sputum containing red cells. There is basal crepitation by auscultation and often whistling rhonchi.

Since the fluid-filled alveoli are not ventilated with air, any blood passing them does not participate in gas exchange. The effect is a functional veno-arterial shunt with hypoxaemia, although hypoxic vasoconstriction tends to reduce its size. Initially, the non-affected alveoli are overventilated and P_{ACO2} is low. Hypercapnia is a late complication when the gas exchange is severely compromised.

Other causes of pulmonary oedema include *decreased colloid osmotic pressure* (hypoproteinaemia, overtransfusion), *increased capillary permeability* (pulmonary oxygen toxicity, radiation damage), and *high-altitude oedema*.

Therapy keypoints:

- Primarily, it is important to find the cause of pulmonary oedema, such as left cardiac failure, and correct the disorder.
- Patients with *chronic cardiac failure* have reduced contractility, which improved by positive inotropic agents such as digoxin.
- Patients with *lung oedema* must sit up erect in bed with the legs over the side and calm down. This reduces venous return and cardiac output, and the effective filtration pressure is reduced.
- Breathing of *air enriched with oxygen* reduces hypoxia and dilatates the lung vessels. The filtration pressure is reduced.
- Effective diuretics *increase the excretion of* Na^+ and thus of water via the kidneys. The loss of fluid also implies oedema fluid.
- *Positive pressure breathing* is thought to minimise the difference between the central and the peripheral venous pressure, so the *venous return* and thus cardiac output is reduced. The blockade of lung capillary bloodflow in the overpressure-phase, and the fear of the patient (increases cardiac output) does not make this treatment the best of choice. The effect is probably similar to the earlier application of bloodletting tourniquet to reduce the pressure gradient from the left to the right atrium.

3b. Pulmonary embolism

is caused by detached parts of thrombi from the venous system. The dislodged thrombus is carried with the venous blood to the pulmonary artery, where the lower lobes are frequently affected, due to their relatively high bloodflow.

The lung tissue is ventilated but not perfused, so the gas exchange suffers and hypoxaemia develops. Destruction of lung tissue of the affected area (pulmonary infarction) is rare, due to the continued oxygen supply by the airways and by the bronchial artery.

The condition can develop into *acute cor pulmonale*, which is sudden failure of the right heart.

Immobilisation by prolonged bed rest, local damage of venous walls with thrombophlebitis, and hypercoagulability of the circulating blood are predisposing conditions.

3c. Pulmonary hypertension

is a condition with a mean pulmonary artery pressure above normal (ie. a pressure above 2 kPa or 15 mmHg).

Pulmonary hypertension is caused by *increased* left atrial pressure (left ventricular failure, mitral valve stenosis), *increased* pulmonary bloodflow (congenital heart disease with left-to-right shunting of blood through septal defects or a persistent ductus arteriosus), and by *increased* resistance of the pulmonary vessels (destruction of the capillary bed in emphysema, obstruction in pulmonary embolism, hypoxic vasoconstriction in chronic bronchitis with emphysema and at high altitude).

Persistent pulmonary hypertension leads to right ventricular hypertrophy and finally to *chronic cor pulmonale*. This is often the final stage of not only chronic *obstructive* lung disease in smokers, but also of the late *restrictive* lung disorder.

Equations

- **The Fick cardiac output equation** states that the cardiac output is calculated from the ratio between alveolar oxygen uptake and arteriovenous oxygen content difference:

  Eq. 14-1: $Q^{\circ} = V^{\circ}_{O2}/(C_{aO2} - C_{v\text{-}O2})$.

- **Fick's law of diffusion** states that the flux of gas transferred across the alveolar-capillary barrier is directly related to the **solubility** (Bunsen's a, Table 13-1) of the gas, the diffusion area (A), the length of the diffusion pathway from the alveoli to the blood (L), and the driving pressure $\mathbf{(P_1 - P_2)}$: $\mathbf{J_{gas} = (D \times a \times A \times 1/L) \times (P_1 - P_2)}$. Marie Krogh incorporated molecular weight (mol. weight), a, A, and L in her lung diffusion capacity (D_L). D_L is equal to a constant, K, multiplied with a, and divided by the square root of the mol. weight. Thus $D_L = K \times a/$ Ömol. weight. This relationship is used on all three gasses: $D_{LC0} = K \times 0.018/Ö28$; $D_{LCO2} = K \times 0.51/Ö44$; and $D_{L02} = K \times 0.022/Ö32$. Thus:

  $D_{L02}/D_{LCO} = [K \times 0.022/Ö32]/(K \times 0.018/Ö28) = 1.14.$

  $D_{LCO2}/D_{LCO} = [K \times 0.51/Ö44] /(K \times 0.018/Ö 28) = 22.6.$

Hereby she eliminated all the unknown variables, and for carbon monoxide, Ficks law of diffusion is simplified to:

Eq. 14-2: $\mathbf{(\ J_{gas} =)}\ V^{\circ}_{CO} = \Delta\ \mathbf{P_{CO} \times D_{LC0}}$

- **The alveolar gas equation** $(\ P_{IO2} - P_{AO2}) = P_{ACO2} * [\ F_{IO2} + (1-F_{IO2})/R\ ]$ in terms of alveolar gas tensions. We can simplify the *alveolar gas equation* for R=1:

  Eq. 14-3: $F_{IO2} - F_{AO2} = F_{ACO2}$ or $\mathbf{P_{IO2} - P_{AO2} = P_{ACO2}}$.

- **The alveolar ventilation equation** describes the hyperbolic relationship between alveolar ventilation ($V°_A$) and F_{ACO2} **:**

  Eq. 14-4: $\boldsymbol{V°_A = V°_{CO2} / F_{ACO2}}$.

F_{ACO2} is equal to [P_{ACO2}/(101.3 - 6.3) kPa], so P_{ACO2} is easily substituted for F_{ACO2}.

- **The final ventilation-perfusion ($\boldsymbol{V°_A / Q°}$) equation**

Without showing the calculations, one **equation** combines Eq.s 14-1 to 14-4:

Eq. 14-5: $\boldsymbol{V°_A / Q° = R(C_{aO2} - C_{v\text{-}CO2}) / F_{ACO2}}$.

The $V°_A / Q°$-ratio is obviously independent of the metabolic rate or oxygen uptake.

$\boldsymbol{V°_A / Q°}$- ratio is the **key variable**, because we all have a certain degree of ventilation - perfusion mismatch, and in almost all cardiopulmonary patients this mismatch is consequential.

- **The total tension** of all three dry gasses is equal to ($P_{O2} + P_{CO2} + P_{N2}$) or the barometric pressure (P_B) minus the tension of water vapour in the alveolar air at 37° C. The total tension at P_B=760 mmHg is thus (760 - 47) = 713 mmHg or (101.3 - 6.3) = 95 kPa.

  Eq. 14-6: $(P_B - 47) = (P_{O2} + P_{CO2} + P_{N2})$.

- The **law of mass balance** is applied to both bloodflow and oxygen flux in the following two equations:

  Eq. 14-7: $Q°_{total} = Q°_{shunt} + Q°_{capillary}$

  Eq. 14-8: $(Q°_{total} \cdot C_{aO2}) = (Q°_{shunt} \cdot \mathbf{C_{v\text{-}O2}}) + (Q°_{capillary} * \mathbf{C_{c'02}})$

where $\mathbf{C_{c'02}}$ is the oxygen concentration in the **pulmonary end capillary blood from ideal lung units** (Fig. 14-8).

- The flow and flux relations implies the following **shunt equation**:

  Eq. 14-9: $\boldsymbol{Q°_{shunt} / Q°_{total} = (C_{aO2} - C_{c'02}) / (C_{v\text{-}O2} - C_{c'02})}$.

- The CO-Oxy-haemoglobin affinity equation:

  Eq. 14-10: $\boldsymbol{C_{aCO}/P_{aCO} = 250 * C_{aO2}/P_{aO2}}$.

CO has a standard affinity for haemoglobin 250 times larger than that of oxygen for haemoglobin: $C_{aCO}/P_{aCO} : C_{aO2}/P_{aO2} = 250 : 1$.

Dalton's law states that the *partial pressure or tension of a single gas in a mixture is equal*

- *to the product of the total pressure and the mole fraction (F).* According to Daltons law the fraction of oxygen in the alveolar air (F_{AO2}) is:

  Eq. 14-11: $\boldsymbol{F_{AO2} = P_{AO2}/(101.3 - 6.3) = P_{AO2}/(760 - 47)}$.

With an alveolar partial pressure of oxygen (P_{AO2}) of 13.3 kPa (or 100 mmHg), the F_{AO2} is **0.14**. There is no interaction between gasses.

- **Henry's law** states that the number of gas molecules dissolved in a fluid is directly proportional to the partial pressure of the gas in air above the fluid. According to Henrys law the concentration (C) of dissolved gas is proportional to its partial pressure (*P*) and the solubility (α or Bunsen's solubility coefficient, Box 13-1):

  Eq. 14-12: $\mathbf{C = P * \alpha}$.

With the pressure given in kPa or mmHg it is necessary to divide by 101.3 kPa or 760 mmHg, respectively, because α is defined at 1 atm.abs. pressure.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. The pulmonary vascular pressure and resistance (*PVR*) is only 1/10 of that of the systemic circulation.

B. The *PVR* is highest in intra-alveolar vessels at high lung volumes.

C. The *PVR* increases when pulmonary arterial pressures increase.

D. The pulmonary circulation is dependent on gravity but the pulmonary ventilation is not.

E. The P_{AO2} has a direct effect on pulmonary circulation.

Case History A

A male person, ages 23 and weight 70 kg, is breathing atmospheric air with traces of carbon monoxide (CO) at one atmosphere. The man is at rest, and has an arteriovenous oxygen content difference of 50 ml per l. An arterial blood sample obtained after equilibrium between alveolar air and pulmonary blood is analysed with the following results: P_{aO2} 13.3 kPa (100 mmHg), C_{aO2} 170 ml STPD per l, C_{aCO} (the concentration of CO in the blood) 28.3 ml STPD per l, and the [haemoglobin] 9.18 mM (148 g per l). The standard affinity between haemoglobin and CO is 260 times greater than the standard affinity between haemoglobin and oxygen. The binding capacity for oxygen and CO is 1.34 ml STPD per g of haemoglobin.

1. *Define the concept standard affinity and P_{50}.*
2. *Calculate the dry CO-fraction in the alveolar air (F_{ACO}).*
3. *Calculate the concentration of oxygen in the mixed venous blood of this patient.*
4. *Calculate the concentration of oxygen in the mixed venous blood of a comparable patient with anaemia (haemoglobin concentration 7.78 mM) and with the same arterio-*

venous oxygen content difference.

5. *Is the oxygen supply to the tissues at the venous end of the capillaries better for the CO-poisoned person than for the anaemia patient?*

Case History B

A 49-year-old female, body weight 61 kg and height 1.7 m, is hospitalised due to severe, progressive dyspnoea. Six years ago the diagnosis of pulmonary sarcoidosis was established by mediastinal lymph node biopsy. The cause of the disease is unknown, and the patient has no history of previous lung disease. When stair climbing the patient has difficulties in reaching the 2. floor.

The spirometric standard values for a female of this age, height and weight are: forced expiratory volume on 1 s (FEV_1) of 2.9 l, and forced vital capacity (FVC) of 3.7 l. The patient has a FEV_1 of 1.3, and a FVC of 1.48 l. The patient has an unforced VC of 1.6 l, with an ERV of 300 ml, tidal volume of 600 ml and an IRV of 700 ml, as compared to a normal VC of 3.9 l.

The normal specific lung compliance (at FRC) is 2 ml per Pascal (Pa); for this patient it is determined to only 0.4 ml per Pa at FRC. The normal single-breath CO diffusing capacity is 3 ml STPD s^{-1} kPa^{-1}, but this patient has only 0.5 ml STPD.

An arterial blood sample shows a P_{aCO2} of 4 kPa (30 mmHg) and a P_{aO2} of 8 kPa (60 mmHg).

***1**. What are the arguments for the diagnosis of restrictive lung disease?*

***2.** Why is the single-breath CO diffusing capacity seriously reduced?*

***3**. Is there any indication of alveolar ventilation-perfusion mismatch?*

Case History C

Following 3 days of fishing in cold weather, a 30 year old man is brought to hospital with high fever (40.8 Centigrade), coughing with chest pains and red coloured sputum. Rales are heard over both lungs and a chest x-ray show large infiltrates in both lungs. A blood gas analysis on an arterial sample reveals P_{aO2} of 50 mmHg and

P_{aCO2} of 26 mmHg. pH_a is 7.38. The RQ is assumed to be 1, and P_B is 760 mmHg.

1. *Calculate the alveolar P_{O2} (P_{AO2}) using the alveolar gas equation.*

2. *Assume a likely value for an ideal gas composition (mean alveolar) just before the man became ill.*

3. *Calculate the alveolar (ideal) - arterial P_{O2} difference. What does this difference mean?*

4. *Calculate the difference between the alveolar ideal P_{ACO2} and the arterial (P_{aCO2}). What does this difference mean?*

Case History D

A male, 44 years of age, is brought to hospital due to severe dyspnoea. He has been smoking

40 cigarettes per day in 30 years. Over the last 10 years an increasing respiratory distress has developed, and the patient is well known at the medical department. The arterial blood gas tensions are measured:

P_{aO2} is 60 mmHg (8 kPa), P_{aCO2} is 35 mmHg (4.7 kPa), and pH_a is 7.44.

An alveolar gas sample reveals a P_{AO2} of 129 and a P_{ACO2} of 28 mmHg.

1. *Calculate the alveolar-arterial P_{O2} difference assuming that the ideal P_{AO2} is 100 mmHg (13.3 kPa).*
2. *Provide a likely diagnosis, which explains his respiratory distress.*
3. *Is there an abnormally high alveolar dead space?*

Case History E

A female surgeon, 56 years old, has smoked 25 cigarettes a day for almost 40 years. Her dyspnoea from stair climbing has increased substantially over the last three years as has her morning cough with abundant green sputum in big lumps. A chest X-ray shows hyperinflation, bronchial expansions and a distinct vascular pattern. The surgeon is examined at the respiratory laboratory including function tests and arterial blood gasses with the following results:

FEV_1 = 1.1 l (normal 2.6 l) ; Forced Vital Capacity (FVC) = 1.9 l s^{-1} (normal 3.4 l s^{-1}); P_{aCO2} = 56 mmHg or 7.5 kPa; pH_a = 7.21; P_{aO2} = 49 mmHg or 6.5 kPa; Base Excess = - 5 mM.

1. *What is the cause of the disease?*

2. *Characterise the acute condition including the acid-base status.*

3 *From where in the upper airways do the big lumps of green sputum arise?*

Try to solve the problems before looking up the answers.

Highlights

- *Any patient - as well as any healthy person - has some degree of ventilation-perfusion mismatch.*
- *The regional ventilation-perfusion-ratio is the key to understanding cardiopulmonary function.*
- *The regional ventilation-perfusion ratio varies theoretically from zero at the lower lung region (only bloodflow) to infinity at the upper region (only ventilation).*
- *The upper alveoli are always more expanded than those of the lower due to the pull of the gravity are, and they did follow the first in-last out principle: During inspiration the first to fill – during expiration the last to empty.*
- *The regional ventilation-perfusion ratios show the lower lung regions to be relatively underventilated (ratio below one), the middle lung regions to be well matched (ideal ratio of 1), and the upper lung regions to be relatively overventilated (ratio above 1 and*

approaching infinity.

- *Pulmonary embolism creates an alveolar dead space. The alveolar ventilation of the region is maintained, but there is no bloodflow, so the $V°_A / Q°$ -ratio of the lung region approaches infinity. In the alveolar dead space, alveolar gas pressures approach the levels of inspired air.*

- *Tracheal occlusion represents an extreme mismatch of venous to arterial shunting of blood, namely perfusion with no ventilation at all (ie, the total ratio for the person approaches zero). The arterial blood gas tensions approach those of venous blood.*

- *The pulmonary vascular system is basically a low-pressure, low-resistance, highly compliant vascular system, which is meant to accommodate the entire cardiac output.*

- *The standard affinity of the haemoglobin-CO reaction is 250 times greater than that of haemoglobin-O_2 .*

- *The single-breath CO diffusing capacity (transfer factor) is normally 3 ml STPD s^{-1} kPa^{-1} at rest and 7.5 during maximal exercise.*

- *Uneven distribution of tidal volume can eventuate from uneven resistance to airflow within the lung (bronchoconstriction, collapse and compression of airways) or from uneven regional lung compliance (insufficient surfactant, loss of elastic recoil as in destruction of alveolar tissue, and increase of elastic recoil as in connective tissue scarring or fibrosis with stiff lungs).*

- *Hypoperfusion can be caused by compression of pulmonary vessels, obliteration of vessels by fibrosis, or blockage by emboli or thrombosis.*

- *Functional shunts arise with any consolidation of alveolar regions that continue to have bloodflow (pneumonia, oedema, haemorrhage, cell necrosis, lack of surfactant).*

- *Patients with lung disorders often have $V°_A / Q°$ -mismatch by a combination of serious veno-arterial shunting in the lower lung regions, and increased alveolar dead space in the upper lung regions.*

- *A healthy person at rest (FRC) has approximately half of the pulmonary capillaries open, but with increasing arterial pressure, previously closed capillaries open (recruitment).*

- *Pulmonary oedema is an emergency caused by filtration of fluid out of the pulmonary capillaries into the interstitial space (interstitial oedema), and eventually in the alveolar spaces (alveolar oedema).*

- *Patients with left cardiac failure (acute myocardial infarct, chronic myocardial failure, mitral stenosis, aortic stenosis, and hypertension) can suffocate, when the alveoli are filled with oedema fluid.*

- *The gas exchange of the chronically ill lung patient is reduced over the years, and abnormal arterial blood gas tensions develop already at rest. This late stage of lung disease is called terminal respiratory insufficiency, due to the grave prognosis.*

Further Reading

West, J.B. *Respiratory Physiology*. Williams & Wilkins, Baltimore. USA, 1999.

Änggård, E. "Nitric oxide: mediator, murderer, and medicine." *Lancet* 343: 1199-1206, 1994.

Return to content

Return to chapter 14

Chapter 15.

Blood Gas Transport And Disorders

Study Objectives

- To *define* the oxygen binding capacity and oxygen saturation of blood, the standard affinity between oxygen and haemoglobin and its reciprocal value (P_{50}), hypoxia, cyanosis, and anaemia.
- To *describe* organs with low and high oxygen extraction from the blood, carbon monoxide poisoning, cyanide poisoning, and the measurement principle for blood gas electrodes.
- To *draw* the dissociation curve for oxyhaemoglobin and for carbon dioxide in whole blood and explain the function of the factors involved.
- To *calculate* blood concentrations from partial pressures of gasses (Henrys law), to calculate one variable when relevant variables are given.
- To *explain* the oxygen, carbon monoxide and carbon dioxide binding to haemoglobin, the Bohr-and Haldane-shifts, and the changes from arterial to venous red cells (increased red cell volume and the water and chloride-bicarbonate shift). To explain the oxygen and carbon dioxide stores in the body, the oxyhaemoglobin dissociation and its importance for the oxygen uptake in the lungs and its delivery to the tissues.
- To *use* these concepts in problem solving and case histories.

Principles

- *Oxygen is transported through the thin lung barrier by diffusion, and reaches the respiring cells via bulk transport in the blood and a final diffusion across the capillary barrier to reach the mitochondria of tissue cells.*
- *Fick's law of diffusion (see Chapter 1).*

Definitions

- **Bohr effect**: The dissociation curve for Oxy-haemoglobin is shifted to the right with increasing carbon dioxide or proton concentration in blood.
- **Cyanosis** refers to the bluish colour of skin, nails, lips and mucous membranes. Visible cyanosis implies more than 50 g of the bluish deoxyhaemoglobin per l of mean capillary blood.
- **Haldane effect**: For similar carbon dioxide tensions, the carbon dioxide binding capacity of oxy-haemoglobin is reduced compared to deoxyhaemoglobin.
- **Normal haemoglobin concentration** (mean for males) is 149 g haemoglobin l^{-1} of blood, which is equal to 9.18 mM (internationally accepted as 100%). The normal range is 130-160 g haemoglobin l^{-1} of blood.

- **Oxygen binding capacity** of haemoglobin is defined as the volume of O_2 that binds to 1 g of haemoglobin at the normal, ambient conditions of 20 kPa (150 mm Hg) in partial pressure: *1.34 ml O_2 STPD g^{-1}*.

- **Oxygen saturation of blood** (S). S_{aO2} is the percentage of haemoglobin present as oxyhaemoglobin, normally 0.97-0.99 in arterial blood. The normal saturation of mixed venous blood is 0.75 at rest.

- **Standard affinity** of haemoglobin for O_2 is defined as the reaction rate between oxygen and haemoglobin, when haemoglobin is 50% oxygenated.

- ***P_{50}* :** The O_2 partial pressure at 50% oxygenation of haemoglobin is a practical estimate of standard affinity. A low P_{50} signals high standard affinity and vice versa.

Essentials

This paragraph deals with *1. Transfer of gasses between blood and alveolar air, 2. Haemoglobin and the dissociation curve, 3. Standard affinity, 4. Blood-tissue exchange, and 5. Venous blood transport.*

1. Transfer between blood and alveolar air

Oxygen is transported from the alveolar air to the red blood cells by *diffusion* across the alveolar-capillary barrier. The transit time for the red cells to pass through the approximately 1 mm long lung capillaries (3/4 s) is virtually always adequate for the haemoglobin to become fully saturated with the oxygen of the alveolar gas (Fig. 15-1). Diffusion-limitation is not present in healthy persons and carbon dioxide diffuses better than oxygen. The arterial blood gas tensions are *ideal informers* about the status of blood and lung gas exchange.

A gas such as *carbon monoxide* (CO) that does not equilibrate across the alveolo-capillary barrier, such that its pressure gradient is maintained while the blood is still in transit through a pulmonary capillary, is purely *diffusion limited*, and is not dependent on the bloodflow per se (no perfusion limitation).

Fig. 15-1: Pulmonary tension profiles for different gasses along the lung capillary (at rest and during exercise).

Gasses, such as hydrogen, nitrogen, nitrous oxide, tritium, ^{133}Xe, and anaesthetic gasses, that equilibrate rapidly across the *alveolo-capillary barrier* are *perfusion limited*. Soon after the blood has entered the pulmonary capillary, the partial pressure for the gas in the blood becomes equal to that in the alveoli. No additional diffusion of this gas will occur during the remaining transit time of the blood in the pulmonary capillary.

The pulmonary transfer of O_2 and CO_2 is *perfusion limited* over a wide range of activity levels. Even though CO_2 is 24-times more soluble in water and diffuses 20 times faster through water than does O_2, the two gasses have essentially the same pulmonary equilibration times (which is 0.25 s when blood has passed one third of the way through the capillary at rest). The equilibrium time is due to the small pressure gradient for CO_2 and the time needed for conversion of bicarbonate and carbamino-compounds to dissolved CO_2.

A perfect matching between alveolar air and capillary blood as described above does not occur throughout the whole lung. Even in healthy persons *alveolar ventilation/perfusion inequalities* disturb the ideal exchange.

A variable part of the venous blood reaches the arterial side without passing ventilated regions

of the lungs. This is termed a functional shunt or *veno-arterial shunting.*

2. Haemoglobin and the dissociation curve

The red cell, erythropoiesis and anaemia are described in Chapter 8.

Oxygen is transported by haemoglobin, but also ***carbon dioxide*** (CO_2) and ***hydrogen ions*** (H^+) are being transported in the blood with a minimal pH change**.**

$$Hb_4(O_2)_4 + CO_2 + H^+ \hat{U} Hb_4CO_2(H^+) + 4O_2$$

According to Henry's Law the concentration of physically dissolved O_2, $[O_2]$ is directly proportional, to its partial pressure in the blood. In the lungs, the concentration is high (due to the high partial pressure); hence the reaction is shifted to the left. This results in the release of CO_2 and the oxygenation of haemoglobin. In the muscles, the $[O_2]$ is falling (due to mitochondrial consumption of O_2); hence the reaction is shifted to the right, which results in uptake of carbon dioxide and the release of O_2 and reduced haemoglobin.

The percentage of the total haemoglobin concentration present as oxyhaemoglobin is termed the ***oxygen saturation*** (S_{O2}) of haemoglobin.

The total CO_2-concentration of blood rises with the P_{aCO2}, and the CO_2-*dissociation curve* is almost linear in the physiological range between 40 and 50 mmHg (Fig. 15-2). In mixed venous blood the oxygen saturation (S_{O2}) is 75% and the CO_2-concentration is larger for a given P_{CO2} than in arterial blood (Fig. 15-2). This is a general relationship called the *Haldane effect*: Oxyhaemoglobin is not able to bind as much CO_2 as oxygen-free haemoglobin. The Haldane effect facilitates the release of carbon dioxide in the lungs (Fig. 15-2).

The reaction of oxygen with haemoglobin follows the law of mass action. The concentration of oxygen in physical solution is proportional to the oxygen partial pressure (Eq. 13-6). Thus, the concentration of physically dissolved oxygen determines the relative amounts of haemoglobin and oxyhaemoglobin or the S_{O2}, which varies from zero to 100%. The Oxy-haemoglobin dissociation curve is shown in Fig. 15-2. As the P_{CO2} of blood increases, the affinity of oxygen for haemoglobin is reduced. Since the pH of blood is closely related to the P_{CO2}, the following statement is also true: As the pH of blood decreases, the affinity of oxygen for haemoglobin is reduced. Hereby, the slope of the oxyhaemoglobin dissociation curve decreases and shifts to the right. This phenomenon is called the *Bohr effect,* and the effect assists the exchange of oxygen in the tissues (Fig. 15-2).

Fig. 15-2: The concentration curves for oxygen and for carbon dioxide in blood as a function of increasing blood gas tensions. The Haldane and the Bohr effects are shown.

The reaction between protons and bicarbonate is shifted to the ***right*** when the H^+-concentration rises:

$$H^+ + HCO_3^- \hat{U} H_2CO_3 * \hat{U} H_2O + CO_2.$$

The last reaction is catalysed by *carbonanhydrase* (*).

In fully oxygenated pulmonary blood, the oxygen concentration is approximately 200 ml STPD per l (S_{O2} = 1.0 or 100%) at an atmospheric oxygen partial pressure of 20 kPa and a P_{aO2} of 13.3 kPa (Fig. 15-3). This fully oxygenated blood is mixed with venous blood that passes through the *physiological shunt* on its way to the left heart. Thus, the oxygen tension falls to 12.6 kPa (95 mmHg) in arterial blood from 13.3 kPa at the pulmonary capillaries and S_{aO2} is only 0.985 (Fig. 15- 3).

The solubility coefficient of oxygen is 0.022 ml STPD per ml per atmosphere (found in Box 13-1). From the dissociation curve, S_{aO2} is read to be 0.985 (Fig. 15-3). The chemically bound O_2: (0.985 × 200) is thus 197 and the physically bound is [(0.022 * 12.6/101.3)*1000] *2.75*, or a total of ***199.75*** ml STPD per l of blood.

Fig. 15-3: The Oxy-haemoglobin dissociation curves. The O_2 - pressure is given in kPa and in mmHg.

3. Standard affinity

Standard affinity of haemoglobin for O_2 binding is defined as the *reaction rate* when haemoglobin is 50% oxygenated. The standard affinity of haemoglobin for oxygen is therefore the slope of the dissociation curve when haemoglobin is 50% oxygenated. A practical expression for standard affinity is the O_2 partial pressure at 50% oxygenation - called P_{50} (ie, the reciprocal value of standard affinity). A metabolite of anaerobic glycolysis called 2,3- DPG (*diphosphoglycerate*) is highly concentrated in erythrocytes (1 mol per mol of Hb_4). Red cells contain a 2,3-DPG mutase and lack mitochondria, so they are confined to anaerobic glycolysis. The highest 2,3-DPG values are found during hypoxia and during exercise. The 2,3-DPG, like H^+ and CO_2, facilitates the unloading of O_2 from haemoglobin by reversible changes of its molecular configuration at low oxygen tensions (see Fig. 8-3).

The following factors shift the oxygen dissociation curve to the *right* (increase P_{50}): 1. Increasing P_{C02} (the Bohr effect), 2. increasing [H^+], 3. increasing 2,3-DPG, 4. increasing temperature. The metabolic activity of the cells augment these factors, and means that less O_2 is bound to haemoglobin at a given P_{aO2} (metabolic activity facilitates unloading of O_2).

The factors that shift the oxygen dissociation curve to the *left* (low P_{50}) are 1. Increasing O_2 tension (the Haldane-effect), 2. decreasing [H^+] and P_{aCO2}, 3. low temperature in the lungs. 4. Increasing CO tension, 5. Reduced 2,3-DPG production in chronic acidosis, and by 6. Foetal haemoglobin. The leftward shift in the lung means that more O_2 is bound to haemoglobin at a given P_{aO2} (facilitates the binding of O_2 to haemoglobin).

Foetal haemoglobin (F) contains 2 gamma-chains in stead of the 2 b-chains of Adult (A) haemoglobin. Haemoglobin F binds 2,3-DPG less tightly than does heamoglobin A, and therefore oxygen more tightly. The sigmoid dissociation curve of haemoglobin F is clearly shifted to the left (low P_{50}) relative to haemoglobin A. Also species differences in the globin part affect the curve.

Tissue diffusion is the rate-limiting process for oxygen transport. The oxyhaemoglobin saturation curve (Fig. 15-3) shows that the P_{O2} must fall appreciably before an important quantity of oxygen dissociates from haemoglobin. The gas tensions of the mixed venous blood provide the best information about the respiratory gas tensions of the tissues. However, mixed venous blood is obtainable only by invasive procedures.

4. Blood-tissue-exchange

Fick's diffusion law governs the transfer of gas. This law states that the amount of gas transferred from blood to a certain mitochondrial site is *directly related* to the partial pressure gradient (D*P* = mean systemic capillary minus mean tissue oxygen tension), area of the systemic capillary barrier (A), and solubility of the gas (a, see Box 13-1). The flux of gas is *inversely related* to the length of the diffusion pathway from the capillaries to the mitochondria (L), and the square root of the molecular weight of the gas.

Marie Krogh incorporated solubility (Bunsen's a), molecular weight (mol. weight), A, and L in

her *lung diffusion constant* (D_L) and by analogy we can define a mitochondrial diffusion constant (D_m) equal to the solubility (a) divided by the square root of the mol. weight.

The diffusion distance from the capillary blood through the tissues to the mitochondria is long in many organs. Here, the conditions are less favourable and diffusion-limitation is sometimes present in cells at the far end of the capillaries, although their mitochondria can maintain oxidative metabolism even at 1 mmHg (0.133 kPa).

The diffusion volume of O_2 per time unit is $\boldsymbol{V°_{O2}}$, and this must equal ($D_m \times DP$) – see Eq. 14-2.

Oxygen equilibrates rapidly across the peripheral tissue barrier, but its transport is *perfusion limited*.

The DP depends upon the bloodflow through each capillary, the density of open capillaries, and their degree of dilatation. The P_{aO_2} is kept high by O_2 released from oxyhaemoglobin, and the steep dissociation curve shows the substantial O_2 delivery to the mitochondria.

Apart from the factors leading to increased P_{O_2}, all factors, which imply a rise in $V°_{O2,}$ are also important.

The force driving oxygen to the most distant mitochondria is often small. Fortunately, mitochondria have the capacity to maintain oxidative metabolism at a tissue tension as low as *133 Pa* or *one mmHg.* Even such a low P_{O2} increases the rate of the respiratory chain events.

Most tissues have a diffusion pathway of 1-10 mm or more. This is particularly critical in *brain and heart tissues*, whereas the even longer diffusion distance in resting muscle tissue is of minor consequence.

The only way in which the diffusion distance can be reduced is by *recruitment* of more capillaries at increasing demand. This is particularly important in skeletal muscles during exercise, where the capillary density increases threefold. Such a rise *increases the systemic, capillary surface area*, simultaneously with the decrease in diffusion distance.

In the muscle capillaries bloodflow is interrupted during the contraction phase, and the tissue-P_{O_2} falls toward zero. At low P_{O_2}, the gradient of the dissociation curve of myoglobin is at its steepest. Hence, myoglobin releases its O_2 readily. During muscular relaxation, bloodflow is restored and myoglobin is rapidly reloaded with oxygen. The myoglobin dissociation curve has a P_{50} some five-fold lower than that of haemoglobin.

The *diffusion constant* is equal to a divided by the square root of the mol. weight. The square root of the molecular weights of oxygen and carbin dioxide (square root of 32/44) is 0.85. This result is obtained by solving the equations for CO_2 and O_2. The ratio of their solubility is 0.51/0.022 = 23.2. The *rate of CO_2 diffusion* is twenty-fold (23.2 × 0.85 = 20) greater than that of O_2.

The flat plateau part going towards right (Fig. 15-3) provides a rapidly available O_2 store. The P_{aO2} can fall from 13.3 to 8.7 kPa (100 to 65 mmHg) with little change in S_{aO2}. Almost the same mass of O_2 will attach to haemoglobin. The steep, middle portion of the curve allows cell mitochondria to extract relatively large quantities of O_2 from haemoglobin with relatively small changes in the P_{aO2} - and P_{aO2} is high enough to maintain diffusive force.

The heart at rest consumes 140 ml O_2 per l of blood out of the 200 ml per l that is supplied by the coronary arteries. The venous oxygen saturation in the coronary sinus is therefore (200 - 140)/200 = 0.30 at rest. Thus, increases in the myocardial O_2 uptake during exercise are

mainly met by an increase in bloodflow.

Intensively working skeletal muscles - just like the myocardium at rest - utilise 140- 170 ml O_2 per l of blood. The main part of the cardiac output is directed for exercising muscles during severe work. The *arterio-venous O_2 content difference* for the total muscle mass will become almost equal to that of the entire body.

The typical capillary passage period for blood is 0.75-1 second (s). Oxygen is rapidly released from haemoglobin. At the venous end of the capillary the most distant mitochondrion has the poorest oxygen supply. Oxygen diffuses through the red cell membrane, blood plasma, and the endothelial cell into the ISF with a rate, which is 20 times slower than that of CO_2 diffusion. The lipid solubility of O_2, CO_2, and N_2 are high. These gasses dissolve in the lipid layer of cellular membranes and diffuse into the cell. Oxygen diffuses to the mitochondria without barrier problems as long as the critical P_{O2} gradient is maintained. Hypoxic tissue may function down to a tissue P_{O2} of only 1 mmHg.

5. Transport in the venous blood

Carbon dioxide is the most important final product of cellular metabolism, since it operates as a *control molecule* in essential regulatory processes linking the size of ventilation to cardiac output and influencing vascular resistance. The venous blood carries carbon dioxide to the lungs, where it is eliminated in the expired air by ventilatory effort. Our blood contains large quantities of carbon dioxide. The [total CO_2] in arterial blood is 500 ml STPD per l (22.3 mM), and in venous blood 540 ml STPD per l (24 mM). The solubility coefficient for CO_2 is 0.51 ml STPD per ml and per 101.3 kPa (see Box 13-1), P_{aCO_2} is 5.3 and the mixed venous P_{CO_2} is 6.1 kPa (Box 14-1).

There are three types of *CO_2 transport* from the cells to the lungs:

a. *The physically dissolved* CO_2 concentration is calculated by application of Henrys law: (0.51 × 1000* D P/101.3). One mmol equals 22.4 ml. The results are for arterial blood: 26.7 ml STPD per l (or 26.7/22.4= 1.19 mM), and for mixed venous blood: 30.7 ml STPD per l (or 30.7/22.4= 1.37 mM). The difference (30.7 - 26.7) equals 4 or 4/22.4 = 0.18 mM dissolved CO_2.

Because CO_2 is 24 times more soluble than O_2, ***dissolved CO_2*** is a more significant form of transport in the venous blood than is dissolved O_2 in the arterial blood.

b. *Carbamino-haemoglobin.* Terminal amine groups, primarily on haemoglobin, react with CO_2: $Hb\text{-}NH_2 + CO_2 \Þ Hb\text{-}NH\text{-}COO^- + H^+$. The concentration of CO_2 binding compounds in arterial blood is 1 mM and in mixed venous blood 1.4 mM, so carbamino compounds account for *0.4 mM* of the CO_2 transported in blood.

c. *Bicarbonate.* Some of the dissolved CO_2 reacts with water, forming carbonic acid, which immediately breaks down to bicarbonate and H^+:

$$CO_2 + H_2O \hat{U} * H_2CO_3 \hat{U} H^+ + HCO_3^-.$$

This reaction is slow in plasma, but more than 10 000 times faster in the erythrocytes because of the presence of the enzyme *carbonanhydrase* (*), which catalyses the hydration of CO_2 to carbonic acid. The H^+ is buffered by the partially oxygenated haemoglobin. The bicarbonate diffuses out of the erythrocytes into plasma, thus preventing an accumulation, which would slow down the hydration of CO_2 according to the law of mass action. Of the total CO_2 in venous blood 2/3 occurs in the form of bicarbonate, and most of this is found in the *plasma*

sink (Fig. 15-5) following exchange with Cl^- (Fig. 15-5).

The actual [bicarbonate] is calculated as follows:

[Bicarbonate] = [Total CO_2 - carbamino CO_2 - dissolved CO_2].

Arterial blood [bicarbonate]: 22.3 - 1 - 1.19 = *20.1 mM.*

Venous blood [bicarbonate]: 24 - 1.4 - 1.37 = *21.2 mM.*

Fig. 15-4: Gas exchange between air and blood. Reactions in plasma and erythrocytes.

As CO_2 is being removed from the blood to the alveolar air, the equilibrium is shifted toward the formation of CO_2 from bicarbonate according to the law of mass action (Fig. 15-4). This process, plus the binding of oxygen to haemoglobin causes both CO_2 and H^+ to dissociate from haemoglobin (Fig. 15-4). The released H^+ then combines with bicarbonate to form more CO_2 for diffusion (Fig. 15-4). Carboanhydrase (CA or *) in the red cell speeds up the CO_2 formation.

The CO_2 from the cells enters the capillary blood, and is buffered primarily by haemoglobin (Fig. 15-5). Within milliseconds the red cell Carboanhydrase catalyses its hydration to form H_2CO_3, which dissociates to form H^+ and bicarbonate. To avoid accumulation of bicarbonate, two out of three newly formed HCO_3^- is exchanged for Cl^- in plasma, through *anion antiporters* in the membrane (Fig. 15-5). This way the membrane potential is sustained. This *bicarbonate-chloride shift* occurs within milliseconds and is vital for the CO_2 transport (Fig. 15-5). The bicarbonate-chloride shift occurs simultaneously with, and probably independent of, the release of Na^+ and HCO_3^- from erythrocytes into plasma probably by *Na^+-H^+ exchange.*

Fig. 15-5: CO_2 diffusion from tissues to blood and oxygen diffusion from blood to tissues.

CO_2 is also rapidly bound to the amino groups of deoxyhaemoglobin (Hb) to form carbamino-groups and H^+ within the erythrocyte (Fig. 15-5).

When oxyhaemoglobin loses its O_2, it can accept H^+ , which combines with the NH moiety on the imidazol ring: O_2-Fe.Hb-NH + H^+ Þ O_2 + Fe.Hb-NH_2^+. This allows more CO_2 to be hydrolysed and be carried in the *plasma sink* as bicarbonate (Fig. 15-5). Thus, at a given P_{CO_2} , the CO_2 concentration at the venous end is greater than in arterial blood (the Haldane effect).

There is a small rise in [CO_2 + H_2CO_3] of 0.2 mM dissolved CO_2 (Fig. 15-5).

Venous erythrocytes contain *more osmotically active particles* than arterial red cells. Consequently, they swell because ***water*** from the plasma enters the venous erythrocytes by *osmosis* (Fig. 15-5). The venous haematocrit (PCV) is larger than the arterial, but the advantage of this water transport is uncertain. This so-called *water shift* is almost solely due to the shift of anions, secondary to the establishment of a new Donnan equilibrium.

Without haemoglobin the transport would be minimal. With haemoglobin most of the *CO_2 transport is isohydric* (it takes place without any change in H^+ concentration). The actual rise in venous H^+ is also buffered by haemoglobin as by any buffer present. However, the function of haemoglobin is not *vital* for the CO_2 transport.

Pathophysiology

This paragraph covers 1. Hypoxia, 2. Cyanosis, and 3. Carbon monoxide poisoning.

1. Hypoxia

Hypoxia denotes *oxygen deficiency* at the mitochondrial sites due to insufficient delivery of oxygen (low P_{aO2}) or inability to utilise oxygen (normal P_{aO2}). *Hypotonic hypoxia* is characterised by a P_{aO2} less than 7.3 kPa (55 mmHg). Below this threshold the ventilation starts to increase by carotid body activity.

Acute hypoxia with low P_{aO2} stimulates the carotid bodies. This triggers a rise in ventilation (primary hyperventilation). The hyperventilation reduces P_{aCO2} and $[H^+]$, which limits the initial rise in ventilation, because it decreases the carotid body and central chemoreceptor stimuli. Symptoms relate to:

- CNS: Headache, nausea, blacks out, cramps and unconsciousness (grey out).
- The peripheral tissues: *Cyanosis* of the skin and the mucous membranes (see below).

Chronic hypoxia increases breathing in another way. The primary hyperventilation leads to an acute respiratory alkalosis. This disorder is partially compensated by renal excretion of bicarbonate. Hereby, the $[H^+]$ returns toward normal. The low [bicarbonate] in the extracellular fluid, including brain interstitial fluid, is partially replaced by lactate from the hypoxic brain. The carbon dioxide response curve is shifted to the left and much steeper than normal. The symptoms of chronic hypoxia are: Nausea, vomiting, lack of appetite, increased ventilation, increased erythropoiesis, increased brain bloodflow, right ventricular failure or cor pulmonale (*lung-heart*) and mountain sickness.

The two types of hypoxia have quite different causes and consequences (Box 15-1).

Box 15-1. Types of hypoxia and their causes.

A **Hypoxia with low P_{aO2} = Hypotonic hypoxia** (P_{aO2} less than 7.3 kPa or 55 mmHg)

- Alveolar ventilation-perfusion mismatch
- Venous-to-arterial shunts (between vessels or within the heart)
- Diffusion impairment
- Low O_2 tension in inspired air (low P_B at altitude or low F_{IO2})

B **Hypoxia with normal P_{aO2}** **(**P_{aO2} between 8 to 14 kPa or 60-105 mmHg**)**

Ischaemic Hypoxia (reduced bloodflow**)**

- Generalised ischaemic hypoxia with low cardiac output
- Stenosis of the coronary arteries (angina pectoris)
- Stenosis of leg arteries (intermittent claudication)

Anaemic Hypoxia (reduced oxygen-binding capacity of blood)

- Anaemia (blood haemoglobin below 8 mM or 130 g l^{-1})
- Met-haemoglobinaemia (reductase deficiency due to sulphonamides)

CO-haemoglobinaemia (cherry red skin)

Histotoxic Hypoxia (insufficient oxygen extraction in tissues)

Cyanide poisoning (mitochondrial block)

ATP insufficiency (dinitrophenol poisoning, hyperthermia exhaustive work, universal cramps)

A. Hypotonic hypoxia. The most frequent cause of *hypotonic hypoxia* is *alveolar ventilation-perfusion mismatch* caused by chronic bronchitis with emphysema (COLD) in smokers. A typical patient is depicted in Fig. 15-6. The mixed venous oxygen tension is low. Hypotonic or hypobaric hypoxia is caused by insufficient oxygen uptake into the blood from the lungs. Insufficient oxygen uptake also occurs in space, during flying, at altitude and during diving. Hypotonic hypoxia is defined as a P_{aO2} of less than 7.3 kPa (55 mmHg). Below this threshold the ventilation starts to increase.

Fig. 15-6: Arterial (a) and mixed venous (v) oxygen tensions on the oxyhaemoglobin dissociation curve of a patient with chronic bronchitis and emphysema.

To optimise the efficiency of gas exchange and achieve normal blood gas levels, the alveoli have to be both adequately ventilated and perfused. Abnormal mismatching leads to abnormal $V°_A /Q°°$ -ratios and to subsequent hypoxaemia and hypercapnia.

Persons with *venous-to-arterial shunts* and persons with *diffusion impairment* all have low P_{aO2} values. Hypoxia also hits persons breathing air with reduced P_{I02} (reduced P_B or reduced F_{I02}).

B. Anaemic hypoxia. The disorder is due to an insufficient oxygen carrying capacity of the haemoglobin. A blood [haemoglobin] below 130 g per l (8 mM) implies reduced working capacity and thus a consequential anaemia.

The normal blood [haemoglobin] is 149-150 g * l^{-1}. Smokers expose their bodies to anaemic hypoxia and to the development of lung cancer. **Anaemic hypoxia** is a common phenomenon (ex: *anaemia and cyanosis; methaemoglobinaemia; cherry-red skin colour in CO-haemoglobinaemia).*

Ischaemic or stagnant *hypoxia* is caused by insufficient bloodflow (cardiac insufficiency or local ischaemia). Stenosis of the coronary arteries leads to *chest pains* or *angina pectoris.* Arteriosclerosis and stenosis of the leg arteries causes *claudication* (ie, intermittent walking due to hypoxic pain). The poor circulation in shock conditions is a *generalised hypoxic ischaemia*.

Histotoxic hypoxia is caused by insufficient capacity for oxygen utility. This is caused by blockade of the mitochondrial metabolism as in cyanide poisoning or by blockade of the ATP-production following supramaximal exercise, where the oxygen utility is excessive.

2. Cyanosis

Cyanosis means *dark blue* in Greek and refers to the *bluish colour* of skin, nails, lips and mucous membranes of healthy persons or patients. The *bluish* deoxyhaemoglobin in the capillaries determines the degree of visible cyanosis. The amount of deoxyhaemoglobin in the middle of the capillary is a likely indicator. Let us assume that 50% of the deoxygenation occurs in the first half of the capillary, so the *mean* capillary oxyhaemoglobin concentration is reached here. The mean capillary blood must contain more than 50 g of the bluish deoxyhaemoglobin per l blood in order to be visible (Fig.15-7). The *mean capillary*

concentration of deoxyhaemoglobin ([Hb] in g l^{-1}) is calculated from Eq. 15-2.

Cyanosis has one of two causes, which may act together: Reduction of the arterial saturation (low S_{aO2}) or increase in the *arterio-venous oxygen content difference.* A large oxygen extraction is seen in disorders with low bloodflow. The standard extraction at rest is 50 ml (STPD) of O_2 per l or 2.23 mmol l^{-1}. Halfway in the capillary 25 ml of oxygen has been extracted per l (assuming linear extraction). This corresponds to 18.66 g bluish deoxyhaemoglobin per l.

The standard extraction relates to a desaturation of the capillary blood of (50* 1.34) = 67 ml of oxygen per l of whole blood.

Fig. 15-7: Capillary model wieh calculations of visible cyanosis in a normal person, anaemia and polycythamia.

A *patient* with a normal [Total-Hb] of 148 g l^{-1} or 9.18 mM and standard extraction, reaches the *cyanosis threshold* already at an arterial saturation around 0.79 or 79% (Eq. 15-2).

Patients with anaemia only have a minor tendency to become cyanotic. A patient with anaemia and a [Total-Hb] of 110 g l^{-1}or 6.8 mM, must reduce S_{aO2} to less than 0.72 (72%) in order to reach the cyanosis threshold. Chronical patients with less than 50 g of haemoglobin have no chance of being cyanotic when alive, because the haemoglobin concentration is too small.

Polycythaemia patients have lots of red cells and bind large amounts of oxygen. Because of their high [Total-Hb] they become easily cyanotic, despite being well oxygenated. Again it is the absolute amount of deoxyhaemoglobin in the mean capillary blood that determines the degree of cyanosis.

Cyanosis is a superficial indicator of hypoxia. A person leaving a cold bath may have dark blue skin and still be well-oxygenated in vital organs.

3. CO poisoning

Combustion in insufficiently ventilated areas leads to accumulation of carbon monoxide (CO). Even minimal concentrations of CO in the inspired air prevent the formation of oxyhaemoglobin. In countries with 5% CO in the domestic gas, inhalation of gas is used for suicide. Following one deep inspiration the victim loses consciousness and is chemically dead even before he falls to the floor. Severe cases survive in coma, which is frequently irreversible. Patients with light carbon monoxide poisoning complain of headache, nausea and vomiting, and have the classical cherry-red skin. The poison can damage the heart and the lungs as well. Prophylaxis is essential, and therapy consists of removing the victim from the CO, and of administering oxygen - if necessary hyperbaric oxygenation.

Equations

- **Fick** proposed that the cardiac output can be calculated as the oxygen uptake divided by the arteriovenous oxygen content difference:

$$\text{Eq. 15-1:}\quad \dot{Q} = \dot{V}_{O2}/(C_{aO2} - C_{vO2})$$

Standard data for a healthy person at rest are an O_2 uptake of 250 ml STPD min^{-1} and an O_2 extraction of 50 ml STPD l^{-1} or 25% of C_{aO2}.

- The **mean capillary concentration** of deoxyhaemoglobin ([Hb] in g l^{-1}) is easily calculated from the blood haemoglobin concentration [Total-Hb]:

Eq.15-2: [mean deoxy-Hb] = [Total-Hb] * (1 - S) + 0.5 * (C - C)/1.34.

aO2 aO2 vO2

The oxygen saturation of the arterial blood is S_{aO2} and [0.5 * (C_{aO2} - C_{vO2})] is 50% of the total oxygen extraction. The constant *1.34* is ml of oxygen per g haemoglobin (the normal oxygen binding capacity).

Self-Assessment

Multiple Choice Questions

I. A 30-year-old female is anaemic with a haemoglobin concentration of 65 g l^{-1}.

Each of the following 5 statements concerning her condition have True/False options:

A. A normal P_{aO_2}

B. A normal S_{aO_2}

C. A rightward shift of the oxyhaemoglobin dissociation curve

D. A smaller oxygen capacity than normal

E. A P_{aO_2} below normal.

II. **Each of the following five statements have True/False options:**

A. Most of the carbon dioxide in the blood is transported in the form of bicarbonate in the plasma.

B. Near the terminal bronchioles the movement of air is accomplished by diffusion.

C. When blood flows through tissue capillaries carbon dioxide is released from haemoglobin.

D: Pulmonary surfactant helps to equalise the distending pressure in alveoli of different sizes by reducing their surface tension.

E. The conducting zones of the airways constitute the anatomic dead space.

III. Each of the following five statements have True/False options:

A. Standard affinity is defined as the reaction rate between oxygen and haemoglobin, when haemoglobin is 50% oxygenated

B. Some CO_2 is transported dissolved in plasma and erythrocytes

C. Some CO_2 is transported as bicarbonate in erythrocytes

D. More than 1% of the CO_2 is transported as carbonic acid

E. Some CO_2 is transported as carbamino compounds in plasma.

Case History A

A female, 24 years of age, has a P_{AO_2} of 13.3 kPa and a F_{ACO_2} of 0.056. Her P_{CO2} in the mixed venous blood is 46 mmHg or 6.1 kPa and her P_{O_2} is 42 mmHg or 6.0 kPa. The solubility coefficients (α) are 0.022 and 0.51 ml STPD per ml for O_2 and CO_2, respectively. The barometric pressure is 101.3 kPa, and the water vapour tension in the alveolar air is 6.2 kPa.

1. *Calculate the partial pressure of carbon dioxide in the alveolar air.*

2. *Develop an equation showing the relation between gas concentration © and partial pressure (P) in a fluid (Henry's law).*

3. *Calculate the concentration of oxygen and carbon dioxide physically dissolved in one litre of arterial blood in equilibrium with the alveolar air of the above female.*

4. *Calculate the concentration of CO_2 and O_2 physically dissolved in her mixed venous blood.*

Case History B

A male, age 25 years, exercises with a maximum oxygen uptake of 4.5 l per min, a C_{aCO2} of 500 ml STPD per l, C_{aO2} of 200 ml STPD per l, and in the mixed venous blood a C_{CO2} of 650 ml STPD per l. The C_{O2} in his coronary sinus is measured at 30 ml STPD per l, and his myocardial oxygen consumption is 420 ml STPD min^{-1}. His RQ is 0.9.

One hour later, the male is at rest with cardiac output of 5450 ml per min, and oxygen uptake 273 ml STPD min^{-1}.

1. *Calculate his carbon dioxide output and his cardiac output during exercise.*

2. *Calculate the coronary bloodflow during exercise.*

3. *What are the energy sources of the heart during exercise and during rest.*

4. *Calculate the concentration of oxygen in his mixed venous blood at rest, and thus the arteriovenous oxygen difference.*

Case History C

A 20-year old person is unconscious after a traffic accident and brought to hospital. No focal lesions are found. A cardio-pulmonary screening reveals the following: Metabolic rate: 10.450 MJ daily, respiratory frequency 14 per min, P_{AO_2} 4.5 kPa and P_{ACO_2} 5.3 kPa. In the mixed expiratory air P_{EO_2} is 15.6 and P_{ECO_2} 4.4 kPa.

Cardiac catheterisation includes the pulmonary artery with mixed venous oxygen and carbon dioxide concentrations: 150 and 542 ml STPD per l of blood.

The gas concentrations of the arterial blood are: C_{aO_2} 200 and C_{aCO_2} 500 ml STPD per l (the radial artery). The dietary energetic equivalent for O_2 is here 20.6 kJ per l STPD. The barometric pressure is 101.3 kPa and the tracheal water vapour tension is 6251 Pa.

1. *Calculate the respiratory quotient (RQ).*

2. *What can be inferred from this RQ value?*

3. *Calculate the oxygen uptake of the patient.*

4. *Calculate the alveolar ventilation and the dead space.*

5. *Calculate the cardiac output of the patient.*

6. *Is the unconsciousness caused by respiratory, cardiovascular or central nervous system malfunction?*

Case History D

A healthy person at rest has an O_2 uptake of 250 ml STPD min^{-1} and an arterio-venous O_2 content difference of 50 ml STPD l^{-1}. With a normal concentration of O_2 in the arterial blood (C_{aO2}) of 200 ml STPD l^{-1} this means that 25% is utilised in the tissues. The body temperature is normal and P_{aO2} is 13 kPa. The solubility coefficient for oxygen in body- warm plasma is 0.022 ml STPD per ml and per atmosphere (101.3 kPa or 760 mmHg).

1. *Calculate the cardiac output at rest.*

2. Calculate the concentration of dissolved oxygen in the arterial plasma of this person.

3. Assume that the person is still utilising 25% of the arterial oxygen, and that all the blood is deprived of red cells. Now the person has to survive without haemoglobin. Calculate the hypothetical cardiac output, which would be necessary to provide the cells with sufficient oxygen.

***4.** Is life without haemoglobin possible for humans ?*

Try to solve the problems before looking up the answers.

Highlights

- *The concentration of physically dissolved oxygen is directly proportional to its partial pressure in the blood.*
- *A metabolite of anaerobic glycolysis called 2,3- DPG (diphosphoglycerate) is highly concentrated in erythrocytes. The highest values are found during hypoxia and during exercise - about 1 mol per mol of Hb_4.*
- *The 2,3-DPG, like H^+ and CO_2, facilitates the unloading of O_2 from haemoglobin by reversible changes of its molecular configuration.*
- *The following factors shift the dissociation curve to the right (increase P_{50}): 1. Increasing P_{C02} (The Bohr effect), 2. increasing $[H^+]$, 3. increasing 2,3-DPG, 4. increasing temperature.*
- *The factors that shift the curve to the left (low P_{50}) are: 1. Increasing O_2 tension, 2. decreasing $[H^+]$ and P_{aCO2}, 3. low temperature in the lungs. 4. Increasing CO tension, 5. Reduced 2,3-DPG production in chronic acidosis, and by 6.Foetal haemoglobin.*
- *The leftward shift in the lung means that more O_2 is bound to haemoglobin at a given P_{aO2} (facilitates the binding of O_2 to haemoglobin).*
- *The arteriovenous O_2 content difference is typically 50 ml STPD l^{-1} or 25% of C_{aO2} at rest with a cardiac output of 5 l min^{-1}.*
- *The pulmonary transfer of O_2 and CO_2 is perfusion-limited over a wide range of activity levels.*
- *The tissue transfer of O_2 and CO_2 is diffusion-limited over a wide range of activity levels.*
- *The oxygen capacity of haemoglobin is defined as the volume of O_2 that binds to 1 g of haemoglobin at the normal, ambient conditions of 20 kPa (150 mmHg) in partial pressure: 1.34 ml O_2 STPD g^{-1}.*
- *Hypoxia denotes oxygen deficiency of the tissues due to insufficient access or insufficient oxygen utility. Hypoxia is present both with low and normal P_{aO2}.*
- *Cyanosis means dark blue in Greek and refers to the bluish colour of skin, nails, lips and mucous membranes of healthy persons or patients. The absolute amount of the bluish deoxyhaemoglobin in the capillaries determines the degree of cyanosis.*

- *Combustion in in sufficiently ventilated areas leads to accumulation of carbon monoxide (CO).*
- *The standard affinity between CO and haemoglobin is 250 fold larger than for the binding of oxygen, so even small concentrations of CO in the inspired air prevent the formation of oxyhaemoglobin.*
- *Inhalation of domestic gas (4-5% CO) is sometimes used for suicide and suicide attempts. - Following one deep inspiration, the victim loses consciousness and is chemically dead even before he falls to the floor. - Some cases survive in coma, which is frequently irreversible.*

Further Reading

Nature. Weekly journal published by Macmillan Magazines Ltd. Porters South, 4 Crinan Str., London N1 9XW.

Krogh, A. (1919) "The number and distribution of capillaries in muscles with calculations of the oxygen pressure head necessary for supplying the tissue." *J. Physiol.* (London) 52: 409.

Nunn, J.F. (1987) "*Applied respiratory physiology.*" 3rd edn. Butterworths, London.

Return to chapter 15

Return to content

Chapter 16.

Control Of Respiration. High Altitude

Study Objectives

- To *define* the following concepts acidaemia, blood-brain-barrier, and different types of hypoxia.
- To *describe* respiratory centres, apneustic and pneumotaxic centres, central & peripheral chemoreceptors, bronchopulmonary stretch receptors, irritant-receptors, J-receptors, slow-receptors, and other receptors affecting cardiopulmonary control.
- To *draw* ventilatory response curves to carbon dioxide at sea level, during sleep, at altitude, during exercise and acidosis.
- To *calculate* gas tensions and ventilatory variables from relevant variables given.
- To *explain* the respiratory autonomy, CNS-effects, adaptation to altitude- hypoxia – hypercapnia –exercise - acid-base disorders. To explain mountain sickness, foetal oxygenation and cardiopulmonary changes at birth.
- To *use* the above concepts in problem solving and case histories.

Principles

- *Dalton´s law: The partial pressure or tension of a single gas in a mixture is equal to the product of the total pressure and the mole fraction (F).*

Definitions

- **Acidaemia** is acid poisoning with a pH below 7.35 in the arterial blood.
- **Apneusis** is a particular respiratory pattern with prolonged inspirations separated by irregular short expirations, following midpontine transection when the vagi are cut.
- **Blood-brain-barrier** consists of tight junctions between the endothelial cells of the capillaries in the CNS. The tight junctions allow only small molecules to pass into the brain tissue, and they are impermeable to H^+ and HCO_3^-, whereas lipid-soluble gasses such as carbon dioxide can pass the plasma membrane.
- **Blood-CSF-barrier** is a tight junction barrier across the chorioid plexus, preventing many large molecules to pass from the blood to the cerebrospinal fluid (CSF).
- **Carotid bodies** are small organs located between the internal and external carotid arteries. The bodies contain *chemosensitive glomus cells,* fixed in a plasma-like tissue-fluid surrounded by abundant sinusoidal capillaries.
- **Circumventricular organs** are discrete structures in the hypothalamus, the pineal gland and the area postrema, with highly fenestrated capillaries that can be easily penetrated by

large and small molecules as well as ions. The circumventricular organs are located close to essential control centres regulating respiration, blood-glucose level, and the osmolality of the extracellular fluid.

- **Hypoxaemia** refers to a condition with low oxygen concentration in the arterial blood including hypobaric or hypotonic hypoxia.

- **Medullary, central chemoreceptors** are located below the ventrolateral surface of the medulla and communicate with the respiratory centre neurons. The main stimulus of these receptors is the hydrogen ion concentration of the brain extracellular fluid, and they respond to acidaemia with hyperventilation.

- **Polycythaemia** refers to a condition with an increase in red cell count, blood haemoglobin and packed cell volume (PCV or haematocrit).

Essentials

This paragraph covers *1. Neural and 2. Chemical control of breathing, 3. Acclimatisation, 4. Alveolar gas oscillations, and 5. Oxygen-altitude>Mt. Everest.*

1. The neural control of respiration

involves three components: ***I***. *The central neurons,* ***II***. *Effectors, and* ***III***. *Sensors*.

I. Central neurons

Results of brainstem transections performed on animals over the past two centuries indicate that the *respiratory controller* or *respiratory centre* (RC) resides in the brainstem. Such transection results also imply the presence of a pneumotactic *off-switch* centre in the nucleus parabrachialis of the rostral pons (Fig. 16-1).

RC is a collective term for a diffuse network of at least two types of neurons located in the *medullary reticular formation*. The *inspiratory* I-neurons fires just before and during inspiration, and they are tonically active in providing inspiration. The *expiratory* E-neurons fire just before and during expiration.

I- and E-neurons are localized in at least two medullary groups, the dorsal and the ventral motor groups.

The dorsal motor group in the *solitary tract nucleus* contains mainly I-neurons. The axonal projections of these neurons terminate in the *cervical* and *thoracic* spinal motor neurons of the phrenic and intercostal nerves. Nucleus tractus solitarius is also concerned with cardiovascular control.

The ventral motor group resides in the nucleus ambiguus, para-ambigualis and retro-ambigualis (ie, the *Bøtzinger complex*). These diffusely located neurons include both I- and E-neurons, and their E-neurons project to expiratory *intercostal motor neurons* in the thoracic segments.

The I-neurons show bursts of *spontaneous activity* (approximately 12-16 times per min) separated by quiet periods, in a pattern mimicking the breathing at rest. The E-neurons, on the other hand, are not self-excitatory. The activity of the I-neurons correlates with the rate and depth of breathing. The *basic activity* of the I-neurons, like that of all pacemakers, is modulated by a multitude of external signals from the periphery and from other areas of the CNS.

The *central inspiratory drive* determines the intensity of the desire to inspire. This drive is measured as the *inspiratory flow rate* or as the *phrenic nerve activity*. The phrenic nerve innervates the diaphragm, which is the most important inspiratory muscle.

The duration of the central inspiratory drive determines the inspiratory time (ie, the active

phase of medullary pacemaker I-neurons). Inspiratory neural activity is terminated by either impulses from the pontine off-switch or from the pulmonary stretch receptors conducted through the vagal nerve.

Inspiratory airflow equals driving pressure divided by resistance. Accordingly, the inspiratory flow rate is a function of the pressure- generating respiratory muscles. The smooth muscles of the upper airways determine upper airway resistance.

The expiratory phase lasts as long as the inspiratory *off-switch* neurons are active, and inspiration is not resumed. Expiration is passive, except at very high levels of ventilation, where expiratory muscles contract rhythmically to augment expiration.

The main stimulus to breathing is CO_2 through the medullary chemoreceptors, located close to the RC.

Normal respiratory rhythm (eupnoea) can persist after removal of the entire brain above the brainstem (ie, in a decerebrated animal). This confirms the presence of a basic inspiratory tonic activity.

Transection made at the midpons, with the vagi intact, cause slowing and deepening of respiration. When the vagi are also sectioned, midpontine transection results in apnoea (Fig. 16-1) or *apneusis* (ie, prolonged inspirations or I-spasms, separated by short expirations).

Removal of the *pontine centres* by transection between the pons and medulla results in a gasping, irregular pattern called *ataxic ventilation* (Fig. 16-1). The neurons of the medulla themselves have a *spontaneous rhythmicity*. Thus, the role of the pontine centres is to make the discharges of the medullary neurons smooth and regular. Transection between the medulla and the spinal cord results in respiratory arrest (apnoea). Respiratory rhytmogenesis is most likely due to a *central pattern generator* located in the brainstem and operating as described above (Fig. 16-1).

A network of interneurons in the brainstem is termed the *reticular activating system,* because it affects our state of wakefulness and increase respiratory activity.

Fig. 16-1: The respiratory rhythm generator with results of brain stem transection. NTS = nucleus tractus solitarius, NAm = nucleus ambigualis, NretroAm= nucleus retroambigualis.

II. Effectors

The *descending brain impulses* that drive the cervical and thoracic motor neurons consist of those arising from voluntary and involuntary cerebral sources. These motor neurons in cranial and spinal nerves lead to the respiratory muscles and to the smooth muscles of the upper airways. The corticospinal tract transfers *voluntary* commands directly from the cortex to the *somatic motor neurons*, or passes the same information through the pontomedullary centres before descending to the motor neurons via the reticulospinal tract. The *involuntary* descending signals arise from the I-neurons in 3 brainstem nuclei. These are the *solitary tract nucleus* (via the phrenic nerve to the diaphragm), the mixed neurons in the *nucleus para- and retro-ambigualis* (via the intercostal nerves to the intercostal muscles), and the *nucleus ambiguus* (via the cranial nerves to the airways including smooth muscles).

The intrafusal muscle fibres can shorten relatively more than the extrafusal (Fig. 4-7). Hence, a *simultaneous* gamma- and a- efferent discharge shortens the respiratory muscles to provide exactly the necessary power. The respiratory muscles (except the diaphragm) show this so-called *gamma-loop servo*. Actually, this is a sensory feedback to ensure homeostasis, but we must look at other sensors for the control system.

III. Sensors

1. *Higher brain centres* (cortex, hypothalamus, and diencephalon). Cortical, voluntary breath holding is possible until the breaking point is reached (ie, the point where apnoea

is disrupted). Similar responses are released by stimulation of the diencephalon. A rise in hypothalamic temperature triggers frequent breathing (*tachypnoea*) via the respiratory centres. During sleep and anaesthesia, breathing provides for metabolic needs primarily via the automatic homeostatic control system described above. During wakefulness, however, the breathing system serves both homeostatic and voluntary, behavioural (nonhomeostatic) needs. The behavioural needs include sucking, swallowing, speech, singing, laughing, crying, defaecation, voiding, breath holding, exercise, hyperventilation, and coughing. Such voluntary acts affect P_{aCO_2}, P_{aO_2}, and $[H^+]$.

2. *Bronchopulmonary stretch receptors* are located in the smooth muscles of the trachea, larger bronchi and also in the lung parenchyma in the alveolar ductules and sacs. The activity of these *smooth muscle receptors* increases markedly with airway distension, and the activity ebbs slowly with time, hence they are called *slowly-adapting* pulmonary receptors. If the tidal volume exceeds *one litre*, these receptors initiate signals that inhibit the inspiratory drive via myelinated vagal fibres, reinforcing the actions of the pontine centres and protecting the lungs from overexpansion. In humans this *Hering-Breuer reflex* plays no part in regulating ventilation during quiet breathing at rest, but the reflex is active during exercise.

3. *Rapidly adapting irritant receptors* are probably *free* or modified vagal nerve endings in the epithelium of the airways. These receptors are stimulated by irritants (smoke, allergens) and by inflammatory mediators such as prostaglandins and histamine. Rapidly adapting irritant receptors mediate the protective and sometimes pathologic responses of *cough* and *bronchospasm*. The efferent limb of this reflex is the motor fibre in the vagus, which initiates bronchospasm.

4. *Juxta-pulmonary capillary* receptors are terminals of non-myelinated vagal fibres (J-receptors). Distension of the interstitial space, as seen in a variety of cardiopulmonary disorders increases the diffusion distance, elicits increased ventilation, tachypnoea, bradycardia, and low arterial pressure via vagal reflexes. These cardiopulmonary disorders are microemboli, pulmonary oedema, pneumonia, fibrosis, atelectasis, and damage from inhaled irritants. *Atelectasis* means alveolar collapse.

5. *Peripheral arterial chemoreceptors* are found in the carotid bodies of humans (the aortic bodies account for only a small effect). The carotid chemoreceptors transmit impulses to CNS via cranial nerves IX and X, and induce rapid changes in ventilation. The glomus cell is sensitive not only for low P_{aO_2} (hypoxaemia), but also for increasing P_{aCO_2} (hypercapnia), hyperkalaemia and acidosis.

6. *Central, medullary chemoreceptors* are essential and particularly important to steady state ventilation.

7. *Thermoreceptors,* stimulated by a *rise* in core temperature (so-called *heat receptors*), elicit tachypnoea and tachycardia (frequent heart rate) via hypothalamus. Other thermoreceptors termed *cold receptors* are stimulated by a fall in shell temperature. Stimulation of these receptors elicits bradycardia via the regulating hypothalamic temperature area. This so-called survival reflex (or “diving bradycardia”) protects the bloodflow through the *brain, heart, and lung* in emergency situations involving breath holding, not only in water, but also on land.

8. *Arterial baroreceptors* in the carotid sinus are *stretch receptors,* stimulated by increased transmural arterial pressure. Stimulation decreases arterial pressure, and inhibits heart rate and ventilation as in the survival reflex.

9. *Receptors in working muscles* stimulate ventilation via type III and IV afferents to the respiratory centres in the medulla. These receptors are involved in the exercise

hyperpnoea**.**

10. ***Protective, vagal reflexes*** are related to vomiting, hick-up and swallowing. These reflexes protect us from inhaling vomit and thus being choked.

2. The chemical control of respiration

The respiratory system exerts its homeostasis by both *peripheral arterial* and by *central medullary chemoreceptors*. Both types of receptors are sensitive to changes in the proton concentration ($[H^+]$) around them. Such changes imply changes in the intracellular $[H^+]$, and changes in the ionic composition (Na^+, K^+, and Ca^{2+}). Both types of receptors are activated by increases in P_{CO_2} independent of $[H^+]$. Acute hypoxia stimulates the *peripheral arterial chemoreceptors*, but maintained hypoxia has a *depressant effect* on both central chemoreceptors and regulatory neurons in the RC.

The carotid and aortic bodies (glomera carotici & aortici) are small organs located in the tissue between the internal and external carotid arteries and at the arch of the aorta, respectively. The bodies contain *chemosensitive glomus cells,* fixed in a plasma-like tissue fluid surrounded by lots of sinusoidal capillaries, which have an extremely large flow (up to 2000 ml $*min^{-1}*100g^{-1}$) of arterial blood.

Fig. 16-2: Glomus cell with a rich arterial blood supply. The IXth cranial nerve is from the carotid bodies and the Xth (vagus) nerve from the aortic bodies. The frequency of action potentials in the glomus nerve increases, when P_{aO2} is below normal.

Glomus cells (type I) are sensitive to ***decreased*** P_{aO_2} (hypoxaemia mediating hyperventilation), but also to *increased* P_{aCO_2} (hypercapnia) and *increased* $[H^+]$ (acidaemia), and $[K^+]$ (hyperkalaemia) in the arterial blood and thus in the fluid surrounding the glomus cells (Fig. 16-2). The carotid bodies are predominant in humans.

The carotid sinus nerve innervates these bodies, which is a branch of the glossopharyngeal nerve (IX). The following hypothesis explains how the 4 glomus cell stimuli operate.

Normally, an oxygen-sensitive Na^+-K^+-pump (*Skou´s enzyme*) maintains the high intracellular K^+ and the low intracellular Na^+ of the glomus cell. *Hypoxia* inhibits this pump, whereby the cell membrane *depolarises*. The stimulation by *hyperkalaemia* is an unspecific *depolarisation*, which occurs in most cell membranes including the glomus cells. *Hypercapnia* and *acidaemia* feed the cell with CO_2, which rapidly converts into H^+ and HCO_3^-. Increased acidity in the cell stimulates the H^+-Na^+–exchanger in the cell membrane, the Na^+-K^+-pump is inhibited and *depolarisation* occurs. Depolarisation triggers a voltage-gated Ca^{2+}-channel. The high Ca^{2+}-influx activates neurosecretory granules and releases an excitatory neurotransmitter that depolarises the nerve fibre membrane and a propagating signal for the RC is transmitted (Fig. 16-2). The glomus cell contains lots of the neurotransmitter, *dopamine*, and the enzyme, *carboanhydrase*.

The short latency and response times of the carotid chemoreceptors and the location of the receptors close to the heart and lung, means rapid ventilatory stimulation. The carotid bodies are therefore responsible for the immediate reactions to changes in P_{aO_2}, P_{aCO_2}, $[H^+]$, and plasma-$[K^+]$.

Stimulation of the carotid bodies by *hypoxaemia, hypercapnia, hyperkalaemia, and acidaemia* increases the impulse frequency of the carotid sinus nerve and the glossopharyngeal nerve to the medullary respiratory centres (RC).

Increased activity from RC to the respiratory striated muscles and the upper airway smooth

muscles causes increased ventilation. This has a homeostatic effect on the initial stimuli - partially relieving *hypoxaemia, hypercapnia and acidaemia.* This is a *negative feedback loop*.

Falling P_{aO_2} stimulates the carotid glomus cells to increase ventilation (Fig. 16-2 right), but the hyperventilation simultaneously reduces P_{aCO2}, which counteracts the stimulation of breathing. Patients with a P_{aO2} below 7.3 kPa are *hypoxic* and suffer from *hypotonic* or hypobaric *hypoxia.* Compensatory hyperventilation maintains adequate oxygenation in the face of acute decreases of P_{IO_2} from the normal 20 kPa (150 mmHg), until P_{IO_2} decreases below 4 kPa (30 mmHg), where consciousness is lost. There is an inverse or hyperbolic relationship between alveolar ventilation and P_{aO_2}. At the peak of Mt. Everest the P_{aO2} is just below 4 kPa in trained mountain climbers, and they can walk only with great difficulty.

Anaemic patients have reduced oxygen content in the arterial blood (C_{aO_2}) in spite of normal P_{aO_2}. Since P_{aO_2} represents the stimulus to the glomus cells of the carotid bodies, and not the oxygen concentration of the arterial blood, such conditions are rarely associated with compensatory hyperventilation. A person with CO *poisoning* can die without increasing his breathing and without having felt shortness of breath (*dyspnoea*). CO acts as metabolic poison at the carotid bodies and increases ventilation to some extent.

Persons without peripheral arterial chemoreceptors can still increase their ventilation, in response to increases in P_{aCO_2} or $[H^+]$. These two stimuli have a particularly strong effect on the central, medullary chemoreceptors.

Three superficial (subpial) areas have been defined on the ventrolateral surface of the medulla, designated L (Loeschke), M (Mitchell) and S (Schläfke) after the scientists that located the *medullary chemoreceptors*. Areas L and M are believed to be chemosensitive, and nerve fibres from L and M converge on area S, where they pass deeper into the medulla to reach the medullary Respiratory Centre (RC). The medullary chemoreceptors are located 100-200 mm below the ventrolateral surface. These receptors communicate with RC neurons located deeper in the medulla. The main stimulus to the chemoreceptors is the $[H^+]$ of the *brain extracellular fluid* (ECF), which is in close proximity to the cerebrospinal fluid (CSF) that bathes these receptors.

The thin layer of *pia* between CSF and the brain ECF is *highly permeable* to CO_2, and even to ions, whereas the *blood-brain barrier* resists *ionic diffusion* (Fig. 16-3). It takes several minutes for changes in blood $[H^+]$ to be reflected in the brain ECF and hours for equilibrium.

The tight junctions between the endothelial cells of the cerebral capillaries (Fig. 16-3) form the blood-brain barrier.

This is why brain ECF has a P_{CO_2} that follows changes in the arterial blood (P_{aCO_2}), and a $[H^+]$ that follows changes in CSF through the thin pia layer (Fig. 16-3).

In *metabolic acid-base disorders,* the maintained changes in brain ECF $[H^+]$ are much smaller than the changes in blood $[H^+]$. The blood-brain barrier has slow *transport proteins* that can move ions in and out of ECF (Fig. 16-3).

Control molecules in area striata, which has widely fenestrated capillaries and no blood-brain barrier, can pass the blood-brain barrier. Breath-holding and hypoventilation leads to *carbon dioxide accumulation* with acute increase of P_{aCO_2}.

Fig. 16-3: The blood brain barrier and the medullary chemoreceptors.

Since CO_2 is lipid soluble it diffuses rapidly across the blood-brain barrier and increases brain ECF $[H^+]$ by hydration to carbonic acid followed by dissociation. Not only this H^+ but also

P_{CO_2} itself stimulate the *medullary chemoreceptors.*

The sensitivity (gain) of the medullary chemoreceptors to changes in $[H^+]$ exceeds that of the carotid chemoreceptors. Most of the steady-state ventilatory rise in response to CO_2 is mediated by the medullary receptors, whereas fast and transient changes are first detected by the *carotid chemoreceptors.*

The quantitative relationship between the ventilation and P_{ACO2} is called the CO_2 *response curve* (Fig. 16-4). The *slope* of such a curve shows the *sensitivity* or the *gain* of the RC. At low P_{ACO2}-values a small rise does not change the ventilation (see the typical *hockey stick* curve in Fig. 16-4A). This is because the *threshold* of the RC has not been reached.

The right curve in panel A shows the rise in ventilation by hypercapnia at normal P_{AO2} (110 mmHg). The left curve shows the response at P_{AO2} 5 kPa (37 mmHg). The slope is steeper which implies an increased sensitivity or gain of the chemosensitive feedback loop during hypoxia. The hypoxia also lowers the threshold for the CO_2 stimulus (Fig. 16-4A).

Fig. 16-4: CO_2 response curves: A shows results from persons exposed to a combination of hypoxia and hypercapnia, B shows the change from a normal control curve to the response of acidotic patients; C shows the change from control to steady state exercise.

The combined effect is *multiplicative.* The shift to the left of the curve is called *reduced threshold* due to the new stimulus, which is hypoxia (Fig. 16-4A). The steeper slope of the hockey stick curve is typical for high altitude acclimatisation.

The right curve in panel B is a normal control: A low pH_a as seen in chronic, metabolic acidosis triggers hyperventilation, so the P_{ACO_2} falls further. This is a reduction of the threshold, but not of the sensitivity to P_{aCO2} of the chemoreceptors. A high blood $[H^+]$ is the new stimulus for the acidotic patients of Fig. 16-4B.

Exercise has a similar effect (Fig. 16-4C). The CO_2 response curve is also shifted to the left, again without a change in the slope. Here, the extra stimuli are signals from the working muscles and from CNS.

3. Acclimatisation

The acclimatisation is the physiologic adaptation of *long lasting or repeated stress*. An example of acclimatisation or adaptation is observed in inhabitants exposed to *chronic hypoxia* at high altitudes. Another example is *physical training.*

Acute acclimatisation:

A person just arriving at high altitude experiences immediate hyperventilation. This high $\dot{V}_E$ is further increased and the high ventilation persists for the rest of his life, if he becomes a permanent high altitude resident. The low PaO2 stimulating the carotid chemoreceptors, which causes the rise in ventilation. Acute ventilatory adaptation is an effect upon the peripheral chemoreceptors. The ventilatory response implies excess clearance of CO_2 with reduced P_{aCO_2} and *acute respiratory alkalosis*, according to the alveolar ventilation equation.

The P_{IO2} in the moist tracheal air will fall with *falling* F_{IO2} or with *falling* atmospheric pressure (P_B) at altitude (Fig. 16-5). Ventilation starts to rise at P_{IO2} values below 13.3 kPa (100 mmHg), which corresponds to a P_{aO2} of 7.5 kPa, and the rise is hyperbolic (Fig. 16-5). The rise in ventilation in itself reduces the permanent P_{aCO_2} stimulus (Fig. 16-5). The data confirm the presence of an *acute respiratory alkalosis* with a high pH (Fig. 16-5 and Chapter 17).

Chronic acclimatisation:

When the low P_{aCO_2} has persisted for five days, adaptive mechanisms in the renal tubule cells are optimised (Chapter 25). The renal compensation consists of *reduced* tubular H^+ secretion and thus *reduced* tubular *bicarbonate reabsorption*. This in turn then elicits a fall in arterial and CSF [bicarbonate] and a rise of the low [H^+]. In the 5- 10 days until adaptation is fully accomplished, the arterial [bicarbonate] is reduced in proportion to the fall in P_{aCO2}. Thus pH_a is only mildly elevated, and the final condition is called a *compensated respiratory alkalosis* (ie, normal to increased pH_a and reduced P_{aCO2}).

The CSF contains fewer buffers and lower buffer concentrations than the blood. High altitude acclimatisation shifts the CO2 response curve to the left (i.e., reduced threshold), and the slope becomes steeper illustrating a higher sensitivity or gain (Fig. 16-4A).

The acclimatised person has similarities to the well-trained athlete with high oxygen capacity (high haematocrit, haemoglobin and erythrocyte count - also similar to chronic polycythaemia). The adaptation is associated with a rise in the circulatory oxygen transport, because both cardiac output and the oxygen capacity of the blood are increased. The erythropoietin release is increased.

Fig. 16-5: Respiratory variables and their changes at altitude.

The acclimatised person has a large blood and plasma volume, growth of the pulmonary arterial wall and the right ventricular muscle mass in response to hypoxic pulmonary vasoconstriction. There is a constant high sympathetic tone, which decreases the renal bloodflow (RBF). Eventually the hypoxic vasoconstriction leads to pulmonary hypertension, and right ventricular failure (cor pulmonale) can develop as in chronic mountain sickness. Stimulation of adrenergic receptors of the pulmonary vessels that causes pulmonary vasodilatation is used in the treatment of mountain sickness.

During climbing at high altitudes, the low atmospheric pressure (P_B) implies a fall in P_{IO2} =(P_B - 6.3)* F_{IO2} and thus in P_{aO_2}. When P_{aO_2} is below 7.3 kPa (55 mmHg), the ventilation must increase progressively, whereby P_{aCO_2} is reduced. The *simplified alveolar gas equation* (Eq. 14-3) can be applied. The solution for P_{ACO_2} is 6 kPa or 45 mmHg.

This theoretical calculation of P_{ACO2} is only mathematically correct. The argument is clearly wrong, when P_{AO2} is below 7.3 kPa, because at this level the hyperventilation will diminish P_{ACO2}.

Ventilatory acclimatisation elicits a long lasting rise in BTPS-ventilation (measured at altitude), inversely proportional to the fall in P_B. A doubling of the BTPS-ventilation at half an atmosphere implies that the ventilation measured in STPD-units is unchanged. According to the *alveolar ventilation equation*, F_{ACO_2} is equal to the ratio between carbon dioxide output and alveolar ventilation **(**both measured at STPD). Since both of these volumes are unchanged after total adaptation, it follows that also F_{ACO_2} must be unchanged. Hence, P_{ACO_2} must fall proportional to the fall in P_B. However, at the top of Mt. Everest the P_B is only 253 mmHg or 1/3 atm, but the ventilation is 5-fold increased.

Following return to sea level a *de-acclimatisation* with falling ventilation and heart rate takes place over the next three weeks. Also the blood pressure and pulmonary vascular resistance is

falling in this period.

Adaptation to high altitude and to physical training is *lost within a few weeks*.

Fig. 16-6: The oxyhaemoglobin curve is shifted towards the right by increased 2,3-DPG content of the red cells.

Altitude residents are exposed to chronic hypoxia. They also have high concentrations of 2,3-DPG in their red cells just as endurance athletes, and their oxyhaemoglobin dissociation curve is shifted towards the right (Fig. 8-3 and 16-6). This phenomenon is also expressed by the term P_{50}, which is the P_{O2} at 50% saturation. The P_{50} denotes the *standard affinity* (Ch. 14) between oxygen and haemoglobin. The standard affinity is high, when P_{50} is low. Normally, the P_{50} is 3.6 kPa or 27 mmHg (Fig. 16-6). With increased P_{50} at altitude, the standard affinity is reduced, and the delivery of oxygen to the tissues is improved at a given oxygen concentration. Blood with only 25% saturation has a P_{O2} of 27 mmHg in altitude residents, much larger than in a sea level resident (see the curve for P_{CO2}= 40 mmHg in Fig. 16-6).

4. Alveolar gas tension oscillations

During a normal respiratory cycle the alveolar gas tensions *oscillate* (Fig. 16-7). During the first part of the inspiration we re-inhale alveolar air contained in the dead space, but soon the tension curves reverse as fresh air is received from the outside. Early in expiration the P_{CO2} starts to increase linearly and the P_{O2} falls. These oscillations around the mean alveolar gas tensions are transferred to the passing blood. The carotid bodies may be sensitive to P_{CO2} oscillations of the arterialised blood.

Fig. 16-7: Oscillations around the mean of the alveolar gas tensions during a normal respiratory cycle performed by a sitting healthy person at rest.

The alveolar and arterial oscillations increase in amplitude and frequency during hypoxia and with tachypnoea and increased tidal volume during exercise. This occurs at high altitude and in all disorders with hypoxia. Accordingly, any stimulus of the carotid bodies by this mechanism must be accentuated.

5. Oxygen - Altitude - Mt Everest

At sea level the ambient or barometric pressure varies around 101.3 kPa dependent upon the meteorological status. Above sea level, the ambient pressure decreases with altitude (Fig. 16-8). The top of Mt. Everest is depicted with mountaineer camps for the last 3 km.

At any given location on earth, the ambient pressure (P_B) equals the gravitational force of the air column per area unit. Classically, P_B is given as the height of the mercury (Hg) column, which exerts the same pressure as the air on an area unit (sea level mean 760 mmHg or 101.3 kPa at 0^oC and 45^o latitude).

Fig. 16-8: The decrease in inspired oxygen tension (P_{IO2}) and in barometric pressure with increasing altitude at 45 degrees latitude.

Since oxygen comprises 20.93% of the air, its partial pressure in moist tracheal air at sea level (water vapour tension 47 mmHg at 37^oC) is easily calculated. The molar fraction is equal to the partial pressure of oxygen in the tracheal, moist, inspired air (P_{IO2}) divided by the maximally possible oxygen pressure in this air: $P_{IO2}/(760-47) = 0.2093$. Accordingly, P_{IO2} is 150 mmHg or (150* 0.1333)= 20 kPa at sea level. The curve for tracheal P_{IO2} in Fig. 16-8 is constructed on the basis of such calculations.

The alveolar gas equation at an RQ of one, states that $P_{AO2} = P_{IO2} - P_{ACO2}$ (Eq. 14-3). At a P_{AO2} of 30 mmHg, most persons lose consciousness, so in theory the lowest critical value of P_{IO2} is: (30 + 40) = 70 mmHg, with a P_{ACO2} of 40 mmHg. However, a hyperventilation drop in P_{ACO2} leads to a similar rise in P_{AO2}, according to the equation above. Therefore, humans can reach the mountain summits without the use of supplementary oxygen (see below).

Altitudes up to 3000 m do not have important adverse effects. The residents of Mexico City live at an altitude of 2200 m without hypoxic problems in general. At an altitude of 3200 m (12000 ft), tracheal P_{IO2} is 94 mmHg and, with P_{ACO2} =40 mmHg, P_{AO2} is expected to be 54 mmHg, as long as the person does not hyperventilate and keep RQ = R =1.

Such a low P_{AO2} is an adequate stimulus to the peripheral chemoreceptors (carotid bodies) and ventilation increases. The accompanying drop in P_{ACO2} allows the mountain climber to reach higher and higher altitudes.

Residents living in the Andes at an altitude corresponding to a P_B of 401 mmHg, inspire air with a tracheal P_{IO2} of 76 mmHg. This altitude is approximately 17500 feet or 5330 m. Mean values of their alveolar gas composition are P_{AO2} 42, and P_{ACO2} 26 mmHg, due to chronic hyperventilation. With an RQ of 0.8 they must have a gas R-value of the same size in order to be in respiratory steady state (Fig.16-9).

Fig. 16-9: Alveolar points for altitude residents in the Andes, for mountain climbers on Mt. Everest, and for sea level residents. The normal ventilation-perfusion ratio curve for sea level residents is shown with green, the same curve for altitude residents in Andes is shown with red, and the black curve is for climbers on the top of Mount Everest.

Depending upon their alveolar ventilation the A *point* moves up and down the R=0.8 diagonal around the value shown in Fig. 16-9. Altitude residents normally have a small alveolo-arterial P_{O2} difference.

Mt. Everest

Mt. Everest is located 8848 m above sea level corresponding to a barometric pressure of 34 kPa or 253 mmHg. A Norwegian group of mountain climbers recorded the pressure to 250 mmHg on the top, when setting the height record of 8848 m (1985). The partial pressure of oxygen in the humidified tracheal air is thus: (253-47)* 0.2093 = *43 mmHg* or 5.75 kPa (43 in Fig. 16-9). How is it possible to ascent to such a summit, without the use of supplementary oxygen?

The A *point* and the ventilation-perfusion-***curve*** is shown in Fig. 16-9 both for sea level, the Andes and Mt. Everest. A calculation of alveolar gas tension can be performed with the simplified alveolar gas equation (Eq. 14-3).

The low oxygen is a strong stimulus to ventilation, and the hyperventilation reduces P_{ACO2} of the acclimatised mountaineer to an average of 8 mmHg or 1 kPa. Accordingly, P_{AO2} is 35 mmHg (4.67 kPa) or less - close to the *critical level of consciousness* of 30 mmHg (4 kPa) and the mission is accomplished with an extremely small margin.

One member of the above mentioned Norwegian group of mountaineers had amnesia for the last 50 m to the summit and down, although he took pictures from the top together with the group; he remembered only the last part of the climb back to camp.

As a matter of fact, a similar simplified calculation seems to predict that the highest mountain to which it is possible to ascend without the use of oxygen does not exist. Let us assume that a P_{AO2} of 30 mmHg and a P_{ACO2} of 8 mmHg defines the critical barometric pressure. Then Eq 14-3 predicts that 38 mmHg is equal to 20.93% of (P_B - 47) mmHg. Accordingly, the critical

P_B is (182 + 47) = 229 mmHg, corresponding to more than 9 km of altitude. Such a postulate can never be proven in real life, because Mt Everest, with 8848 m, is the highest mountain on Earth. The simplified calculations are rough estimates, because the critical conscious level of 30 mmHg is in the arterial blood - not in the alveolar air.

The prediction has recently been tested in hypobaric chambers, where indians from the Andes remained well at a pressure corresponding to more than 9 km of altitude.

Pathophysiology

This paragraph covers *1. Shock*, *2. Hypermetabolism*, *3. Chronic anaemia* *and* *4. Mountain sickness.*

1. Shock

Patients with falling cardiac output and imminent shock, are forced to increase their arterio-venous oxygen content difference (ie, tissue extraction), since the necessary cellular O_2 uptake does not fall appreciably in the case shown in Fig. 16-10. The C_{AO2} is 165 and the concentration in the mixed venous blood is 80 ml*l^{-1} (assuming a normal haemoglobin concentration). Thus, the normal resting oxygen content difference of 50 is increased to a tissue extraction of 85 ml*l^{-1} (Fig.16-10).

Fig. 16-10: The arterial (a) and mixed venous point on the oxyhaemoglobin saturation curve of a shock patient with falling cardiac output.

For other aspects of shock - see Chapter 12.

2. Hypermetabolism

Exhausted athletes may develop *hypermetabolism*, where ATP degradation exceeds the production due to decoupling of ATP synthetase, whereby heat is released instead of ATP in the respiratory chain. The decoupling is perhaps due to a thermal change in the protein configuration of ATP synthetase. These athletes are unable to utilise the high oxygen tension (Fig. 16-11). The tissue extraction of oxygen is insufficient - only 25 ml l^{-1} (Fig. 16-11).

Fig. 16-11: Arterial (a) and mixed venous (v) oxygen tensions on the oxyhaemoglobin dissociation curve of an athlete with hypermetabolism, and of a severely anaemic patient.

3. Chronic anaemia

Chronic anaemia may develop so slowly that the patient can still work, although deprived of half the haemoglobin content of the blood (Fig. 16-11). The arterio-venous oxygen content difference is normal (50 ml l^{-1} in Fig. 16-11), and the venous oxygen tension gradient, responsible for delivery of oxygen to the most hypoxic tissues, is relatively high.

Patients with anaemia, chronic pulmonary disease and cyanotic heart disease, have high erythrocyte concentrations of 2,3-DPG (see Fig. 8-3). This molecule shifts the oxyhaemoglobin dissociation curve to the right, just as it occurs in hypercapnia and at increased body temperature. The 2,3-DPG diminishes the haemoglobin affinity for oxygen, so the binding in the lungs is impaired, while the peripheral release of oxygen is facilitated. The effect is beneficial for the tissue oxygenation of the anaemic patient, because of a larger P_{O2} gradient at a given blood oxygen content.

4. Mountain sickness

This concept covers several disorders, which are divided into acute mountain sickness (AMS) and chronic mountain sickness (CMS).

Acute Mountain Sickness

Acute Mountain Sickness has 1. A frequent benign and 2. A rare malign form.

1. Benignant Acute Mountain Sickness hits up to 40% of the 30 million persons yearly attracted by the high altitude environment above 3500 m. Symptoms occur within 6-12 hours of arrival and include frontal headache, malaise, nausea, lassitude and insomnia - mainly cerebral symptoms.

The *normal response of a healthy person* to hypobaric hypoxia is a rise in ventilation within seconds, and a rise in urine flow developing over 10-20 min. The hypoxic pulmonary vasoconstriction increases the right atrial pressure and releases the hormone, *atrial natriuretic peptide,* from the atrial walls to the blood. Atrial natriuretic peptide increases the excretion of Na^+ and water in the urine. Persons with high normal blood levels of the hormone are less prone to develop acute mountain sickness.

The *typical patient* reacts – at the time the disease develops - with a relatively small hyperventilation due to hyposensitive carotid bodies. The drop in oxygen tension does not affect the carotid bodies in these patients as much as in healthy persons. The P_{aCO2} is not reduced as much as in the healthy subject. The patient secretes more H^+ and reabsorbs more bicarbonate and Na^+ than normal. This results in *Na^+ retention* and in *water accumulation*. The relatively high P_{aCO2} also leads to a relatively high cerebral bloodflow. This is in contrast to the healthy mountaineer who has a relatively low P_{aCO2}, which causes less cerebral vasodilatation and less cerebral disorder.

Slow acclimatisation by slow ascent is the best prophylactic advice to mountaineers. Since part of the mechanism of acute mountain sickness is obviously related to the degree of respiratory alkalosis, it is no surprise that prophylactic treatment with *carboanhydrase inhibitors* (eg, acetazolamide) reduces the symptoms, because of the increased bicarbonate elimination in the kidneys, and because of the increased tissue P_{CO2}, which stimulates ventilation ("artificial acclimatisation").

2. Malignant Acute Mountain Sickness includes high-altitude pulmonary oedema (HAPO) and high-altitude cerebral oedema (HACO). They are rather infrequent, life-threatening forms with an incidence of 1% among mountaineers.

HAPO is established when a dyspnoeic subject has bloodstained frothy sputum or substantial chest rales. The salt- and water-accumulation increases the haemodynamic pressure and reduces the colloid-osmotic pressure by dilution. Both changes tend to produce oedema. This is a condition of imminent death from cardiovascular and respiratory failure, unless treated rapidly.

HACO is established when a drowsy mountaineer develops ataxia, nystagmus, papiloedema (retina is brain tissue) and coma. This condition is probably caused by extreme salt and water retention.

HAPO and HACO must be treated urgently with oxygen and descent to a lower altitude as fast as possible. The glucocorticoids, dexamethasone and betamethasone, exhibit minimal mineralocorticoid effect, maximal anti-inflammatory activity. These drugs are the choice in case of acute brain and lung oedema.

Chronic Mountain Sickness

Chronic Mountain Sickness occurs in long-term residents of altitudes above 4000 m - the Andes and elsewhere. The number of red cells in the blood (polycythaemia) and the haematocrit develops to exceptionally high values. The high haematocrit and viscosity of the blood tends toward formation of emboli. Turbulence in the circulation tends to form atherosclerotic plaques due to the high cardiac output with increased linear velocity.

Pulmonary vasoconstriction maintained over years leads to pulmonary hypertension and enlargement of the right ventricle, resulting in congestive heart failure of the right heart - Latin: *cor pulmonale* (see Chapter 10 and 14).

A large fraction of the cardiac output is shunted through vessels without alveoli or vessels with hypoventilated alveoli, so the patient is cyanotic with congested ear lobes and finger clubbing.

The treatment is descent to lower altitude.

Equations - see Chapter 14.

- The alveolar gas equation
- The alveolar ventilation equation
- The final ventilation-perfusion-equation

Self-Assessment

16. Multiple Choice Questions

I. Each of the following five answers have True/False options:

Statement: **The hyperventilatory response to hypoxia and hypoxaemia is mediated by the:**

A. Bronchopulmonary mechanoreceptors

B. Chemoreceptors in the carotid bodies

C. Central chemoreceptors

D. Irritant airway receptors

E. Arterial baroreceptors.

II. Each of the following five statements have True/False options:

A. The primary muscle of inspiration is the diaphragm.

B. The airflow resistance is highest in the small terminal bronchioles.

C. The intrathoracic pressure is less than atmospheric pressure.

D. The VC is defined as the maximum volume of air that can be inhaled following a maximum expiration.

E. The lung compliance is reduced in persons with emphysema.

III. Each of the following five statements have True/False options:

A. Polycythaemia patients have lots of red cells and a large oxygen binding capacity ([Total-Hb] * 1.34).

B. Patients with a high haemoglobin concentration show cyanosis as frequent as do patients with anaemia.

C. The typical patient with acute mountain sickness reacts with a high diuresis.

D. The 60 g of reduced haemoglobin in one litre of average capillary blood is the

approximate threshold for the bedside diagnosis of visible cyanosis.

E. Patients with anaemia, chronic pulmonary disease and cyanotic heart disease, have high erythrocyte concentrations of 2,3-diphosphoglycerate (2,3-DPG).

Case History A

A healthy female, with an anatomic V_D of 0.12 l and a respiratory frequency (f) of 14, is flying in an open-cockpit aeroplane at 2000 m. Here, her P_{IO2} is 16 kPa (120 mmHg) and her alveolar ventilation is increased to 5.6 l STPD per min (from 4.2 at the ground). She has had a mixed diet meal before take-off, and her R- value remains at 0.8. Her P_{AO_2} is restored to 13.3 kPa (100 mmHg) by the rise in ventilation.

1. *Calculate her P_{ACO_2} by the simplified alveolar air equation.*
2. *Calculate her expired minute ventilation.*

Case History B

A group of altitude residents (Sherpas) reside for a month at an altitude of 6 km with a P_B of 410 mmHg (5.47 kPa). After three weeks their average alveolar gas tensions were P_{AO2} 40 (5.3 kPa) and P_{ACO2} 22 mmHg (2.9 kPa). The average metabolic rate at rest was 5000 J per min or 84 Watts.
The respiratory energy equivalent for oxygen on a mixed diet is 20 kJ per l STPD. A P_{AO2} of 30 mmHg (4 kPa) is the threshold for loss of consciousness.

1. *Calculate the alveolar minute ventilation in STPD volume units and at the altitude (in l BTPS).*
2. *Calculate the lowest barometric pressure (P_B) at which man can survive by breathing air. Assume that P_{ACO2} is 8 mmHg (1 kPa) at the summit.*

Case History C

A polio patient, 11 years old (body weight 35 kg), is connected to a pressure-controlled respirator, because the boys respiratory muscles are paralysed. He receives atmospheric air through a tracheal tube by positive-pressure ventilation (input pressure 800 Pascal, Pa) at a frequency (f) of 10 per min. The equipment dead space and the dead space of the small boy added together amounts to 130 ml BTPS. At FRC the specific lung compliance is 1.5 ml BTPS per Pa, and the specific compliance of the thoracic cage is 0.8 ml BTPS per Pa. The specific total standard compliance of a healthy adult is 1 ml BTPS per Pa. The barometric pressure (P_B) is 770 mmHg (102.64 kPa), and the F_{ACO2} is 0.059. The arteriovenous oxygen content difference is 55 ml STPD l^{-1}. The patient is in respiratory steady state with a respiratory exchange quotient of 1.

1. *Calculate the specific total standard compliance of the polio patient.*
2. *Calculate the size of the alveolar ventilation.*
3. *Calculate the carbon dioxide output in ml STPD.*
4. *Calculate the cardiac output.*

Try to solve the problems before looking up the answers.

Highlights

- *Medullary chemoreceptors communicate with the respiratory centre and respond to acidaemia or hypercapnia with increased ventilation.*

- *The glomus cells of the carotid body respond to hypoxaemia, hypercapnia, acidaemia, and hyperkalaemia with increased ventilation.*
- *Chronic respiratory acidosis is compensated renally by bicarbonate reabsorption.*
- *Respiratory insufficiency in chronic obstructive lung disease is endangered when breathing pure oxygen.*
- *Chronic hypoxia adaptation is by increased erythropoiesis, increased 2,3-DPG production, and reduced ventilatory response to acute hypoxia.*
- *Ventilatory altitude acclimatisation is a hypoxic chemoreflex.*
- *The respiratory altitude alkalosis is compensated renally in days.*
- *Altitude acclimatisation is similar to hypoxia-training and endurance training.*
- *The normal response of a healthy person to hypobaric hypoxia is a rise in ventilation within seconds, and a rise in diuresis developing over 10-20 min. The hypoxic pulmonary vasoconstriction increases the right atrial pressure and releases the hormone, atrial natriuretic peptide from the atrial walls to the blood. Atrial natriuretic peptide increases the excretion of Na^+ and water in the urine.*
- *Persons with high normal blood levels of atrial natriuretic peptide are less prone to develop acute mountain sickness.*
- *The typical mountain sick patient reacts with a relatively small hyperventilation due to hyposensitive carotid bodies. The fall in oxygen tension does not affect the carotid bodies in these patients as much as in healthy persons. The P_{aCO2} is not reduced as much as in the healthy subject. The patient secretes more H^+ and reabsorbs more bicarbonate and Na^+ than the normal. This results in Na^+ retention and in water accumulation. The relatively high P_{aCO2} also leads to a relatively high cerebral bloodflow.*
- *This is in contrast to the healthy mountaineer who has a relatively low P_{aCO2} , which causes less cerebral vasodilatation and less cerebral disorder.*
- *Chronic Mountain Sickness occurs in long-term residents of altitudes above 4000 m - the Andes and elsewhere. The number of red cells in the blood (polycythaemia) and the haematocrit develops to exceptionally high values. The high haematocrit and viscosity of the blood increase the risk of emboli. Turbulence in the circulation tends to form atherosclerotic plaques.*
- *Pulmonary vasoconstriction maintained over years leads to pulmonary hypertension and enlargement of the right ventricle, resulting in congestive heart failure of the right heart - Latin: cor pulmonale.*

Further Reading

Gonzales C, L Almaras, A Obeso, and R Rigual (1994). Carotid body chemoreceptors; From natural stimuli to sensory discharges. *Physiol. Rev.* 74: 829-898.

Hultgren, HN (1997) *High altitude medicine*. Stanford, California.

Return to chapter 16

Return to content

Chapter 17

The Acid-Base Balance And Disorders

Study Objectives:

- To *define* pH, pK, P_{aCO2}, acids, bases, buffers, buffer bases (titratable base), actual bicarbonate and non-carbonic buffers, the extended extracellular fluid volume (ECV extended with red cells), and the base excess of the extended ECV.

- To *describe* the daily acid-base balance, titratable acidity and the net excretion of acid in the urine, the base excess of the blood and extended ECV, and the buffers of the intracellular fluid volume.

- To *calculate* the third variable of the Henderson-Hasselbalch equation from two measured variables.

- To *draw* the acid-base charts (ie, a double logarithmic plot of the Henderson-Hasselbalch equation).

- To *explain* the acid-base chart, metabolic processes involved in the regulation of pH, tubular H^+-secretion, and renal bicarbonate reabsorption. To describe intracellular buffers and pH, transport of H^+ across the cell membrane. To describe four acute disorders of the acid-base status (primary respiratory and metabolic acidosis and alkalosis) and their chronic counterparts with therapeutic suggestions.

- To *use* the acid-base variables in problem solving and diagnosis of acid-base disorders.

Principles

- *The mass action equation. Blood plasma is an open system where gasses are in equilibrium both with the alveolar air in the lungs and with the extracellular fluid volume (ECV). ECV and all red cells work as a buffer unit called extended ECV. Bicarbonate - carbonic acid is the strongest buffer and haemoglobin is the strongest non-carbonic buffer. The pH of arterial blood plasma is a function of P_{aCO2} and of the concentration of titratable buffer bases in the extended ECV (Box 17-1). The concentration of titratable base (Base Excess, BE) is measurable, and BE remains zero during acute changes of P_{aCO2}.*

- *The disorders of the acid-base balance are described by the following two interrelated chemical reactions:*

$$CO_2 + H_2O \; \text{⌠*} \; H_2CO_3 \; \text{⌠} \; H^+ + HCO_3^- \text{ and}$$

$$[\text{Acid form}] \cap H^+ + [\text{Non-carbonic buffers}^-]$$

Non-carbonic buffers - or non-bicarbonate buffers - refer to all other relevant buffers than the

*carbonic acid-bicarbonate buffer system. The enzyme carbo-anhydrase is shown with an asterisk: *.*

Definitions

- **Acid** (HB) is defined as a compound that can release a hydrogen ion or proton (H^+), whereas a **base** (B^-) can bind H^+.

- **Acidosis** (acidaemia) is defined as a disorder with accumulation of acids in the extended ECV. The pH measured in the arterial blood is less than 7.35.

- **Alkalosis** (alkalaemia) is defined as a condition with accumulation of bases in the extended ECV. The pH of the arterial blood is larger than 7.45.

- **Actual bicarbonate concentration** is the concentration calculated from P_{aCO2} and pH measurements in the arterial blood sample. The value can be calculated with Eq. 17-6 or read from an acid-base chart with a bicarbonate axis (slope -1 in Fig. 17-3).

- **The acid-base buffer capacity** of a system is defined as the amount of strong acid or base added to one litre (l) of the system (ie, mmol per l or mM) in order to change the pH one unit.

- A **buffer** is a corresponding acid-base pair, which acts as a pH-stabiliser.

- **Buffer base (BB)** refers to the total concentration of all carbonic buffer anions plus the non-carbonic anions.

- **Carbonic acid** refers to carbon dioxide.

- **Carbonic buffers** refer to the carbonic acid-bicarbonate buffer system.

- **Glomerular filtration** is due to a hydrostatic/colloid osmotic pressure gradient - the Starling forces

- **The acid-base buffer capacity** of the **extended ECV** is normally 61 mM per pH unit as an average (Box 17-1).

- **Non-carbonic buffers (non-bicarbonate buffers)** refer to all other buffers than the carbonic acid-bicarbonate buffer system in the extended ECV (Box 17-1).

- **The Base Excess (BE)** of the extracellular fluid (extended ECV) is calculated as the concentration difference of buffer bases in the actual sample and in the same sample following titration with strong acid or base and equilibration to standard conditions (pH 7.4, $\boldsymbol{P_{aCO2}}$ 5.3 kPa or 40 mmHg). The difference is given in mmol of strong acid or base, which must be added to 1 l of the sample. The base excess remains normal (zero) in acute respiratory acid-base disorders

- **Extended extracellular fluid volume** (*extended ECV*) is the extracellular fluid volume (ECV) plus the erythrocyte volume. A useful model of the *extended ECV* is an arterial blood sample diluted three folds with its own plasma (ie, 2+1). Extended ECV behave as a functional buffer unit. As a model extended ECV is considered to be all red cells distributed evenly in the extracellular fluid volume. The *extended ECV* of a healthy standard person covers

approximately 20% of body weight. *The principal buffers* of the *extended ECV* in a healthy person are shown below (Box 17-1). Extended ECV serves as an early distribution volume in acid-base disorders.

- **Intracellular fluid volume** (ICV minus red cells) comprises approximately 40% of the body volume. *The principal buffers* of ICV are proteins, phosphate and bicarbonate. The transport across cell membranes take hours to equilibrate and the intracellular buffer effect is involved in the delayed distribution. The intracellular buffer effect is larger in disorders with negative base excess (acidosis with pH approaching pK for the best buffers) than with positive base excess. OH^- has a molecular weight 17 times larger than that of H^+, so H^+ diffuses faster through the cell membrane, since the diffusion rate depends on the hydrated ionic radius

- **pH** is defined as the negative logarithm to the proton concentration. In this Chapter the pH is measured with glass electrodes in the arterial blood sample, and used in the extended ECV

- **pK** is equal to pH, when the concentration of acid and of base is the same.

- **Partial pressure of carbon dioxide** in the arterial blood sample (P_{aCO2}) is measured with glass electrodes, and P_{aCO2} refers to the extended ECV in this Chapter.

- **Metabolic acidosis** is acidosis with *negative* Base Excess. - Negative Base Excess means that the numeric difference between the measured buffer base (BB) and the normal buffer base (NBB) is negative.

- **Metabolic alkalosis** is alkalosis with *positive* Base Excess. - Positive Base Excess means that the numeric difference between the measured buffer base (BB) and the normal buffer base (NBB) is positive

- **Metabolic acid-base disorders** are caused by changes of the titratable base of the extended ECV (Base Excess is positive or negative).

- **Nephron** - the functional renal unit. Each nephron consists of a glomerulus, a proximal convoluted tubule, a Henle loop and a distal tubule ending in a collecting duct with several other nephrons.

- **Respiratory acid-base disorders** are caused by alterations of P_{aCO2}, without any change of the amount of titratable base in the extended ECV (BE is zero).

- **Respiratory acidosis** is an acidosis with a P_{aCO2} above 6.4 kPa or 48 mmHg at rest

- **Respiratory alkalosis** is an alkalosis with a P_{aCO2} below 4.4 kPa or 33 mmHg at rest.

- **Titratable phosphate acidity** in the daily urine is the amount of base (mmol) needed to titrate an acidic urine (phosphoric acid) back to the pH of plasma and glomerular filtrate (pH = 7.4 and P_{CO2} of 5.3 kPa).

- **Tubular reabsorption** is the movement of water and solute from the tubular lumen to the tubule cells and to the peritubular capillary network

Tubular secretion represents the net addition of solute to the tubular lumen.

-

Essentials

This paragraph deals with *1. Proton concentration & pH, 2. Buffer capacity and the bicarbonate buffer , 3. Buffer capacities, 4. From the H-H-equation to the acid-base chart, 5. Extended ECV & Base excess, 6. Metabolic acid-base production, 7. Renal acid-base control, 8. Intracellular buffers*.

1. Proton concentration and pH

Normally, the $[H^+]$ of arterial plasma of humans at rest is maintained by the lungs and kidneys within the range of 40 ± 5 nM, corresponding to a pH of 7.36-7.44. The $[H^+]$ of the arterial plasma is also maintained during strenuous conditions and diseases: 16-160 nM or pH 6.8-7.8 (Fig. 17-1). Proton concentrations in nM differ by a factor 10^{-6} from the normal standard-$[Na^+]$, which is around 140 mM.

A pH of 6.8-6.9 is not sustainable for long, and the patient is dying in a state of coma. In contrast, a healthy athlete can survive a pH of 6.85 following supramaximal exercise with headache as the only consequence.

A pH approaching 7.8 implies dissociation of all buffer proteins in order to produce H^+ (Eq. 17-7), and negative albumin charges capture Ca^{2+}, so the extracellular Ca^{2+} falls and the patient dies in *tetanic cramps* including *laryngeal spasms* (Fig. 17-1). Tetanic cramps are spontaneous prolonged or continuous muscular contractions (for explanation see later)

Fig. 17-1: Relationship between pH and the proton concentration

2. Buffer capacity and the bicarbonate buffer

The acid-base buffer capacity of a system is defined as the amount (mmol) of strong acid or base added to one litre (l) of the system (ie, mmol per l or mM) in order to change the pH one unit

The importance of the CO_2/bicarbonate buffer is easily realised when comparing the addition of one mmol of strong acid to one litre of a closed and an open system with 24 mM bicarbonate each (pH 7.4; P_{CO2} 5.3 kPa).

Fig. 17-2: The CO_2/bicarbonate buffer working in a closed and an open system.

In the closed system, the base concentration is reduced by 1 mM to 23, and the acid concentration is increased by one mM, because the reaction is shifted towards formation of CO_2 causing a high P_{CO2} (Fig. 17-2). Accordingly, the acid concentration is now: $([0.225 \times 5.3] + 1) = 0.225 \times newP_{CO2}$. Thus the $newP_{CO2}$ is 2.19/0.225 or 9.73 kPa. The pH of the closed system is changed to *7.02* (Fig. 17-2). The *buffer capacity* is 1/(7.4 - 7.02) = *2.6 mM* per pH unit, which is negligible.

In an open system such as the body, the ventilation simply eliminates excess carbon dioxide and the P_{CO2} is kept constant: 5.3 kPa or 40 mmHg. The chemoreceptors are bathed in extracellular fluid. Any rise in the P_{CO2} of the extracellular fluid is sensed by the chemoreceptors and releases a proportionate rise in ventilation. Thus, the **new pH** is **7.38** (Fig. 17-2). The *buffer capacity* is at least: 1/(7.4 - 7.38) = *50 mM per pH unit*, which is an essential

capacity (see Box 17-1)

3. Buffer capacities

The *buffer capacity* of a system is already defined as the amount of strong acid or base added to one litre (l) of the system in order to change the pH one unit

The most important buffer system for the regulation of $[H^+]$ in extended ECV is the carbonic acid (CO_2)/bicarbonate system (82%), although its pK is 6.1 and not so close to ideal as for the primary/secondary phosphate system (Box 17-1).

The *CO_2/bicarbonate buffer* is distributed in the *extended ECV* (up to 15 l of CO_2 is bound as bicarbonate). Its buffer capacity is *high*, since P_{aCO2} is rapidly maintained normal by *respiration,* and the kidneys control bicarbonate excretion.

Box 17-1: Principal buffer concentrations and their contribution to the buffer capacity of the extended ECV in a healthy person.

	Net-anion Concentration (mM, mean)	Buffer capacity (mM per pH unit)
Bicarbonate	24 (67%)	50 (82%) at constant PaCO2
Non-carbonic buffers	12 (33%)	11 (18%)
Haemoglobin	7	9
Plasma Protein	4	2
Phosphate	1	0.4
Total	36 (100%)	61 (100%)

Of the *non-carbonic buffer bases* (18%), the haemoglobin has the dominating buffer value with 9 mM per pH unit in the *extended ECV*, leaving only about 2 mM per pH unit for proteins (essentially albumin) and phosphate (Box 17-1).

4. From The H-H-equation to the Acid-Base Chart

The three variables in the final Henderson-Hasselbalch equation (pH_a, $\boldsymbol{P_{aCO2}}$, and actual bicarbonate in Eq. 17-6) are easy to handle. Rearrangement of Eq. 17-6, with constant bicarbonate concentration, shows the following:

$$\log P_{aCO2} = - pH + k \quad \text{or} \quad y = -x + k,$$

where **k** stands for three constants (6.1 + log [constant bicarbonate] - log 0.03).

Thus, y is a linear function of x, and the slope of the line is -1. This slope is called an *iso-bicarbonate line*, and the slope reflects the buffer capacity (rise in carbonic acid per pH unit) of the bicarbonate buffer in water.

Siggaard-Andersen has developed a *double logarithmic plot*, where factor 10 has the same distance (decade equidistant). The pH is chosen as the abscissa (pH = 6.8 - 7.8) and log P_{aCO2} is the ordinate (10-100 mmHg in Fig.17-3).

Fig. 17-3: The pH-log $\boldsymbol{P_{aCO2}}$ plot (decade equidistant).

All points on each iso-bicarbonate line in the plot with a slope of -1 have the same bicarbonate concentration. An example is the *iso-bicarbonate line* relating pH and log P_{aCO2} in a pure bicarbonate solution in water (15 mM) - shown by a stippled line with the slope -1 in Fig. 17-3. Here the bicarbonate concentration is constant (included in **k**) namely 15 mM in all points. Such iso-bicarbonate lines are pure mathematics. When two variables are given, the third

variable is readable from the plot (or easy to calculate according to the equation).

Originally, the equilibration line for a blood sample was determined by equilibration with high and low P_{CO2} as shown by the points **a** and **b,** crossing a *bicarbonate axis* in mM running through P_{aCO2} 40 mmHg (= 5.3 kPa in Fig 17-3). The *actual bicarbonate concentration* is readable at this bicarbonate axis. As an example, let us assume that the point **b**, characterises the actual pH and P_{aCO2} measured. An iso-bicarbonate line is drawn with the slope -1 (use a piece of folded paper) and the actual bicarbonate for the point **b** is read on the *bicarbonate axis* to be 15 mM. This is simply a fast geometrical calculation

Adding the non-carbonic buffers of whole blood (thick red line between **a** and **b** in Fig. 17-3) or of extended ECV (purple line) to a pure bicarbonate solution in water, naturally increases the buffer capacity. These lines are therefore steeper than the iso-bicarbonate lines (-1). The most important of the non-carbonic buffers is haemoglobin in the red cells (Box 17-1).

The actual bicarbonate concentration is dependent of variations in P_{aCO2}. When P_{aCO2} increases, carbonic acid is buffered by non-carbonic buffers causing a rise in actual bicarbonate (see the two interrelated reactions in *Principles*). The elevated actual bicarbonate is interpretable as a *metabolic alkalosis,* also when a *respiratory acidosis* is the real cause. The interpretation is a serious error

Accordingly, we have to look for a better index to separate metabolic from respiratory acid-base disturbances

The steepest equilibration line illustrates the buffer capacity for whole-blood buffers (Fig. 17-3 **a** to **b**). The slope of an *equilibration line* for all the buffers in the extended ECV is slightly smaller than that of whole-blood, because the whole-blood buffers are diluted in the large extended ECV (Fig. 17-3, purple line).

Base Excess (BE) of the extracellular fluid (extended ECV) is calculated as the concentration difference of buffer bases in the actual sample and in the same sample following equilibration to standard conditions (pH 7.4, P_{aCO2} 5.3 kPa or 40 mmHg and well oxygenated). A healthy person with these actual values for pH and P_{aCO2} must have a BE of *zero*.

As P_{CO2} rises due to hypoventilation the [bicarbonate] rises, but BE of the extended ECV remains zero. The slope of the BE- *zero line* depends primarily on the concentration of haemoglobin, less on the plasma protein concentration, and minimally on the effect of inorganic phosphate (Box 17-1). The position of the BE-*line* depends on the BE: adding strong acid shifts the line to the left, and adding base shifts the line to the right. The series of almost parallel *iso*-BE-*lines,* with negative and positive values around the BE-*zero line,* has been determined experimentally. Each iso-BE-*line* is an equilibration line for a standard person in a given metabolic condition.

The iso-BE-*line* zero depicts the condition with normal concentration and buffer capacity of buffer bases, and the other almost parallel lines represent abnormal conditions with either positive or negative base excess. With pH, and P_{CO2} measured, it is easy to read the precalculated *base excess* from the *base excess axis* at the top of the acid-base chart (Fig. 17-13).

5. Extended ECV & Base excess

The extended ECV is the extracellular fluid volume plus the red cells. An abstraction is to

think of all red cells distributed in the total extracellular fluid volume. A useful model of the *extended ECV* is an arterial blood sample diluted three folds with its own plasma (2+1).

The BE is independent of P_{aCO2}, since any change in P_{aCO2} implies opposite molar changes of the *bicarbonate* and the *non-carbonic buffer concentrations.* Hereby, there is no change in BE. Thus, the base excess is constant (normally equal to zero) during acute changes in P_{aCO2} by hyper- or hypoventilation.

The buffer capacity of the carbon dioxide- bicarbonate buffer is high, since *respiration and bicarbonate* excretion by the kidneys rapidly maintain P_{aCO2}.

The *base excess* is a key variable in diagnosing *metabolic disorders*. The base excess tells the whole story, whereas the bicarbonate buffer only tells part of the story (Box 17-1) - and sometimes in error. This is what makes *Base Excess* the diagnostic choice. BE alterations from zero are therefore used to diagnose *metabolic*, as opposed to *respiratory* acid-base disorders.

6. Metabolic acid-base production

The daily metabolic production of carbon dioxide is up to 24 mol. The CO_2 is a potential acid as H_2CO_3, and because the lungs eliminate it, it is called a volatile acid.

The hepatic production of non-volatile acid depends upon its daily amino acid load from intestinal absorption including amino acids from the daily degradation of intestinal mucosal cells. Persons on a normal mixed diet produce up to *100 mmol fixed* or non-volatile acids daily mainly from amino acids (plus organic acids as uric acid, phosphoric acid, and minimal amounts of HCl and sulphuric acid)

Persons on a mixed diet absorb approximately 50 mmol organic bases daily ($RCOO^-$ in Box 17-2). A maximum of 20 mmol of these organic anions is excreted in a 24-hour urine, because at least 30 mmol react with H^+ produced by the liver. As stated above, the hepatic oxidation of sulphydryl groups in amino acids and hydrolysis of phosphate esters from lipids produces 100 mmol H^+ daily, of which 60-70 mmol is excreted in a 24 hour urine (Box 17-2).

Box 17-2: Absorption, renal excretion and metabolic acid-base production (mmol daily)

		Non-volatile acids	Non-volatile bases
Input	Gut absorption:	0	50 (RCOO-)
	Metabolic production:	Up to 100	Self-elimination
	Total:	100	50
Output (renal excretion):	NH4+/ NH3 -buffer	30 (400 in acidosis)	NH3 in basic urine
	Phosphate (mainly from diet)	30 (60 in acidosis)	
	RCOO-		20
	Bicarbonate		1-3 (100 in alkalosis)

	Total:	60	21

As stated above, the remaining 30 mmol of metabolic H^+ is eliminated daily by oxidation of 30 mmol organic bases from the gut ($RCOO^-$).

About 30 mmol NH_4^+ is excreted in the daily urine, but the excretion is controlled during acid-base disorders. During acidosis, the high [H^+] stimulates hepatic glutamate- and renal NH_4^+ -production, and the renal NH_4^+ -excretion increases towards 400 mmol daily. During alkalosis the urea production accounts for the nitrogen elimination and only negligible amounts of NH_4^+ is excreted

Diets rich in vegetables and fruits generate non-volatile bases (salts of organic acids such as oxalate and citrate), which must be metabolised or excreted im the kidneys. In the body only amino acids are metabolised to bases (ammonia). Urinary excretion of the base form of ammonia takes place only in extremely basic urine.

Normally, there is no non-volatile acid-base production from carbohydrates or lipids.

7. Renal acid-base control

The kidneys prevent the loss of enormous amounts of bicarbonate in the urine, produce the most important urinary buffer (NH_3/NH_4^+ -buffer), and excrete the titratable phosphate acidity. - The free H^+ is not a major issue, since even acid urine with a pH =5 only eliminate 1/100 mmol per l

The details of paragraph 7.1 to 7.3 are best understood following a careful study of Chapter 25.

7.1 Bicarbonate reabsorption

Normally, the daily filtration flux of bicarbonate amounts to 4500 mmol (ie, filtration of 180 l of plasma with a mean concentration of 25 mM). Most of the filtered bicarbonate flux is reabsorbed already in the proximal tubules, where the luminal membrane contains a *Na^+-H^+-antiporter* (Fig. 17-4). The bicarbonate reabsorption is accomplished by means of *H^+-secretion*. Most of the H^+ secreted in the proximal tubules is derived from the *Na^+-H^+-exchange* through the antiporter. When the tubular fluid reaches the collecting ducts an important H^+-secretion is mediated by a *proton-K^+-ATPase* in the intercalated cells (see below).

Any change in the filtered bicarbonate flux is matched by a similar change in proximal bicarbonate reabsorption (ie, *glomerulo-tubular balance*). A change in Na^+ - homeostasis alters the bicarbonate reabsorption secondarily.

The *Na^+-K^+-pump* in the basolateral membrane provides the energy for the secretion of H^+ into the tubular fluid. This secretion serves to reabsorb the filtered bicarbonate, which is thus not excreted in the urine. The daily tubular secretion of H^+ is enormous, because we excrete 70 mmol of non-volatile acid and also have to match almost all of the total filtration flux of bicarbonate. At the brush border of the proximal tubule cell, *carboanhydrase* (CA) catalyses

the reaction, so CO_2 is formed and can enter the cell easily by diffusion (Fig. 17-4).

Fig. 17-4: Reabsorption of bicarbonate in the proximal and distal parts of the nephron.

Also within the cell CA facilitates the production of ($H^+ + HCO_3^-$). For each bicarbonate produced in the cell from the CO_2 of the tubular fluid, one bicarbonate ion diffuses to the interstitial phase and the renal venous blood back to the body (Fig. 17-4).

The cells of the thick ascending limb of the Henle loop also reabsorb bicarbonate by the same mechanism as in the proximal tubule

The small residue of bicarbonate enters the distal tubules, where it is reabsorbed almost totally through a special mechanism independent of Na^+ . In the *intercalated cells* of the collecting ducts, the reabsorption is dependent on a *proton-K^+-ATPase* (Fig. 17-4). The bicarbonate ion crosses the basolateral cell membrane in exchange of chloride through a *chloride-bicarbonate antiporter*. - This special mechanism is most likely ineffective in *distal renal tubular acidosis*.

Acidosis, which involves the intracellular space and stimulates production of *proton-K^+-ATP-ases*, also favours H^+ -secretion. Hereby, bicarbonate reabsorption is stimulated, whereas alkalosis inhibits bicarbonate reabsorption by the opposite mechanisms.

Aldosterone stimulates the *proton-K^+-ATPases* of the intercalated cells and the *Na^+-reabsorption/ K^+ -secretion* of the principal cells. Both effects favour H^+ -secretion and thus bicarbonate reabsorption.

7.2 The NH_3 /NH_4^+ -buffer

As stated above, the most important urinary buffer is NH_3/NH_4^+. This is because the synthesis of NH_3/NH_4^+ in the tubule cells is controlled by the acid-base status of the body. Acidosis stimulates NH_4^+-production from renal glutamate, whereas alkalosis stimulates hepatic urea production from hepatic glutamate

In a normally dieting adult person the amino acid load is from 90 g of protein (16% nitrogen) daily. The (90* 0.16)= 14.4 g nitrogen (1 mol) corresponds to 1000 mmol of NH_3 /NH_4^+. According to the equation:

$$2\ NH_4^+ + 2\ HCO_3^- \text{®}\ CO(NH_2)_2\ + CO_2 + 3\ H_2O$$

One mol of nitrogen daily produces 500 mmol urea, which is equal to the typical urinary urea excretion. The daily urea filtration flux is 900 mmol (5 mM * 180 l of plasma each day). The degree of reabsorption of the water-soluble urea depends upon the tubular flow rate

Under normal conditions only a small amount of nitrogen is used to produce hepatic glutamate, which is an excellent atoxic ammonia store that can transfer ammonia to the proximal tubules of the kidneys in cases of acidosis.

Fig. 17-5: Glutamate metabolism and renal ammonia production. The renal handling of ammonia and the excretion by diffusion trapping is shown.

In the proximal tubules of the kidneys, renal glutamate produces NH_4^+ and a-ketoglutarate.

+

One molecule of NH_4 is produced by deamination of one glutamine molecule by the enzyme, glutaminase, and a second by oxidative deamination of glutamic acid forming a -ketoglutarate that is metabolised. The NH_4^+ in the proximal tubule cells is in equilibrium with minimal amounts of NH_3 at the relatively low pH. The NH_4^+-secretion into the tubular fluid makes use of the *Na^+-H^+-antiporter*, where NH_4^+ substitutes H^+. The NH_4^+ passes with the tubular fluid to the thick ascending limb of the Henle loop, where a major portion is reabsorbed and accumulated in the interstitial fluid (Fig. 17-5).

Secretion of NH_4^+ in the collecting ducts involves a special mechanism. The NH_3 is lipid soluble and easily passes any membrane, so it reaches the tubular fluid of the collecting ducts and form NH_4^+ at the low pH (Fig. 17-5). The charged molecule cannot pass the membrane and it is trapped in the tubular fluid and eliminated in the urine. This *diffusion trapping* of charged molecules such as NH_4^+ is called *non ionic diffusion* - a general elimination principle for many charged metabolites and drugs. Excretion of NH_4^+ reduces the excretion of other positive ions.

The a -ketoglutarate is metabolised into bicarbonate. Bicarbonate of the extracellular fluid reacts with H^+ from hepatic phosphoric and sulphuric acid to form carbon dioxide and water. The $H_2PO_4^-$ (and a minimal amount of SO_4^{2-}) is excreted in the urine. On a mixed diet the production and excretion of non-volatile acids and bases results in a net excretion of acids equal to the daily net production of non-volatile acids.

With a urine pH of 6.5, organic acids such as lactic acid, b-hydroxybutyric acid, pyruvic acid etc., are present in the base form ($RCOO^-$ of Fig. 17-6). Most of the phosphoric acid is $H_2PO_4^-$, and almost all ammonia is in the NH_4^+ form.

Fig. 17-6: Excretion flux for organic bases ($RCOO^-$), titratable acid ($H_2PO_4^-$), and NH_4^+ in normal daily urine.

This is not so in an alkaline urine. At a urinary pH of 8, there is 5% NH_3 of the total, just as in ileum and colon, where the pH is also 8.

The high pK (=9.3) of NH_3/NH_4^+ has the consequence that in gastric juice with a pH of 1, the (pH-pK)- difference is -8.3, so virtually all ammonia must be NH_4^+. Even in body fluids with a pH of 7.3 the NH_3/NH_4^+ ratio is 1/100.

7.3 The titratable phosphate activity

The dominating buffer system in the urine is secondary/primary phosphate. This is because of its urinary concentration and of its pK (6.8) being close to urinary pH (6.5).

A healthy person has a renal filtration flux of phosphate of 180 mmol daily (1 mM*180 litres of plasma filtered daily). Phosphate is a threshold substance, which is reabsorbed, in the proximal tubules, where parathyroid hormone (PTH) inhibits phosphate reabsorption (Fig. 17-7). With 30 mmol left in the tubular fluid, the *secondary/primary phosphate-ratio* is 24/6 mmol as calculated in Fig. 17-7. Secretion of H^+ during the passage of the fluid through the

renal tubules converts HPO_4^{2-} to the acid form, $H_2PO_4^-$. Thus, in the final urine, the base/acid-ratio is 10/20 (Fig. 17-7). *Titratable phosphate acidity* in the daily urine is the amount of base (mmol) needed to titrate an acidic daily urine back to the pH of plasma and glomerular filtrate (pH 7.4). Weak acids are not titrated, because they are minimally dissociated at pH 6.5 to 7.4. Normally, the titratable phosphate acidity is 30 mmol in a 24 hour urine. In our example above, the distal H^+-secretion has titrated 14 mmol of HPO_4^{2-} to $H_2PO_4^-$ (Fig. 17-7). *Acidosis* increases the urinary *titratable phosphate acidity* (towards 50 mmol daily) in order to get rid of the acid.

Fig. 17-7: Titratable phosphate acidity produced in the distal tubules.

The small amount of bicarbonate in the daily urine (zero to 3 mmol) hardly affects the measured titratable phosphate acidity, and the ammonia buffer is not titrated in acid urine

8. Intracellular buffers

The importance of the *buffer capacity* of the *extended ECV*, and that of the intracellular fluid is comparable in the majority of acute conditions, with extended ECV as the initial distribution volume and the intracellular fluid participating importantly after hours. This is because the alteration of the cellular transport processes takes time.

The $[H^+]$ of the *intracellular fluid volume* (ICV) is higher than that of the extended ECV. The intracellular pH is precisely controlled in cells with different functions and needs, and the range of values is 7.0 -7.4 (Fig. 17-8).

Fig. 17-8: The proton concentrations and buffer bases of the extended ECV and of the ICV.

The buffer bases within the cells are proteins, phosphate and bicarbonate. The precise intracellular control is necessary for the pH-sensitive cellular processes with pH-optima for the enzyme systems. The active transport of H^+ out of the cell is a *coupled Na^+/H^+ -exchange*. The energy for this exchange is delivered by the Na^+-K^+-pump, which maintains the Na^+ -gradient across the cell membrane (Fig. 17-8). Carbohydrate- and K^+-containing meals, insulin, hyperkalaemia, adrenaline and aldosterone stimulate the Na^+-K^+-pump.

A bicarbonate-transport protein (capnoforin in the red cell and in many other cell membranes) transfers bicarbonate to the extended ECV by *bicarbonate/chloride exchange* (Fig. 17-8).

In disorders with extracellular accumulation of carbon dioxide, CO_2 diffuses rapidly into the cells. This causes a shift towards the right in Eq. 17-4. The intracellular buffers buffer the H^+. Bicarbonate leaves the cells both directly via capnoforin and via other membrane channels (Fig. 17-8), whereby intracellular bicarbonate falls. Intracellular acid accumulation accompanied by hyperkalaemia may develop, if not compensated by the lungs. The K^+-output follows the bicarbonate exit

Non-volatile acid is also buffered intracellularly during metabolic acidosis by movement of H^+ into the cell, where it reacts with proteins, phosphate and bicarbonate. During metabolic alkalosis movement of H^+ out of the cells in exchange of Na^+ (Na^+-influx in Fig. 17-8) also buffers non-volatile base.

Pathophysiology

Humans can suffer from 4 acid-base disorders: *1. Respiratory acidosis, 2 Respiratory alkalosis, 3 Metabolic acidosis and 4 Metabolic alkalosis.*

Acidosis (acidaemia) is defined as a disorder with pH in the arterial blood (pH_a) less than 7.35, and *alkalosis* (alkalaemia or baseosis) is defined as a condition with a pH_a larger than 7.45.

Each of these two disorders has *respiratory* and *metabolic* forms.

Respiratory acid-base disorders are caused by primary changes of P_{CO2}, and compensated by altered renal excretion of acid in a matter of days.

Metabolic acid-base disorders are caused by primary changes in BE, and compensated partially by the lungs in a matter of hours. The final correction of metabolic disorders is always renal and takes several days.

Let us now return to the *four primary* acute disorders and their *four chronic* forms:

1. Respiratory Acidosis

is caused by *hypoventilation* (or breathing of CO_2 containing air). Hypoventilation is associated with an *impaired ability* to eliminate CO_2, whereby P_{aCO2} increases and the accumulated CO_2 reduces the arterial pH.

For each mol of bicarbonate produced, one mol of non-carbonic buffer base is eliminated, which means that Base Excess (BE) is unchanged *zero* (Fig. 17-9). The slope of the BE -zero line depicts the buffer-base capacity of the extended ECV.

Any primary respiratory disorder is compensated *renally* over *days*. This is because the high intracellular $[H^+]$ increases the glutaminase synthesis and activity, the renal ammonia production, the urinary H^+-excretion (mainly NH_4^+ but also $H_2PO_4^{-)}$, with a virtually complete reabsorption of filtered bicarbonate. Hereby, BE becomes positive during compensation (arrow in Fig. 17-9)

Fig. 17-9: Acute respiratory acidosis with a base excess of zero, and its compensation.

In severe cases of chronic CO_2 accumulation, artificial ventilation is necessary. This is often the case in the terminal phase of chronic obstructive lung disease (ie, chronic bronchitis and emphysema). Other causes of respiratory acidosis are asthma, pulmonary cancer or tuberculosis, polio, drug overdose, anaesthesia, strangulation, near drowning and myasthenia gravis.

***2. Respiratory Alkalosis* is caused by hyperventilation.**

The hyperventilation is disproportionately high compared to the CO_2 production, whereby the P_{aCO2} falls and the pH increases (Fig. 17-10). When the alveolar ventilation is doubled, the P_{aCO2} is halved. This is a typical reaction to high altitude. As the P_{IO2} falls with increasing altitude, the P_{aO2} eventually falls below 55 mmHg, which stimulates the chemoreceptors to hyperventilation (CO_2 -wash-out).

Other typical cases are the anxious patient during an attack of asthma or the hysterical hyperventilation in neurotic patients. These patients often experience *tetanic cramps* (see below)

Hyperventilation before underwater swimming eliminates the CO_2 stimulus and shifts the oxyhaemoglobin dissociation curve towards the left. Hereby, oxygen is bound firmly to haemoglobin. When the P_{aO2} falls below 30 mmHg (4 kPa), blackout and grey-out occurs. Loss of consciousness below water is often fatal.

Fig. 17-10: Acute respiratory alkalosis and its compensation.

For each mol of bicarbonate eliminated, one mol non-carbonic buffer base is formed, which means that BE is maintained at *zero* (Fig. 17-10). As long as *no* non-carbonic acid or base is added to the extracellular fluid volume, the extracellular base excess remains unchanged (zero)

Acute respiratory alkalosis is compensated by *increased* renal excretion of *bicarbonate*, which is the result of decreased tubular H^+ -secretion. This is because the low P_{aCO2} reduces the tubular H^+-secretion, and the alkalosis inhibits formation and secretion of NH_4^+ . The renal mechanisms affect the production and activity of cellular enzymes and hormones, so it takes days to become effective.

After a few days the renal compensation of the respiratory alkalosis is complete, and the pH is normal. This is called *totally compensated respiratory alkalosis*. - In cases of asthma-anxiety or hysteria with *hyperventilation tetany*, simple rebreathing from a bag cures the disorder within minutes.

Dissociation of protein molecules occurs in all types of alkalosis in order to liberate H^+. The dissociation leads to *tetany*: Protein + Ca^{2+} ® Ca-proteinate + 2 H^+. The equilibrium dislocates towards the right in alkalosis. The falling extracellular [Ca^{2+}] activates Na^+-Ca^{2+}-pumps and opens Na^+-channels in the cell membranes of neurons, muscle cells and the myocardial syncytium. The Na^+-influx reduces the membrane potential and increases the excitability of the tissues, which causes tetanic cramps (almost continuous muscular contractions).

3. Metabolic Acidosis

is caused by *accumulation* of strong acids in the extended ECV. Metabolic acidosis is diagnosed by negative base excess, because both types of buffer bases are reduced. In Fig. 17-11 both types of buffers and the total concentration of buffer bases of the extended ECV are clearly reduced and the BE is -15 mM.

The iso-base excess-*line* (-15 mM) is steeper than the base excess-zero line, whereas the **iso**-base excess-*line* (+15 mM) is less steep than the BE-zero-line, due to the relative low pK-values of essential buffers.

Strong acids accumulate, because of excess production or impaired H^+ -excretion.

Hunger (*hunger diabetes*), diabetic ketoacidosis, lactic acid accumulation or high protein intake with increased production of hydrochloric and sulphuric acid, cause *excess production*.

Fig. 17-11: Acute metabolic acidosis: The base excess is reduced. The compensation is shown with an arrow.

Impaired renal H^+ excretion is related to increased loss of bicarbonate in the urine (due to renal failure). Diarrhoea causes acidosis by *loss of bicarbonate* with the faeces. Any loss of bicarbonate in the urine or faeces is equivalent to an addition of H^+ to the extracellular fluid

(Fig. 17-11).

Lactic acidosis is caused by increased lactic acid production during exercise, shock, anoxia or following cardiac arrest. Another type of *lactic acidosis* is caused by decreased hepatic lactate metabolism - often drug-induced.

Renal tubular acidosis is a damage of tubular cells caused by drugs or immunological reactions - or it may be inherited. The impaired H^+ secretion reduces the tubular bicarbonate reabsorption. Kidney disease with destruction of a large number of nephrons reduces the tubular capacity to excrete H^+ and NH_4^+ in the urine. The chloride-bicarbonate antiporter of the intercalated cells of the collecting ducts is probably *ineffective* (see above).

The patient with metabolic acidosis suffers from dyspnoea (deep and frequent *Kussmaull respiration*). This hyperventilation is a *respiratory* compensation, which develops over hours as a reduction in P_{aCO2} (Fig. 17-11). This compensation is caused by the chemoreceptors, which are surrounded by the extended ECV and stimulated by its hydrogen ion concentration to react with hyperventilation.

Acidosis shifts the oxygen dissociation curve to the right (ie, the Bohr effect), increasing the delivery of oxygen to the tissues. Acidosis stimulates K^+ -loss from the cellular pool, because of K^+ -efflux from the cells (Fig. 17-4). *Chronic acidosis*, however, inhibits 2,3-DPG production, which tends to shift the oxygen dissociation curve back towards the left

When renal function is normal, K^+ -loss from the cellular pool may lead to K^+ -deficiency. When renal K^+ -secretion is impaired, the cellular K^+ -efflux may lead to hyperkalaemia.

Oxygen enriched air is administered in cases of *lactic acidosis* with poor tissue bloodflow. Insulin must be given in diabetic ketoacidosis.

A patient with *metabolic acidosis* and a base excess of -15 to -25 mM may have to be treated with bicarbonate infusion. The *primary strategy* is to eliminate the lack of base of the extended ECV. This strategy is accomplished by infusion of approximately X mmol bicarbonate (X= negative BE in mM multiplied by extended ECV). The extended ECV is approximately 20% of the body weight in kg or l. Careful monitoring of acid-base variables is necessary. Two errors are possible. Hours later, H^+-ions from the cells enter the extended ECV and a further bicarbonate infusion is necessary. Correction of the primary disease (insulin to diabetic ketoacidosis) may in itself cure the acidosis by combustion of keto-acids to bicarbonate with the danger of overinfusion with bicarbonate.

Continuous control of acid-base variables over days is therefore important.

Rapid infusion of bicarbonate may be dangerous. A rapid decrease in $[Ca^{2+}]$ releases *tetanic cramps* (see alkalosis above). Administration of an overshoot of Na^+-bicarbonate leads to volume expansion, pulmonary oedema, and a new equilibrium with too much CO_2, which diffuses into the cells, worsening the *intracellular acidosis*, and causing bicarbonate efflux accompanied by *hyperkalaemia*.

The *final* correction of a *metabolic* acidosis is always *renal*.

4. Metabolic Alkalosis

is caused by a *primary accumulation* of strong bases in the extended ECV. Both the

[bicarbonate] and the [non-carbonic buffer base] is increased, so the BE is increased (Fig. 17-12). The actual [bicarbonate] is often above 27 mM, which is the renal plasma concentration threshold for reabsorption. Above this threshold lots of bicarbonate is lost in the urine.

Vomiting (loss of gastric acid and volume depletion), *increased* metabolism of lactate and citrate (turns into bicarbonate and water), and *excessive intake of bases* towards gastric ulcer can cause this form of alkalosis. Long-term use of thiazides and loop diuretics, K^+-deficiency, and excess secretion of mineralocorticoid increase the H^+-secretion in exchange for Na^+ in the distal tubules. This increases *renal bicarbonate reabsorption* and leads to metabolic alkalosis.

Fig. 17-12: Acute metabolic alkalosis: The base excess is increased. The hypoventilatory compensation is shown with an arrow.

The *hypoventilatory* compensation reduces pH, but raises P_{aCO2} (Fig. 17-12). The compensation is never total, since the rise in P_{aCO2} and the fall in P_{aO2} in itself oppose the hypoventilation.

The *delayed* renal correction increases bicarbonate excretion by reducing its reabsorption. This correction is also counteracted by the rise in P_{aCO2}, which stimulates tubular bicarbonate reabsorption. The *final* renal correction of metabolic disorders takes *several days*.

Careful monitoring with replacement of Na^+ and K^+ is essential, in order to improve the renal excretion of bicarbonate. In metabolic alkalosis the kidneys increase H^+-secretion less than the bicarbonate filtration, whereby the bicarbonate excretion is increased. - Metabolic alkalosis combined with hypokalaemia and reduced ECV often has *increased* H^+-secretion, whereby the bicarbonate excretion is reduced. Such cases must be treated with NaCl and KCl.

Only rarely is it necessary to infuse acid, when deficits of NaCl, K^+ and Mg^{2+} are corrected.

A patient with *metabolic alkalosis* is therefore rarely treated with infusion of acids. In the very few cases, the acid of choice is an *ammonium chloride solution*, which produce H^+, when NH_3 is used for carbamide (urea) production in the liver. Metabolic alkalosis is difficult to compensate by the body and difficult to treat. Each OH^- has a molecular weight 17 times larger than H^+, so OH^- passes the membrane channels comparatively slowly in *metabolic alkalosis* compared to the H^+ -transfer of metabolic acidosis. This results in delayed intracellular transfer and buffering of the alkalosis.

Cerebral insufficiency is common in alkalosis, and the respiratory centre is depressed. Alkalosis displaces the oxyhaemoglobin dissociation curve to the left, impairing the delivery of oxygen to the tissues. Dissociation of protein molecules may lead to tetany. Protein anions bind Ca^{2+} during alkalosis, reduces the free serum [Ca^{2+}] and triggers tetanic cramps (see explanation).

The nine van Slyke conditions

These *four primary* acid-base disturbances and their *four compensated or chronic* types constitute, together with the normal condition, the *nine* van Slyke conditions (Fig. 17-13). Plotting the measured pH and P_{CO2} in the acid-base chart allows estimation of the base excess (BE), and combined with the case history, the correct diagnosis can be reached.

Fig. 17-13: The Siggaard-Andersen acid-base chart with the 9 conditions of van Slyke

(modified with permission).

Please observe that each point on the chart can be reached in several ways. Van Slyke, who made the first apparatus to diagnose acid-base disturbances, emphasised the importance of the case history and common clinical sense. The importance is illustrated by the following three examples. The first example is a case of respiratory acidosis due to *chronic obstructive lung disease* (COLD in Fig. 17-13) and a metabolic acidosis due to *diabetic coma*. Without the case history it is difficult to diagnose the primary and secondary events in the development of the patients condition. The second example is a case of *respiratory alkalosis* due to acute mountain sickness (AMS), complicated by a metabolic alkalosis (AMS in Fig. 17-13), because the patient is loosing acid by vomiting. The third example is a serious case of birth anoxia with a combined respiratory and metabolic acidosis. In this case instantaneous intubation and oxygenation with 50-40-30 % oxygen saved the newborn.

The case history, not the acid-base variables, is the only source to the sequence of events.

Equations/Reactions (Eq)

- **Acid** (HB) is defined as a compound that can release a hydrogen ion or proton (H^+), whereas a **base** (B^-) can take up H^+. The H^+-concentration is symbolised as $[H^+]$ just as other concentrations:

  Eq 17-1: HB ® $H^+ + B^-$, where the dissociation constant (K´) is defined by:

  $K' = [H^+]^* [B^-]/[HB]$.

- This is the **mass action equation**, which can be logarithmically rearranged to:

  Eq 17-2: $pH = pK' + \log ([B^-]/[HB])$;

The pK´ is equal to the pH, when the acid is 50% dissociated.

- For the carbonic acid system:

  Eq 17-3: $CO_2 + H_2O$ Û $^*H_2CO_3$ Û $H^+ + HCO_3^-$

  Eq 17-4: pH = pK + log([bicarbonate]/[dissolved CO_2]).

- This is the Henderson-Hasselbalch-Equation - a logarithmic rearrangement of the mass action equation.

- The law of Henry: Dissolved CO_2 (ml l^{-1} of blood) is in equilibrium with and equal to P_{CO2} multiplied by the **solubility coefficient** α according to Henrys law: **[dissolved CO_2]** = α***** P_{CO2} (for the definition and size of α for CO_2 see Box 13-1):

  Eq 17-5: $(0.51 \times 1000 \times P_{CO2})/(760 \times 22.22) =$ **$(P_{CO2} \times 0.03)$ mM**

with P_{CO2} in mmHg.. One mmol of ideal gas equals 22.4 ml STPD and since CO2 deviates slightly from an ideal gas its molar volume is only 22.22 l per mol. All the constants combine to *0.03*.

With kPa as unit for P_{CO2} (760 mmHg equals 101.3 kPa):

$0.03 \times 760/101.3 = 0.225$; with this new constant the dissolved CO_2 is equal to: ($P_{CO2} \times 0.225$) mM.

- At normal ionic strength and temperature the pK is 6.1

Eq 17-6: pH = 6.1 + log([bicarbonate]/(0.03 × PCO2))

with P_{CO2} in mmHg or

pH = 6.1 + log([bicarbonate]/(0.225 × P_{CO2})) with P_{CO2} in kPa**.**

This (Eq 17-6) is the *final* Henderson-Hasselbalch-Equation.

- Proteins dissociate during alkalosis:

Eq 17-7: Protein + Ca^{2+} ® Ca-proteinates + 2 H^+, and the equilibrium dislocates towards the right. Of all the proton-binding groups on proteins, the calcium-binding groups are only a minor fraction.

Self-Assessment

17. Multiple Choice Questions

I. A patient who has been vomiting for several days is in hospital with the following findings in a sample of arterial blood: pH = 7.60; Base Excess = +20 mM; and P_{aCO2} = 6.7 kPa (50 mmHg). Each of the following disorders and statements has True/False options:

A. Compensated respiratory alkalosis.

B. Metabolic alkalosis

C. Hypoventilation

D. Increased base deficit

E. Increased bicarbonate concentration.

II. **The following five statements have True/False options**:

A. In connection with alkalosis the free serum [Ca^{2+}] is reduced, as proteins liberate H^+ and bind more Ca^{2+}.

B. Metabolic acidosis is diagnosed by a negative base excess (base deficit).

C. The titratable acid of the urine is the amount of acid needed to titrate the acid urine to the pH of plasma.

D. The treatment of a patient with acute metabolic alkalosis and a base excess of 20 mM is hypoventilation in a respirator.

E. Acute Mountain Sickness often causes vomiting.

Case History A.

A female patient, 25 years of age, is undergoing a routine examination at an outpatient ward. Her pH is 7.42, and P_{aCO2} is 5 kPa (37 mmHg).

The patient is suffering from a disorder with acute attacks of anxiety, panic, dyspnoea and tetanic cramps.

A week later the patient is admitted to the emergency ward of the hospital with severe hyperpnoea and tetanic cramps. She has abnormal blood gas tensions: P_{aCO2} 2.7 kPa (20 mmHg), P_{aO2} 14.6 kPa (110 mmHg), and pH is 7.64.

1. *Read her Base Excess of the extended ECV at the routine check by the help of Fig. 17-13.*

2. *Read her Base Excess of extended ECV during emergency conditions by the help of Fig. 17-13.*

3. *What is the most likely diagnosis?*

Case History B

A 56-year-old male with chronic bronchitis and emphysema suddenly developed respiratory failure during flying in a pressurised intercontinental plane. The cabins of commercial aeroplanes are pressurised to a P_B of 80 kPa corresponding to 2000 m. The patient was bluish in colour and the cabin personnel gave him a mask for breathing with pure (100%) oxygen, which was continued until landing. The patient was then admitted to hospital. An arterial blood sample (without supplementary oxygen) revealed the following: P_{aO2} 6.7 kPa (50 mmHg), P_{aCO2} 8 kPa (60 mm Hg) and pH_a 7.44. The saturated water vapour pressure at body temperature is 6.26 kPa.

1 *Calculate P_{IO2} in the moist tracheal air of the patient, when he breathes the air of the pressurised cabin just before respiratory failure. Is the oxygen tension insufficient?*

2 *Calculate the P_{AO2} of the patient at the hospital assuming P_{ACO2} to be equal to P_{aCO2} and RQ =1.*

3. *Read the base excess of the extended ECV by the help of Fig. 17-13 and describe the acid-base disorder.*

Case History C

Two male persons with a normal arterial pH (pH_a = 7.40) are in hospital. One person starts to vomit excessively, which increases his pH to 7.80. The other person, a fit athlete, does a cardiopulmonary exercise test, which reduces his pH to 7.00.

1. *Calculate the three concentrations of H^+ (in nmol l^{-1} = 10^{-9} mol l^{-1}) at these conditions.*

2. *The two changes in pH from the normal pH = 7.4 are similar. Why are the related changes in H^+ concentration not identical?*

3. *Describe other conditions with metabolic acidaemia and metabolic alkalaemia.*

Try to solve the problems before looking up the answers.

Highlights

- *The pH of the intra- and extra-cellular body fluids is maintained within narrow limits by the co-ordinated function of the respiratory and the renal system.*
- *Persons on a high-protein diet produce 50-100 mmol fixed or non-volatile acids and 24 mol of volatile carbonic acid daily.*
- *Produced and ingested acid-base is normally excreted, so the acid-base balance is maintained.*
- *The kidneys prevent loss of bicarbonate. Of the renally ultrafiltered bicarbonate (4500 mmol daily) we reabsorb 90% already in the proximal tubules, and hardly any bicarbonate is excreted in the urine.*
- *Urinary buffers (ammonia and phosphate) are essential for the excretion mechanism of non-volatile acid.*
- *Acidosis (acidaemia) is defined as a disorder with pH in the arterial blood (pH_a) less than 7.35, and alkalosis (alkalaemia) is defined as a condition with pH_a larger than 7.45.*
- *Respiratory Acidosis is caused by hypoventilation (or breathing of CO_2 containing air). Hypoventilation is associated with an impaired ability to eliminate CO_2, whereby P_{aCO2} increases and the accumulated CO_2 reduces the arterial pH.*
- *Respiratory Alkalosis is caused by hyperventilation, which is disproportionately high compared to the CO_2 production, whereby the P_{aCO2} falls and the pH increases.*
- *Metabolic Acidosis is caused by accumulation of strong acids in the extended ECV. Metabolic acidosis is diagnosed by negative base excess, because both types of buffer bases are reduced.*
- *A primary accumulation of strong bases in the extended ECV due to vomiting, increased metabolism of lactate/citrate and excessive intake of bases cause metabolic alkalosis.*
- *Respiratory acid-base disorders are caused by primary changes of PCO2 , and are compensated by altered renal excretion of acid in a matter of days.*
- *Metabolic acid-base disorders are caused by primary changes in Base Excess, and are compensated partially by the lungs in a matter of hours. The final correction of metabolic disorders is always renal and takes several days, because production and activation of cellular enzymes and hormones are involved.*
- *Alkalosis dissociates protein molecules, which bind ionised calcium. The hypocalcaemia opens Na+ - channels, and the influx increases the excitability of neuromuscular tissues, which releases tetanic cramps.*

Further Reading

Astrup, P. and J.W. Severinghaus. "The history of blood gases, acids and bases." *Munksgaard*, Copenhagen, 1986.

Koeppen, B M. Renal regulation of acid-base balance. *Am. J. Physiol.* 275 (Adv. Physiol. Educ. 20): S132-S141, 1998.

Siggaard-Andersen, O. The acid-base status of the blood. *Munksgaard*, Copenhagen, 1974.

Siggaard-Andersen, O & OH Gøthgen (1995) Oxygen and acid-base parameters of arterial and mixed venous blood. Relevant versus redundant. *Acta Anaesthesiol Scand 39*. Suppl 107, 21-27.

Siggaard-Andersen, O and N Fogh-Andersen (1995) Base excess and buffer base (strong ion difference) as measure of a non-respiratory acid-base disturbance. *Acta Anaesthesiol Scand 39*. Suppl 107, 123-128.

Return to top

Return to content

Chapter 18.

Exercise, Sports And Doping

Study Objectives

- To *define* factors of importance to oxygen uptake, cardiac output, and ventilation during exercise.
- To *describe* the rise in ventilation, oxygen uptake and cardiac output during increasing exercise intensity, the concepts anaerobic threshold, oxygen deficiency and oxygen debt.
- To *calculate* the relationship between the major variables.
- To *explain* the metabolism and limits of exercise, typical sport injuries, doping, the effects of training including health consequences, and hypotheses of the cardiopulmonary regulation.
- To *use* the concepts in problem solving and case histories.

Principles

- *The human body has a redundancy of overlapping cardio-pulmonary control systems during exercise.*
- *The redundancy-hypothesis, with neural factors dominating at the start of work and peripheral feedback control during steady state, is a possible explanation of the hyperpnoea of exercise and the related increase in cardiovascular activity .*

Definitions

- **Anaerobic threshold.** This is the exercise level above which the energy requirements can be satisfied only by the combined aerobic metabolism and anaerobic glycolysis. Lactic acid is produced and stimulates the peripheral chemoreceptors. Hereby, ventilation starts to increase out of proportion to the rise in oxygen uptake.
- **Blood doping.** Blood boosting is an artificial improvement of performance through an increase in the haemoglobin binding capacity. Blood doping (one litre) definitely improves the oxygen transport with the blood and also the maximal oxygen uptake, which is beneficial to distance runners.
- **Doping**: Athletes who use drugs or other means with the intention to improve performance artificially are doped by definition.
- **Endurance capacity** or *fitness number* is given as the *maximal oxygen uptake* in ml of oxygen STPD min^{-1} kg^{-1}.
- **Energy equivalent** of oxygen on a mixed diet is defined as the heat energy liberated in the

body per litre of oxygen used (20 kJ of energy per litre at an RQ of 0.8).

- **Flow Units** (FU) measure relative bloodflow as the number of ml of blood passing an organ per 100 g of tissue and per min.
- **Mean Arterial Pressure** (MAP) is the arterial blood pressure measured as the sum of the diastolic pressure plus 1/3 of the pulse pressure (see below).
- **Mechanical efficiency** is the ratio between external work and the total energy used during work.
- **Oxygen debt** is defined as the extra volume of oxygen that is needed to restore all the energetic systems to their normal state after exercise.
- **Oxygen deficiency** is defined as the difference in oxygen volume between an ideal, hypothetical oxygen uptake and the actual uptake in real life. The missing oxygen volume at the initiation of exercise is the oxygen deficit.
- **Pseudo-doping.** Many drugs reputedly increase athletic performance, but the fact remains that such effects rarely show up in double-blind controlled trials. - On the contrary, serious side effects occur with a biologically high and statistically significant frequency.

Essentials

This paragraph deals with *1. Athletes and training, 2. Fitness testing, 3. Limits of exercise performance, 4. The anaerobic threshold, 5. Ventilation and oxygen uptake, 6. Cardiopulmonary control, and 7. Oxygen debt and deficiency.*

1. Athletes and training

At the start of exercise, signals from the brain and from the working muscles bombard the cardiopulmonary control centres in the brainstem. Both cardiac output and ventilation increase, the a-adrenergic vasoconstrictor tone of the muscular arterioles falls abruptly, whereas the vascular resistance increases in inactive tissues. The systolic blood pressure increases, whereas the MAP only rises minimally during dynamic exercise. The total peripheral vascular resistance (*TPVR*) falls during moderate exercise to 0.25-0.3 of the level at rest, because of the massive vasodilatation in the muscular arterioles of almost 35 kg muscle mass. This is why the major portion of cardiac output passes through the skeletal muscles (Fig. 18-1) and why the diastolic pressure often decreases during exercise. The coronary bloodflow increases, and at some intensities of exercise we see increases in the skin bloodflow (Fig. 18-1).

Fig. 18-1: Distribution of cardiac output during exercise. HBF means hepatic bloodflow, and CBF is cerebral bloodflow.

A top athlete increases his cardiac output from 5 to 30-40 l of blood per min, when going from rest to maximal dynamic exercise (Fig. 18-1). However, the muscle bloodflow can rise 25 fold in the total muscle mass. Accordingly, the total muscular oxygen uptake rises 85 fold from rest to maximal exercise (see Box 8-1 with calculations).

Training improves the capacity for oxygen transport to the muscular mitochondria, and improves their ability to use oxygen. After long-term endurance training the athlete typically has a lower resting heart rate, a greater stroke volume, and a lower *TPVR* than before. The maximum oxygen uptake progressively increases with long-term training, and the extraction of oxygen from the blood is increased. The lung diffusion capacity for oxygen probably increases

by endurance training. The capillary density of skeletal muscles, the number of mitochondria, the activity of their oxidative enzymes, ATPase activity, lipase activity and myoglobin content all increase with endurance training. Endurance training also produces a rise in ventricular diastolic volume. Strength training (weight lifting) produces a rise in left ventricular wall thickness without any important increase in volume

During dynamic exercise the stroke volume increases as does heart rate, and the residual ventricular volume decreases (Fig. 18-2).

Fig. 18-2: The pressure-volume loop of the left ventricle in a healthy male at rest (red curve) and during dynamic exercise (blue curve).

Although the peak ventricular pressure during systole rises considerably and thus the arterial peak pressure, the diastolic pressure falls because of the massive fall in total peripheral vascular resistance. The contractility of the heart is depicted as the slope of the pressure-volume curve. The contractility increases considerably from rest to exercise (Fig. 18-2).

2. Fitness testing

A simple objective method of estimating the *endurance capacity* or *fitness number* (maximal oxygen uptake, $\dot{V}_{O2}$max) in a person, is to measure the heart rate (HR) at a standardised work on a cycle ergometer. The test rests on the assumption, that there is a linear increase in HR with increasing oxygen uptake or work rate (Fig. 18-3). The net mechanical efficiency is relatively constant in each individual (approximately 20%). On a mixed diet the energy equivalent for oxygen is 20 kJ per l (STPD), so it is easy to calculate the volume of oxygen corresponding to any maximal work rate extrapolated from Fig. 18-3.

The test subject wears light clothes, and is not allowed to smoke, eat or work at least three hours before the measurement, which takes place in a comfortable not too warm room.

Fig. 18-3: The relationship between work intensity and steady state heart rate at work. The fitness number is in ml STPD oxygen per min and per kg of body weight.

The work intensity on the ergometer is chosen to produce a heart rate between 130-150 beats per min, and must be continued for at least 5 min in order to secure respiratory steady state. Respiratory steady state means that the pulmonary oxygen uptake is equal to the oxygen uptake of the tissues. This implies that ventilation and heart rate at work is also stable.

As a standard work rate of 100 W is chosen for females, and 150 W is standard for untrained males. An optimally performed fitness test results in a heart frequency of 130-160 beats per min. The submaximal test is designed by P.-O. Åstrand (and is sometimes performed at 2 work rates - Fig. 18-3). Since the rise in HR is linearly correlated to the work rate, a line is drawn through the points (Fig. 18-3). The line is extended until it reaches the horizontal line (maximal HR). Here, the *maximal heart rate* is the mean of the maximal heart rate of persons of the same age and sex.

The rise in mean arterial pressure (MAP) during dynamic exercise is often minimal, because of the arteriolar dilatation with a large, rhythmic bloodflow through the working muscles. The *TPVR* is typically reduced to 1/4 of the value at rest.

In contrast, static exercise often results in a doubling of the MAP, because a large muscle mass is contracting and the contraction is maintained. Static work is typically accomplished with a low cardiac output, so the *TPVR* is relatively high. This is dangerous to elderly people with known or unknown degrees of atherosclerosis.

3. Limits of exercise performance

The limitation of performance (measured as oxygen uptake) depends upon the type of work and upon the person. Several factors are involved. The mass balance principle provides an overview (see Eq. 18-1).

Both the maximum cardiac output and the maximum arterio-venous O_2 content difference are limiting factors. Healthy persons have redundant ventilation and diffusion capacity in their lungs imposing no limitation.

3.1. ***Pulmonary ventilation***. *Increasing work rate* (with 15 W more each min) leads to a marked increase in pulmonary ventilation without any ceiling being reached even at maximal oxygen uptake ($V°_{O2}$max). The steeper rise in ventilation is shown by its deviation relative to the thin line towards the right (Fig. 18-4). Light exercise often increases ventilation by an increased tidal volume (V_T). With increasing work rate also the respiratory frequency must rise from 10 towards 50 respiratory cycles per min. The tidal volume can increase to half the value of the vital capacity (6 l), which corresponds to an exercise ventilation of (3*50=) 150 l per min. *At exhaustion* the ventilation is much greater than at the point, where maximal oxygen uptake is already reached. At this maximum many individuals can increase ventilation further voluntarily. The alveolar gas tensions, P_{AO2} and P_{ACO2}, are essentially maintained during most work rates. At maximal work rate the P_{AO2} increases and P_{ACO2} decreases 5-10%. This fact illustrates effective gas exchange or adequate ventilation during non-exhausting exercise. Thus ventilation is not the limiting factor in these healthy persons.

3.2. The oxygen utilisation in the tissues is not a likely limitation in healthy people. A group of skiers increased their maximal oxygen uptake further, when they started arm work during continued running. Obviously, the *maximal oxygen uptake* measured during running is not always maximal.

3.3 Pulmonary diffusion capacity for oxygen with failure of the lungs to fully oxygenate blood. This is certainly not a limiting factor in healthy persons with a redundant lung diffusion capacity. The arterial blood is fully saturated with oxygen even during the most strenuous exercise at sea level. The lung diffusion capacity (D_{L02}) increases, because the number of open lung capillaries is increased, the surface area increases and the barrier-thickness is reduced. In addition, O_2 transport is boosted further by the rise in cardiac output from 5 to 30 l of blood per min

Oedema or interstitial pulmonary fibrosis leads to thickening of the alveolar-capillary barrier, which will impede O_2 exchange. The reason is that the pulmonary vascular volume is reduced (reduced capillary transit time), and thus the diffusion equilibrium point is moved towards the end of the capillary. If patients with lung diseases try to exercise, this problem is further aggravated by the still more reduced capillary transit time. Thus exercise would impose a significant diffusion limitation on O_2 transfer.

3.4. Cardiac output. Limited transport capacity for oxygen caused by limited peripheral bloodflow is the only logical explanation. Limitations in reducing *TPVR* or in the pumping capacity of the heart could cause the limited muscle bloodflow. When work is maintained at peak cardiac output and maximal oxygen uptake, the blood pressure falls as more vasodilatation occurs and there are no signs of even a slight relative increase in the low *TPVR*.

- The major limitation to exercise in well-trained athletes is the heart's pumping capacity in delivering oxygen to the working muscles.

4. The Anaerobic threshold

The *anaerobic threshold* (AT) is the exercise level at which the energy requirements can be satisfied only by the combined aerobic metabolism and anaerobic glycolysis. The lactic acid formed in the muscle cells diffuse into the blood and causes a metabolic acidosis, which stimulates the peripheral chemoreceptors. Hereby, ventilation starts to increase out of proportion to the rise in oxygen uptake (Fig. 18-4).

Just after the AT is passed, the ventilation increases proportional to the increase in carbon dioxide output (ie, so-called *normo-capnic* buffering). Accordingly, ventilation increases linearly with carbon dioxide output but out of proportion with the oxygen uptake. The carbon dioxide output and ventilation will increase faster than oxygen uptake, because bicarbonate react with the lactic acid produced, so CO_2 is liberated, added to the metabolic CO_2 production and eliminated by hyperventilation, causing P_{aCO2} to fall (ie, hyperventilation)

The rise in blood [lactate] is gradual, and Fig. 18-4 does not show any sign of a lactic acid threshold at the anaerobic threshold. Note the total rise in plasma [lactate] of 10 mM, which is equal to the fall in plasma [bicarbonate] from 24 to 14 mM (Fig. 18-4).

Exercise levels above the maximal aerobic capacity is called *supra-maximal work*. Here, the anoxia leading to a metabolic or lactic acidosis contribute with a large ventilatory drive, as shown in the steep component of $\dot{V}_E$ (Fig. 18-4 and 18-5).

Fig. 18-4: Ventilation and arterial blood concentrations (pH, lactate and bicarbonate) at rest and during an incremental work test on a cycle ergometer up to 100%.

Lactate is produced even at light exercise, but only minimal amounts are liberated to the blood (Fig. 18-4). Untrained subjects at any oxygen uptake, have higher ventilation and heart rate than the trained. The AT in untrained persons is often about 50% of maximal oxygen uptake, whereas the AT of athletes approaches 80%. Patients with heart disease increase their blood [lactate] at a minimal activity.

The oxyhaemoglobin dissociation curve is moved progressively to the right as exercise intensity increases due to the rise in 2,3-diphosphoglycerate (DPG) concentration (Fig. 8-3) and to the rise in temperature.

Above the AT, when oxidative metabolism is high, extra mechanical output is financed by anaerobic energy generation. The end product is lactic acid (Fig. 18-4). The *lactic acidosis* causes a further shift to the right of the oxyhaemoglobin dissociation curve easing oxygen delivery to the mitochondria. Lactate, nitric oxide and adenosine also dilatate muscle vessels and increase the number of open capillaries, thus improving the diffusion of oxygen from capillary blood to the mitochondria.

5. Ventilation and oxygen uptake

Results from an untrained person with a maximal oxygen uptake of 2.7 l STPD min^{-1} (AT: 1.3 l STPD min^{-1}), and from a top athlete with 6 l STPD min^{-1} (AT: 3.6 l STPD min^{-1}) are shown in Fig. 18-5. Several studies have shown oxygen uptake to remain at maximal level despite increasing work rates, and with carbon dioxide output increasing too. These curves also illustrate that ventilation - in these persons -is not the limiting factor for maximal oxygen

uptake.

If the athlete is suddenly breathing oxygen instead of atmospheric air, while working at a high level (5-6 l STPD min^{-1}), a drastic fall in ventilation will occur within 30 s. This is not a chemoreceptor response, since there is no stimulus. The oxygen breathing reduces the blood [lactate], but not within 30 s. Oxygen breathing abruptly increases the diffusion gradient and thus the rate of diffusion from haemoglobin to the muscle mitochondria. Exhaustive exercise with a severe metabolic acidosis may cause the steep rise in ventilation without a further rise in oxygen uptake (Fig. 18-5). In this case, oxygen seems to diffuse at a reduced rate from haemoglobin to the muscle mitochondria.

Fig. 18-5: Ventilation and oxygen uptake in an untrained person with a maximum oxygen uptake of 2.7 l per min. Results from a top athlete, with a $V°_{O2}max$ of 6 l min^{-1} breathing air (•) or oxygen (o) is shown for comparison.

Strenuous exercise is also associated with a rise in plasma concentration of catecholamines, dehydration and a rise in core temperature approaching 41^{o} C. The sensitivity of most receptors is increased in an overheated body. Increased activity of the *arterial chemoreceptors* causes hyperventilation in exercise situations where plasma-K^{+} is high and P_{aO2} is dangerously low. The athlete approaches exhaustion and collapse.

6. Cardiopulmonary control

The proportional increase in ventilation and cardiac output with increasing oxygen uptake suggests a common control system. The integrator consists of sensory and motor cortical areas, and the brain stem neighbour-centres for respiratory and cardiovascular control. The link between the respiratory and the circulatory control system is probably established in the neural network of the brain stem centres.

The nucleus of the tractus solitarius is the site of central projection of both chemoreceptors and baroreceptors. The respiratory and the cardiovascular systems are connected during most forms of dynamic exercise (Fig. 18-6), but they can also operate differently. There is a sharp rise in ventilation within the first breath at the on-set of exercise, and cardiac output also increases abruptly (Fig. 18-6). Both variables increase progressively over minutes until a steady state is reached. At the offset of exercise, ventilation and cardiac output falls instantly (Fig. 18-6).

The cardiopulmonary adjustments to exercise comprise an integration of I. neural and II. humoral factors.

I. The neural factors consist of: 1) Signals from the brain, 2) Reflexes originating in the contracting muscles, and 3) the central & peripheral chemoreceptors.

1. Signals from the brain to the active muscles passes the reticular activating system (RAS) in the reticular formation of the medulla, which includes the respiratory (RC) and cardiovascular centres. This signal transfer is called irradiation from the motor cortex to the RC, and proposed as an explanation of the exercise hyperpnoea. The mesencephalon and hypothalamus are also involved in the Krogh irradiation hypothesis now called central command. Cortical activation of the sympathetic nervous system accelerates the heart, increases myocardial contractility, dilatate the muscular arterioles and contract other vascular beds such as the splancnic region. Speculative mechanisms as irradiation or central command are so-called feedforward hypotheses.

Fig. 18-6: The exercise hyperpnoea and the rise in cardiac output follow the same pattern.

2 Afferent signals from proprioceptors in the active muscles through thin myelinated and unmyelinated fibres in the spinal nerves (type III and small unmyelinated type IV) to RC are the best-documented feedback hypothesis.

3. Central and peripheral chemoreceptors are sensitive to the final product of metabolism, carbon dioxide. The carbon dioxide molecule is most likely the controlled variable, perhaps as P_{aCO2}. The pH, P_{aO2}, and P_{aCO2} are normal during moderate steady state exercise, where the central chemoreceptors dominate. However, during transitions from rest to exercise and during severe exercise the peripheral chemoreceptors are stimulated. Stimulation of peripheral chemoreceptors increases the rate and depth of respiration and causes vasoconstriction.

II. The humoral factors that influence skeletal muscle bloodflow, cardiac output and ventilation are metabolic vasodilatators and hormones. Neural and chemical control mechanisms oppose each other. During muscular activity the local vasodilatators supervene. The local vasodilatators have not been identified. Ischaemic mitochondria in fast oxidative muscle fibres release many vasodilatators such as adenosine, AMP, and ADP. However, it is possible to block many of the neural and humoral factors without disturbing the proportional exercise hyperpnoea and the rise in cardiac output**.** These experiences suggest that the human body have a redundancy of overlapping control systems. The *redundancy-hypothesis*, with neural factors dominating at the start of work and peripheral feedback control during steady state, is a logical compromise.

7. Oxygen debt and deficiency

The *O_2 deficit* is defined as the difference in O_2 volume between an ideal, hypothetical O_2 uptake and an actual uptake as it occurs in real life (see Fig. 18-7). The missing O_2 volume is the oxygen deficit.

The energy demand increases instantaneously at the start of a working period, but the actual O_2 uptake via the lung lags behind for 2 min. The oxygen demand deficit is provided for by the O_2 stores (oxymyoglobin) and by anaerobic energy.

Fig. 18-7: The oxygen deficit and the oxygen debt at exercise.

The *oxygen debt* is defined as the extra volume of O_2 that is needed to restore all the energetic systems to their normal state after exercise (Fig. 18-7). The non-lactic O_2 debt following moderate work is characterised by maintained blood lactate concentration around the normal resting value of 1 mM. The non-lactic debt is maximally 3 l, used for regeneration of the Phosphocreatine and for refilling the O_2 stores. The lactacide O_2 debt following supramaximal work (100-400 m dash) can amount to 20 l and the blood [lactate] to as high as 20-30 mM. This O_2 debt is used for oxidation of 75% of the lactate produced, and for the formation of 25% of the lactate to glycogen in the liver. Restoration of Phosphocreatine etc following activity, is a process referred to as repayment of the O_2 debt. However, it is very uneconomical, since the debt is often twice as high as the O_2 deficit.

Pathophysiology

The pathophysiology of sports is related to the ultimate limits of human performance. Severe exercise for prolonged periods, such as a 20-fold rise in metabolic rate in a marathon runner,

sometimes result in life-threatening conditions: *Histotoxic hypoxia* with blockage of ATP production, dehydration, hyperthermia and metabolic acidosis with a pH_a below 6.9.

Following a short paragraph on *1. Muscle fatigue*, two consequences of aggressive attitudes in competitions are dealt with here: *2. Sport injuries* and *3. Doping*. The final point is *4. Fit for life*.

1. Muscle fatigue

Muscular contraction releases a great ionic leak (Na^+ -influx and a K^+-outflux) through the skeletal muscle membrane, which elicits the action potential (Fig. 18-8). Thus the muscle cell loses K^+ and gains Na^+ during intensive exercise. Contraction stimulates the Na^+-K^+-pump acutely, and training increases its activity. Still, at high intensity exercise the ionic leaks can exceed the capacity of the Na^+-K^+-pump for intracellular restoration.

During intensive exercise the Osmolarity of the contracting muscle cells increases together with the capillary hydrostatic pressure. As a consequence, the ECV and plasma volume can fall by 20% within a few min. The plasma [K^+] can rise to 8 mM due to efflux from the contracting muscle cells and from red blood cells into a reduced plasma volume. Training reduces exercise-induced hyperkalaemia.

Muscle fatigue following prolonged muscle activation increases proportional to the performance and to the loss of muscle glycogen. The insufficient and uncoordinated muscle contractions are due to the lack of glycogen and to failing neuromuscular transmission. *Exhaustion* of the stores of neurotransmitters in presynaptic terminals can occur within seconds to minutes of repetitive stimulation. Weight lifting, football dash and 100 m dash use up the phosphagen system within seconds.

Exhaustion often causes a serious drawback in the systematic practice of an athlete. The body stores are totally depleted, and deleterious consequences may occur.

Fig. 18-8: Skeletal muscle cell maintaining homeostasis by the activity of Na^+ -K^+ -pumps.

During exercise the striated muscle cells loose K^+ to the ECV and the blood. The Na^+ -K^+ -pump contains Na^+ -K^+ -ATPases, which are temporarily inefficient in maintaining homeostasis during exercise (Fig. 18-8). The rise in extracellular K^+ is probably related to muscular fatigue and dependent upon the maximal work capacity. Following exercise there is an extremely rapid homeostatic control in healthy well-trained persons. The activity of the Na^+ -K^+ -ATPases seems optimised in well-trained persons - not necessarily the concentration of Na^+ -K^+ -ATPases in skeletal muscle biopsies.

Even minor diseases, such as a common cold, may reduce cardiac output in an endurance athlete, thus causing muscle ischaemia during the usual practise and extreme muscle fatigue. Isolated muscular fatigue is thus due to depletion of ATP stores, whereby the actin-myosin filaments form a fixed binding and develop rigor or cramps. Neuromuscular fatigue is probably caused by progressive depletion of acetylcholine stores during prolonged, high frequency muscular activity.

Fatigue can never be fully explained by a simple rise in plasma-[K^+] only. Many other signals are integrated in the CNS before a person feels fatigued.

Endurance athletics in a hot and humid environment can increase the temperature of the body core to more than 41 ^{o}C. Such a level is dangerous to the brain and CNS symptoms and signs develop severe fatigue, headache, dizziness, nausea, confusion, staggering gait, unconsciousness, and profuse sweating. When the victim suddenly faints, this is termed heat stroke, which can be fatal.

2. Sport injuries

Five typical categories of sport injuries are considered here.

1. ***Runners*** are almost always damaged when working at a *too high velocity* or high velocity combined with turning or jumping. The force applied to the feet of a 75 kg person while walking is around (Gravity acceleration * body weight) = (9.807 m s^{-2} * 75 kg) = 750 kg m s^{-2} or 750 Newton. The force applied to the feet while running is 3-4 fold larger

Four typical injuries of runners are shown in Fig. 18-9.

Fig. 18-9:Two athletes showing four frequent leg and foot injuries attended by running.

The typical injuries are 1) muscle fibre lesions (myopathy with tender muscles), 2) tendosynovitis (shin splint) of the tibial posterior muscle, 3) tendinitis or rupture of the Achilles tendon, and 4) subluxation of the peroneus muscle tendon. 5) Dome fractures are osteochondral fractures from the talus with pain during running. This often occurs as a complication after a foot distortion, which does not heal. 6) Stress fractures are consequences of walking long distances but are also found after distance running and basketball.

These injuries occur during activities (athletes, ball players) with acceleration and deceleration by running or jumping in different directions. Quite often, the athlete is damaged following a break in the training. Even a few days of absence are enough. The athlete starts out too rapidly in order to compensate for the break in the training schedule.

2. ***Brain injuries*** *(boxing)* are known from serious accidents during many types of sport - in particular boxing. Even the elegant boxing legend, Muhammad Ali, was seriously injured during a long - although rather successful - carrier.

Acute brain damage or brain contusion includes deeper brain structures with neuronal damage, increased intracranial pressure and brain ischaemia (Fig. 18-10). Head injury during boxing can result in epidural haematoma (cranial fracture with rupture of the middle meningeal artery). The boxer hits the floor, is unconscious, wakes up and appears in good condition. Suddenly, he collapses again, and develops hemiplegia or die. The development of subdural haematoma is insidious venous bleeding sometimes with a latency of weeks between the head injury and the clinical phenomena (Chapter 7). CT scanning confirms the diagnosis. In chronic subdural haematoma there is a slow development of headache, drowsiness, confusion, sensory losses, hemiparesis, stupor and coma.

Fig. 18-10: Professional boxer with typical damages from the carrier.

Incomplete recovery from brain damage impairs higher cerebral function, with damages of locomotion (hemiplegia), and of psychological functions (Fig 18-10). The end result for the so-called punch-drunk boxer is chronic traumatic encephalopathy with dementia, post-traumatic epilepsy and other neurological disorders (Chapter 4).

*3.****Ball play damages***. Cruciate ligament lesions are common from ball play (ie, handball, football, baseball, basket and volleyball).

Basketball players often land on the toe tip from height and eventually develop exostoses. The exostosis hallucis is called basketball toe. The nail is tender and the exostosis has to be surgically removed.

The tibial anterior muscle originates on the tibia and passes to the navicular bone. Tendinitis in the tendon of this muscle leads to oedema, pain and crepitation.

Fig. 18-11: Soccer, baseball and basketball players are shown with typical injuries from the sport.

Baseball finger or mallet finger is an avulsion of an extensor tendon of the finger usually including a small flake of bone (Fig. 18-11).

Foot distortion (distorsio pedis) frequently includes rupture of the talofibular- calcanofibular- and bifurcate ligament or even fracture (Fig. 18-11).

Orthopaedic specialists must handle Malleole and other complicated fractures.

Turf toe is overextension of the basal joint of the large toe - frequently during ball play. In this case the large toe is protected with spica plast.

4. Skiing injuries range from trivial to fatal. The incidence of knee sprains is high, because improvements of binding design seem to be unsuccessful. The ski acting as a moment arm (Fig. 18-12) magnifies external rotation of the knee. Slalom skiing is the type of skiing with most fractures. The medial collateral ligament of the knee often ruptures.

Fig. 18-12: Typical skiing and tennis injuries are shown in a male and a female.

Another common ski injury is the *skiers thumb*. During a fall the ski pole and the wrist strap tend to concentrate forces to extend the thumb at the mid phalangeal joint until the ligaments burst.

5.Tennis injuries are haematoma subungualis (tennis toe) with bleeding under the nail of the big toe. This is a painful condition - not reserved for tennis players only. The haematoma pressure is relieved by puncture through the nail. The so-called tennis fracture is a fracture of the base of the 5.th metatarsal bone (Fig. 18-12)

The tarsal tunnel syndrome is also frequent in tennis players with pains along the medial side of the foot and toes. This involves the tibial posterior nerve in the channel behind the inner Malleole.

Tennis elbow is a painful disease of the aponeurotic fibres through which the common extensor origin is attached to the lateral humerus epicondyle. Tennis players from the strain use the name tennis elbow (Fig. 18-12); only few of the sufferers actually play tennis.

Conclusion:

*The demand of **fast progress** is linked to competitive sports. A better strategy is to practice at a **relaxed level**, until stamina is developed and hard training is tolerated. Relaxed training is often so comfortable that it becomes a lifestyle. Tender muscles are avoided by prewarming, and a careful **muscle stretch program** following exercise.*

3. Doping

Doping derives from the word dope, which means a stimulating drug. Athletes, who use drugs or other means with the intention to improve performance artificially, are doped by definition

The list of forbidden drugs counts more than 3500, and it is still growing.

Pseudo-doping

Many drugs reputedly increase athletic performance, but the fact remains that such effects rarely show up in double-blind controlled trials. On the contrary, serious side effects occur with a biologically high and statistically significant frequency.

Pseudo-doping with Ginseng and a multitude of other extracts and substances is often quite harmless, and - just as many potent drugs - without proven beneficial effect on athletic performance.

Anabolic steroids

Anabolic steroids are used to increase muscle strength in females and in male athletes with a poor natural testosterone production - possibly a pure placebo effect. Compared to placebo in double-blind studies there is no detectable steroid-effect on the maximal oxygen uptake, size of the muscles or erythropoiesis. However, both steroids and placebo improves the mood and motivation, so both groups trained more and were eating more than before.

As an example, the muscles of body builders are extremely large, but not necessarily equally strong (Fig. 18-13). Some side-effects of dope are lesions of muscle fibres, hypogonadism, liver disorders, and psychosocial deroute (see illustration for further information).

The reversible side effects and irreversible sequel are indisputable. Doping addicts have a high risk of cardiovascular diseases (arterial hypertension, atherosclerosis, heart attacks and strokes), muscular disorders, liver disease, and - in males - testicular failure. Both the sperm formation and the testosterone production are suffering, often irreversibly.

Body building is considered to be the most doping related discipline - in particular by the use of anabolic steroids - and the results are often monstrous (Fig. 18-13).

Fig. 18-13: A body builder, a Sumo wrestler and an obese super-heavy weight champion with a world record (235 kg). All have serious health problems.

In wrestling, discos and super-heavy weight lifting the use of anabolic steroids is frequently disclosed.

A previous world record holder in super-heavy weight lifting developed extreme adiposity when increasing his natural body weight from 80 to 183 kg. The use of steroids resulted in muscular lesions and severe psycho-social crises. The adiposity developed into restrictive lung disease and arthrosis in the knees and other articulations. The athlete was actually a patient with a normal thoracic skeleton, but the lungs were compressed by fat accumulation. During his career he developed the Pickwick syndrome (ie, a fat patient with reduced ventilation, somnolence, sleep apnoea, secondary polycythaemia and cyanosis).

The Japanese Sumo wrestlers have the same problems created by the required extreme adiposity, and many excellent wrestlers have obvious difficulties in walking.

Blood doping

Blood boosting is an artificial improvement of performance through an increase in the haemoglobin binding capacity. Blood doping (one litre of the athletes own blood) definitely improves the oxygen transport with the blood and also the maximal oxygen uptake, which is beneficial to distance diciplines.

Approximately 6 weeks before the competition (Olympic Games or World Championship) the athlete deposits 1000 ml of his own blood as separated red blood cells. The haemoglobin

binding capacity is regained by maintained training, and a few hours before the competition, he receives a blood transfusion with his own erythrocytes. Of course, a sudden improvement of the maximum oxygen capacity of more than 10 %, is unfair in endurance disciplines (long distance running, cycling, skiing etc), but it may cause viscosity problems and thrombus formation (see below).

High altitude training and erythropoietin

High altitude training is a physiological method to obtain the same increase in haemoglobin as in blood doping. The idea is to obtain an advantage not present for most of the other competitors, and thus it is unethical, but impossible to disclose. Training at high altitude implies a larger degree of hypoxia than the same sea level training, so two hypoxic metabolites are produced: *Erythropoietin* and *2,3-DPG*. Erythropoietin increases erythropoiesis and thus the haemoglobin concentration, whereas 2,3-DPG form haemoglobin in the deoxy-conformation and increases the P_{O2} gradient when delivering oxygen to the muscle cells.

A serious development occurred following the introduction of industrially produced human erythropoietin (EPO). Natural production is increased, if there is hypoxia in the kidneys.

Erythropoietin is clearly beneficial to endurance athletes, but most types of doping have deleterious effects. The synthesised erythropoietin, when administered to athletes, definitely stimulates the red bone marrow to increase the production of erythrocytes. The effect on the maximum oxygen capacity is indisputable, but the price is often death, because of fluid loss, Haemo-concentration, drastically increased blood viscosity and thrombus formation all over the circulatory system. The death of a whole group of young racing bicyclists, within a short period of time, was probably caused by erythropoietin.

Stimulants

Ephedrine, amphetamine and other psychomotor CNS stimulants are still used by athletes in the hope of increased velocity (so-called speed). Amphetamine or speed pills have improved results in running, bicycling, swimming, weight-throwing and other disciplines compared to placebo. The same stimulants have increased blood pressure and heart rate in athletes exercising heavily in hot climates, until they died from cerebral bleeding or ventricular fibrillation. This has taken place several times in the history of Tour de France.

Cocaine and coffeine seem to suppress natural fatigue, and is also on the doping list. Suppression of natural fatigue leads to exhaustion and circulatory collapse sometimes with cerebral bleeding and ventricular fibrillation.

β- **Adrenergic blocker>**

b-Adrenergic blockers are drugs that reduce heart rate (negative chronotropic effect) and the force of contraction (negative inotropic effect). Both mechanisms reduce the myocardial oxygen demand. In precision sports, where relaxation without tremor is essential, these drugs have a proven beneficial effect in double-blind controlled clinical trials, and they are therefore on the doping list. Precision sports include archery, standard pistol, skeet shooting, rifle shooting, ski jumping, billiards, etc.

Participation is ski-shooting competition is hardly advantageous on b-blockers, because the abuser gets too tired to accomplish endurance performance.

Diuretics

The athletes in disciplines with specific weight classes - such as boxing, wrestling, weight

lifting etc - reputedly use diuretics in order to cause a rapid weight loss, with the advantage of competing against smaller persons. Uncontrolled use disturbs the normal distribution of ions in the cells and body fluids, and reduces the blood volume and increases viscosity. In extreme cases there is circulatory collapse and death.

Peptidergic hormones

Gonadotropins - in particular the luteotropic hormone (LH) - stimulate release of testosterone from the Leydig interstitial cells of the testes. Human chorion Gonadotropin (hCG) also binds to the Leydig cells and releases testosterone in males.

Corticotropin (ACTH) from corticotropic cells of the adenohypophysis stimulates production and secretion of adrenal cortical hormones (mainly glucocorticoids).

Somatotropin (human growth hormone, HGH) from somatotropic cells of the adenohypophysis increases and regulates growth, partly directly and partly through evoking the release of somatomedins from the liver. HGH increases protein synthesis, lipolysis and blood glucose. HGH induces gigantism in growing individuals and acromegaly in adults. Uncontrolled use may lead to cardiomyopathy, diabetes, adiposity, articular pain, hypertension and early death.

Monstrous growth of the shoulders and bodies of female swimmers is disclosed by vision alone, and the sight is clearly different from a naturally top- trained female.

Some of the female track runners have written history by winning WM and the Olympics for females year after year, although they looked like a male.

Pregnancy/abortion as doping

Pregnancy seems to increase muscle strength in female athletes. Female top athletes - just following the period, where they gave birth to their first child - have set several world records. Of course, this is acceptable as a natural and unintended event.

However, in some countries female athletes have become pregnant for 2-3 months, in order to improve their performance just following an abortion.

Genetic doping In countries where the political will, is not balanced by ethics, recombinant DNA technique may be used in the future to clone groups of individuals with remarkable talents for special athletic performances.

Smoking has acute and deleterious effects on both the cardiovascular and the respiratory system, but athletes have used it. The substances involved are not at the doping list. The CO blocks off part of the haemoglobin, and limits the transport capacity for oxygen to all mitochondria, which is especially inhibitory to the heart and the skeletal muscles. Nicotine constricts terminal bronchioli and arterioles in many vascular beds. Nicotine also paralyses the cilia of the epithelial cells of the respiratory tract. Chronic smoking leads to life-long chronic bronchitis and emphysema or to lung cancer.

4. Fit for life

A high endurance capacity or fitness is healthy. The mortality increases with low endurance capacity in males (Fig. 18-14). A similar pattern is recorded for females. An endurance capacity (fitness number, $V°_{O2}$max) of 34 ml O_2 min^{-1} kg^{-1} or more seem compatible with a reasonable health status and mortality risk.

Fig. 18-14: The endurance capacity ($V°_{O2}$max) in relation to mortality. The total mortality is given as Number of deaths per year per 10 000 males.

Physical inactivity with an endurance capacity (fitness number) below 34 ml min^{-1} kg^{-1} is a risk factor for the development of atherosclerosis, other risk factors and sudden death in males - and probably also in postmenopausal females.

Equations

- The principle of mass balance states that cardiac output is equal to the oxygen uptake ($V°_{O2}$) divided by the arteriovenous oxygen content difference:

  Eq. 18-1: $Q° = V°_{O2} / (C_{aO2} - C_{v\text{-}O2})$ - or $V°_{O2}\,max = Q°°\,max * (C_{aO2} - C_{v\text{-}O2})$.

- The following is valid for exercising healthy males (up to 70% of their maximal oxygen capacity, $V°_{O2}max$):

  Eq. 18-2: $Q°$ litre $min^{-1} = 3.07 + 6.01 \times V°_{O2}$.

- This calculation of $Q°$ allows for estimation of the rarely available mixed venous carbon dioxide concentration ($C_{v\text{-}CO2}$) from Fick's principle:

  Eq. 18-3: $C\bar{v}CO2 - C_{aCO2} = V°_{CO2} / Q°$; or $C_{v\text{-}CO2} = C_{aCO2} + V°_{CO2} / Q°$.

- The diffusion-limited oxygen uptake ($V°_{O2}$) equals the product of lung diffusion capacity (D_{L02}) and the mean alveolar oxygen tension gradient (DP_{O2}):

  Eq. 18-4: $V°_{O2} = (D_{L02} \times DP_{O2})$.

- Poiseuilles law relates bloodflow ($Q°$) to total peripheral vascular resistance (*TPVR*) and the mean arterial driving pressure (MAP):

  Eq. 18-5: $Q° = MAP/TPVR$.

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A. Fatigue can never be fully explained by a simple rise in plasma-[K^+] only.

B. Fitness tests rest on the assumption, that there is an exponential increase in HR with increasing oxygen uptake or work rate.

C. Erythropoietin clearly improves the endurance capacity of athletes.

D. Erythropoietin is produced and secreted from the kidneys.

E. At maximal works the lung diffusion capacity for oxygen rises to 9 ml STPD s^{-1} kPa^{-1} .

Case History A

A 32-year old marathon runner (body weight 60 kg) is examined on a treadmill. He is running at a velocity of 16 km per hour. The following variables are measured:

$Q\circ$= 25 l per min, C_{aO2} = 200 and $C_{v\text{-}O2}$ = 40 ml (STPD) per l of blood. The concentration of lactic acid is measured in the blood every second min for 18 minutes. The level increases from 1.1 to a steady state value of 15 mM.

The total work rate of the heart (pressure-volume work and kinetic work) is 16 Watts or 16 J per second, and the mechanical efficiency of the heart work is assumed to be 20%.

The energy equivalent for oxygen on a mixed diet is 20 kJ per l (STPD) of oxygen.

1. *Calculate the oxygen uptake of the heart during this work.*
2. *Assume that the arterio-venous oxygen content difference for the heart is equal to that of the whole body. Calculate the coronary bloodflow during running.*
3. *Has the athlete accumulated an important oxygen debt during the 18 minutes of running?*

Case History B

A male world record holder in 100-m dash is examined on a treadmill with a velocity capacity up to 35 km per hour. His body weight is 78 kg. While standing relaxed on the treadmill before exercise, his oxygen uptake is measured to 300 ml STPD min^{-1}. The ventilatory exchange quotient (R) in respiratory steady state, is measured to 1.0. At a given signal the athlete jumps on the running treadmill and performs a 20-second dash similar to a 200 m dash on the track. The inspired, atmospheric air has an oxygen fraction (F_{IO2}) of 0.2093 and a carbon dioxide fraction (F_{ICO2}) of 0.0003. As soon as he jumps off the treadmill, he is connected to a system of rubber bags, where his expired air is collected over the next hour until his resting oxygen uptake is re-established. All the expired air in the rubber bags is mixed and analysed. The mixed expired air fractions are F_{EO2} = 0.1746 and F_{ICO2} = 0.035. The volume of mixed expired air is measured at ATPS, and by calculation corrected to a STPD volume of 1090 litres.

1. *Calculate the oxygen debt repaid over the 1-hour post-exercise period.*
2. *What assumptions is maid in order to perform this calculation?*
3. *Calculate the ventilatory exchange quotient in the post-exercise period and compare the result to the pre-exercise R-value.*
4. *What is the basis of oxygen debt?*

Case History C

A well-trained male long distance skier, weight 74 kg, has a $V\circ_{O2}max$ of 6 l STPD min^{-1} .The maximal arterio-venous oxygen content difference is 150 ml STPD l^{-1} of blood.

During maximal work his lung diffusion capacity for oxygen (D_{LO2}) rises to 9 ml STPD s^{-1}

kPa^{-1}.

1. *Calculate maximum cardiac output and describe the consequences for the lung perfusion.*

2. *Calculate the fitness number and explain what it means.*

3. *Define the mean oxygen tension gradient for lung diffusion and calculate its size. Explain how the gradient can increase to this extent.*

Case History D

A female 20 years of age, with a body weight of 62 kg, is exercising on a bicycle ergometer during steady state. Her cardiac output ($\dot{Q}$) is *measured to 25 l* min^{-1} *by the mass balance principle with carbon dioxide as indicator, and her arteriovenous* O_2 *content difference is measured to 170 ml STPD* l^{-1}.

1. *Calculate her oxygen-uptake per min (*$\dot{V}_{O2}$*).*

2. *What assumption must be made in order to calculate her fitness?*

Case History E

An adult male has a lung diffusion capacity for oxygen of 22 ml STPD per min and per mmHg, and a mean alveolar O_2 *tension gradient of 12 mmHg. His oxygen concentration in the arterial blood (*C_{aO2}*) is 200 ml per l and the renal bloodflow (RBF) is 1200 ml per min. The renal* O_2 *consumption is 15 ml per min.*

1. *Calculate his oxygen uptake in ml STPD per min.*

2. *How is it possible for this person to increase the oxygen uptake to 4900 ml STPD per min?*

3. *Calculate the arteriovenous oxygen content difference in the kidneys.*

4. *Is the oxygen delivery to the kidneys redundant?*

Try to solve the problems before looking up the answers.

Highlights

- *The total muscular oxygen uptake can rise by a factor of 80 from rest to maximal exercise.*
- *At the start of exercise, signals from the brain and from the working muscles bombard the cardiopulmonary control centres in the brainstem. Both cardiac output and ventilation increase, the α-adrenergic tone of the muscular arterioles falls abruptly, whereas the vascular resistance increases in inactive tissues.*
- *The circumventricular organs of the brain contain many fenestrations, and they are located close to the control centres of the hypothalamus and the brainstem.*
- *Training improves the capacity for oxygen transport to the muscular mitochondria, and improves their ability to use oxygen. After long-term endurance training the athlete typically has a lower resting heart rate, a greater stroke volume, and a lower peripheral resistance than before.*

- *The $V°_{O2}$max progressively increases by endurance training, and also the extraction of oxygen from the blood increases.*
- *At maximal work the lung diffusion capacity for oxygen (D_{LO2}) rises to 9 ml STPD s^{-1} kPa^{-1} (from 3.6 at rest). The lung diffusion capacity for oxygen increases by endurance training.*
- *According to the redundancy hypothesis, the rise in cardiac output and in ventilation during exercise is caused by an integration of neural and humoral factors.*
- *The demand of fast progress is linked to competitive sports. A better strategy is to practice at a relaxed level, until stamina is developed and hard training is tolerated.*
- *Relaxed training is often so comfortable that it becomes a lifestyle. Tender muscles are avoided by prewarming, and a careful muscle stretch program following exercise.*
- *Amphetamine or speed pills and other CNS stimulants have improved results in running, bicycling, swimming, weight-throwing and other disciplines compared to placebo. The same stimulants have increased blood pressure and heart rate in athletes exercising heavily in hot climates, until they died from cerebral bleeding or ventricular fibrillation.*
- *Pregnancy/Abortion as doping. Pregnancy seems to increase muscle strength in female athletes. Female top athletes have set world records, just following the period, where they gave birth to their first child. - In some countries female athletes have become pregnant for 2-3 months, in order to improve their performance just following an abortion.*
- *Blood doping definitely improves the oxygen transport with the blood and also the maximal oxygen uptake.*
- *Doping with erythropoietin stimulates the red bone marrow to increase the production of erythrocytes. The beneficial effect on the maximum oxygen uptake is indisputable, but the prize has been death because of thrombus formation.*
- *Doping addicts have a high risk of cardiovascular diseases (arterial hypertension, atherosclerosis, heart attacks and strokes), muscular disorders, liver disease, and - in males - testicular insufficiency. Both the sperm formation and the testosterone production are suffering, often irreversibly.*
- *Lack of fitness is a risk factor for the development of atherosclerosis and for sudden death.*

Further Reading

Medicine and Science in Sports and Exercise. Monthly journal published by the Am. College of Sports Medicine. Williams & Wilkins Co, 428 East Preston Street, Baltimore MD 21202-3993, USA.

Apps DK, Cohen BB and CM Steel. Biochemistry. Bailliere Tindall, London, 1994.

Wood, S C, and R C Boach: Sports and Exercise Medicine. Marcel Dekker Inc, N.Y. 1994.

Katzung BG. Basic & Clinical Pharmacology. Appleton & Lange, Stanford, Connecticut, 1998.

Return to chapter 18

Return to content

Chapter 19

Flying, Space And Diving

Study Objectives

- To *define* the law of ideal gasses, Boyle-Mariotte´s law and Newton's law of universal gravitation.
- To *describe* essential topics related to flying, space life and diving, hyperbaric oxygenation and recompression therapy, inert gas narcosis and oxygen toxicity.
- To *calculate* one variable from relevant variables given.
- To *explain* decompression sickness, hypobarotrauma (squeeze), and alveolar rupture in the hypo- and hyperbaric environment.
- To *use* the above concepts in problem solving and in case histories.

Principles

- *Newton's law of universal gravitation: Every particle in the universe attracts every other particle with a force directly proportional to the product of the two masses and inversely proportional to the square of the distance between them.*
- *Haldane's law: Any saturation dive to less than 10 m does not cause decompression sickness and allows ascent without decompression stops.*
- *Flight decompression. Fast ascent during flight rarely causes decompression problems at altitudes below 10 km (30 000 feet).*

Definitions

- **Absolute atmospheric pressure** is the pressure measured in *atmospheres absolute* (ie, one *atm.abs.* at the surface of earth, and 2-3-4 atm abs at 10-20-30 m of seawater depth).
- **Decompression** is the gradual fall in ambient pressure during ascent both in air and in water.
- **Decompression sickness** refers to the formation of growing bubbles during decompression causing tissue ischaemia and necrosis.
- **Gravity units (G)** refers to the gravity of the earth (G = 9.807 or approximately 10 m s^{-2}).
- **Hyperbarotrauma** or *air embolism* is a major problem of submarine escape. As the submariner escapes and ascends from the vehicle, the gas volume of the lungs expands and ruptures pulmonary vessels. Air enters the tissues, the thoracic cavity and the circulatory system. This causes air embolism and death.

- **Inert gas narcosis** describes an intoxication syndrome with euphoria or anxiety, gross errors in performance and eventually unconsciousness. Nitrogen starts to be narcotic at depths of 40 m or more.

- **Pressure** is measured as force per area unit. The pressure of a 10 m high sea water column resting on 1 m^2 is *1 atmosphere* equal to 101.3 kPa or 760 mmHg.

- **Recompression therapy** of decompression sickness takes place in a pressure chamber, where sufficient pressure can be established to eliminate the bubbles causing the disease.

- **Squeeze** or *hypobarotrauma* is caused by a negative pressure difference across the wall of a non-collapsible air space (middle ear, sinuses, compressed lungs, hollow teeth etc.).

Essentials

High-altitude flight, space flight, and diving encounters disorders due to change of total pressure (cf. barotrauma, decompression sickness), forces of acceleration and weightlessness, fall in oxygen pressure (oxygen deficiency), fall in ambient temperature, and increased radiation.

This paragraph deals with *1. Commercial aircraft flying, 2. Satellite flying, and 3. Diving.*

1. Commercial aircraft flying

Commercial aeroplanes fly at an altitude of approximately 33 000 ft or 10 300 m, where the ambient pressure is 26 kPa or 195 mm Hg and the oxygen tension in the outside cold air is around 40 mmHg - less than on Mt. Everest. The aeroplanes have cabins pressurised to 80 kPa (corresponding to an altitude of 2000 m or 6000 ft) or sometimes only to an altitude of 2750 m. These levels cause *hypoxic dyspnoea* only in cardiorespiratory patients. The main problem for the average traveller is *squeeze* (hypobarotrauma) of sinuses and teeth during landing.

In case of *depressurisation* of an aircraft, P_{AO2} is rapidly reduced, and oxygen delivery must be sustained with an oxygen mask. Hypoxia is the acute danger; decompression sickness (see below) is avoided by fast descent to lower altitudes (higher pressures).

The rhythmicity of human organ functions is coupled to the daily periodicity of the earth's rotation. The unknown endogenous periodicities, correspond approximately (Latin: Circa) to the duration of a day (Latin: Dies), so they are called *circadian rhythms.* Many circadian rhythms are synchronised to the 24-hour cycle (the biological clock) by *external signals* such as light and darkness or social habits.

The most important diurnal rhythm is the waking-sleeping cycle, with a fall in body temperature, respiratory rate and heart rate at the onset of sleep. Flight across time zones causes jet lag, which is a discrepancy between the circadian rhythms and the external signals. An eastward flight causes greater jet lag than westward flights. The adjustment of circadian rhythms takes one day for every two hours of time shift when flying east.

Parachuting*:* The accelerator force of gravity has the same size at a given latitude. Consider a body falling freely in an environment without air molecules along a straight line perpendicular to the surface of the earth. In the absence of air resistance, all bodies fall towards the earth with the same acceleration regardless of the size and composition of the body. The magnitude of this accelleration, which is represented by the gravity unit (G), is 9.8 or approximately 10 m s^{-2}. If a parachuter was only exposed to ***gravity*** he would fall towards the earth with the acceleration 10 m s^{-2}. Following the first 10 s, the velocity is approximately (G* t) = (10 m s^{-2}

*10 s) or *100 m per s.*

In real life there is no such thing as a free fall from an aeroplane, because the air resistance soon reduces the velocity and outbalance the acceleration. Here, the parachuter reaches the *terminal velocity* determined by the relation between G and air resistance. As the parachute opens, the parachuter experiences an *opening shock load* of 500-600 kg. The area of the standard parachute reduces the terminal velocity to 1/9. The force of impact at landing is $(1/9)^2$ = 1/81 of the landing force without parachute. This is equal to a jump from a height of 2-3 m.

2. Satellite flying

Space flight requires pressure suits or cabins ensuring adequate oxygenation and pressurisation to protect against hypoxia and decompression sickness. More than 20 km out in the stratosphere (Fig. 16-9), the barometric pressure is less than the water vapour tension in alveolar air (47 mmHg or 6.3 kPa), and the oxygen tension is rapidly approaching zero. Without pressurised cabins the blood of the astronaut would boil. In the stratosphere, instantaneous decompression by pressure equalisation with the environment causes a monstrous form of decompression sickness with air emboli all over the body. The body volume of humans increases by a factor of at least three by air expansion in tissues and blood.

Recycling techniques have been developed for the reuse of O_2. The spacecraft - just as submarines - must carry along enough carbon dioxide absorbent to prevent death from *carbon dioxide poisoning*.

Lethal radiation dose is rapidly reached when flying around the earth in the *Van Allen radiation belts*, consisting of *high-energy* level protons and electrons. Therefore, commercial flying take places essentially below the inner belt (500 km). Start and landing of long distance space flights take place as close to the magnetic poles - with minimum radiation energy exposure - as possible.

Forces of acceleration (G) cause pooling of blood in the lower extremities, a critical drop in arterial blood pressure with *orthostatic hypotension* or *collapse* due to reduced venous return, but also cause nausea, vomiting, and spatial disorientation.

The start acceleration of a spacecraft is approximately linear, and the astronaut is exposed to a tremendous acceleration - often close to 8 times the gravity of the earth (8 G) at the first stage of a 3-stage blast-off. The space shuttle implies a 3 G start and a 1.2 G landing.

The body of the astronaut is located transverse to the axis of acceleration in order to prevent pooling of blood in the legs. The astronaut often carries an *anti-G suit* with the same purpose.

Gravity acts on both the astronaut and the spacecraft in space, so astronauts are floating inside the satellite. *Weightlessness*, caused by the absence of gravitational forces, produces diverse reactions. The space pilot experience *space sickness* (nausea and vomiting over the first days), falling total blood volume, muscular atrophy, Ca^{2+} loss from the bones, and obstipation, brought about by the lack of stimuli. After living in outer space for weeks, *adaptive difficulties* develop upon return to the life on earth. These difficulties are low working capacity and a tendency to faint while standing. The bone loss continues long after the return to earth, because stimulation of bone formation requires physical activity in a gravity field.

3. Diving

This paragraph deals with *3a. Senses in water*, *3b. Inactive gasses*, *and 3c. Gas stores of man.*

3a. Senses in water

The light intensity decreases rapidly with depth, and 100 m below sea level, it is permanently dark.

The ratio of the speed of light in free space to its speed in a medium is called the *index of refraction* of the medium. The power of a lens is proportional to the change in index of refraction and inversely proportional to the radius of its curvature.

Fig. 19-1: Light refraction in the reduced eye in air and in water. The effect of goggles is remarkable and shown below.

The light refraction is shown in a schematised eye, where the air-cornea interface contributes 43 Diopters and the lens 15 to the focusing of parallel light rays upon the retina (Fig. 19-1). Normally, it is the difference in index of refraction between the air (1.00) and the cornea (1.38), which accounts for the major refraction at the air - cornea interface (Fig. 19-1). When water (1.33) is substituted for air, this difference becomes so small that parallel light rays are now focused behind the retina. Only maximal accommodation of the lens can refract parallel light rays but not enough to focus them on the retina. Under water the unprotected eye becomes farsighted (hypermetropic) and the image is distorted.

Goggles restore the air-cornea interface. In spite of the high index of refraction of glass (1.5), the parallel light rays are not refracted, because the radius of a glass plate is infinite. Light rays not falling perpendicular to the glass are refracted, which explains why objects look larger under water.

Sound propagates more rapidly in soft tissue and water than in air (1540 m s^{-1} as opposed to 340 m s^{-1} in air). In air the detectable phase- or time lag between the two ears helps us to determine the direction of the sound source. Therefore *underwater sound sources* appear much nearer than they actually are. Because of the shortened delay between the two ears, localisation of sound sources becomes extremely difficult.

The voice of a diver speaking in a helium-oxygen atmosphere resembles the voice of Disney's Donald Duck, because the high tones dominate in this light medium.

3b. Inactive gasses

Inactive gasses are eliminated by diffusion to the peripheral blood and by convection with the blood to the lungs. During diving the ambient pressure will rise by *one atmosphere* for every *10 m of seawater* depth you pass.

Fig. 19-2: Desaturation curves showing elimination of nitrogen from tissues with different half-life. – The depth and bottom time for stage decompression is shown to the right.

Inactive gasses (ie, N_2, H_2, He, Ne, Ar etc) contribute with aa alveolar partial pressure when present in the breathing medium, and the blood of the lung capillaries is in immediate equilibrium with the partial pressure in the alveolar air. Atmospheric air is the commonly used breathing gas for diving. The *solubility coefficient* for N_2 in blood is low (ie, 0.012 - see Box 13-1), but five times larger in fat tissue, which has a modest perfusion rate. A large cardiac output is necessary to transport substantial amounts of N_2 to the tissues. During saturation dives it takes a long tome before the nitrogen tension of fat tissue reaches equilibrium with P_{aN2}.

The human body is a complex of tissue types, each with an individual, exponential uptake rate or half-life for saturation (Fig. 19-2). Often a model with *half-life tissues* of 5-, 10-, 20-, 40-,

80-, 120-, 240-, 480- minutes (of half-life) is used.

No decompression diving limits (cf. decompression tables) are limits of depth and time spend at the depth. A work period of one hour at 60 m of depth requires a decompression period of three hours on the steps scheduled in the decompression tables during ascent.

Any saturation dive to less than 10 m allows ascent without decompression stops (Haldane's law). Fast ascent during flight rarely causes problems at altitudes below 10 km (30 000 feet). At altitudes above 20 km (where the ambient pressure is less than 6.3 kPa or 47 mmHg) the blood of a space traveller with normal body temperature will boil, if the pressure within his suit is lost. This is because the partial pressure of saturated water vapour at body temperature is 6.3 kPa or 47 mmHg.

3c. Gas stores

The gas stores of the human body are easy to calculate from standard values, and the calculations are shown in Fig. 19-3.

The resting person has a lung volume of 2700 ml STPD (standard Expiratory Reserve Volume added to Residual Volume), and the blood volume consists of 2 l of arterial and 3 l of venous blood. The extracellular fluid volume (ECV) and intracellular fluid volume is 13 and 26 l, respectively. The oxygen content of the arterial and venous blood is 200 and 150 ml STPD per l, respectively. The bicarbonate concentration extra- and intra-cellularly is 24 and 10 mM, respectively. The carbon dioxide content of the arterial and venous blood is 500 and 550 ml STPD, respectively.

Oxygen stores

The oxygen stores of the lungs are 18 mmol, and of the blood 38 mmol (Fig. 19-3).

We also have small O_2 stores (7 mmol) in myoglobin and in the tissues, a total of (18+38+7) = 63 mmol.

Hypoxic brain damage can occur after 5 min of apnoea and after 5 s of cardiac arrest (black out and grey out).

Fig. 19-3: Comparison of the small oxygen stores to the large exchangeable carbon dioxide stores of man.

Carbon dioxide stores

The CO_2 stores of the lungs are 8 mmol, and those of the blood 118 mmol. The ECV of an adult person contains 312 mmol bicarbonate. Intracellularly there is 260 mmol mobile bicarbonate. These exchangeable CO_2 stores comprises (8 +118+312+260) = 698 or 700 mmol in total (Fig. 19-3). The size is subject to changes by alterations of ventilation or acid-base-status. Besides, there are large amounts of CO_2 fixed in bones.

Nitrogen stores

The human *tissue stores of nitrogen* (N_2) are calculated in Fig. 19-4. A normal weight person contains 15 kg of fat and 42 litres of water (total water). Nitrogen is five times as soluble in lipid tissue (solubility coefficient 0.065, see Chapter 13, Box 13-1) as in water (0.013). The blood of the pulmonary capillaries is saturated with the nitrogen at the pressure prevailing in the alveolar air.

Fig. 19-4: Partial pressures and bubble formation. The nitrogen content of a diver at sea

level before diving (left) is compared to that depth being saturated with N_2 (middle), and to the content following sudden ascent (right).

At sea level before diving, the diver contains about 400 ml of N_2 in the total water phase and 725 ml in the fat tissue (Fig. 19-4, left). The partial pressure of nitrogen is 566 mmHg in all tissues.

When the diver reaches the bottom, there is a rise in ambient pressure to 3040 mmHg, and the new N_2-gradient (2504 - 566 mmHg) will increase the uptake of N_2 by the pulmonary blood. It takes hours before all fat-containing body tissues are saturated with dissolved N_2 at the new pressure, because of the low bloodflow and poor capillary network of fat tissue. The total water phase is in equilibrium within an hour. The amount of nitrogen is now 1800 ml in the total water phase and about 3200 ml in the fat tissue (Fig. 19-4, middle). At depth the compression of the external pressure keeps all gasses dissolved in the body.

When the diver suddenly surfaces (sudden decompression to sea level), the total pressure is suddenly only 1 atm.abs. or 760 mmHg, whereas the pressure inside the body tissues approaches that of the alveolar air, when he left the bottom (3040 out of which 2504 mmHg was nitrogen). Obviously, most of the total pressure in the tissues is due to dissolved nitrogen, and without counter-pressure the N_2 -molecules are released from the dissolved state in the form of bubbles. In the case shown in Fig. 19-4, the diver is supposed to eliminate a large surplus of N_2 (5000 - 1133 ml) during the rapid ascent to the surface. This is not possible for the circulation to cope with and bubbles are formed in blood and other tissues (see decompression sickness below).

Pathophysiology

This paragraph deals with *1.Drowning*, *2. Skin diving and disorders*, *3. Disorders from hard hat and SCUBA diving, and 4. Hyperbaric oxygenation therapy*.

1. Drowning.

In more than 90% of drowning and *near drowning* the lungs are *flooded with water.*

Fresh water is hypotonic and rapidly absorbed, diluting plasma to become hypotonic and causing bursting of the red cells (*haemolysis*). The victim often dies within 3 min.

Seawater is hypertonic and draws fluid from the blood plasma into the lung alveoli and interstitial fluid, so that plasma volume decreases. This causes *haemoconcentration* and shock. The victim often dies within 6 min.

The *rational therapy*, if early enough, is immediate resuscitation and treatment of the respiratory and circulatory disorders including haemolysis or haemoconcentration.

2. Skin diving and disorders

The simplest type of diving is *breath-hold diving* or *naked skin diving*. This is performed without equipment or with only a snorkel and a mask. This is the most popular form of diving performed as underwater sports, fishing, photography, archaeology etc, and involving diving tribes as the Amas of Japan and Korea, and the Polynesian pearl divers.

Breathing through a snorkel has limitations. The maximum distance between the water surface and the middle of the immersed thorax, is 0.35 m (Fig. 19-5).

Fig. 19-5: Swimming and skin diving with snorkel just below the surface.

This is because the water pressure creates a *pressure gradient* in the pulmonary vascular system, which is a *low-pressure system.* The intra-alveolar pressure is equal to the atmospheric pressure, so the thoracic vessels are distended with blood, and the high extra-thoracic pressure restrains the respiration. Thus, the length of the snorkel cannot be extended beyond that range. The snorkel is sometimes provided with a ball valve that closes automatically at diving.

Snorkel breathing implies a dead space problem. This is because the sum of the snorkel- and person- dead space (here 400 ml) is comparable in size to the normal tidal volume of 500 ml. Thus it is necessary to increase the tidal volume.

Breath-hold diving is dangerous for several reasons: *2a. Hypoxia*, *2b. Cardiac arrhythmias*, *2c. Air embolism*, *2d. Squeeze*.

2a. Hypoxia

Prior hyperventilation with only three deep inspirations increases P_{AO2} to 17.3 kPa (130 mmHg), whereas P_{ACO2} is only 2.6 kPa (20 mmHg). A well-trained swimmer can use up the oxygen stores during forceful underwater swimming, before sufficient carbon dioxide is produced to awake the desire for breathing. Hypoxia in itself is insufficient to trigger inspiration. Cerebral P_{O2} can fall below 4 kPa or 30 mmHg (zone of unconsciousness or grey out) without a sufficient P_{aCO2} is build up to force the swimmer to the surface. Drowning is likely, if this occurs underwater. This sequence of events has sometimes been called *shallow water blackout.*

Underwater swimming *record attempts* should be discouraged, as they are dangerous.

Unconsciousness from hypoxia and subsequent drowning has occurred among expert pool swimmers trying to set distance records in underwater swimming (Fig. 19-6).

Fig. 19-6: Record attempts in apnoea underwater swimming is life threatening (+). Point I is the normal inspiratory point.

Hyperventilation prior to the dive can cause *acute respiratory alkalosis* with dizziness and convulsions even before the dive.

The initial hyperventilation to the point A_1 (hyperventilation) eliminates the CO_2 drive to respire, so while exercising underwater they slide into hypoxic unconsciousness (point A_2 in Fig. 19-6) and death. The data for the alveolar points are obtained from a medical student, who tried to keep his breath for as long as possible (191 s) lying supine in air - and fainted. He was immediately rescued by neck extension, whereby he expired and started spontaneous breathing again - making artificial ventilation unnecessary.

Deep diving with long bottom time results in *anoxia during ascent.* The P_{AO2} falls so drastically during ascent that oxygen transport is *reversed* in the alveoli.

2b. Cardiac arrhythmias

In cold water especially, a *cutaneo-visceral reflex* from cold receptors of the skin seem able to elicit *malignant* cardiac arrhythmia's or a slow heart rate termed *diving bradycardia* or both. The latter can be beneficial to the breath-hold diver by reduction of the oxygen demand of the myocardium, but the slow heart rate may develop into a ***vasovagal syncope***, which is lethal under water. Heart rates around 30 beats per min have been observed in small children, who have been victims of cold water near drowning for periods up to 40 min. Such heart rates are

similar to those found during hypothermic surgery. These arrhythmias and bradycardias are triggered from cold receptors in the skin and from general hypothermia.

The *human diving response* is a natural vagal reduction of heart rate occurring during breath holding on land as well as under water (*diving bradycardia*). Individuals with a high vagal excitability can experience sino-atrial blockage, which may develop into a severe Adam Stokes syndrome. This is fatal under water.

2c. Submarine escape training and emergency escape

Escape from a submerged submarine is accomplished from more than 100 m depth without the use of equipment. The submariner ascends through the water, while exhaling continuously in a controlled manner. He is trained to follow the ascent rate of his own exhaled air bubbles. The art of the procedure is to exhale in such a way that the pulmonary pressure is only slightly elevated when surfacing.

Fig. 19-7: Principal features of breath-hold diving in a submarine-escape training tank.

At 90 m of depth the ambient pressure is 10 atm.abs., so with a *total lung capacity* of 6 l, the lungs of the submariner contain the same number of air molecules as those in 60 l of air at 1 atm.abs. according to Boyle-Mariotte´s law. If the submariner panics and try to keep all air in his lungs, the gas will expand the lungs until they rupture, and the air enters the surrounding tissues, pleural cavity and blood vessels. The *air embolism* can kill the submariners even before they reach the surface. When training such a controlled procedure in submarine escape training facilities (Fig. 19-7), it is important to have a *recompression tank* close to the surface, and means to place a victim of air embolism at high pressure within seconds.

Air embolism is a *hyperbarotrauma* or barotrauma of ascent (Fig. 19-7). This danger is present whenever a person breathes compressed air at depth. An emergency escape even from 3 m of depth may lead to death, if the subject does not exhale. The lung volume increases during ascent according to Boyle-Mariotte´s law. The low-pressure vessels of the thoracic cavity are thin-walled. They are compressed and the alveoli dilatate until they burst, whereby the air dissects its way into the tissues (*subcutaneous emphysema*) and into the pleural cavity (*pneumothorax*). Alveolar air enters the blood through damaged vessel walls. The blood stream to the brain, heart and lungs are blocked by air (air embolism), and death ensues within minutes due to the blocked bloodflow to brain and heart.

2d. Squeeze or hypobarotrauma

is a function of *Boyle-Mariotte´s law*: For a certain amount of gas the product of pressure and volume is constant at constant temperature.

The hypobarotrauma is caused by a negative pressure difference across the wall of a non-collapsible air space in the body (middle ear, sinuses, hollow teeth, compressed lungs etc). A breath-hold diver who dives too deep will eventually bleed into his alveoli, and the presence of **lung squeeze** is evident, when he surfaces with **bloodstained froth** around the mouth.

When a healthy person, with a total lung capacity (TLC) of 6 l and a residual volume (RV) of 1.5 l at the surface, breath-hold dive to 30 m or 4 atm.abs., his TLC is compressed to 1/4 or 1.5 l, which is equal to his RV at the surface. By calculation this depth is assumed to be the maximal diving depth. In such a calculation is implied that the RV at the *surface* is equal to the RV at the *bottom*. However, this is not true.

The *world record* is beyond 105 m, which is unbelievable, when compared to the maximum calculated above. The following two phenomena explain this world record. *RV decreases* with increasing diving depth, because the diaphragm is pushed upwards as the piston in a syringe,

and *blood* is pushed into the pulmonary circulation. At some further depth, the capillaries will rupture and blood/oedema fluid reaches the alveoli (alveolar squeeze). This is a *hypobarotrauma* or barotrauma of descent, and it has been observed in some of the record holders.

A diver who has caught a cold, will descend with sinus openings closed, and develop pain with depth as the pressure in the occluded sinuses becomes more negative compared to the surroundings (ie, *sinus squeeze*).

3. Disorders from hard hat and SCUBA diving

Use of a diving helmet with a hose, through which atmospheric air is pumped to the diver at depth, is the classical form of diving. The *depth limit* is around *50 m* for air breathing, because of the danger of gas narcosis and of decompression sickness at prolonged dives (Fig. 19-8). With a *hose-and-pump system* the inspired air is pressurised to match the surrounding pressure, and the expired air is given off to the water. Such a device is of no use in secret operations underwater.

Fig. 19-8: SCUBA and hard hat diving from ship or from personal transfer vehicles. Pathophysiological barriers of diving depth are given.

Self-Contained underwater breathing apparatus *(SCUBA)* is an open circuit system with a *demand valve* at the mouth. Usually the diver receives compressed air from 2 tanks carried on his back (Fig. 19-8). First the pressure of the air stream is reduced from the initial value of 100-200 atm.abs. inside the tank, to a pressure somewhat higher than the ambient water pressure. Then the air passes an inhalation demand valve due to the negative pulmonary pressure during inspiration, and finally the expired air is exhaled into the water at a positive pressure from the expiratory muscles. Most SCUBA dives are performed to depths of 30-40 m (Fig. 19-8). SCUBA divers release bubbles in the water when breathing, so they are useless for secret tasks.

The problems related to these types of diving are *3a. Decompression sickness*, *3b. Inert gas narcosis*, *3c. Body squeeze*, *3d. Oxygen toxicity*, *and* *3e. Carbon dioxide toxicity.*

3a. Decompression sickness

If the diver ascents from the bottom to the surface too rapidly, inert gases stored in the tissue form bubbles in the blood and tissues, just as bubbles are formed when a bottle of soda water is opened. The intravascular bubbles cause tissue ischaemia and necrosis. Besides there is also serious extravascular damage. The formation of bubbles is increased by exercise, just as more bubbles are formed when the bottle of soda water is shaken up. These phenomena are gathered in the concept *decompression sickness*.

Symptoms and signs can be present in all tissues of the body. The main problems are caused by bobbles blocking the blood supply. Pains occur in the muscles and in the joints (bends). *Life-threatening bubbles* may block the pulmonary capillaries, which trigger thoracic pain called *chokes* with alarming dyspnoea, pulmonary oedema and often death. The *CNS* symptoms and signs are dizziness, paralysis, collapse and unconsciousness.

The *ascent* from deep dives must be slow and systematically in stages.

Stage decompression, according to decompression tables, prevents decompression sickness in the majority of seemingly healthy persons. Stage decompression with a rate below 18 m per min between stages, allows most people to ascent without decompression sickness (bends, caisson disease). All dives no deeper than 10 m allows the diver to emerge without stage stop (Haldane’s law). The limit for compressed air diving should be 50 m (150 feet).

Rational therapy requires *immediate recompression* in a pressure chamber, where sufficient pressure can be established to eliminate the bubbles causing the disease.

The nitrogen gradient from the divers body to the air in the decompression tank is preferably increased (with oxygen enriched air) for rapid removal of nitrogen from the body. Recompression can be succesful even hours after the dive.

Diving at great depth (80-300 m) makes it necessary to live in large habitats at depth for longer periods. These divers are saturated with the inert gas (He, Ar etc) to which they are exposed, and this type of diving is termed *saturation diving*. Helium is used together with a small O_2 fraction, in order to avoid acute O_2 poisoning (see below). At 200 m of depth only 1% O_2 is necessary in the *helium-oxygen* mixture (so-called *heli-ox-mixture*). The oxygen pressure in the inspired air is (21 atm.abs. * 0.01) = 0.21 atm.abs. or 160 mmHg.

Dives deeper than 50 m are performed with helium instead of nitrogen as the inert gas. The advantages of helium is a low solubility coefficient in the tissues, a low narcotic-toxic effect, a rapid diffusion rate out of the tissues, and a minimal airway resistance due to its low density. The breathing resistance due to the high density of nitrogen at great depth makes it almost impossible to perform manual work with nitrogen as the inert gas.

Cases of decompression sickness have occurred as a result of breath-hold diving in submarine escape training tanks and among the skin-diving pearl divers of the *Tuamotu Archipelago* in the South Pacific. When severe it comprises loss of consciousness and paralysis of one or more limbs, and is often fatal. Repetitive skin diving to great depths for prolonged periods in warm and ideal waters must be avoided.

3b. Inert gas narcosis

Rapture of the depth - or *nitrogen narcosis* - appears at 40 m when breathing compressed air. The diver becomes euphoric with behaviour similar to alcoholic intoxication (ie, lack of judgement and concentration, incoordination, anxiety). The inert gas narcosis increases in intensity with depth according to *Martini's law*: Each 10 m of diving depth changes the behaviour as much as one drink. Therefore, the limit for compressed air diving should be 50 m (Fig. 19-8). At a depth of 90 m or more a typical narcotic condition develops (anaesthesia and unconsciousness). Inert (inactive) gases are lipid-soluble and dissolve easily in fatty tissues, cell membranes and intracellular structures, where they bind to active sites or receptors. The gases modify neuronal activity and nerve conduction velocity as well as ionic transport across cell membranes. Argon has a larger narcotic effect than nitrogen (larger lipid-solubility, larger energy content or van der Waal-forces), and nitrogen is a much stronger anaesthetic than helium.

3c. Whole body squeeze

was a dramatic event of the hard hat diving period. Accidental disrupture of the air hose to the diver working at depth suddenly exposed the diver to an internal pressure of 1 atm.abs. while the ambient water pressure could be 6 atm.abs or more (Fig. 19-8). Before the diver was rescued the soft tissues of his body was virtually compressed into the helmet. Such accidents lead to the development of contra-valves in the helmet.

3d. Oxygen poisoning

Active oxygen or *oxygen free radicals* (such as the superoxide O_2- and hydrogen peroxide) are continuously produced in the mitochondria from the dissolved oxygen. As long as the oxygen tension of the tissues is normal, the production equals the removal by tissue enzymes. These enzymes can be inactivated following breathing of 100% oxygen above 2 atm.abs for longer

periods. Therefore *oxygen free radicals* accumulate to a degree that is lethal for cells in particular brain neurons.

The *acute cerebral* oxygen intoxication is presented with fasciculation of the mimic face muscles, vertigo, universal cramps and coma. The retinal cells of the eye are actually brain cells. Acute oxygen intoxication in premature babies may cause retinal vasoconstriction, vessel wall proliferation and retinolysis - so-called *retrolental fibroplasia.*

Closed or recirculation systems with pure oxygen are used by *frogmen* for secret tasks underwater. The acute oxygen toxicity limits the ***O_2 diver*** to dives not deeper than 7 m (Fig. 19-8), and duration less than 75 min. This depth is equal to a total pressure of (1.7 * 760) or 1292 mmHg (= 172 kPa). The P_{IO2} of the saturated tracheal air is (1292 - 47)= 1245 mmHg, implying a P_{AO2} which is toxic to the CNS. Recirculation systems with pure oxygen are so hazardous that they should be prohibited for sport divers.

Chronic oxygen toxicity frequently develops when a patient has been breathing 80% oxygen or more at one atm.abs. for more than 12 h.

Firstly, the patient develops pulmonary hypertension with typical *retrosternal* pain, dyspnoea, coughing, and pulmonary oedema with bloodstained frothy sputum.

Secondly, the patient develops atelectasis, because the surfactant is inactivated by the toxic free radicals. The exposure of lung tissue to high oxygen tension is direct and total without the protection of the haemoglobin or any other oxygen buffer system.

Some patients actually die of hypoxia in spite of the high oxygen tension!

3e. Carbon dioxide poisoning

Equipment with a large dead space permits accumulation of carbon dioxide with rebreathing. This is a problem with hard hat diving and SCUBA diving with large facemasks (Fig. 19-8).

The highest permissible F_{ICO2} is 0.005 or 0.5%. Slightly higher concentrations are harmful. Firstly, the diver increases ventilation, but soon after exposure to 5% or more, the respiratory centre of the medulla is suppressed and ventilation becomes insufficient. An *acute respiratory acidosis* develops, and finally *carbon dioxide narcosis* (anaesthesia and unconsciousness) is present.

4. Hyperbaric oxygenation therapy

Hyperbaric oxygenation therapy is used for disorders with either local or global oxygen deficiency with oxygen at approximately 2 atm.abs. These conditions are accomplished in pressure tanks with atmospheric air, such as those used for recompression treatment of decompression sickness. Oxygen is administered through a mask at ambient pressure. Acute O_2 poisoning is the risk.

Anaerobic bacilli cause *Clostridial infections*. Clostridium botulinum causing botulism, and clostridium tetani causing tetanus, both produce *neurotoxins*, whereas clostridium perfringens and septicum - causing gas gangrene - produce *enzymes*. The clostridia bacilli are spore forming and survive for years in our surroundings. The bacilli require anaerobic conditions to grow. See infectious disorders in Chapter 33.

The botulism- neurotoxins cause cholinergic and neuromuscular blockade leading to strabismus and *respiratory insufficiency* (mortality rate: 70%). - The tetanus-neurotoxins are *tetanospasmin* causes muscle spasms ending in neuromuscular blockade and death. - *Gas*

gangrene in lacerated wounds is a life-threatening condition.

The treatment of these three Clostridial disorders with *antitoxins* is sometimes supported by *hyperbaric oxygenation.* The *oxygen free radicals* are supposed to inhibit the growth of the clostridia. The therapeutic results - often some benefit in moribund patients - are controversial.

Air embolism, decompression sickness, CO-poisoning and leprosy have been treated with some benefit with hyperbaric oxygen.

Equations

(See Chapter 13: Eq. 13-3 to 13-6)

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A: The normal length of a snorkel is 0.6 m.

B: A P_{AO2} below 30 mmHg is the usual threshold for unconsciousness.

C: Air embolism is always easy to separate from decompression sickness.

D: The botulism- neurotoxins cause adrenergic and neuromuscular blockade.

E: The acute oxygen toxicity limits the O_2 diver to dives not deeper than 7 m.

II. Each of the following five statements have True/False options:

A. The cornea-air interface becomes a cornea-water interface, when in water without goggles. The refractive index for water and for cornea itself is much the same, eliminating most of the refractive power in air. A skin diver without goggles is hypermetropic under water.

B. Sound propagates more rapidly in soft tissue and water than in air (440 m s^{-1} as opposed to 340 m s^{-1}). Therefore underwater sound sources appear much nearer than they actually are.

C. The solubility coefficient for nitrogen in fat is 10 times larger than in blood.

D. Slightly higher concentrations of carbon dioxide than 0.5% in the inspired air are harmful to humans.

E. Lethal radiation dose is rapidly reached, when flying around the earth in the Van Allen radiation belts, consisting of high-energy level protons and electrons. Therefore, commercial flying takes place essentially below the inner belt (500 km).

Case History A

A male farmer, 18 years of age, is stabbed through the left foot while working as a stableman. He develops a fulminate infection with clostridium tetani, which only grows under anaerobic conditions and produce a potent toxin causing permanent depolarisation of the motor end plate. He is brought to hospital in a moribund condition, which is not improved by antitoxin and antibiotics. The farmer is transferred to a pressure chamber, where hyperbaric oxygen

therapy is provided at 3-atmosphere pressure for 20 min. The haemoglobin concentration of the patient is 170 g per l. P_B is one atmosphere (760 mmHg), P_{ACO2} is 5.3 kPa (40 mmHg), and the alveolar water vapour tension (P_{water}) is 6.5 kPa (47 mmHg). The Bunsen solubility coefficient (α) for oxygen in the blood is 0.022 ml STPD per ml and per mmHg. One mmHg equals 133.3 Pascal.

1. *Calculate the concentration of physically dissolved O_2 in the blood leaving the lung capillaries of the patient.*
2. *Calculate the concentration of chemically bound O_2 in the blood.*
3. *Does excessive oxygenation lead to gas transport problems?*

Case History B

A tall, lean diver, with a body weight of 80 kg, is working 30 m below the surface of the sea in a standard divers suit for one hour. The relative density of seawater is 1033 kg m^{-3}. The diver is breathing atmospheric air at the environmental pressure, delivered by a pump system to his helmet. The ventilation is effective, and the diver has normal alveolar gas tensions (P_{aO2} is 100 mmHg and P_{aCO2} 40 mmHg). The pressure at the surface of the sea is 1 atmospheres absolute (atm abs) or 101.3 kPa or 760 mmHg and the water vapour pressure of body warm alveolar air is 47 mmHg or 6.25 kPa.

1. *Calculate the total pressure in kPa and in atm abs at 30 m of depth.*
2. *Calculate the partial pressure of nitrogen in the alveolar air of the diver at the surface and at the bottom.*
3. *Is it advisable to drag the diver to the surface without step decompression?*

Try to solve the problems before looking up the answers.

Highlights

- *Fast ascent during flight rarely cause problems at altitudes below 10 km (30 000 feet).*
- *The aeroplanes have cabins pressurised to 80 kPa (ie an altitude of 2000 m or 6000 ft) or sometimes only to an altitude of 2750 m. These levels cause hypoxic dyspnoea only in cardiorespiratory patients.*
- *The main problem for the average traveller is jet lag, squeeze (hypobarotrauma) during take-off or Hyperbarotrauma during landing.*
- *The area of the parachute reduces the terminal velocity to 1/9. The force of impact at landing is $(1/9)^2$ = 1/81 of the landing force without parachute. This is equal to a jump from a height of 2-3 m.*
- *The space pilot experience space sickness (nausea and vomiting), falling total blood volume, muscular atrophy, Ca^{2+} loss from the bones, and after living in outer space for weeks, adaptive difficulties to life on earth upon return. These difficulties are low working capacity and a tendency to faint. The bone loss continues long after the return to earth, because stimulation of bone formation requires physical activity in a gravity field.*

- *At altitudes above 20 km, where the ambient pressure is less than 47 mmHg or 6.5 kPa, the blood of the space traveller will boil, if he is suddenly exposed because the pressure within his suit is lost.*
- *Any saturation dive to less than 10 m allows ascent without decompression stops (Haldane's law).*
- *If decompression is too rapid, inert gasses stored in the tissues form bubbles in the blood and tissues (decompression sickness).*
- *Drowning. Fresh water drowning kills the victim rapidly following haemolysis. Seawater is hypertonic and draws blood from plasma to the alveoli causing haemoconcentration and shock.*
- *Shallow water blackout is a condition where the oxygen stores are used up during forceful underwater swimming before sufficient carbon dioxide is produced.*
- *Carbon dioxide poisoning results in acute respiratory acidosis and finally in carbon dioxide narcosis.*
- *Oxygen poisoning is caused by active oxygen or oxygen free radicals (such as the superoxide O_2- and hydrogen peroxide).*
- *Rapture of the depth - or nitrogen narcosis - appears at 40 m when breathing compressed air. The diver becomes euphoric with behaviour similar to alcoholic intoxication (ie, lack of judgement and concentration, incoordination, anxiety).*
- *Hypobarotrauma (squeeze) is a barotrauma of descent and a function of Boyle-Mariottes law.*
- *Hyperbarotrauma (air embolism) is a barotrauma of ascent during emergency escape. The danger is present whenever a person breathes compressed air at depth.*
- *Clostridial infections are treated with antitoxins and sometimes supplied with hyperbaric oxygenation. Accumulation of oxygen free radicals kills the microorganisms. Some cases of air embolism, decompression sickness, CO-poisoning and leprosy have been treated.*

Further Reading

Bennett, PB, and Elliott, DH (1993) The Physiology and Medicine of Diving. Philadelphia: WB Saunders Co.

Paulev, P-E, Pokorski, M, Honda, Y, Ahn, B, Masuda, A, Kobayashi, T, Nishibayashi, Y, Sakakibara, Y, Tanaka, M, and Nakamura, W. Facial Cold Receptors and the Survival Reflex "Diving Bradycardia" in Man. *Jpn. J. Physiol.*40: 701-712, 1990.

Undersea & Hyperbaric Medicine. Journal published by the Underwater and Hyperbaric Medical Society, 10531 Metropolitan Av., Kensington MD 20895, USA.

West JB. *Respiratory Physiology – the essentials*. Williams & Wilkins, Baltimore, USA, 1999.

Return to chapter 19

Return to content

Chapter 20

Metabolism & Nutritional Disorders

Study Objectives

- To *define* heat energy, basal metabolic rate, Gibbs energy for ATP-formation, mechanical efficiency, metabolic rate, and the energy equivalent for oxygen.
- To *describe* direct and indirect calorimetry, factors influencing metabolic rate and basal metabolic rate, and conditions with unsteady respiratory state.
- To *draw* a curve for the combustion rate of alcohol.
- To *calculate* a metabolic variable from relevant variables given.
- To *explain* the alcohol metabolism and toxicity. To explain the control of appetite, dietary thermogenesis, energy balance, net combustion and RQ relations. To explain the first law of thermodynamics applied to humans.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The first law of thermodynamics. The internal energy of a system can change for any transition between two equilibrium states and is equal to the heat exchanged by the system and the work done by or on the system. – As a consequence, the metabolic heat energy transfer equals the heat loss plus the stored heat energy.*
- *The surface law: The basal metabolic rate (BMR) per body surface area is much more uniform than the BMR per kg of body weight in individuals of the same species but of different form and size. The best expression for comparison is the BMR per kg of lean body mass. The lean body mass is the fat free mass.*
- *Van´t Hoof´s rule: The rate of energy conversion in chemical reactions increases in proportion to the rise in temperature.*

Definitions

- **Basal metabolic rate** *(BMR)* is defined as the metabolic rate measured with the subject awake in the morning, fasting, at neutral ambient temperature and resting horizontally in the respiratory steady state.
- **Body mass index** (BMI) is the weight of the person in kg divided by the height (in m) squared. The normal range is 19-25 kg per square metre.
- **Brocas index** is the predicted body weight in kg, which equals the height of the person in cm minus 100 for males and 110 for females.
- **Dietary thermogenesis** is the increase in metabolic rate following food intake.

- **Energy balance** is a condition, where the energy input equals the energy output, so the energy stores of the body are unchanged.
- **Gibbs energy** is the free chemical energy in food, which is available for life.
- **Heat energy** is energy transfer caused by a temperature gradient.
- **Ideal weight** refers to the weight associated with the highest statistical life expectancy. The ideal weight is determined with the Brocas index or with prediction tables.
- **Inactivity** is defined as a low *endurance capacity* (ie, a maximal oxygen uptake below 34 ml * min^{-1}*kg^{-1}). Inactivity is probably involved in development of *life-style risk factors.*
- **Lean body mass** is the fat free body mass.
- **Marasmus** is the result of universal starvation in a child, who has low body weight, muscle wasting, and the look of an old person.
- **Mechanical efficiency** is the ratio between external work and the total energy used during work.
- **Metabolism** is defined as the sum of all chemical processes in which energy is made available and consumed in the body.
- **Metabolic rate** *(MR)* is defined as the decrease in internal energy (enthalpy) of the body in a given time period. The metabolic rate refers to the measurement in energy units with direct or indirect calorimetry.
- **Net metabolism** is the stochiometric sum of the net reactions in the body.
- **Nitrogen balance** is a condition, where the nitrogen input from absorbed amino acids equals the nitrogen output in the urine.
- **Obesity** implies the excess storage of fat, and is defined as an actual body weight exceeding the ideal weight by more than 20%, or by a body mass index above 30 kg per square metre (WHO).
- **Protein deficiency** (kwashiorkor) is starvation in children, which subsists on a protein-poor diet rich in carbohydrates.
- **Respiratory Quotient (RQ)** and **ventilatory exchange ratio (*R*)** is defined in Chapter 14.
- **Teratogens** refer to all chemical, physical and biological agents that cause developmental abnormalities (teraton means monster).
- **Vitamins** are *essential organic catalysts* in the diet, necessary for normal metabolic functions in humans, but not synthesized in the human body.
- **Respiratory steady state** is a condition where RQ equals *R*.

Essentials

This paragraph deals with *1. Energy exchange, 2. Metabolism, 3. Alcohol, 4. The respiratory quotient and R, 5. Net mechanical efficiency, 6. Energy sources, 7. Direct calorimetry, 8. Native diets, and 9.*

Control of energy balance.

1. Energy exchange

It is generally believed that nutrients are necessary in order to *produce* energy in the human body. However, this is *impossible*. The *first law of thermodynamics* states that energy can neither be created nor destroyed but is *transferred* from one form to another or from one place to another.

Life is thermodynamically the maintenance of an infinite row of non-equilibrium reactions in such a way that appear to be in a stationary condition, a *steady state.* Real life is *chaos*, a steady state only maintained as long as we derive chemical energy from food. Only part of the dietary energy is available for ATP formation in humans. Cellulose, for example, passes the digestive tract without being absorbed. The absorbable chemical energy passes through the intestinal mucosa, and is in the body transformed to energy rich phosphate bindings in ATP (*Gibbs-energy*, D*G*).

ATP is broken down to ADP during muscular contractions. Muscular contractions stimulate the oxidation of fatty acids and carbohydrates in the muscle cells which liberate more energy for rephosphorylation of ADP to ATP. The energy is used for the maintenance of chemical syntheses, electrochemical potentials and for the net-transport of substances across membranes.

The *Gibbs energy* is the free chemical energy available in food. However, 75% is lost as heat energy, and the *mechanical efficiency* of exercise is therefore only 25%. The ratio between *external work* (*W*') and the total energy used during work (-DU) is called the *mechanical efficiency*. In this case DU equals D*G*. The *mechanical efficiency* is always less than one and often only 0.25 as stated above. The energy, which is not transferred to external work, is released as heat energy (-*Q*) or is accumulated in the body as heat. At the onset of exercise 50% of the total energy from hydrolysis of ATP is converted into mechanical energy in the myofibrils. The remaining 50% are lost as *initial heat*. As shown above the mechanical efficiency is only 25%, however. This is because energy recapturing recovery processes (oxidative regeneration of ATP etc) occur outside the myofibrils. Hereby, half of the energy is dissipated as so-called *recovery heat*.

Heat energy is *low prize energy*. In contrast to ATP energy, it is not available for work in the body. The sum of heat energy generated and work performed is constant and equal to the Gibbs energy.

When no work is performed *W*' is zero, and all body reactions are reflected by the liberated heat energy (-*Q*), which is equal to the decrease in Gibbs energy (-D*G*).

When the pressure-volume work is zero, we have a special energy concept: the *heat content* or *enthalpy*, *H*, which sums up all energy. The sum of liberated heat energy (*Q*) and liberated work (-*W*') is thus equal to the fall in enthalpy (Eq. 20-1).

The decrease in enthalpy of the human body (-D*H*) is equal to the fall in potential, chemical energy stored in the body.

The *decrease* in Gibbs energy covers almost the total energy, except for the pressure-volume work. Since oxygen consumption is almost equal to the carbon dioxide output, the pulmonary volume change is negligible and this work is negligible.

2. Metabolism

The *metabolism* of a person is defined as the sum of all chemical reactions in which energy is made available and consumed in the body. The bindings between hydrogen and carbon in nutrients are a source of energy for animals. Such substances are changed into *metabolic end products* (eliminated as bilirubin, urobilin, urea, uric acid, creatinine etc.) and to *metabolic intermediary products* (ie, products that participate in other chemical reactions). The *net metabolism* is the sum stochiometry of

the single net reactions in the body.

The decrease in enthalpy in a given time period (-D*H*/min) is the *metabolic rate (MR).*

The *oxidation* of fuel (carbohydrates, glycerol, fatty acids) to CO_2 and water is the primary pathway for generation of energy and subsequent heat energy liberation. Protein can also serve as an important energy source during prolonged exercise, but it must first be broken down to amino acids, who are then partially oxidised (to CO_2, water, NH_4^+ etc). The daily production of *metabolic water* is 350 g and of urea 30 g.

Diabetes mellitus and hunger (*hunger diabetes*) are conditions where fatty acids can produce ketone bodies.

During forceful exercise, energy is obtained primarily from *non-oxidative sources* (glycolysis). There is, therefore, a net formation of lactic acid from glycogen. Following anaerobic exercise the *lactate elimination* accounts for an extra O_2 consumption called *oxygen debt* (Chapter 18).

Oxidation of alcohol can contribute to metabolism. The energetic value of alcohol is 30 kJ/g. An adult person of *70 kg* body weight can combust *7 g* of alcohol per hour (see calculation below). The chemical energy liberated is (7 × 30) = *210 kJ* per hour or 70% of his resting MR (300 kJ per hour or 83 Watts).

Most of the chemical reactions in our body are *degradative* or *catabolic* - they break a molecule down to smaller units. These reactions are often also *exothermic* (heat releasing) and *exergonic* (the content of Gibbs energy decreases during these reactions). The *synthetic* or *anabolic* reactions (the formation of protein from amino acids) are obviously coupled to these degradative reactions. Synthetic reactions are most often also *endothermic* and *endergonic*.

3. Alcohol

Alcohol diffuses easily in the human body. 20% of the intake by drinking is already absorbed in the stomach. The absorption is *fast* and is stimulated by CO_2 (champagne).

Alcohol distributes in the *total water* of the body within one hour. The distribution volume depends upon the fat mass, because fat tissue only contains 10% of water. The Swedish scientist Widmark called the fraction of the body weight, which is distribution volume for alcohol, **r**. The values for **r** varies considerably, but the *mean*-**r** for females is 0.55 and for males it is 0.68 kg per kg of body weight.

The blood alcohol concentration is measured in permille (ie, one g of alcohol per kg of distribution volume). The most important elimination of alcohol is by *oxidation*. The rate of alcohol oxidation is constant (b = 0.0025 permille per min) and is independent of the blood alcohol concentration. The absolute amount of alcohol *eliminated* per minute is: (b × r × body weight) - see Eq. 20-4.

The constant rate is due to the primary, partial oxidation to acetate via acetaldehyde in the liver by *alcohol dehydrogenase*: $C_2H_5OH + 0_2 \text{ « } CH_3COOH + H_20$. Acetate is broken down in nearly all tissues. The total oxidation of alcohol: $C_2H_5OH + 3\ 0_2$ Þ $2\ CO_2 + 3\ H_20$ implies an RQ of 2/3. A healthy person with a metabolic rate (*MR*) of one mol O_2 per hour can partially oxidise almost 1/6 mol of alcohol per hour, by using almost 1/6 of his *MR* in the liver (one mol = 46 g alcohol; 46/6 or about *7 g* alcohol per hour).

Fig. 20-1: Absorption and oxidation of alcohol.

If this standard person receives an alcohol infusion of 7 g per hour and has a normal hepatic bloodflow of 90 l per hour (1.5 × 60 min), his maximal alcohol elimination rate corresponds to a blood [alcohol] of (7/90) = 0.08 g per l. This is a blood [alcohol] threshold below, which the oxidation rate decreases with time (Fig. 20-1).

The excretion of alcohol molecules takes place through *expiratory air, urine and sweat*. This excretion is generally considered to be negligible compared to the oxidation. This is actually true at rest (Box 20-1). A resting athlete with a blood [alcohol] of one permille or 1 g per kg has a small alcohol partial fraction in pulmonary blood and alveolar air (1/2000- 1/2100). With an alveolar ventilation of 5 l BTPS per min at rest, this person excretes *0.15 g* each hour via expiratory air (Box 20-1). The concentration of alcohol in plasma water, sweat and urine is 20% (1.2 fold) higher than the blood [alcohol]. A resting athlete with a diuresis of one ml/min or 0.060 l per hour excretes alcohol by renal ultrafiltration at a rate of only 0.072-g per hour. If the person also has a sweat loss of 0.1 l each hour, he further excretes *0.12 g* per hour. The total excretion at rest is only *0.342 g* each hour (Box 20-1).

Box 20-1: Oxidation of alcohol is generally agreed to be the single essential elimination method Look here for excretion, which is considered to be negligible. The person is a male athlete with a blood alcohol of one permille.

Excretion of alcohol by Route	Resting condition g each hour		Exercise g each hour	
1. Expiratory air	(1×5×1/2000*60)	= 0.15	(1*80*60*1/2000)	=2.4
2. Urine	(1.2*0.060)	= 0.072	-	
3. Sweat	(1.2*0.1)	= 0.12	(1.2*4)	= 4.8
Total	**0.342**		**7.2**	

However, during *one hour of exercise* in a warm climate, when ventilation is 80 l per min, and when water loss is 4 l per hour (sweat and evaporation), the total alcohol excretion of the athlete is *7.2 g* per hour. This total *excretion* is actually larger than the amount broken down by *maximal oxidation: 7 g* each hour (calculated above).

The rate (b) increases with *increasing* temperature, with *increased* metabolism (thyroid hormones, dinitrophenol), and decreases under the influence of enzyme inhibitors.

Hepatic alcohol dehydrogenase metabolises alcohol to acetaldehyde, which is oxidised to acetate by aldehyde dehydrogenase in the mitochondria. Acetate is then oxidised to carbon dioxide and water, primarily in the peripheral tissues. Fructose increases b. Both enzymes are dependent on nicotinamide adenine dinucleotide (NAD^+). One mole of alcohol oxidised to acetate produces 2 moles of reduced nicotinamide adenine dinucleotide (NADH).

The sum [NAD^+ + NADH] is constant. Hepatic alcohol oxidation causes [NADH] to rise, so that NADH inhibition becomes the rate-limiting factor for oxidation.

There are two other enzymes, apart from *hepatic alcohol dehydrogenase*, that can oxidise alcohol. These are catalase and MEOS (Microsomal Ethanol Oxidation System). A small amount of alcohol dehydrogenase is found in the gastric mucosa.

4. The respiratory quotient and R

RQ is the *hypothetical, metabolic ratio* between carbon dioxide output and oxygen consumption of all the cells of the body. RQ is an indicator of the type of foodstuff metabolised.

R is the *ventilatory ratio* between CO_2 output and O_2 uptake for the person quantified by gas exchange equipment.

Respiratory *steady state* is a condition where *R* equals RQ, and the gas stores of the body are unchanged. See Chapter 16.

Compared to the oxidation of carbohydrates (RQ = 1), fat oxidation has a distinctly low RQ (0.7), and protein is oxidised with a RQ of 0.8.

Carbohydrates are rich in oxygen compared to the minimum in fats. Overfeeding with carbohydrates results in a partial conversion to fat. The corresponding release of oxygen from carbohydrates diminishes the oxygen uptake, and *R* becomes larger than one.

The diminished glucose metabolism during fasting and in diabetics lowers the *R* towards 0.7, because of the increased conversion rate of fat.

Hyperventilation decreases the amount of exchangeable CO_2 in the large body stores, without altering oxygen uptake. The tissues and blood cannot store additional oxygen. As a consequence the $V°_{CO2} / V°_{A}$ -ratio (= F_{ACO_2}) is reduced. This implies a fall in P_{ACO2} and in P_{aCO2}. *R* is distinctly increased during hyperventilation - often up to 2-3.

Hypoventilation reduces *R* towards zero at apnoea.

Metabolic acidosis is characterized by low pH and by negative *base excess* in the extracellular fluid (Chapter 17). A high pH and positive base excess characterise *metabolic alkalosis*. Metabolic acidosis is compensated by hyperventilation implying a rise in *R*, and metabolic alkalosis is compensated by hypoventilation with a fall in *R*.

R does not change when a person on a mixed diet (RQ = 0.83), or when a person on a high fat diet (RQ = 0.7) exercises moderately, because the fat combustion dominates.

R will fall, however, when a person on *carbohydrate rich* diet (RQ = 0.96-1) works for hours.

Strenuously heavy exercise implies a substantial, initial rise in *R* (*R*>3), because the lactate liberated will release CO_2, which is then eliminated in the lungs in much larger volumes than oxygen is taken up.

Glycogen: $(C_6H_{10}0_5)n + 6n\ 0_2 = 6n\ CO_2 + 5n\ H_20$, that is RQ = 1.

Glucose: $C_6H_{12}0_6 + 6\ 0_2 = 6\ CO_2 + 6\ H_20$, that is RQ = 1.

The enthalpy released per mol of glucose is 2826 kJ. One mol of glucose has a mass of 180 g, and 6 mols of oxygen have a volume of $(6 \times 22.4) = 134.3$ l STPD. The *enthalpy* per g of glucose is thus 2826/180 = 15.7 kJ/g, and the *energy equivalent*, which expresses the energy with respect to the oxygen consumed, is 2826/134.3 = 21 kJ per l STPD.

The dietary protein-nitrogen is equal to the nitrogen excretion in the urine when the person is in nitrogen balance. Protein-retention during growth, training, protein-rich diet, pregnancy and reconvalescens are called *positive* nitrogen balance (not urea accumulation in uraemia). Protein-loss

during inactivity, bed rest, fever, blood loss, burns and lesions is called *negative* nitrogen balance.

5. Net mechanical efficiency

The *net mechanical efficiency (E_{net})* is the ratio of *external work rate* (N × m/s = J/s) to net chemical *energy expenditure* (J/s or Watts) during work. E_{net} is 20-25% in isolated muscles and also in humans during aerobic cycling. Its size increases with the amount of training, because the untrained individual does not use the muscles effectively. Legwork has the largest E_{net}, since arm work necessitates fixation of the shoulder belt. The work rate is measurable with a cycle-ergometer (Eq. 20-5).

6. Energy sources

The predominant source of energy is *oxidation* of fuel in the mitochondrion. Hereby, high-energy compounds such as creatine phosphate and ATP are formed. Glucose is oxidised by nicotinamide-adenine-dinucleotide (NAD^+), so by glycolysis two pyruvate molecules are formed in the cytosol, transported to the mitochondrion, and transformed to a *Co-enzyme-A derivative (acetyl-CoA),* which then is involved in the*Tri-Carboxylic Acid (TCA)* cycle (Fig. 20-2).

Provided a certain oxygen flux from the lungs to the mitochondria is present, the electron transport chain (the glycero-phosphate shuttle) will reoxidize ($NADH+H^+$) and $FADH_2$ to NAD^+ and FAD (Fig. 20-2).

In the glycolysis, one glucose molecule is converted to 2 molecules of pyruvate, with the other products being 2 ATP and (2 NADH + H^+).

Through the oxidation of pyruvate in the TCA-cycle, three ($NADH+H^+$), one $FADH_2$, and one GTP are formed. If *complete oxidation* occurs in the glycerophosphate shuttle of the mitochondrion, one NADH equals 3 ATP, and one $FADH_2$ equals 2 ATP. Since the NAD^+ reduced in the glycolysis is cytosolic, it usually equals 2 ATP only, depending on the shuttle used.

When pyruvate is transformed to acetyl-CoA, one molecule of ($NADH^+$+H^+) is formed.

The total production by use of the *glycero-phosphate shuttle* in oxidative phosphorylation is *36* ATP molecules per glucose molecule (6 from the glycolysis, 6 from the transformation and 24 from the TCA cycle). - If the *malate-aspartate shuttle* is used, a total of *38* ATP molecules are formed per molecule of glucose oxidised.

Oxidation of one glucose molecule typically implies the use of six oxygen molecules. Accordingly, the P: O_2 ratio is 36/6 = 6, which is equal to a P:O ratio of 36/12 = *3*. The free fatty acids (FFA) from the cytosol (intramuscular or extramuscular origin) are transformed to acetyl-CoA (Fig. 20-2).

The pyruvate production rises with increasing glycolysis rate, and pyruvate is the substrate for alanine production. Alanine is liberated to the blood and its concentration increases linearly with [pyruvate] during rest and exercise.

During anaerobic conditions - an insufficient oxygen supply - (NADH + H^+) is reoxidized by the pyruvate- lactate reaction, and the glycolysis continues. The anaerobic ATP production does not block the aerobic ATP production, but functions as an *emergency supply*.

Fig. 20-2: Biochemical pathways for ATP production.

The largest rise in blood [lactate] takes place at work intensities above 50% of the maximum oxygen

capacity. Lactic acid is a fixed acid - in contrast to the volatile H_2CO_3 - produced during exercise, and in a muscle cell with a pH of 7 such an acid is essentially totally dissociated (pK = 3.9). Since the proton associated with lactate production reacts immediately with bicarbonate within the cell, its CO_2 production must increase by one mol CO_2 for each mol of *bicarbonate* buffering *lactic acid*.

Lactate accumulates in the muscles and blood, if the glycolysis proceeds at a rate faster than pyruvate can be utilised by the mitochondria, or if ($NADH + H^+$) is not reoxidized rapidly enough.

We posses *100 mmol* of glucose (stored as glycogen) per kg of wet muscle weight, or 3.5 mol in the muscle tissue. Muscle tissue does not contain glucose-6-phosphatase. Our normal 5-l of circulating blood only contains 5 mM, or as a total 25 mmol (5 g) of glucose. During exercise the muscle uptake of glucose increase considerably, but the blood [glucose] does not fall. The blood [glucose] is kept normal by an increased flux of glucose from the liver (Fig. 20-3).

Fig. 20-3: A schematic overview of carbohydrate metabolism.

1. With increasing intensity and duration of exercise, the sympatho-adrenergic activity and the blood [catecholamines] increase. This is a strong stimulus to the *hepatic glucose production*. The liver contains 50-100 g of *dynamic glycogen*. This liver glycogen is easily broken down into glucose by *glycogenolysis* and released to the blood. Any fall in blood [glucose] during exercise will increase the blood [glucagon] and decrease [insulin] toward zero. Glucagon is bound to hepatocyte receptors, and via cAMP a *glycogenolytic cascade* is started (Fig. 20-3). Hereby the hepatocytes produce large amounts of glucose, sparing muscle glycogen and delaying the onset of fatigue. The lack of insulin inhibits the glucose transport across the cell membranes.

2. Glucose is also produced by *gluconeogenesis* in the liver from glycerol, lactate, pyruvate, and glucogenic amino acids. The gluconeogenesis is stimulated by pituitary ACTH and by cortisol from the adrenal cortex.

   With prolonged exercise the blood [glucose] will fall at the end, when hepatic and muscle glycogen stores are depleted, and the *compensating gluconeogenesis* is also running out of energy sources.

3. Complete exhaustion is delayed considerably in trained athletes, because they utilise lipids, so the glycogen stores are spared by *oxidation* of free fatty acids (FFA).

Skeletal muscles contain *lipid stores* (20-g triglycerides/kg wet weight or 700 g in a person with 35-kg muscles). A standard 70-kg man also contains *extramuscular fat stores* of triglycerides (15 kg).

Sympathetic activity and catecholamines *increase* lipolysis (i.e., hydrolysis of the stored adipose tissue to FFA and glycerol) via activation of *adenylcyclase*, increase in cAMP, phosphorylation and activation of the *hormone sensitive lipase*. Increased blood [lactate] and *glucose intake* reduces lipolysis during exercise.

The fat stores are the *ideal energy stores* of the body, because a large quantity of ATP is available per g; this is due to the relatively low oxygen content of lipids - the point being that the necessary oxygen is inhaled at request.

At rest we have a slow turnover of muscle protein, but during exercise alanine is released in appreciable amounts by transamination of pyruvate in the muscle cells, and the blood [alanine] is doubled - without any important change in other amino acids. Alanine is produced via the *pyruvate-alanine cycle*, and the amino groups are from valine, leucine and isoleucine. The blood to the liver transports the muscle alanine, where its carbon skeleton is used in the *gluconeogenesis*. The blood

[alanine] also stimulates the pancreatic islet-cells to *increased* glucagon secretion. Glucagon activates the *glycogenolytic cascade* (see above) in the liver cells, further stimulating glucose output from the liver. These are the two factors in the *alanine-liver cycle* of exercise.

The ventilatory, the cardiovascular and the metabolic systems are coupled, and determined by the following factors:

The primary factor is the *size* of P_{aO_2}, but the *blood oxygen store* is of similar importance in keeping P_{aO_2} as high as possible. The blood oxygen store depends upon the haemoglobin concentration, the haemoglobin-oxygen affinity incl. 2,3-DPG, temperature, and P_{aCO_2}.

The *total oxygen flux* to a certain population of mitochondria also depends upon the bloodflow (ie, cardiac output, muscle bloodflow, lung perfusion etc.).

Indirect measures of enthalpy (*MR* in kJ/min) are easily applicable both at rest and in an exercise setting. Expired air is collected in a Douglas bag (volumetric principle) for subsequent air analysis, and the *volume of oxygen* consumed per min is calculated. It is convenient also to determine the carbon dioxide production in the same period, because their ratio is the respiratory quotient (RQ).

A person on a mixed diet has a RQ of 0.83 and a heat energy yield of 20 kJ per l or 0.45 kJ/mmol of O_2. The metabolic rate (in kJ/min) is calculated by multiplying the *estimated* volume (l/min) of O_2 consumed with *20* kJ per l. The heat energy *yield* varies with RQ and is found in a table (see Symbols). A metabolic ratemeter - a spirometer (Fig. 13-1) with CO_2 absorber - is practical for determination of oxygen uptake.

A more detailed calculation of the metabolic rate is performed as shown with Eq. 20-6 and ->.

Disadvantages of indirect calorimetry are that it ignores the O_2 debt, and that the method depends upon maintained nitrogen balance and gas stores.

7. Direct calorimetry

The total output of heat energy from the body is most precisely measured in a whole-body calorimeter. The Atwater-Rosa-Benedict's *human calorimeter* has been used to verify the *first law of thermodynamics* in humans. The heat energy delivered from the chamber is only equal to the metabolic rate (*MR)*, provided the external work is zero, and neither equipment nor the human body alters temperature.

Fig. 20-4: The human calorimeter combined with a metabolic ratemeter.

The major single factor is *muscular activity*, which can increase *MR* with a factor of 20 even for hours in marathon running. Inactive persons can have a daily *MR* of 9600 kJ, whereas heavy occupational labour requires 20 000 kJ (20 MJ).

Dietary intake can increase *MR* by 20-30% (see Specific Dynamic Activity, below).

Increased *energy demand* in heart and lung diseases, rapid growing cancer will increase MR importantly. Energy is also lost in other disease states such as proteinuria, glucosuria, ketonuria, diarrhoea, and exudate loss (of plasma) through lesions in the skin or in the mucosa. An extra physiologic *energy loss* takes place during pregnancy and during nursing.

Deposition of heat energy in the body (as in fever and hyperthermia) can increase *MR*.

No work is done under basal conditions, so that all energy is ultimately liberated in the body as heat

energy. The liver and the resting skeletal muscles account for half of the basal metabolic rate.

Measurement of the *basal metabolic rate (BMR)* requires the subject to be awake in the morning, fasting and resting horizontally. The ambient temperature must be *neutral*, which is the temperature at which compensatory activities are minimal. Prediction tables for *BMR* in different races are available, and the variables are age, sex, height, and weight and thyroid function.

BMR is rarely used for diagnosis of thyroid disease, because radioimmunoassays (Chapter 26) for thyroid hormone analysis are specific and uncomplicated in use.

The surface law states that the *BMR per body surface area* is much more uniform than the *BMR* per kg of body weight in individuals of the same species but of different form and size. The best expression is the BMR per kg of lean body mass. The lean body mass is the fat free mass. Among different animal species the large animals (elephants) have the smallest relative surface area (ie, surface area per kg), so elephants must have small BMR per surface area compared to mice. This is because the surface-volume ratio decreases with increasing body weight. Besides, small animals also have a thin body shell. The body surface area is estimated with Eq. 20-8.

BMR decreases with age in both sexes (Fig. 20-5).

Fig. 20-5: The basal metabolic rate in females and males decreasing with age.

The *female* surface-related *BMR* values are approximately *10% below* the male values throughout life. Let us compare a female and a male both 21 years of age. The female has a Height of 1.68 m, weight 58 kg and a surface area of 1.66 m^2, whereas the male values are:1.8 m, 76 kg and 1.95 m^2. Calculations from the values read at Fig. 20-5 result in *BMR*s of 70 and 90 Watts, respectively. Now, let the couple live for 50 healthy years maintaining height and weight. At the age of 71, their *BMR*s are reduced to 60 and 75 Watts, respectively.

Intake of meals as such *increases* metabolic rate. This is the *specific dynamic activity of the diet (SDA)* or *dietary thermogenesis*. SDA is less than 10% of the intake energy for carbohydrates and for fat, but 30% for proteins (Fig. 20-6).

Fig. 20-6: Dietary thermogenesis or so-called specific dynamic activity of foods.

Glucose loaded person forms glycogen and fatty acids out of glucose within an hour, even before the glucose can be oxidised. Accordingly, the SDA caused by glucose can be due to an obligate formation of glycogen and fatty acids. The thermogenic response to carbohydrate seems to include a *muscular* component activated by adrenaline via b_2-receptors and a *non-myogenic* component activated by noradrenaline (NA) via b_1-receptors.

Proteins have no SDA in hepatectomized animals, so hepatic intermediary processes must cause the SDA of proteins. These intermediary processes include formation of *urea* from NH_4^+, breakdown of *amino acids* etc.

In general, SDA can also be related to *mass action* due to increases supply of nutrients, and to temperature increase by the activity (increases the rate of all enzymatic processes).

8. Native Diets

Native diets in Africa and the Orient are rich in fibre, which are plant substances (ie cellulose, hemicellulose A & B, and lignins) resistant to digestion. Dietary fibre has been used in an attempt to cure obesity. Constipation with or without diverticulosis/ diverticulitis of the colon also responds to dietary fibre.

The most widespread *dietary fibre* is *cellulose*, which is a major component of plant cell walls. Cellulose is a linear glucose polymer, but human intestinal enzymes cannot hydrolyse its b-1,4-linkages.

Hemicellulose A is a heteropolymer with linkages between glucose, galactose, mannose, xylose and arabinose (ie, gums or mucilages). *Mucilages* delay gastric emptying and decrease the rate of intestinal absorption.

Hemicellulose B or *pectin* binds water in the gastrointestinal tract, but in addition salts minerals and heavy metals. Hemicellulose A and B seem to lower LDL concentrations, while maintaining HDL concentrations.

Lignins in natural fibres are cross-linked polymers of oxygenated phenylpropane entities. Lignin provides bulk for the faeces because they are difficult to degrade.

Dietary fibres reduce postprandial blood glucose and insulin concentration.

Delay in gastric emptying caused by some dietary fibres reduces symptoms of the dumping syndrome. This is an unwanted consequence of large *gastrectomies*. Following removal of the major part of the stomach, the food pass quickly down the small intestine and elicit distension by nutrients and osmosis, causing a massive sympathetic activity with discomfort.

Dietary fibres seem to prevent hiatus hernia by softening the food bolus and decrease of the swallowing effort. Softening of the faecal bulk with decreased defaecation strain seems to reduce the frequency of haemorrhoids.

Overconsumption of dietary fibre can produce adverse effects with increased flatulence, diarrhoea and intestinal discomfort.

Fig. 20-7: Continuous fasting leads to numerous serious complications or death.

Fasting is a *total stop* of food intake. After 12 hours of fasting, conditions are optimal to measure *BMR* or to analyse the chemical composition of blood (fasting blood values are predictable and easy to interpret). The 12 hours are the methodological criterion for the correct minimum *BMR*, but continued fasting for day's results in a much lower value (65% of *BMR*). Following the first 2 weeks of hunger is the normal body weight reduced to 85%, whereas the resting *MR* is stable at 65% of *BMR*, which is constant to the end of the fasting period (either voluntary or by death).

Glycogen stores are broken down in a few days, since only small stores prevail in the liver and muscles. Then urine nitrogen *increases* as a sign of renewed protein combustion (*gluconeogenesis*). In general the *fat* combustion dominates, until the fat stores are used. Healthy people contain 5-15 kg of fat, but monstrous amounts have been recorded in a 540-kg male from *Guinness Book of Records*.

Oxidation of fat stores - including the partial hepatic oxidation to ketonic bodies - implies development of ketoacidosis and a diabetic glucose tolerance test (Chapter 27). Such a *hunger diabetes* with ketonaemia and ketonuria as in diabetes, have been found in healthy individuals even after only 24 hours of fasting or after extremely fatty meals.

Serious illnesses develop after a few weeks of fasting, because the cell structure proteins are broken down (Fig. 20-7). The proteins of the cell nuclei produce uric acid, which accumulate in the heart (*cardiac disease*) and in the articulations (*uric acid arthritis* or *podagra*).

9. Control of Energy Balance

Energy balance is a condition, where the energy input equals energy output, so the energy stores of

the body are unchanged. A person with a body weight of 70 kg contains 550 MJ of combustible energy (entalphy), and if allowed to eat naturally, at least 10 MJ is consumed every day. If the person is fasting for some days he will lose body weight and his metabolic rate will fall to 6.6 MJ daily, so a certain *input control* is hereby documented. The loss in body weight is rapidly compensated when feeding is resumed. If enough food is available the person automatically eat more and more (towards a doubling) with increasing workload (MJ/day), so also a certain *output control* is documented. The internal feedback signals operating in this output and input control are uncertain.

Signals from gastrointestinal centres inhibit the *feeding* or *hunger* centre in the lateral hypothalamic area through afferent nerves (Fig. 20-8). Chyme in the duodenum containing HCl and fatty acids liberate enterogastrones to the blood (ie, intestinal hormones that inhibit gastric activity and emptying). The enterogastrone family consists of secretin, somatostatin, cholecystokinin (CCK) and gastric inhibitory peptide (GIP). Enterogastrones reduce gastric activity, stimulate the *satiety centre* (Fig. 20-8), and increase the production of bicarbonate-rich bile and pancreatic juice. A glucose-rich chyme in the duodenum liberates members of the incretin family to the blood. The incretin family consists of gut glucagon, glucagon-like peptide 1 & 2, and GIP. incretin produce a rapid rise in insulin secretion, which causes the energy stores to increase. The hunger and satiety centres operate reciprocally.

The lipostatic theory explains the constant body weight by liberation of a lipostatic, satiety peptide called *leptin* (ie, thin) from fat tissue. The plasma concentration of leptin is recorded by hypothalamic satiety centres, and seems to reflect the size of the body fat stores or the body fat percentage. Obese patients, often with excessive high plasma leptin concentrations, reduce their leptin concentrations by Banting. Some patients may lack the normal sensitivity to leptin. The leptin molecule is large (16 kDa) and it probably must pass the large fenestrae of the circumventricular organs in order to reach the hypothalamic control centres (Fig. 20-8). The plasma leptin concentration is highest at night.

Also thermoregulatory signals from cold and heat receptors and the plasma concentrations of nutrients may stimulate the satiety centre.

In workers, a minimal work activity threshold must be passed in order to trigger the hypothalamic weight control, but above this threshold eating increases proportional with the workload, and the body weight is constant. If the work rate is extremely high as in marathon training, the hypothalamic control is broken, and the dietary intake and the body weight cannot cope with the high combustion. The body weight falls drastically, which is a certain sign of overtraining.

The cybernetics of appetite control is not only feedback factors. As in all human behaviours, cerebral feedforward factors can dominate. Cerebral feedforward factors are exercise habits, eating habits, *social inheritage*, and they can be of extreme importance to the individual.

Fig. 20-8: A schematic overview of the regulation of food intake. The hypothalamic-feeding centre is located in the lateral region, whereas the satiety centre is medially located in the hypothalamus.

The hypothalamus controls food intake and metabolism, mainly by *autonomic* effects on the islets of Langerhans (secreting insulin, glucagon, pancreatic polypeptide, and gastrin), *hepatocytes* and *adipocytes*. Neuroendocrine-behavioural disturbances seem to be involved in abnormal eating patterns such as *anorexia nervosa, bulimia nervosa* and *obesity*. We seem to regulate our appetite by a *combination* of negative feedback and essential feedforward factors.

Vitamins are *essential organic catalysts* in the diet, necessary for normal metabolic functions in humans, but not synthesized in the human body. *Essential catalysts* refer to the fact that lack of the compound in the diet results in a clearly demonstrable *disorder* in humans.

Vitamins A, D, and K are lipid soluble, so they follow the lipid absorption to the liver, where they are stored. Accordingly, any type of lipid malabsorption results in vitamin deficiency of these vitamins.

The vitamin B complex (B_1, B_2, B_6, B_7, B_{12}, folate) and vitamin C are water-soluble. Since they are only stored in minimal amounts, vitamin deficiency develops rapidly. Exceptional is the *enormous* vitamin B_{12} store in the human liver, so pernicious anaemia takes years to develop.

Pathophysiology

This paragraph deals with *1. Starvation with marasmus, 2. Vitamin deficiencies, 3. Alcohol intoxication, 4. Obesity, and 5. Hyperuricaemia and gout.*

1. Starvation with marasmus

Lack of all elements in the diet of a child - or *universal starvation* - leads to *marasmus*. Marasmus is often complicated by deficiencies in vitamins and essential minerals.

Marasmus is common throughout the third world, because when breast-feeding stops, the child must try to survive on an insufficient diet. The body weight decreases, the fat stores disappear, muscle wasting leads to thin limbs that do not grow in length, and infants and children look like aged persons (Fig. 20-9).

Fig. 20-9: The child has lived through periods of universal starvation alternating with periods on a diet mainly consisting of cassava. The result is a combination of marasmus and kwashiorkor.

The abdomen is tremendously distended, because of *hepatomegaly*, flaccid abdominal muscles and possibly oedemal fluid in the abdominal cavity (*ascites*). The short half-life of the intestinal mucosal cells makes them especially sensitive to lack of nutrients, so villous atrophy develops with malabsorption and diarrhoea.

Marasmus is unexpected in the rich part of the world. However, this is not so. Malignant tumours and severe cardiopulmonary disease imply an enormous loss of energy, and terminal weight loss is unavoidable even where the diet is supposed to be sufficient (hospitals and other institutions).

The basal metabolic rate is low and the core temperature is also controlled on a lower than normal level. The heart rate and arterial blood pressure is also low. Haemopoiesis is deficient and anaemia prevails. The immune defence system is impaired and the patient suffers from numerous infections.

Children suffer from growth failure, and brain development is probably affected.

Children fed with a diet deficient in protein alone (essential amino acids) develop *protein deficiency* or *kwashiorkor*. Following breast-feeding, these children subsist on a protein-poor diet rich in carbohydrates (eg, *cassava*). Thin limbs, hypoproteinaemia and ascites (Fig. 20-9) characterise kwashiorkor. The large liver is fatty, because there is carbohydrates enough to provide the hepatocytes with lipids, but the lack of protein makes the production of lipid transporting proteins (*apoproteins*) inadequate.

In spite of the effort from all international institutions concerned, the fraction of the global population falling below the minimum food intake defined by WHO, is increasing, and has done so for years.

2. Vitamin deficiencies

Vitamin A (retinol) and its analogues are termed *retinoids*. Vitamin A occurs naturally as retinoids

or as a precursor, b-carotene, in vegetables. Infants fed with cooked milk in developing countries, and adults suffering from chronic disorders with fat malabsorption may develop vitamin A deficiency. Retinoids have the following three effects:

1. Retinoids are an important constituent of the photosensitive pigment of the retina and enhance night vision. The 11-cis-retinal is the aldehyde of vitamin A_1. Retinal combines with the glycoprotein *opsin* to form *rhodopsin* (visual purple) in the retina during darkness.

   Vitamin A deficiency implies a massive fall in the number of rhodopsin molecules in the outer segment of the rods. This impedes dark adaptation and *night blindness* occurs.

2. Retinoids stimulate cellular growth and differentiation. Retinoids convert keratin-producing cells into mucus-producing cells, transcribe new mRNA and encode for new cell proteins, so a more differentiated cell type develops.

3. Vitamin A promotes growth of the skeleton.

Lack of vitamin A causes diminished vision in dim light, followed by *night blindness*, and eventually *blindness*. Lack of vitamin A leads to *squamous metaplasia* of the conjunctiva and glandular epithelium. The tear ducts are occluded by the metaplasia, causing eye dryness (ie, *xerophthalmia*), and occluded sebaceous glands cause follicular hyperkeratosis. Vitamin A deficiency causes *growth retardation.*

Retinoids are used in cystic acne and in psoriasis. Vitamin A in normal dosage has been proposed as an anticarcinogen. Excess intake may be *teratogenic*.

Thiamine - as the majority of the B complex - is found in green vegetables, milk and liver.

Thiamine is the co-factor for many enzymes in the glycolytic pathway. Thus lack of thiamine leads to inadequate glucose metabolism with accumulation of vasodilatating lactate and pyruvate. The peripheral vasodilatation leads to oedema. The increased work of the heart eventually develops into *cardiac failure,* which increases the venous stasis and worsens the oedematous state.

Thiamine deficiency (beri-beri) is found in persons consuming polished rice (classical beri-beri), in chronic alcoholics, and in marasmus (see above).

Dry beri-beri is symmetrical polyneuropathy (ie, paresthesia, weakness, heaviness, and paresis of the legs). CNS involvement with ischaemic damage results in the **Wernicke-***Korsakoff syndrome* (ie, ataxia, confusion and ophtalmoplegia).

Wet beri-beri describes thiamine deficiency with oedemas of the legs, pleural effusions and ascites. - These disorders respond immediately to thiamine treatment.

Vitamin B_2 (riboflavin) deficiency

Riboflavin is widely distributed in animal and vegetable foods. Riboflavin is destroyed by ultraviolet light but thermostable and not destroyed by cooking. In the human body riboflavin is converted into flavin mono- and di-nucleotides. These compounds are of crucial importance in the electron transport chain.

Riboflavin deficiency is frequently only part of a *combined* vitamin B deficiency, but the classical manifestations are lesions around the natural openings: *1. interstitial keratitis* of the cornea with vascularisation, *2. seborrhoic dermatitis* (face, vulva, and scrotum), *3. angular stomatitis* (ie, *cheilosis* or fissures at the angles of the mouth), and *4. glossitis.*

These lesions respond to riboflavin usually given parenterally as a *vitamin B complex*.

Vitamin B_6 deficiency

Vitamin B_6 activity is found in three compounds found in both vegetable and animal foods: pyrodoxine, pyridoxal, and pyridoxamine. Pyridoxal phosphate is a co-enzyme for transaminases, carboxylases (formation of the neurotransmitter GABA) and other enzymes. Drugs like the antituberculosis drug, isoniazid, and the copper-chelating agent, penicillamine, are B_6-antagonists.

Certain types of polyneuropathy including the CNS and anaemia with saturated iron-stores respond to vitamin B_6 therapy.

Vitamin B_7 deficiency

Niacin deficiency or *pellagra* (ie, roughs skin) is recognized by the combination of the 3 diagnostic Ds: Dermatitis, diarrhoea, and dementia.

Light exposed areas exhibit dermatitis with rough scales. The diarrhoea is colonic and watery. Cortical atrophia and degeneration of myelinated tracts in the spinal cord cause the dementia.

Niacin is involved in the formation of nicotinamide adenine dinucleotide (NAD) and its phosphate (NADP). These molecules are important in many oxidation/reduction reactions of the intermediary metabolism.

We consume niacin found in different types of grains (poor content in maize), and our endogenous synthesis is from tryptophan found in meat, eggs and milk.

Pellagra is seen in malnourished alcoholics, food faddists, and in patients with the *carcinoid syndrome*, where most of the tryptophan is used for serotonin synthesis.

Vitamin B_{12} deficiency

Vitamin B_{12} is almost ubiquitary in animal foods (meat, fish, eggs and milk) but not in vegetables, so dietary deficiency is only found in extremely rare cases of vegetarianism, starvation or anorexia nervosa. Malabsorption disorders (ie, pancreatitis, coeliac disease) seldomly result in biologically consequential B_{12} deficiency.

Vitamin B_{12} deficiency causes *pernicious anaemia* (Chapter 8). Below is described the absorption of the vitamin in the terminal ileum (Fig. 20-10).

Fig. 20-10: The mechanism of normal vitamin B_{12} (cobalamin) absorption in the terminal ileum and storage in the liver.

The *intrinsic factor-cobalamin complex* is resistant to pancreatic proteases, and is normally carried along the gastrointestinal tract to specific *receptor proteins* on the mucosal surface of the terminal ileum. The complex is recognized and bound to the receptor. The free vitamin B_{12} enters the enterocyte, and the intrinsic factor remains in the lumen. Vitamin B_{12} exits from the enterocyte by facilitated or active transport, and appears in the portal blood bound to the glycoprotein, *transcobalamin II* (Fig. 20-10). The hepatocytes clear the portal blood for vitamin B_{12} by receptor-mediated endocytosis. The hepatic vitamin B_{12} store is enormous in healthy individuals. An average value of 5 mg stored vitamin B_{12} must be compared to a daily requirement of 1 mg.

Transcobalamin II is the main carrier in delivering vitamin B_{12} to the red bone marrow, although most of the vitamin B_{12} is bound to transcobalamin I and III.

Folic acid deficiency

Foliates are present in *leafy green vegetables* such as spinach and broccoli, and in organs such as kidney and liver. Excessive cooking destroys much of the food folate. Pregnancy increases the requirement for folate up to tenfold. Folate is absorbed in the small intestine, and transported to the cells via the blood plasma.

Folate deficiency with poor diet for a few months' results in megaloblastic anaemia and glossitis because the stores of folate are small compared to the enormous liver storage of vitamin B_{12} (Chapter 8).

Vitamin C deficiency (scurvy)

Vitamin C or ascorbic acid is a reducing substance found in fresh fruit and vegetables. Humans cannot synthesise ascorbic acid from glucose as several animals. Ascorbic acid contributes in controlling the redox potential of the cells.

Ascorbic acid is necessary the hydroxylation of *proline* to *hydroxyproline*. This is the single process necessary for the production of *collagen* in all tissues including the vessel walls.

Vitamin C deficiency (scurvy or scorbutus) is found among food faddists and in developing countries, where infants are fed with excessively boiled milk.

The patient with scurvy can only produce *abnormal collagen* without sufficient tensile strength. The capillaries become fragile and bleedings are frequent. They are recognized as bruises of the skin, as haemarthron, as subperiosteal bleedings and eventually bleeding anaemia develops (Chapter 8). Infections are prolonged and the healing of wounds is poor. Infections of the gingiva (gingivitis) leads to loose teeth, and the lack of normal collagen in growing bones results in arrested bone growth.

Bottle fed infants must receive daily fruit juice, and for poorly fed adults fresh fruits and vegetables are the best preventive means of avoiding scurvy.

There is no advantage in the daily intake of *large doses* of vitamin C to prevent or improve common cold or cancer. In one controlled clinical trial there was an accumulation of cases with kidney stones. *Rebound scurvy* may occur following a sudden stop of the intake of large doses of vitamin C.

Vitamin D deficiency is described in Chapter 30.

Hypervitaminosis D is caused by excess consumption of vitamin preparations. This leads to hypercalcaemia, nephrolithiasis, nephrocalcinosis and *ectopic calcification* of other organs including premature arteriosclerosis.

Vitamin K deficiency

Vitamin K occurs in two forms in nature. Vitamin K_1 is produced in plants, and intestinal bacteria in animals synthesise vitamin K_2.

Insufficient dietary intake of vitamin K is infrequent, and occurs occasionally in the chronically ill patient such as cases of *anorexia nervosa*.

Fat malabsorption is accompanied by *vitamin K deficiency*, because vitamin K is fat-soluble.

Newborn babies sometimes suffer from vitamin K deficiency, because the molecule only crosses the placental barrier with difficulty, and because the sterile gut of the baby cannot produce vitamin K_2.

Destruction of the intestinal bacteria by long term antibiotic treatment may also lead to vitamin K deficiency.

Vitamin K deficiency can lead to *terminal bleeding*. This is because vitamin K normally activates four clotting factors: prothrombin, factor VII, factor IX, and factor X. These four proteins probably receive Ca^{2+} binding properties from vitamin K (see Chapter 8).

Therapy with intramuscularly administered vitamin K is rapid and effective.

Vitamin E deficiency

Vitamin E (a-tochoferol) is found in fish, fish oil and vegetable oil from Soya beans and corn. Vitamin E is an *antioxidant*. Vitamin E protects the phospholipids of the plasma membrane against peroxidation by free radicals produced by the cell metabolism.

Prolonged vitamin E deficiency is rare, but leads to CNS lesions, haemolytic anaemia, and muscle disorders. Patients with fat malabsorption or patients receiving parenteral nutrition may develop vitamin E deficiency.

3. Alcohol intoxication

The sequence of events in *acute alcohol intoxication* proceeds with an increasing sense of warmth, flushing of the face, dilated pupils, dizziness and euphoria. There is a general sense of well being with unjustified optimism and the feeling of increased strength and energy. The subject shows a boisterous behaviour with increased psychomotor activity, which is clumsy, and social inhibitions are dissolved.

Negative consequences of alcohol abuse are arrests, automobile accidents, and deleterious effects upon job performance and chronic health problems.

Alcohol interferes with the arrangement of molecules (ion channels, receptors, the GABA-benzodiazepine-channel etc.) in the lipid bilayers of the cell membrane. With increasing intoxication the symptoms and signs of CNS depression become apparent. The subject becomes drowsy, argumentative, angry or weepy, and eventually he is vomiting and complaining of diplopia. Later an examination reveals areflexia, loss of muscular tension, loss of sphincter control, rapid heart rate and respiratory frequency, decreasing arterial pressure and mean arterial pressure leading to shock. The subject develops hypothermia and increasing stupor, anaesthesia, coma or death.

The intoxication depresses the myocardium and dilates the peripheral vessels. This is why the MAP is falling together with cardiac performance.

Some alcoholics benefit from treatment with disulfiram (Antabuse) or similar drugs. Antabuse inhibits aldehyde dehydrogenase, which results in poisoning from accumulated acetaldehyde.

4. Overeating with obesity

Overeating is related to social patterns and constitutional family traits.

Obesity or *adiposity* implies the excess storage of fat, and is defined by WHO as an *actual body weight* exceeding the *ideal weight* by more than *20%* (if not explained by an above-average muscle and bone mass). The diagnosis is frequently set by inspection of the undressed patient (Fig. 20-11). The fatty stores of the patient in Fig. 20-11 are clinically acceptable. The *ideal weight* is the weight

associated with the highest statistical life expectancy. The *Broca index* is a popular and easy method of determining the recommended weight. Brocas index is the predicted body weight in kg, which equals the height of the person in cm minus 100 for males and 110 for females.

Fig. 20-11: The ideal weight and clinically acceptable fatty stores.

Obesity is also established in another way by the help of the *body mass index (BMI).* BMI is the weight of the person in kg divided by the height (in m) squared. The normal range is 19-25 kg per square metre (Fig. 20-11). Marginal overweight is defined as a BMI between 25 and 30 kg* m^{-2}. Obesity is defined as BMI above 30 kg* m^{-2}, which corresponds to body weights 20% above ideal weight.

Obesity results from a long-term excess of nutritional intake relative to the energy liberation. There are at least three types of obesity: genetic, overeating and inactivity induced.

1. *Genetic obesity*. Genes account for quite some cases of obesity. Genes seem to be causative in 2/3 of all cases of obesity in a lifetime. Genetic movement oeconomists may explain many cases of obesity including familiar obesity, but weight gain does not occur in all pairs of mono- or di-zygotic twins. - Hyperplastic fatness (too many adipose cells) is often found in babies from the rich part of the world, whereas many adults have hypertrophic obesity, which is caused by too large adipose cells.

2. *Overeating.* Intake of poor food, dominated by sweet-fat combinations, explains other cases of obesity. Sugar and fat eaters need not eat very much in order to develop obesity, if they live with a marginal motility pattern. In any type of obesity the low physical activity or inactive life style is typical.

3. *Inactivity.* The major factor in obesity is *physical inactivity*. Persons who exercise can increase their metabolic rate by a factor of 10-20 several hours a day. The second choice of obese persons is to reduce the dietary intake of nutrients. A reduction to half the usual amount of food would be a short, heroic and probably futile project, as well as inefficient, when compared to a metabolic factor of 10-20 during exercise. *Inactivity,* defined as a *low fitness number* (ie, a maximal oxygen uptake below 34 ml * min^{-1}*kg^{-1}), is probably involved in western life-style obesity. Obesity is the fate of people dominated by their parasympathetic activities and minimising the use of the sympathetic nervous system. The mortality of a male population increases dramatically with falling fitness (Fig. 18-13).

Obese people have a small dietary thermogenesis, because they avoid physical activity with a high metabolic rate in skeletal muscle and in adipocytes. Decreased sensitivity to leptin has been described in obese patients.

Rare cases of obesity are caused by hormonal, metabolic diseases (insulinoms, hypercorticism, diseases of the thyroid gland, hypothalamic lesions etc). For persons with insulinoms the *high food intake* (hyperphagia) can be a question of life and death.

Obesity and interrelated risk factors

Obesity is clearly a *risk factor*. A risk factor is an epidemiological term for conditions statistically correlated with shortened life expectancy. Obesity paves the way for maturity-onset (type II) diabetes, atherosclerosis, hypertension, myocardial infarction, and stroke.

Type II diabetes is linked to obesity, perhaps because the increased weight or the rise in blood glucose concentration stimulates insulin secretion. A period with high concentration of insulin in the

blood plasma decreases the number of insulin receptors on the membranes of muscle and adipose cells (insulin resistance or glucose intolerance). More insulin must be produced from the beta cells of the pancreatic islets, and finally the b-cells are exhausted and a diabetic condition without insulin production developed. Body weight reduction ameliorates the glucose intolerance.

Atherosclerosis, acute myocardial infarction, hypertension, hypercholesterol-aemia, gall-stones, low concentrations of HDL, hyperuricaemia, gout, osteoarthritis of the hip and knee joint, and restrictive lung disease are all related to obesity (Fig. 20-12).

Fig. 20-12: Adipose patient with numerous complications.

There is also an increased incidence of depression, psychological and social problems, intertriginous dermatitis, hernias, impotence, and thrombophlebitis in obese patients.

Amenorrhoea and oligomenorrhoea, reduced fertility is common among premenopausal obese females.

Therapy of established obesity is a frustraneous and highly resistant task. The current therapy of adiposity, including drugs, diet and behavorial modification with exercise, is ineffective - often due to the lack of motivation.

1. Anti-obesity drugs are either centrally or peripherally active.

The *centrally* active drugs either act on catecholamine neurotransmitters (amphetamines), or they act on serotoninergic neurons in the CNS. Initially, all these drugs reduce food intake, and some of them also increase the metabolic rate.

Amphetamines, “holiday pills”, were the first centrally active anti-obesity drugs developed, but there abuse potential is a definite contraindication.

The *peripherally* active anti-obesity drugs, such as *acarbiose*, have only modest effect in controlled clinical trials. Acarbiose is an amylase inhibitor, which reduce the digestion of sucrose. The lipase inhibitor, *tetra-hydro-lipostatin*, blocks the intestinal digestion of lipids, but has only a marginal effect on obesity. From a physiological point of view there is little perspective in new development of anti-obese drugs, because any reduction in nutritional input - even a 10% reduction - is futile.

2. Diet. Most popular, weekly journals publish up to three miraculous diets for obese persons in each issue. This is a contradiction with only marginal possibilities of success. Even a 5-10% reduction in dietary input is experienced as self-torture and tolerated only for a short time. There is no alternative to a healthy mixed diet with a sufficient amount of dietary fibre (see above).

3. Behavioural therapy with exercise. The single factor that can cure obesity is a *balanced degree* of exercise. Exercise can increase metabolic rate from typically 70 Watts at rest to 700 (eg, walking, dancing, sexual intercourse) or 1400 Watts (long distance running). Thus, the most important determinator (factor 10-20) is the increase in metabolic rate from any type of self-induced locomotion.

Running is an alternative for younger and middle-aged persons. Older people must walk, if they have the capacity for it. They may prefer walking in a hilly environment in a relaxed way, so the heart is stimulated. Callisthenics, dancing, skiing, swimming, tennis, cycling, golf are alternatives for persons motivated. The important point is to chose the type of exercise, which is enjoyable - in a relaxed way - for the individual concerned.

Increased awareness of nutritional and fitness intervention in improving the health status of large population groups is essential.

5. Hyperuricaemia and gout

Excessive *production* or inefficient *excretion* of uric acid causes hyperuricaemia.

Hyperuricaemia is a condition with an abnormally high concentration of uric acid in the blood plasma and ECF (above 0.42 mM). Normal values of serum-urate are 0.2-0.42 mM. The saturation threshold over which urate crystals precipitate is around 0.42 mM for tissues with an acid pH.

Hyperuricaemia is asymptomatic for varying periods. When the hyperuricaemia becomes clinically important through recurrent attacks of painful acute arthritis, the condition is called *uric arthritis* or *gout* (Fig. 20-13).

Fig. 20-13: Gout and its complications are shown to the right. Blockage of the uric acid synthesis by allopurinol is shown to the left.

Gout can be primary or secondary.

1. Primary gout. Either increased metabolic urate production, inefficient renal excretion of urate or the two in combination causes genetic or idiopathic gout. Primary metabolic gout relates to two inherited, X-linked enzyme disorders: Hypoxanthine-Guanine-Phospho-Ribosyl-Transferase (HGPRT) deficiency with increased purine synthesis or an abnormally high activity of Phospho-Ribosyl-Pyro-Phosphate -Synthetase (PRPPS). The inherited disorder is exacerbated by diets high in purine or nucleic acids. During purine degradation large quantities of NH_4^+ are liberated. The acidosis leads to crystallisation of urate.

2. Secondary or acquired gout can also be both metabolic and renal. Intercurrent disease with lysis of cell nuclei and release of nucleic acids increases urate production (eg cancer, psoriasis, and excessive weight loss). Impaired renal excretion leads to secondary renal gout.

Most forms of metabolic gout are a result of overproduction of uric acid caused by accelerated purine synthesis from amino acids, formate and CO_2, whereas dietary purines play a minor role. *Xanthine oxidase* oxidises hypoxanthine to xanthine and xanthine to uric acid.

Supersaturated body fluids precipitate thin urate crystals in acid environments. The result is an inflammatory reaction, where leucocytes migrate to the crystals and surround them for phagocytosis.

The acute attack of gout typically occurs in a male with severe pain in the big toe. The pain attack is also called *podagra.* The pain responds to therapy and the patient is asymptomatic for a variable period. Following a series of acute attacks of gout, the pain is persistent, because the urate crystals are permanently present in the joints and other tissues. This is called *chronic gout.*

Toes, ankles and knees are frequently affected. Symptoms and signs of gout include hyperuricaemia, tophi and painful arthritis. Symptoms and signs of gout include hyperuricaemia, tophi and painful arthritis, with extremely tender and swollen joints.

Complications to gout are increased risk of atherosclerosis, hypertension and renal disease including renal calcification and uric acid stones in the ureter.

Persons with hyperuricaemia are at risk, and should be treated with allupurinol (100-300 mg daily) until the plasma-urate is brought down to normal levels (see effect below).

Allupurinol is a *xanthine oxidase inhibitor.* Allupurinol is an analogue to hypoxanthine, but xanthine oxidase (XO) prefers allupurinol as a substrate, so allupurinol is oxidised to oxypurinol (Fig. 20-13). Oxypurinol (alloxanthine) blocks XO, because it binds to the active site on the enzyme. Thus, urate

production is inhibited and xanthine/hypoxanthine is accumulated in the blood and ECF. This is fortunate, because these substances are water-soluble and easily excreted in the urine - just as alloxanthine. This is in sharp contrast to the less soluble urate. Uric acid is filtered in the renal glomeruli. Urate is reabsorbed in the proximal tubules by a Na^+ -substrate cotransport with a capacity, which is normally far greater than the amount of urate in the glomerular filtrate. Accordingly, the normal urate secretion takes place by active secretion of urate ions in the distal tubules. The *organic acid-base secretory system* transfer urate ions from the blood to the tubular fluid, but the system has a low capacity for urate.

Patients with *renal gout* suffer from abnormally low distal tubular secretion of uric acid. These patients are treated with uricosuric agents such as *probenecid* and sulfinpyrazone. These molecules compete for the proximal Na^+ -substrate cotransport, so less urate is reabsorbed.

Colchicine is a drug that binds to *tubulin*, a protein in the microtubules of the leucocytes. Hereby, the microtubules disintegrate, which inactivates the leucocytes. Thus, the colchicine prevents the focal infiltration of leucocytes to the damaged tissue and blocks their usual liberation of lactic acid which would further precipitate urate crystals. This effect is slow and not purely beneficial.

Weight loss is often indicated during treatment of gout, but a rapid weight loss is risky in hyperuraemic patients, because the cellular destruction liberates nucleic acids and may elicit an acute attack of gout.

Equations

- **The first law of thermodynamics** states that the sum of liberated heat energy (-*Q*) and liberated work (-*W*') of a system is equal to the fall in internal energy (enthalpy) or heat content (*H*). The decrease in enthalpy of the human body (-D*H*) is equal to the fall in potential, chemical energy stored in the body:

  Eq. 20-1: $(-DH) = (-Q) + (-W')$.

- **Entropy** is the tendency of atoms, molecules and their energies to spread in a maximum space. The *Gibbs energy* (*G*) is the difference between enthalpy (*H*) and entropy (*S*) when multiplied with the absolute temperature (T):

  Eq. 20-2: $G = H - T \times S$.

G determines if a certain reaction occurs, since *G* is minimal at equilibrium. According to the formula, entropy is important at high temperatures, and energy is most important at low temperatures.

- **The Fick cardiac output equation:**

  Eq. 20-3: $Q^\circ = V^\circ_{O2/(} C_{aO2} - Cv^-_{O2)}$

- **Elimination of alcohol by oxidation**. The rate of oxidation is constant (b = 0.0025 permille per min) and is independent of the blood [alcohol]. The absolute amount of alcohol eliminated per minute is thus:

  Eq. 20-4: Alcohol oxidation (g/min) = $(b \times r \times$ body weight$)$

  The fraction of the body weight which is distribution pool for alcohol is called **r** (mean-**r** for females is 0.55 and for males 0.68 kg per kg body weight).

- **Calculation of work rate** on a bicycle ergometer: A measurable blocking force is applied to a wheel with a given radius (r) and with a given rotation-frequency (RPM). The *work rate* or power (force × velocity) is now determined, because the force is known (N) and the distance per s is (2 × p × r RPM)/60. The work rate is thus measured in J/s or Watts. Work rate = Force * Distance, or

  Eq. 20-5: Work rate (Watts) = N* (2*p*r *RPM)/60.

- **Calculation of metabolism** by *indirect calorimetry*: The oxygen uptake and carbon dioxide output is masured volumetrically or gravimetrically together with determination of the nitrogen content in 24 hours urine from the person examined. The urine nitrogen expresses the protein combustion, since protein contains 16% nitrogen. Subtraction of the gas volumes for protein combustion (see data in Symbols) from the total, result in residual volumes of oxygen and carbon dioxide only related to the fat (F g/min) and carbohydrate (C g/min) combustion. Thus F and C can be calculated by solution of two equations with these two unknowns (Eq. 20-6 and -7). By multiplication with the nutritive equivalents for O_2 and for CO_2 (mmol gas per g in Symbols) the mass balance states:

  Eq. 20-6: F and C related O_2 uptake = (37 mmol/g × C g/min) + (91 mmol/g × F g/min)

  Eq. 20-7: F and C related CO_2 liberation = (37 mmol/g × C g/min) + (64 mmol/g × F g/min).

  Now the mass of protein, fat and carbohydrate combusted per min is found, and a very precise indirect measure of MR in kJ/min is obtained by multiplication with their energy equivalents (Se Symbols).

- **Body surface area (*BSA*)** is estimated with the approximation formula of the DuBois family:

  Eq. 20-8: $BSA\ (cm^2) = WEIGHT^{0.425}\ (kg) \times HEIGHT^{0.725}\ (cm) \times 71.84$

The *BMR* is normally 45 Watts/m^2, so an adult with a body surface area of 1.8 m^2 has a *BMR* of 80 Watts. This corresponds to a daily *BMR* of ($60 \times 1440\ min \times 80 \times 10^{-3}$) =6912 kJ.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Vitamins are essential organic catalysts synthesized in the human body.

B. Antibiotics in meats or milk may induce allergy or antibiotic resistance in humans.

C. Obesity is the consequence of inactivity alone, and genes are not involved.

D. Either increased metabolic urate production, inefficient renal excretion of urate or the combination causes primary or idiopathic gout.

E. During therapy with large doses of ascorbic acid, withdrawal may cause symptoms of scurvy.

II. Each of the following five statements have True/False options:

A. Prolonged use of some anti-obesity drugs may imply serious abuse.

B. Osteomalacia does not involve the organic bone matrix.

C. Vitamin A deficiency implies a massive fall in the number of rhodopsin molecules in the outer segments of the rods. This impedes dark adaptation and night blindness occurs.

D. The epiphyseal plate of the growing skeleton is sufficiently mineralised in rachitis.

E. Vitamin K can easily cross the placental barrier.

Case History A

At 7 p.m. a male alcoholic drinks 150 ml of whisky (40 w/v%). His body weight is 58 kg (due to hepatic failure and malnutrition). The fraction of his body weight, which is distribution pool for alcohol, is 0.60. The rate of oxidation of alcohol is reduced to 80% of normal. The rate of oxidation in a healthy person is 0.0025 o/oo per min. At 8 p.m. the patient is involved in a traffic accident and at 9 p.m. his blood [alcohol] is measured to 1.36 g kg^{-1} (o/oo). The patient states to the police that he has been drinking only the whisky at 7 pm.

1. *Calculate the blood [alcohol] at the time of the traffic accident.*

2. *Is the statement concerning alcohol intake correct?*

Case History B

Following an earthquake, three adult females are confined under the ruins of their house in an airtight space of 8 m^3. The initial pressure is 752 mmHg, the temperature of the water-saturated air is 20 ^oC (water vapour tension 20 mmHg) and the composition of the atmospheric air is normal. The average oxygen uptake is 190 ml STPD min^{-1}, and the average RQ is 0.83. Assume that P_{IO2} = 50 mmHg is the survival threshold.

1 *Calculate the time period, where they have oxygen enough for survival, provided carbon dioxide could disappear?*

2. *Calculate the carbon dioxide output per min.*

3 *Calculate the theoretical amount of CO_2, which should accumulate in the time period for survival from 1. Calculate the theoretical CO_2 fraction in the airtight space.*

4. *Explain the consequences of CO_2 accumulation.*

Case History C

A male, 23 years of age, has a daily metabolic rate of 12 600 kJ (12.6 MJ) and he is eating a mixed diet resulting in a RQ of 0.83. The enthalpy equivalent for oxygen is 0.46 kJ/mmol. His arterial pH is 7.30 and pK for ammonia is 9.3.

1. *Calculate the ratio between ammonia and NH_4^+ in his blood.*

2. *Calculate his daily carbon dioxide output in mol.*

3. *Calculate the amount of carbon dioxide eliminated per day in combination with ammonia, assuming one mol/day of ammonia to be involved in the urea production.*

Case History D

A 70 kg male, 22 years of age, is in a room, where the temperature is 20^{o} C and P_B are 101.3 kPa. The P_{AO_2} is 14.1 kPa, and P_{AN_2} is 78 kPa. The carbon dioxide output is 660 ml STPD per min. The urinary nitrogen excretion is 10 mg per min. The pressure of water vapour in the alveoli is 6.2 kPa, and F_{IO_2} is 0.2093.

1. *Calculate the alveolar ventilation and F_{ACO_2}. Estimate P_{aCO_2}, and provide reasoning for a possible hyperventilation in this condition.*
2. *Calculate the oxygen uptake for this person.*
3. *Is the pulmonary exchange quotient different from the respiratory quotient (RQ)?*

Try to solve the problems before looking up the answers.

Highlights

- *The first law of thermodynamics states that energy can neither be created nor destroyed but is only transferred from one form to another or from one place to another.*
- *Heat energy is low prize energy. In contrast to ATP energy, it is not available for work in the body. The sum of heat energy generated and work performed is constant and equal to the Gibbs energy.*
- *The oxidation of fuel (carbohydrates, glycerol, fatty acids) to CO_2 and water is the primary pathway for generation of energy and subsequent heat energy liberation. Protein can also serve as an important energy source during prolonged exercise.*
- *Alcohol distributes in the total water phase of the body (60% of body weight) within one hour.*
- *Hepatic alcohol dehydrogenase, catalase and MEOS (Microsomal Ethanol Oxidation System) can oxidise alcohol.*
- *The most important elimination of alcohol is by oxidation. The rate of oxidation is constant (β = 0.0025 permille per min) and is independent of the blood alcohol concentration.*
- *The blood alcohol concentration is measured in the unit g per kg (permille) of distribution volume. The absolute amount of alcohol eliminated per minute is thus: ($\beta \times r \times$ body weight).*
- *An adult person can oxidise 7 g of alcohol per hour, and at rest only negligible amounts are excreted in sweat, urine and expired air.*
- *During severe exercise by an athlete working in a hot climate, the total excretion of alcohol can be larger than the maximal oxidation (7 g each hour).*
- *The net mechanical efficiency (E_{net}) is the ratio of external work rate (N × m/s = J/s) to net chemical energy expenditure (J/s or Watts) during work. E_{net} is 20-25% in isolated muscles and also in humans during aerobic cycling.*
- *The total production by use of the glycero-phosphate shuttle in oxidative phosphorylation is 36 ATP per glucose molecule (6 from the glycolysis, 6 from the transformation and 24 from the TCA cycle).*

- *The total output of heat energy from the body is most precisely measured in a whole-body calorimeter. The Atwater-Rosa-Benedict's human calorimeter has been used to verify the first law of thermodynamics in humans.*

- *The metabolic rate can increase by a factor of 10-20 during steady state exercise.*

- *Hepatic intermediary processes cause the specific dynamic activity (dietary thermogenesis) of proteins: Breakdown of amino acids, formation of urea etc. The dietary thermogenesis in general is related to mass action and temperature increase when eating.*

- *Serious illnesses develop after a few weeks of fasting, because the cell structure proteins are broken down. The proteins of the cell nuclei produce uric acid, which accumulate in the heart (cardiac disease) and in the articulations (uric acid arthritis or podagra).*

- *The lipostatic theory explains the constant body weight by liberation of a lipostatic, satiety peptide called leptin (ie, thin) from fat tissue. The plasma concentration of leptin is recorded by hypothalamic control centres, and seems to reflect the body fat percentage. Obese patients reduce their plasma leptin concentration by Banting and may lack the normal sensitivity to leptin.*

- *The help of the body mass index (BMI) establishes obesity. BMI is the weight of the person in kg divided by the height (in m) squared. The normal range is 19-25 kg* m^{-2}. Marginal overweight is defined as a BMI between 25 and 30 kg* m^{-2}. Obesity is defined as BMI above 30 kg* m^{-2}, which corresponds to body weights 20% above ideal weight.*

- *Obesity results from a long-term excess of nutritional intake relative to the energy output. There are at least three types of obesity: genetic, over-eating and inactivity induced obesity.*

- *Atherosclerosis, acute myocardial infarction, diabetes, hypertension, hypercholesterol-aemia, gall-stones, low concentrations of HDL, hyperuricaemia, gout, osteoarthritis of the hip and knee joint, and restrictive lung disease are all related to obesity.*

Further Reading

The Journal of Nutrition. Monthly journal published by the Am. Institute of Nutrition, 9650 Rockville Pike, Bethesda, MD 20814-3990, USA.

Silverstone, T (1992) *Appetite suppressants: A review*. DRUGS 43: 820-836.

Maffei, M et al. “Leptin levels in humans and rodents: Measurement of plasma leptin and of RNA in obese and weight-reduced subjects.” *Nature Med* 11: 1155-61, 1995.

Katzung BG. *Basic & Clinical Pharmacology*. Appleton & Lange, Stanford, Connecticut, 1998.

Return to chapter 20

Return to content

Chapter 21

Thermo-Regulation, Temperature And Radiation

This Chapter is written following discussions with my colleague, Leif Vanggaard, MD, Arctic Institute, Copenhagen.

Study ObjectivesF

- To *define* body core and body shell, heat balance, heat exchange (conduction, conversion, evaporation and radiation), hyperthermia, hypothermia, mean body temperature, heat capacity, and thermal steady state.
- To *describe* fever (pyrogens), benignant and malignant hyperthermia, heat exhaustion, heat syncope, heat stroke, sun stroke, and hypothermia.
- To *describe* radiation sickness.
- To *calculate* one thermal variable, when relevant variables are given.
- To *explain* the concepts heat exchange, thermogenesis by food and shivering, the human temperature control system and its function at different environmental temperatures.
- To *use* the above concepts in problem solving and case histories.

Principles

- ***Newton's law of cooling****: The dry heat loss is proportional to the temperature difference between the human body (shell) and the surroundings.*
- ***The total energy of a system*** *is conserved in an interaction, not the kinetic energy or the mass (Einstein). If the mass changes during an interaction, there is a resultant change in kinetic energy, so that the total energy remains constant. – Heat energy is proportional to molecular movement rates – "heat energy equals movement."*
- ***Stefan-Boltzmanns rule****: The higher the temperature of an object, the more it radiates. The energy radiated from an object is proportional to the fourth power of its Kelvin temperature. – The energy radiating from an object and received by the human body is proportional to the temperature difference between the object and the skin (see Eq. 21-4). This is because human life implies relatively small temperature gradients.*

Definitions

- **Body core** consists of the thermoregulated deeper parts of the body and the proximal extremity portions of warm-blooded animals including man.
- **Body shell** refers to those outer parts of the body (skin and subcutaneous tissue) that change temperature at cold exposure.
- **Conductance** changes of the shell are used as a measure of skin bloodflow.

- **Conductive heat loss** describes a direct transfer of heat energy by contact between two bodies of different temperature (eg, skin and objects).
- **Convective heat loss** is defined as the heat loss by contact between the surface (skin) and a moving medium (air or water).
- **Evaporative heat loss** is defined as the heat loss by evaporation from the body surface or lungs.
- **Fever** occurs when the *core temperature of the body* is raised above normal steady state levels. The body reacts as if it is too cold. Fever implies a disorder resulting in shivering combined with vasoconstriction, headache, dedolation, and general discomfort (eg, malaria).
- **Heat flow** is defined as energy exchanged due to a temperature difference. Heat flow is transmitted along a temperature gradient.
- **Heat capacity** is the amount of heat required to produce a temperature increase for a given amount of substance.
- **Heat energy balance** in a resting person is a condition, where the heat production is equal to the heat loss. Thus the body temperature is constant and the heat storage is zero (*thermal steady state*). Usually, there is no internal heat energy flux between body core and shell.
- **Hyperthermia** is an increase in core temperature above normal.
- **Hypothermia** refers to a clinical condition with a lowered core temperature (below 35 ^{o}C).
- **Mean body temperature** is defined according to Eq. 21-1 (see end of Chapter).
- **Non-shivering thermogenesis** is a rise in metabolism, which is not related to muscular activity (shivering or exercise).
- **Insensible perspiration** (leakage of the skin) is the small cutaneous evaporation loss, which is unrelated to sweat gland function.
- **Insulation** refers to resistance to heat transfer.
- **Radiative heat loss** is a transfer of heat energy between 2 separate objects at different temperature. Heat energy is transferred via electromagnetic waves (photons). This heat transfer does not require a medium, and the temperature of any intervening medium is immaterial.
- **Shell temperature** is the temperature of the outer parts of the body (measured on the skin surface) and related to cold environments.
- **Shivering** is a reflex myogenic response to cold with asynchronous or balanced muscle contractions performing no external work.
- **Specific heat capacity** is the relationship between heat energy exchanged per weight unit of a substance and the corresponding temperature change. The *specific heat capacity* of ***water*** is 4.18 and of the *human body* (blood and tissues) 3.49 kJ kg^{-1} ^{o}C^{-1}, respectively. The specific heat capacity of atmospheric air is 1.3 kJ (m^{3})$^{-1}$ ^{o}C^{-1}.

- **Temperature** is the measurement of heat energy content.

Essentials

This paragraph deals with

1. The temperatures of the body, 2. Body responses to cold, 3. Body responses to heat, 4. Emotional sweating, 5. Metabolic Rate and environmental temperature, 6. Temperature control, 7. The human thermo-control system, and 8. Thermoregulatory effectors.

1. The temperatures of the body

The human body consists of a peripheral shell and a central core (Fig. 21-1). The *heat content* (***H*** or *enthalpy*) of the human body is reflected by its temperature. By definition a thermometer only measures the temperature of the thermometer, so its location is essential. The mean core temperature is 37 ^{o}C in healthy adults at rest, but small children have larger diurnal variations.

The *skin* is the *main heat exchanger* of the body. The skin temperature is determined by the core temperature and by the environment (temperature, humidity, air velocity). Thus the shell temperature is governed by the needs of the body to exchange heat energy.

Fig. 21-1: Heat transfers, body cores and shells temperatures of a naked person standing in cold and warm air, respectively.

The *shell temperature* is measured on the *skin surface* and at the *hands and feet* to approach the room temperature of 19^{o}C in a person standing in a cold room for hours (Fig. 21-1, left). The shell temperature is several degrees lower than the temperature in the central core. The limbs have both a longitudinal and a radial temperature gradient. The shell temperature and the size of the shell vary with the environmental temperature and the termal state of the person. A naked person, standing on a cold floor in 19^{o}C air has a small core and a thick shell compared to the same person in a warm environment (Fig. 21-1). The shell temperature of the skin and distal extremities is difficult to evaluate. The best estimate is measurement of the infrared heat radiation flux with a radiometer.

The *core temperature* is the rather constant temperature in the *deeper parts of the body* and in the proximal extremity portions (see the red stippled lines of Fig. 21-1). However, the core temperature may vary several Centigrades between different regions depending on the cellular activity. The brain has a radial temperature gradient between its deep and superficial parts. In a sense, the temperature of the mixed venous blood represents an essential core temperature.

The rectal temperature

A high core temperature is found to be constant in the rectum about 10-15 cm from the anus. When measuring the *rectal temperature* a standard depth of 5-10 cm is used clinically. The venous plexus around the rectum communicate with the cutaneous blood in the anal region. The rectal temperature falls when the feet are cold, because cold blood passes the rectum in the veins from the legs for the same reason. The rectal temperature rises during heavy work involving the legs.

Parents should be advised to measure the rectal temperature in disease suspect children. The rectal temperature is a reliable estimate of the core temperature in resting persons.

Sublingual (oral) or axillary temperatures are unreliable measures of the core temperature - often more than half a degree lower than the rectal temperature.

The cranial temperature (tympanic and nasal)

The main control of temperature is performed by the anterior hypothalamus, which has a high bloodflow. Within the cranium the hypothalamus lies over the Circle of Willis, which supplies it with blood, and close to the cavernous sinus which drains it. Hypothalamus elicits heat loss

responses when stimulated by heat. The tympanic membrane and areas in the nasal cavity (the anterior ethmoidal region, part of the sphenoid sinus) are supplied with blood from the internal carotid artery just like the hypothalamus. These *cranial* locations then serve as a substitute for the measurement of the inaccessible hypothalamic temperature.

Intake of 250 g of ice releases an abrupt fall in the nasal temperature in a warm person, whereas the change in rectal temperature is smaller and delayed (Fig. 21-2). The cranial core temperature is more dynamic than the rectal.

Fig. 21-2: Intake of ice reduces the temperature in a warm person resting at 45^{o}C.

In sports and in surgical hypothermia dynamic measurements of core temperature are essential. The cranial temperature is often preferred. During forceful movements the thermistor may be displaced. In such situations an *oesophageal* location is applied at heart level. This is an approximative measure of the temperature of the mixed venous blood of the right heart located close to the thermistor.

The *mean body temperature* is defined according to Eq. 21-1. The storage of heat energy in the body can be calculated according to its heat capacity (3.49 kJ* kg^{-1}*^{o}C^{-1}), the body weight (kg) and the change in mean body temperature in the period (Eq. 21-2).

According to the first law of thermodynamics, the storage of heat energy equals the metabolic energy change minus the heat loss (Eq. 21-3). Quantification of thermodynamics in humans is possible using equations 21-1 to 21-7 (later in this chapter).

The body is in heat energy balance, when the storage is zero. However, the core temperature may change with internal fluxes of heat energy between core and shell without storage or loss of heat energy at a constant activity.

Venous blood draining active muscles and the liver is likely to be warmer than pulmonary venous blood, since this has undergone evaporative cooling in the alveoli. A patient with high fever can be in *thermal steady state,* with a high constant heat production, if both core- and shell-temperatures are constant, and no internal energy flux occurs.

Warm-blooded animals, *homeotherms* such as humans, can change their metabolism in order to keep their heat production equal to the heat loss. Such animals have a *temperature control system* and thereby maintain a *rather constant core temperature*. Warm-blooded animals live with the advantage of an unchanged cell activity and temperature in their core. However, the human core temperature falls during the oestrogen phase of the menstrual cycle and during sleep (circadian rhythm). The lowest temperature is between 18 at night and 6 o'clock in the morning (Fig. 21-3). The temperature cycle is part of the circadian periodicity. Our biological clock seems to be synchronised with the rotation of the globe. Also meals, light and temperature plays a role.

Ovulation releases a sharp rise in morning temperature. Progesterone effects seem to explain the higher temperature in the last phase of the menstrual cycle (Fig. 21-3).

Fig. 21-3: Variations of the core temperature during 24 hour (above), and variations related to phases of the menstrual cycle (below).

Cold-blooded animals (*poikilotherms*) live with a behavioural temperature rhythm, but have no autonomic temperature control. The core- and shell-temperatures vary with the environment and the cellular activity. Reptiles, premature and low weight-premature newborn babies are cold-blooded. These babies have no thermoregulation (see later). However, their capacity for heat production is 5-10 times as great per unit weight as that of adults.

Humans have a *warm-blooded* (homeothermic) *core* and a *cold-blooded* (poikilothermic) *shell* in a cold environment.

Persons exposed to general anaesthesia, alcohol, and certain drugs lose the autonomic

thermoregulation. Cold-blooded animals must live with varying core and shell temperature, whereby the rate of their cellular activities varies with the surrounding temperature (Fig. 21-4).

Fig. 21-4: The body core temperature and the environmental body temperature for a warm-blooded animal (cat) and a cold-blooded animal (lizard).

a) Convection. The convective heat loss is calculated by Eq. 21-7. A healthy person in sports clothes experiences thermal comfort at three times the resting metabolic rate (3 MET), when the surrounding temperature is 20^{o}C, the humidity is 50% and the wind velocity is 0.5 m*s-1.

Diving (water has a high thermal conductivity) illustrates the importance of conduction andconvection in heat energy transfer.

The *dry diving suit* excludes water from contact with the skin and traps low-conductance air in insulating clothing worn inside the watertight sealing.

The *wet suit* traps water next to the skin but prevents its circulation. The water is warmed through contact with the skin, and the high insulation of the foam rubber wet diving suit, with its many pockets of trapped air, minimises the rate of heat energy loss to the surrounding water. Air is a poor heat conductor and thus a good insulator. During deep diving high pressures compress these air pockets and thus reduce the insulation properties of wet diving suits.

b) Radiation describes a transfer of energy between objects in the form of electromagnetic waves (photons). This includes ultraviolet and visible (sun light) radiation from the outside and from the body infrared or warm heat radiation.

Radiative heat transfer can be calculated for a naked person according to Eq. 21-4.

When the skin temperature (T_{skin}) is less than the temperature of the surrounding objects, heat is gained by radiation.

At wintertime, heat can be lost through a window glass by radiation from the body to the cold environment irrespective of the room temperature. This is because the skin temperature is higher than the outside temperature.

c) Conduction. Sitting on a cold stone is a typical example of conduction loss, just as standing on a cold floor (Fig. 21-1). – Conduction heat can also be gained, although it is really possible to walk on glowing coals with speed and a thick epidermal horn layer.

d) Evaporative heat loss- see sweat secretion below.

2. Body-responses to cold

Cutaneous vasoconstriction lowers skin temperature, and thereby reduces the *conductive-convective heat loss* that is determined by the temperature gradient from the skin surface to the environment. Cutaneous vasoconstriction directs the peripheral venous blood back to the body core through the deep veins and the commitant veins. These veins are located around the arteries with warm blood, so that the venous blood receives part of the heat energy from the arterial blood - so-called *counter current heat exchange* (Fig. 21-5). The vasoconstriction is so effective, that the bloodflow through the arterio-venous anastomoses in the fingers and toes can fall to below one percent of the flow at normal temperature. The cooling of the shell is immediate, and the size of the shell increases (Fig. 21-1). Obviously, the shell is large for a naked person in cold air. The resistance vessels of the hands may open periodically to nourish the tissues, but the high viscosity of the cold blood can endanger the tissue nutrition and result in trench foot.

The *arterio-venous shunts* of the hands and feet are closed, so the bloodflow to the limbs is a nutritive minimum.

The deep arteries and veins of the limbs lie in parallel, so the arterial bloodflow loses heat to

the incoming venous blood partially surrounding the arteries (Fig. 21-5). This is a typical *counter-current heat exchange*. In a cold environment, where vasoconstriction and heat exchange produces cold extremities, the total insulation is increased at the expense of reduced neuromuscular efficiency.

Fig. 21-5: Counter-current exchange in a human arm conserving heat energy in a cold climate (left). Superficial venous cooling ribs eliminate heat energy in a warm climate (right).

In a *warm climate* the high bloodflow of the extremities ensures an optimal temperature of the deeper structures (eg, the neuromuscular system). The temperature of the arterial blood is maintained (Fig. 21-5, right) and the arterio-venous anastomoses are wide open conveying warm blood to the superficial veins. The superficial veins also act as **cooling ribs** and transfer large amounts of heat to the skin surface, where it is eliminated from the body by convection, conduction and evaporation (Fig. 21-5, right).

Shivering is a *reflex myogenic response* to cold with *asynchronous* or balanced muscle contractions elicited from the hypothalamus via cutaneous receptors. The activity in agonist and antagonist muscles balance, so there is no external work. Without outside work, all energy is liberated as metabolic heat energy. Heat production is also increased by thyroid gland activity and by release of catecholamines from the adrenal medulla.

External work, such as running, is helpful in maintaining body temperature when feeling cold. Cold increases the motivation for *warm-up exercises* and illustrates the voluntary, cortical (feedforward influence) on temperature homeostasis. The core temperature increases proportionally to the work intensity during prolonged steady state work (Fig. 21-6). The mean skin temperature falls with increasing work intensity at 20^{o}C, because the sweat evaporation cools the skin.

Fig. 21-6: Muscular and oesophageal temperature during steady state exercise. The levels of exercise range from zero to 100% of the maximum oxygen uptake.

The temperature in the active muscles determines the level of the rectal temperature. Following marathon rectal temperatures of more than 41^{o}C have been measured and heat strokes have occurred. A marathon is even more difficult to accomplish in warm, humid environments and strong sun may cause sunstroke (see later).

People may adapt to prolonged exposure to cold by increasing their basal metabolic rate up to 50% higher than normal. This *metabolic adaptation* is found in Inuits (Eskimos) and other people continuously subject to cold.

The environmental temperature, where we maintain our autonomic temperature control, is in the range of zero to 45^{o}C. Below and above this range we adapt to the environment by behaviour (adding or removing clothing, warm or cold bath, sun or shadow). A core temperature above 44^{o}C starts protein denaturation in all cells and is incompatible with life. Below 32^{o}C humans lose consciousness and below 28^{o}C the frequency of malignant cardiac arrhythmia's is increasing, ending with ventricular fibrillation and death at a core temperature below 23 ^{o}C (Fig. 21-7).

Fig. 21-7: Environmental temperature variations and temperature control. Lack of vital signs in the clinic (respiration, heart rate, EEG) must not be taken as death. Treatment must be instituted until death signs are developed.

3. Body-responses to heat

***Sweat secretion*.** Three million sweat glands produce sweat at a rate of up to 2 litres *per hour* or more during exercise in extreme warm conditions. If not compensated by drinking, such high sweat rates lead to circulatory failure and shock. Sweat resembles a dilute ultrafiltrate of

plasma. Healthy humans cannot maintain their body temperature, if the environmental air reaches body temperature and the air is saturated with water vapour. Primary sweat is secreted as an isosmotic fluid into the sweat duct, and subsequent NaCl reabsorption results in the final hypo-osmotic sweat. *Thermal sweating* is abolished by atropine, proving that the postganglionic fibres are cholinergic. Cholinergic drugs provoke sweating just as adrenergic agonists do. *Evaporation* of water on the body surface eliminates 2428-2436 J g^{-1} at mean shell temperatures of 30-32^{o}C. Evaporation of a large volume of sweat per time unit ($V°_{sweat}$) implies a substantial loss of heat according to Eq. 21-5.

Normally, the skin temperature falls with increasing work intensity, because the sweat evaporation cools the skin (Fig. 21-6). Danger occurs when the average skin temperature and the body core temperature converge towards the same value.

Condensation of water on the skin gains heat energy, which is stored in the body. This is what happens in a Sauna.

Vasodilatation of skin vessels in warm environments results in increased cardiac output. The *arterio-venous anastomoses* in the hands and feet are open, and the bloodflow can rise up to at least 10 folds. The shell is minimal, when a naked person is in warm air (Fig. 21-1, right). The *skin bloodflow*, mainly in the extremities, determines the amount of heat energy, which is carried from the body core to be lost on the surface. The heat energy is transported from the large body core to the skin by *convection in the blood*. A substantial part of the heat energy is lost through the superficial veins of the extremities acting as *cooling ribs* (Fig. 21-5). The blood of the superficial veins is thus arterialized, when the person is warm.

A piece of steak has the same composition as human skin but of course no blood flow and no sweat evaporation. Thus the steak will be cooked at an air temperature that humans can survive. A person can stay in a room with dry air at 128^{o}C for up to 10 min during which time the steak is partially cooked.

4. Emotional sweating

This is a *paradoxical response* in contrast to the *thermal sweating* of thermoregulation. Emotional stress elicits *vasoconstriction* in the hands and feet combined with *profuse sweat secretion* on the palmar and plantar skin surfaces.

5. Metabolic Rate and environmental temperature

The total heat loss consists of the evaporative heat loss and the dry heat loss (ie, the sum of convective, conductive and radiative loss). *Newton's law of cooling* states that the dry heat loss is proportional to the temperature difference between the human body (shell) and the surroundings.

Let us look at a healthy, lightly dressed sitting person in thermal balance. His heat loss is plotted as the ordinate and the environmental temperature as the abscissa (Fig. 21-8). The two types of heat loss are added in order to provide the total heat energy loss. At a room temperature of 37^{o}C there is a dry heat loss of zero, and below there is an increasing dry heat loss. Above 37^{o}C the heat loss turns into heat input. Obviously, the dry heat loss also depends upon conduction and convection of heat inside the body by contact and by the perfusion. At extremely low environmental temperature the dry heat loss becomes larger than the metabolic heat liberation and the body is cooled down.

The person is in *thermal steady state*, and the metabolic rate is almost constant in the *thermoneutral zone* between 20 and 30^{o} C (Fig. 21-8). The *law of metabolic reduction* reflects the tendency for heat production to match the rate of heat loss. The thermoneutral zone, where minimal compensatory activity is required, is separated in the ***lower vasomotor*** and the *upper sudomotor control zone.*

In the lower, comfortable zone (20-26^{o} C) the *total heat dissipation* is maintained equal to the metabolic rate by cutaneous, vasomotor alterations. The small evaporation loss is termed leakage of the skin or *insensible perspiration*, which is unrelated to sweat gland function and rather constant at the basal metabolic rate.

Fig. 21-8: Metabolic rate and environmental temperature in a fasting dressed human at rest. The wet and the dry heat loss, as well as the metabolic heat and the basal metabolic rate (*BMR*) is measured in Watts.

In the *upper sudomotor zone* above 26^{o} C environmental temperature, the bloodflow through the skin rises, as does sweat secretion and evaporation. At 37^{o} C the rise in energy loss occurs via evaporation (Fig. 21-8).

When the environmental temperature falls the metabolic rate increases - first by increasing muscle tone and then by shivering. The *chemical* or *metabolic temperature control* is in the environmental region from 20^{o} C and below (Fig. 21-8), where shivering, decreased bloodflow through skin and non-myogenic heat production take place. Here, metabolism controls the core temperature by increasing metabolic rate with falling temperature in the environment. Above 20^{o} C, the *physical temperature control* takes over, as an autonomic capacity for alterations in heat loss. In this thermoneutral zone the body temperature is kept constant almost without either heat-producing mechanisms or sweat secretion. - The *thermal comfort* for light clothed, seated persons is about an air temperature of 26^{o}C when the humidity is 50%, 30^{o}C for nude persons, and about 36^{o}C sitting in water to the neck.

In the *metabolic zone*, the total heat loss rises with falling environmental temperature, but below 5 ^{o}C in the environment, the dry heat loss exceeds the metabolic rate, and the body is cooled down (Fig. 21-8). This is the *zone of hypothermia*, where cold death is inevitable without treatment.

The *zone of hyperthermia* begins at an environmental temperature of 37 ^{o}C, where humans soon reach the maximal capacity for evaporation and there is an unbalanced heat influx to the body ending in *heat death.*

The metabolism varies with the shell and the core temperature. These relations were elucidated by series of similar experiments. A person was placed in stirred water to the neck, where the water temperature thus defined the mean shell temperature. By pre-treatment in other baths an array of core-and shell- temperature combinations were obtained and measured simultaneously with the metabolic rate (*metabolic heat liberation*). The core temperature was reduced by intake of up to 2 l of crushed ice in water.

At a shell temperature of 20 ^{o}C, intake of ice water reduces the core temperature below 37.1 ^{o}C (here called the set point) and the metabolic rate increases by shivering (Fig. 21-9) with falling core temperature.

Fig. 21-9: The metabolic rate as a function of the core and the shell temperature.

Warmer skin makes the set point fall and the rise in metabolic rate per ^{o}C (of core temperature fall) is less steep. Core temperatures above 37.1 ^{o}C all have a low metabolic rate, regardless of the shell values. At a shell temperature of 30 ^{o}C the set point decreases from 37.1 to 36.7 ^{o}C. The hypothalamus functions as a *thermostat* using the rest of the body to stabilise its own temperature. With rising core temperature the metabolic rate is maintained low and the body tries to cool down. When the body core temperature falls below set point, the metabolic rate increases by shivering and the heat energy storage as well. The cutaneous cold receptors are maximally active at 20 ^{o}C and they are silent above 33 ^{o}C - they can only trigger shivering, when the core temperature is below set point. Shivering ceases immediately, when we take a

warm shower and the skin is warmed. The hypothalamic, preoptic heat receptors inhibit shivering, and shivering is totally blocked above the set point.

Another series of experiments were directed towards heat loss. The heat loss from evaporated sweat was recorded during exercise at different shell temperatures.

Fig. 21-10: Evaporative heat loss as a function of the core and the mean skin temperature during different intensities of exercise.

Cold stimulates cutaneous cold receptors with connections to the hypothalamus (Fig. 21-12).

At mean skin temperatures of 33-39 ^{o}C, where the cutaneous cold receptors are silent, the person is unable to sweat before the core temperature is above 36.9 ^{o}C (the set point of Fig. 21-10). Below the set point, the evaporation is low (perspiratio insensibilis). With increasing muscle activity and core temperature the evaporative heat loss may rise towards 20 kJ each min.

Mean skin temperatures below 33 ^{o}C reduce the evaporative heat loss. The rising hypothalamic temperature releases the sweat secretion, but the local secretion is inhibited by the cutaneous cold receptors.

Heat has a direct effect on preoptic heat receptors in the hypothalamus. At increased core temperature, the preoptic heat receptors totally block shivering, although the hypothalamic centre also receives shivering signals from cutaneous cold receptors in cold surroundings with mean skin temperatures below 33 ^{o}C. The falling mean skin temperature can also act directly to lower cutaneous bloodflow and sweat secretion.

6. Temperature control

Humans have a rather *constant core temperature* although the metabolism and environmental temperature may vary considerably. This implies that *control* is exerted.

Fig. 21-11: Thermoregulation by dynamic gain and set point systems.

Information about the environmental temperature is provided by *peripheral thermo-sensors*, which are located in the skin, abdominal organs, and muscles. Internal or blood temperature is monitored by *central thermo-sensors* in the preoptic hypothalamus and the medulla.

A rise in *hypothalamic temperature* causes vasodilatation in the skin and reduces muscular tone. The person looses motivation for physical activity and reduces clothing. Then thermal sweat is observed, and after some time, reduced activity of the adrenal cortex and of the thyroid gland is also observed.

A *fall in hypothalamic temperature* by cooling of the shell and core releases cutaneous vasoconstriction together with increased muscular tone and shivering. There is a sympathetic activation with secretion of catecholamines, oxidation of fatty acids and glucose, and increased secretion of the thyroid and adrenal gland. The muscle tone is increased, and shivering is triggered reflexly as asynchronous muscle contractions without external work (movements), so all the metabolic energy is released as heat. The capacity for muscular thermogenesis by shivering is high. Up to five folds basal metabolic rate is observed, which corresponds to heavy industrial work.

Shivering may be suppressed voluntarily at the beginning. Transmission signals for shivering passes the rubrospinal pathways to a- and g-motor neurons of antagonist muscles.

Dynamic gain and set point control

A: *A dynamic gain system* responds *continuously to feedback signals* - regardless of the core temperature. With rising tissue temperature, the neural activity of heat sensors increases linearly, whereas the activity of cold sensors decreases (Fig. 21-11A).

This combined sensory input to the hypothalamus increases the core temperature and thus increases the activity of *heat loss effectors,* while inhibiting *heat production effectors* (Fig. 21-11 A). This determines the *reference signal*.

The dynamic gain system has a floating *reference signal* moving with the continuous heat loss and heat production (Fig. 21-11A).

B: *A set-point system* does not respond to a rise in temperature before a certain set point is reached. The set point is the core temperature at which neither heat loss mechanisms or heat production mechanisms are active.

When a thermal disorder reaches a certain *set-point* in the hypothalamus, signals passes to the effectors. The desired core temperature (t_{set} = *set point temperature*) is compared to the actual value (t_{core} in Fig. 21-11B). The caudal hypothalamus works as a thermostat. Error signals - a deviation from t_{set} - evoke responses that tend to restore core and hypothalamic temperature toward the set point. When the actual core and hypothalamic temperature rises above the desired set point such as 37 ^{o}C, effectors are turned on, and the compensatory heat energy loss is almost linear (Fig. 21-11B). These compensatory mechanisms (vasodilatation, sweat, reduced muscle tone) do not turn off until the temperature drops to the set point (ie, an all-or-non response).

When the actual core and hypothalamic temperature is just below the set point, the compensatory mechanisms (vasoconstriction and shivering) are relatively inactive.

The hypothalamic set point change with the physiological conditions and is elevated in fever by *pyrogens* from microorganisms. The rise in metabolism is mainly accomplished by shivering.

7. The human thermo-control system

Human temperature control exhibits both *dynamic gain* and *set point* characteristics. The control system implies widespread cutaneous and deep sensors. Their afferents converge towards the hypothalamic integrator, which acts as a *thermostat.* The hypothalamus also contains thermosensors in the preoptic region, and inhibitory neurons perform *crossing inhibition* (Fig. 21-12). The *central heat drive* from the *preoptic hypothalamus* is maintained (Fig. 21-12). The stability of the core temperature is maintained by the *large heat capacity of the body mass,* and by the deep thermosensors, which are dominant.

Shivering is released from cutaneus cold sensors firing maximally at 20 ^{o}C. These cold sensors are silent above 33 ^{o}C. The cold shell (Fig. 21-12) activates deep cold sensors in the preoptic hypothalamus. This increases heat production by shivering. The preoptic thermostat simultaneously reduces heat loss by crossing inhibition.

Fig. 21-12: The hypothalamic thermostat and its connections.

Sweat secretion is released by preoptic warm sensors as soon as their temperature is 37 ^{o}C or above (t_{set}). Cutaneous cold sensors inhibit sweat secretion at shell temperatures below 33 ^{o}C, since they are silent above this temperature.

In conclusion, preoptic warm sensors show set point characteristics below the set point, and preoptic cold sensors show set point characteristics above the set point.

Apart from that, preoptic sensors show *dynamic gain*: With rising tissue temperature, the neural activity of heat sensors increases linearly, whereas the signal frequency of cold sensors increases with falling temperature (Fig. 21-11**A**).

Cutaneous sympathetic vasodilatation is probably also released by preoptic warm sensors above set point (Fig. 21-12). A fall in skin temperature below 33 ^{o}C will reduce skin

bloodflow by crossing inhibition (Fig. 21-12).

Alcohol seems to off set the thermocontrol mainly by inhibition of the hypothalamic thermostat (Fig. 21-12).

Newborns, down to premature babies above 1000 g, possess certain thermoregulatory functions. The newborn can increase thermogenesis by a factor of three without shivering shows vasomotor reactions, sweat secretion, and reduces the surface area in cold air. However the baby has special problems: The surface-volume ratio is 3-fold higher than that of an adult. The baby has a thin shell due to the thin subcutaneous fat layer, so even a maximal vasoconstriction cannot limit heat loss with a capacity comparable to that of an adult. Newborns are specially equipped to perform non-shivering thermogenesis (chemical thermogenesis). Non-shivering thermogenesis is any rise in metabolism, which is not related to shivering. In babies this form of thermogenesis is particularly large in their brown adipose tissue. This tissue is abundant around vital organs, in neck and mediastinum, between scapulae and in the armpits. Brown adipose tissue cells contains multilocular droplets in the fat phase and many mitochondria. The tissue receives sympathetic innervation and is stimulated by catecholamines and thyroid hormones. Cooling increases the bloodflow and temperature of brown adipose tissue. Noradrenaline injections cause vasodilatation via b-receptors, and increase the metabolism to the same extent. The cause of the increased metabolism is an increased cell membrane permeability for Na^+- and K^+-ions, whereby the Na^+- K^+-pump (high ATP demand) is activated.

Fig. 21-13: Thermocontrol in the newborn.

Newborns are in thermal balance at a minimal metabolism only when the surrounding temperature is high (32-34 ^oC). With other words, newborns have an extremely high lower threshold for the thermoneutral zone, namely 32 ^oC. The threshold for maximal use of shivering is approximately 23 ^oC compared to a naked adult about 5 ^oC. In general, these threshold values for newborn increase with falling body weight in premature. Premature below 1000 g has hardly any thermoneutral zone (Fig. 21-13). They are actually cold-blooded and their temperature control is maintained with a *couveuse*.

8. Thermoregulatory effectors

The *sympathetic* and the *somatomotor nervous system* participate in thermoregulation (Fig. 21-14).

Noradrenergic sympathetic neurons control the bloodflow through fingers, hands, ears, lips and nose. *Arterioles* contract and *arteriovenous anastomoses* close (thermal insulation) following an increase in sympathetic tone, and dilatate following a decrease in tone. When arterioles and arteriovenous anastomoses open, the bloodflow is markedly increased and thus the *convective heat loss* from the skin is increased.

Fig. 21-14: The thermoregulatory feedback system.

Cholinergic sympathetic fibres control sweat secretion. The vasodilatator *bradykinin* is liberated in the skin. Thus, profuse sweat secretion is always accompanied by vasodilatation.

Sympathetic activation releases *thyroid hormones* from the thyroid gland and *catecholamines* from the adrenal medulla. These hormones liberate fatty acids and glucose for combustion. A reduced sympathetic tone also reduces the activity of the adrenal and the thyroid gland.

The thermogenic response to cold also involves a *non-myogenic* or *non-shivering component* probably in adipocytes. Non-shivering heat production is controlled by the sympathetic nervous system via adrenergic b-receptors. The noradrenaline (NA in Fig. 21-14) released at the nerve terminals close to the adipocytes, stimulates the liberation of free fatty acids and their subsequent oxidation. *Non-myogenic heat production* includes a contribution from the

brown fat of babies, but is insignificant in adults.

Shivering is induced by way of the motor system. The *central shivering pathway* passes from the hypothalamus to the motor neurons in the spinal cord. Shivering is abolished by blockade of the neuromuscular end plate with curare.

Thermoregulatory behaviour such as fanning and adding or removing clothing is effective in changing the thermal insulation. Several layers of clothing with trapped air act as a good insulator.

In healthy persons, heat energy is liberated from cellular metabolism and transferred to the environment through the skin by sweating and vasodilatation. Sweating occurs during exercise, and its evaporation is the most important mechanism in maintaining the core temperature as close to 37^{o} C as possible. The thermoregulatory centre in the hypothalamus controls all processes.

Fig. 21-15: Heat and cold adaptation

Cold adaptation is found among Australian aborigines and Inuits in Greenland. Inuits have relatively more sweat glands in the face and less on the body.

Aborigines can sleep naked on the ground even at low temperature. Inuits have a basal metabolic rate 50% higher than persons living in a temperate climate do. The threshold for shivering is shifted towards the left in cold-adapted persons, but they maintain normal function at the new set point. Very old people may show the same phenomenon, and live with a core temperature of 35 ^{o}C without shivering. Obviously, cold adaptation implies non-shivering thermogenesis, which is oeconomic metabolic heat liberation (Fig. 21-15).

Heat acclimatization is actually a *sweat gland adaptation*, and a 2-week process following arrival to a hot climate such as the tropics or a desert. Gradually sweat-evaporation is increased and the NaCl loss is reduced. The sweat secretion capacity may reach 4 l each hour with a thin sweat. The adaptation is caused by increased aldosterone secretion from the adrenal cortex. Aldosterone increases the reabsorption of NaCl and the secretion of K^{+} from the sweat, during its passage of the sweat gland tubules. The larger sweat loss the thirstier one feels. This is because the large sweat secretion reduces the time period for NaCl- reabsorption in the sweat gland tubules. The resulting high NaCl concentration in the plasma implies thirst, so heat adapted persons have to drink a lot. Thirst is an extremely late indicator of dehydration during work in a warm climate.

Tropic inhabitants are of course heat-adapted. They have an increased core temperature (set point) and their threshold for sweat and vasodilatation is typically 0.5^{o}C higher than that of a person living in the *temperate zone* (Fig. 21-15).

Pathophysiology

This paragraph deals with *1. Heat cramps, 2. Heat exhaustion and heat stroke, 3. Malignant hyperthermia, 4. Hypothermia, 5. Frostbite, and 6. Fever and hyperthermia.*

A paragraph concerning nuclear energy radiation is given at the end.

1. Heat Cramps

Painful cramps in the leg muscles occur following exercise, when athletes run too fast in a hot climate. Heat cramps respond to salt and water replenishment in a normal diet, and the cramps are probably caused by hyponatraemia. During prolonged sweating, the runner is losing salt and water. If only the water loss is replenished, the result is water-induced hyponatraemia - a parallel to the classical *miner's cramps.*

2. Heat Exhaustion and Heat Stroke

When the water-salt balance is at risk in a hot climate, it is always a threat to the circulation. Profuse sweat secretion, in a subject who is not acclimatized, results in salt- and water depletion, with a daily loss of more than one mol of NaCl and more than 6 l of water. Within a period of one-hour strenuous working endurance athletes have lost up to 8 l of water.

The falling extracellular fluid volume and increasing body and brain temperature to above 40^{o}C elicit severe symptoms and signs. As the volume and salt depletion develops, the sweat production goes down in spite of extreme vasodilatation. The falling blood pressure stimulates the high-pressure baroreceptors resulting in a rising heart rate. The dehydration with imminent shock frequently results in cerebral, renal and hepatic failure. The low brain bloodflow through an overheated brain leads to fatigue (*heat exhaustion),* confusion and unconsciousness or syncope (*heat syncope*). The confusion may develop into a veritable *delirium* (an acute impairment of consciousness) with brain oedema.

Heat stroke - in the sun it is called *sunstroke* - is heat collapse that occurs suddenly, hereby creating a life-threatening condition. This heat collapse often occurs in warm, humid environments, when an unacclimatized subject exercises. The subject - without sweating (hypothalamic failure) -suddenly falls into coma, if not preceded by a short period of confusion and delirium. The condition is fatal, if not relieved by rapid cooling.

3. Malignant hyperthermia

Malignant hyperpyrexia is often caused by a *genetic defect* (autosomal dominant) in the sarcoplasmic reticulum of skeletal muscles. *General anaesthesia* (often halogen-substituted ethane) triggers an allergic reaction with sudden opening of Ca^{2+} - channels in the muscle cells. The following influx of Ca^{2+} elicits generalized and maintained muscle contraction (rigidity), which liberates enormous quantities of heat energy. This condition is life threatening and often results in sudden death during or just after anaesthesia.

4. Hypothermia

Hypothermia is a fall in core temperature to values below 35 ^{o}C. Hypothermic subjects lose consciousness, when the core temperature falls below 32 ^{o}C - a potentially lethal condition called *severe hypothermia*.

During anaesthesia and surgery the core temperature of the patient falls and stabilise around 34-35 ^{o}C after 4-5 hours. Such a *surgery hypothermia* increases the risk of cardiac complications, bleeding tendency and prolonged wound healing.

The condition is a prominent cause of death in climbers and skiers, as well as in persons being immersed in cold water or living in the Antarctis. The climbers and skiers are often exposed to a cold, wet and windy environment carrying insufficient clothing. As the neuromuscular function suffers, they can no longer move. General hypothermia develops and they die.

In mild hypothermia the subject can still take action to rewarm by exercise and clothing, but as consciousness is lost the core temperature falls further, because shivering is abolished. The patient feels cold to touch, is in a developing coma, the circulation and respiration fall, as does the metabolic rate. The dissociation curve of oxyhaemoglobin is moved to the left, and the solubility of gasses increases as the blood temperature falls.

The heart is the target organ in hypothermia. Below 30 ^{o}C, spontaneous remission is practically impossible, and death ensues from ventricular fibrillation occurring around 28 ^{o}C.

Careful monitoring of all vital functions is required during rewarming, which is performed either passively (rewarming by the patients own metabolism) with insulation and space blankets or actively by a warm water bath, while monitoring the patient. A good strategy in treating hypothermic victims is the slogan: "*They are not dead until they are warm and dead*."

An arterial blood sample from a hypothermic patient is routinely analysed at 37 ^{o}C. The P_{aO2} and pH_a is falsely higher, the P_{aCO2} is falsely lower, than in the circulating blood of the hypothermic patient. This is because more CO_2 is bound also as carbamino-haemoglobin, and less O_2 is bound in the cold blood. *Base Excess* is defined at 37 ^{o}C and thus a true metabolic variable (see Chapter 17).

Artificially induced hypothermia is used in brain- and heart-surgery, where the usual thermocontrol is inactivated by general anaesthesia. The procedure becomes dangerous, when it elicits *ventricular fibrillation*.

5. Frostbite

Frostbite is a local cold injury due to the formation of extracellular ice crystals in the skin and other tissues. This leads to extracellular dehydration and *hyperosmosis*, whereby the cells lose water until they die. As the skin temperature falls *below + 7 ^{o}C* the subjects have lost their sensory functions, and thus do not recognise the developing frostbite.

The most effective treatment is *warming* of the still frozen area by immersion in 40-42 ^{o}C hot water following hospitalisation. The patient experiences severe pain (above + *7 ^{o}C*) and morphine must be administered. Tissue which is no longer frozen must be treated sterile.

6. Fever and hyperthermia

Fever occurs when the *core temperature of the body* is raised above normal steady state levels. The body reacts as if it is too cold, while the temperature rises up to the new higher set point. Fever attacks imply shivering combined with vasoconstriction, headache, dedolation pains, and general discomfort. Fever is the result of one of two phenomena: the set point may be set to a higher level, or the efficacy of the temperature control system may be impaired. Fever implies *hyperthermia;* however many cases of hyperthermia do not constitute fever. Fever results from the action of *endogenous pyrogens* on the *hypothalamic heat control centre* (they increase the set point for the core temperature via prostaglandins). Exogenous pyrogens from microbes cause these *endogenous polypeptides* to be released from the defence cells of the body (ie, the reticuloendothelial system, RES). Antipyretic drugs inhibit cyclo-oxygenase activity, hereby interfering with the synthesis of prostaglandins and thromboxanes.

Following an attack of fever, vasodilatation and sweat evaporation reduces the core temperature.

Physiologic hyperthermia is an increase in core temperature caused by extreme heat stressor exercise. During work the body temperature rises up to 39^{o}C without clinical consequences.Hereby, the heat loss capacity is exceeded. During hyperthermia the *heat loss effectors* are strained to the utmost. The high body temperatures of *exercise* activate cooling mechanisms and elicit sweat loss, which strive to return the core temperature to its normal level.In extreme hyperthermia the core temperature may rise to more than 41^{o} C (heat stroke). Irreversible protein denaturation occurs above 44^{o}C with brain oedema and destroyed thermoneurons in the hypothalamus. Clinically, the brain damage is shown with disorientation, lack of sweat secretion, delirium and *universal cramps* before death.

Ionizing Radiation Hazards

Ionizing radiation implies destruction of tissue molecules.

Dosages of absorbed radiation is measured in Joules per kg (1 J kg $^{-1}$ is known as a ***gray, Gy***). *One Gy* is equivalent to *100 rads*. Radioactivity is measured as the number of degradations per second in *bequerels* or *Bq*. One degradation per s from radioactive material equals one Bq.

The absorbed dose of radiation is balanced for damaging ionizing tissue effect, since different types of radiation have different density of ionization. A dose equivalent termed sievert (Sv) causes a rather large damage. The annual background radiation is 2.5 milli-Sv (*2.5 mSv*). Non-penetrating radiation occurs from alpha- and beta particles. These particles are stopped by paper, but when they enter a tissue - such as the bone marrow – they stay there and spoil everything.

Penetrating radiation consists of either gamma rays (neutron) or X-rays. Survivors from nuclear power plant accidents, with whole body absorption greater than 100 rads, are threatened by acute and chronic radiation sickness.

Acute radiation sickness appears as vomiting and malaise following exposure to 1 Gy (100 rad) or more. Lymphocyte production is reduced immediately, soon followed by leucopenia and thrombocytopenia with bleeding. The villi of the gastrointestinal tract are destroyed, absorption of nutrients is impaired, and new villous cells are not produced. Diarrhoea, often with blood loss, results in dehydration and anaemia. The skin is red and blistering, and the hair is loosed. The immunodeficiency system is destroyed, and secondary infections have a high mortality - especially pulmonary infections. *Cerebral oedema* may kill the victim within hours when exposed to 35 Gy or more.

Contamination with radioactive iodine is treated by immediate intake of potassium iodide, which block the major part of the thyroid absorption. The treatment of seriously exposed radiation victims is supportive and frustrane.

Chronic radiation sickness or *late radiation damage* implies an increased rate of mutagenesis, which includes a high frequency of leucaemia, cancer of the brain, the thyroid and the salivary glands, infertility and *cataract* (ie, an eye disease where the vision is blurred by an opaque lens). The radiation liberates large amounts of highly reactive ions in the cells. The reactive ions rupture DNA strands and cause mutations with production of cancer cells. Cancer cells multiply exponentially and their energy demand approach the total nutritive energy available, causing malnutrition and death.

The nuclear power plant accident in *Tjernobyl*, Ukraine, had many victims from radiation exposure. The nearby town, Pripatja, previously with 40,000 residents is now abandoned.

The frequency of thyroid cancer among children in Ukraine is more than 100 times the expected. Most of the children and young persons exposed have developed leucaemia or cancer. Their airway epithelium and respiration is seriously affected, and they have an unusual high frequency of cerebral haemorrhage.

Equations

- The mean body temperature (T_{body}) is calculated with the following equation:

$$\text{Eq. 21-1: } T_{body} = (0.7 \times T_{core} + 0.3 \times T_{shell}),$$

where the mean shell temperature (T_{shell}) is estimated from a series of representative skin temperatures. The obvious assumption in this equation is that 70% of the body weight is core (T_{core}), and the balance is shell, but the size of the shell varies with the environmental temperature.

- The storage of heat energy in the body is calculated as follows:

$$\text{Eq. 21-2: } \Delta H_{STORE} = 3.49 * \text{BODY WEIGHT} * (T_2 - T_1)$$

with the unit kJ per hour.

- According to the first law of thermodynamics, the following relation is valid:

Eq. 21-3: $\Delta H_{STORE} = \Delta H_{METABOLIC} - (\Delta H_{RADIATION} + \Delta H_{CON} + \Delta H_{EVAPORATION})$

where ΔH_{CON} is the change in heat loss by convection and conduction.

- Radiative heat loss from a warm object (T_{obj}) can be calculated for a naked person with a known skin temperature (T_{skin}):

Eq. 21-4: $\Delta H_{RADIATION} = 0.5 \times A \times (T_{skin} - T_{obj})$

where 0.5 is kJ min^{-1} m^{-2} K^{-1}, A is the area of the human body (radiating or receiving radiation), and T_{obj} is the temperature of the object exchanging energy. This equation is an approximation of Stefan-Boltzmanns rule.

- Evaporation of water on the body surface eliminates 2430 J g^{-1}. Evaporation of large volume rates of sweat ($V°_{sweat}$) implies a substantial loss of evaporative energy (J/min) according to the equation:

Eq. 21-5: $\Delta H_{EVAPORATION}$ (J min^{-1}) = 2430 (J g^{-1}) × $V°_{sweat}$ (g min^{-1}).

- Changes in heat conduction through the shell ($conductance_{shell}$) reflect changes in skin bloodflow. The conductance can be calculated by the formula:

Eq. 21-6: $Conductance_{shell}$ = Metabolic Rate/($T_{core} - T_{shell}$).

A large temperature difference implies an effective isolation, whereas a small difference implies a low isolation capacity.

- The heat loss of a naked person resting in quiet air by convection (including a minor part by conduction without movement) is given by the following equation:

Eq. 21-7: $H_{CON} = 0.5 * (T_{shell} - T_{air})$ in kJ per min.

This equation is an approximation of Newton's law of cooling.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. Humans can survive at a temperature that would cook a piece of steak, because the steak cannot dissipate core heat energy.

B. Heat conductance of air at high pressure exceeds that at low pressure, whereby more heat energy is lost from the divers body by conduction through air inside a diving bell deep under water than at one atmosphere of pressure.

C. Temperature homeostasis is present when heat energy production equals heat loss.

D. Basal metabolic rate (BMR) is lower before than after a meal.

E. The neutral environmental temperature defines the level, where the resting metabolic rate is minimal.

II. Each of the following five statements have True/False options.

A. Water-induced hyponatraemia is called miners cramps.

B. Substantial influx of Ca^{2+} to the cells is probably involved in some cases of malignant hyperthermia.

C. Hypothermia is a fall in core temperature below $32^{o}C$.

D. One degradation per second from radioactive material equals one bequerels (1 Bq), which is also equal to one curie.

E. Acute radiation reduces leucocyte and thrombocyte production.

Case History A

A male with a body weight of 70 kg stops his malaria prophylaxis with primaquine when leaving the endemic area. Three weeks later, a sudden attack of fever increases his core temperature from 37 to $40^{o}C$ within 30 min. The heat capacity of the human body is 3.47 kJ/kg and per ^{o}C. The metabolic rate of the subject increased substantially during the rise in temperature. Following the cold stage with uncontrollable shivering, the patient develops a delirious condition with severe headache. Two hours later the patient develops a profuse sweating. Partially evaporation of the water in 32 ml sweat/min (25% evaporates) occurs from the body surface (eliminating 2,436 J/g), during which body temperature drops.

Calculate the extra heat energy stored in his body after 30 min.

1. *Calculate the smallest possible metabolic rate in the 30-min period.*

2. *What causes the extra heat energy stored in the body?*

3. *Calculate the reduction in body temperature by ingestion of one L of ice water, when the fever is at its highest level.*

4. *Calculate the time it takes to lose the accumulated heat energy by evaporation.*

Case History B

A male sedentary person, weight 70 kg, has a daily food intake of 400 g carbohydrate (17.5 kJ/g), 100 g fat (39 kJ/g), and 100 g protein (17 kJ/g). There is a metabolic water formation of 32 mg/kJ. The man excretes 2100 ml of water (i.e., 1200 ml in the urine, 100 ml in faeces, and 800 ml through lungs and skin).

1. *Calculate the metabolic rate (in kJ/day or MJ/day).*

2. *The man is in water balance by a water intake of 1,700 ml as a total. Explain this apparent imbalance.*

Case History C

A 79-year old male is found apparently dead in the snow following a winter storm, where all traffic was arrested by snow. His muscles are stiff, and the heart rate is not palpable. The

tendon reflexes are depressed, and the pupillary and other brainstem reflexes are lost.

The body is placed in a chapel at the hospital until the funeral. The next day the personnel are disturbed by noises from the chapel. Obviously, the man is alive.

1. What has awakened the man?

2. Suggest a likely core temperature, at the time where the man was admitted to the hospital.

Case History D

A 20-year old person, with a body surface area of 1.8 m^2, is at rest with a metabolic rate (MR, or heat energy production and transfer) of 80 Watts. His rectal temperature is 37 ^oC and his mean skin temperature is 33 ^oC. - Suddenly, a malaria attack develops with a 7-fold rise in MR, and the patient reach a temperature plateau of 40 ^{oC}, with a mean skin temperature of 34 ^oC.

The conductance of the shell (C_{shell}) is a measure of skin bloodflow. This can be calculated by the formula: $C_{shell} = MR/(T_{core} - T_{shell})$.

1 *Calculate the conductance of the shell at rest.*

2. *Calculate the conductance of the shell at the fever plateau.*

3. *How does the body accomplish this increase?*

Try to solve the problems before looking up the answers..

Highlights

- ***The core temperature*** *is 37 ^oC in healthy adults at rest, but small children have larger diurnal variations.*
- ***Parents*** *should be advised to measure the rectal temperature in disease suspect children. The rectal temperature is a reliable estimate of the core temperature in resting persons.*
- ***A constant core temperature*** *does not guarantee heat energy balance, because the energy content of the shell may change with the surroundings.*
- ***The human core temperature*** *falls during the oestrogen phase of the menstrual cycle and during sleep. The lowest temperature is between 18 at night and 6 o'clock in the morning. Progesterone effects seem to explain the higher temperature in the last phase of the menstrual cycle*
- ***The temperature cycle*** *is part of the circadian periodicity. Our biological clock seems to be synchronised with the rotation of the globe. Also meals, light and temperature plays a role.*
- ***The thermal comfort point*** *for light-clothed, seated persons is about 26^oC when the humidity is 50%, 30^oC for nude persons, and about 36^oC sitting in water to the neck.*
- ***Vasodilatatio****n in warm environments permits increased skin bloodflow as the cardiac output also begins to rise. The arterio-venous anastomoses are dilatated, and the bloodflow*

can rise to at least 10 folds in hands and feet.

- *In the **metabolic zone**, the total heat loss rises with falling temperature, but below 5 ^{o}C the dry heat loss exceeds the metabolic rate, and the body is cooled down. This is the zone of hypothermia, where cold death is inevitable without therapy.*

- ***Sweat secretion**. Three million sweat glands produce sweat at a rate of up to 2 litres per hour or more during extreme conditions. Such high sweat rates lead to circulatory failure and shock. Sweat resembles a dilute ultrafiltrate of plasma.*

- ***The zone of hyperthermia** begins at an environmental temperature of 37 ^{o}C in humid air, where we soon reach the maximal capacity for sweat secretion, and there is a heat influx to the body ending in heat death.*

- ***The hypothalamus** functions as a thermostat using the rest of the body to stabilise its own temperature.*

- ***A dynamic gain system** responds continuously to feedback signals, regardless of the size of the core temperature. With rising tissue temperature, the neural activity of heat sensors increases linearly, whereas the activity of cold sensors decreases*

- ***A set-point system** does not respond to a rise in temperature before a certain set point is reached. The set point is the core temperature at which neither heat loss mechanisms or heat production mechanisms are active.*

- ***Human temperature control** exhibits both dynamic gain and set point characteristics. The control system implies widespread cutaneous and deep sensors. Their afferents converge towards the hypothalamic integrator and thermostat.*

- ***The thermogenic response** to cold also involves a non-myogenic or non-shivering component probably in adipocytes. Non-shivering heat production is controlled by the sympathetic nervous system via adrenergic β-receptors.*

- ***Non-shivering thermogenesis** is any rise in metabolism, which is not related to muscular activity. In babies this form of thermogenesis is particularly large in their brown adipose tissue. The brown adipose tissue of babies is abundant around vital organs, in neck and mediastinum, between scapulae and in the armpits.*

- ***Newborns** are in thermal balance at a minimal metabolism only when the surrounding temperature is high (32-34 ^{o}C).*

- ***Thermoregulatory behaviour** such as fanning and adding or removing clothing is effective in changing the thermal insulation. Several layers of clothing with trapped air are good insulators.*

- ***Heat acclimatization** is actually a sweat gland adaptation, and a 2-week process following arrival to a hot climate. Gradually sweat-evaporation is increased and the NaCl loss is reduced. The adaptation is caused by increased aldosterone secretion from the adrenal cortex.*

- ***Hypovolaemia** with low brain bloodflow through an overheated brain leads to fatigue (heat exhaustion), confusion and unconsciousness or syncope (heat syncope). The confusion may*

develop into a veritable delirium with brain oedema.

- ***Heat stroke*** *- in the sun called sunstroke - is heat collapse with high brain temperature that occurs suddenly, hereby creating a life-threatening condition.*

- ***Hypothermia*** *is a fall in core temperature to values below 35 ^{o}C. Hypothermic subjects lose consciousness, when the core temperature falls below 32 ^{o}C - a potentially lethal condition called severe hypothermia.*

- ***Artificially induced hypothermia*** *is used in brain- and heart-surgery, where the usual thermocontrol is inactivated by general anaesthesia.*

- ***Penetrating radiation*** *consists of either gamma rays (neutrons) or X-rays. Survivors from nuclear power plant accidents with whole body absorption greater than 100 rads are threatened by acute and chronic radiation sickness.*

- ***Acute radiation sickness*** *appears as vomiting and malaise following exposure to 1 Gy (100 rad) or more. Lymphocyte production is reduced immediately, soon followed by leucopenia and thrombocytopenia with bleeding. The villi of the gastrointestinal tract are destroyed, absorption of nutrients is impaired, and new villous cells are not produced. Diarrhoea, often with blood loss, results in dehydration. The skin is red and blistering, and the hair is loosed.*

- ***Chronic radiation sickness*** *or late radiation damage implies an increased rate of mutagenesis, which includes a high frequency of leucaemia, cancer of the brain, the thyroid and the salivary glands, infertility and cataract.*

Further Reading

Scientific American. Monthly journal published by Scientific American Inc., 415 Madison Avenue, N.Y., USA.

Benzinger, T. H. "The human thermostat." *Sci. Am.* 204: 134-147, 1961.

Hong, S. K. *News in Physiol. Sci.* 2: 79, 1987.

Wasserman, D.H. and A.D. Cherrington. "Hepatic fuel metabolism during muscular work: Role and regulation." *Am. J. Physiol.* 260: E811, 1991.

Return to top

Return to content

Chapter 22.

Gastrointestinal Function And Disorders

Study Objectives

- To *define* concepts such as achlorhydria, enterogastrones, haematemesis, incretins, macrolide, malabsorption, melaena, migrating motor complex, paracrine secretion, peptide hormone families, peptic ulcer disease, peristalsis, segmentation, slow waves, and spike potentials.

- To *describe* the extrinsic and intrinsic enteric nervous system including neurotransmitters and gastrointestinal hormones, cholesterol and lipid metabolism,

- To *explain* gastrointestinal motility, gastrointestinal secretion (saliva, gastric juice, pancreatic juice, bile), digestion and intestinal absorption of nutrients, vitamins, water and iron. To *explain* the pathophysiology of common gastrointestinal disorders including malabsorption of carbohydrate, amino acids and fat, osmotic and secretory diarrhoea, and iron deficiency.

- To *use* the above concepts in problem solving and case histories.

Principles

- ***The central autonomic nervous system*** *(hypothalamus and brain stem) mediates its influence on the gastrointestinal function through the intrinsic, enteric nervous system (the so-called "little brain").*

- ***Cannons law of the gut****: The peristalsis of the small intestine always proceeds in the oral- aboral direction.*

Definitions

- **Achlorhydria** refers to absence of HCl production in the stomach

- **Defaecation** is a reflex act involving colon, rectum, anal sphincters and many striated muscles (diaphragm, abdominal and pelvic muscles). The motor pathway is the pelvic nerves. Defaecation implies a temporal release of anal continence brought about by a reflex. The coordinating centre is in the sacral spinal cord.

- **Enterogastrones** are enterogastric inhibitory hormones liberated from the duodenal mucosa by acid chyme (ie, cholecystokinin: CCK, gastric inhibitory peptide: GIP, secretin, somatostatin, neurotensin and vasoactive intestinal peptide: VIP).

- **Haematemesis** is defined as vomiting of whole blood or blood clots.

- **Incretins** are hormones, which increase *insulin secretion* from the b-cells of the pancreatic islets much earlier and to a greater extent, than when the blood glucose concentration is elevated by intravenous infusion (GIP, glicentin, glucagon-like peptides-1 and -2).

- **Intrinsic, enteric nervous system** refers to the large number of neuronal connections in the gut wall, in particular the *submucosal Meissner plexus*, which regulates the digestive glands, and the *myenteric Auerbach plexus*, primarily connected with gut motility.

- **Macrolides** are antibiotics, which bind to and prevent translocation on bacterial ribosomes.

- **Malabsorption** describes the condition resulting from inefficient absorption of nutrients by the gastrointestinal tract.

- **Melaena** is defined as passage of dark tarry stools (coal-black, shiny, sticky, and foul smelling).

- **Migrating motor complex** refers to a gastric sequence of events, where contractions occur each 90 min during fasting. There is a quiet period (I) followed by a period of irregular contraction (II), and culminated with a *peristaltic rush* (III) accompanied by increased gastric, pancreatic and biliary secretion.

- **NANC neurons** are non-adrenergic, non-cholinergic postganglionic neurons, which liberate gastrin-releasing peptide (GRP) to the gastrin producing G-cells.

- **Nitric oxide** (NO) is a possible neurotransmitter between the preganglionic and the NANC postganglionic neurons.

- **Paracrine secretion** is the release of signal molecules to neighbour cells.

- **Peptide hormone families** are groups of hormones that exhibit *sequence homology:* They possess a common amino acid sequence, such as the *gastrin family*, which has sequence homology in their terminal penta-peptide. *Peptide hormones* have *autocrine* and *paracrine* functions in the gastrointestinal tract.

- **Peristalsis** is a propagating contraction of successive sections of circular smooth muscle preceded by a dilatation. The dilatated intestinal wall is drawn over its content in this reflex mechanism, which transports the content aborally and is called the *law of the gut.*

- **Segmentation** divides the small intestine into many segments by localised circular smooth muscle contractions. Segmentation mixes the intestinal content and propagate it at a slow rate, which allows sufficient time for digestion and absorption.

- **Slow waves** (*basic electrical rhythm*) are slow gastrointestinal depolarisation's occurring at a frequency of 3-18 per min. The slow

waves change the resting membrane potential of smooth muscles from -50 to -40 mV.

- **Spike potentials** are *periodic fast waves of depolarisation* that most often follow a slow wave, and then always initiate gastric contractions (elicited by a rise in cytosolic [Ca^{2+}]).

- **Vaso-active intestinal peptide** (VIP) is a vasodilatator in line with adenosine, ATP, NO. The increased bloodflow increases intestinal secretion.

Essentials

This paragraph deals with *1. The autonomic and enteric nervous system, 2. The cephalic, gastric and intestinal digestive phase, 3. Mastication and swallowing, 4. Gastric and intestinal motility, 5. Vomiting, 6. Colonic motility and defecation, 7. Gastrointestinal hormones, 8. Saliva, 9. Gastric secretion, and 10. Intestinal digestion and absorption.*

1. The autonomic and the enteric nervous system

The digestive system is innervated with nerve fibres of both the sympathetic and parasympathetic divisions, although the *parasympathetic control* dominates (Fig. 22-1). Movements of the gastrointestinal tract are brought about by smooth muscle activity. There is an outer longitudinal layer, an inner circular layer, and a submucosal muscle layer (muscularis mucosae) with both circular and longitudinal fibres that moves the villi of the mucosa. The inner surface is lined with mucosal epithelium (Fig. 22-1). The outer muscle layer is covered by the serosa, which is continuous with the mesentery containing blood vessels, lymph vessels and nerve fibres.

The main CNS centres regulating digestive functions are located in the brain stem, where the sensory taste fibres from gustatory, tactile and olfactory receptors terminate on the cell bodies of the motor vagal and salivary nuclei. Many afferent, sensory fibres in the vagus nerve inform the *central autonomic system* about the condition of the gut and its content. The higher cortical and olfactory centres influence these *brain stem motor centres* and their parasympathetic outflow.

The parasympathetic system *increases* digestive activity (secretion and motility), and the sympathetic system has a net *inhibitory* effect. The generally inhibitory digestive effects of the *sympathetic nervous system* are caused indirectly by vasoconstriction, which reduces bloodflow in the digestive tract.

The vagus nerve innervates the gastrointestinal tract down to the transverse colon and contains both efferent and afferent fibres. The last part of the gastrointestinal tract receives parasympathetic innervation from the pelvic nerves.

The efferent parasympathetic fibres enhance digestive activities by stimulating local neurons of the *intrinsic, enteric nervous system* located in the gut wall (Fig. 22-1).

Fig. 22-1: The autonomic innervation of the gastrointestinal system and the structure of the enteric wall. – A sensory neuron to the CNS is shown to the left.

The intrinsic, enteric nervous system consists of two sets of nerve plexi. The *submucosal Meissner plexus* mainly regulates the digestive glands, whereas the *myenteric Auerbach plexus*, located within the muscle layers, is primarily connected with gut motility (Fig. 22-1). The nerve plexi contain local sensory and motor neurons as well as interneurons for communication. Motor neurons in the myenteric plexus release acetylcholine and Substance P. Acetylcholine contracts smooth muscle cells, when bound to muscarinic receptors. Inhibitory motor neurons release *vasoactive intestinal peptide* (VIP) and *nitric oxide* (NO). These molecules relax smooth muscle cells.

Sensory neurons are connected to mucosal chemoreceptors, which detect different chemical substances in the gut lumen, and to stretch receptors, which respond to the tension in the gut wall, caused by the food and chyme. The short effector neurons increase digestive gland secretion and induce smooth muscle contraction. The large number of neuronal connections constitutes the intrinsic, enteric nervous system, mediating brain influence on digestive functions. The enteric nervous system is also called the *little brain*.

2. The cephalic, gastric and intestinal digestive phase

The secretion related to a meal occurs in three phases (Box 22-1).

2a. The cephalic phase is elicited even before food arrives to the stomach. The thought, smell, sight, or taste of food signals to the limbic system (including the hypothalamus) that elicits an unconditioned reflex secretion with intensity dependent upon the appetite.

Box 22-1: The secretion related to a meal from salivary, gastric and exocrine pancreatic glands.	
Cephalic phase	
	Unconditioned reflexes secrete saliva, gastric and pancreatic juice
	Conditioned reflexes (the thought of food) also.
Gastric phase (distension of the stomach)	
	Vagal reflexes - cholinergic, muscarinic receptors
	Intrinsic peptidergic neurons (VIP, GRP)
	Histamine
	Somatostatin (multipotent inhibitor)
	Gastrin
Intestinal phase	
	Gastrin from duodenal G-cells increases gastric secretion
	Secretin (from S-cells) and bulbogastrone inhibit gastrin-stimulated acid secretion
	Cholecystokinin (CCK) and gastric inhibitory peptide (GIP) inhibit gastrin release from G-cells, and acid secretion by the parietal cells

	All these entero-gastric *inhibitory* hormones are called *enterogastrones*

2b. The gastric phase

is brought about when food enters and distends the stomach. Distension stimulates stretch receptors and peptide sensitive chemoreceptors. They provide afferent signals for both long, central vago-vagal reflex loops as well as local, enteric reflexes. Signals in these fibres reach cholinergic, muscarinic receptors on the basolateral membrane of the parietal cells.

Distension of the body of the stomach can release gastrin from the antral mucosa by vagal reflexes. Most of the daily gastric secretion of 1.5 l is accounted for by the gastric phase.

2c. The intestinal phase is elicited by duodenal and jejunal mechanisms that both stimulate and inhibit gastric acid secretion. Gastric secretion and motility are at first increased to promote further digestion and emptying. This fills the duodenum with acidic and fatty chyme. Acid chyme reaching the duodenum with peptides and amino acids releases gastrin from duodenal G-cells, which increases gastric secretion. Normally, the inhibitory intestinal mechanisms dominate, when the pH of the chyme is low. Acid chyme in the duodenum causes release of secretin (from S-cells) and of bulbogastrone (Box 22-1).

3. Mastication and swallowing

The process of chewing or *mastication* requires co-ordination of the chewing muscles, the cheeks, the palate and the tongue. Chewing is normally a reflex action. The forces involved in grinding and cutting the food are enormous, and sufficient to fragment cellulose membranes. Finally, the food is mixed with saliva and formed into a bolus. The bolus is pushed back into the pharynx, when the tongue is pressed against the hard palate.

Fig. 22-2: Swallowing of a food bolus in three steps (OES stands for the upper Oesophageal sphincter).

The gastrointestinal tract moves ingested materials and secretions from the mouth to the anus. These movements, as well as nonpropulsive contractions, are called *motility*.

Gastrointestinal sphincters possess adrenergic a_1-receptors. Stimulation of these receptors results in contraction.

Swallowing (*deglutition*) begins as a voluntary process by which the tongue pushes a portion of the food back against the soft palate (Fig. 22-2). Elevation of the soft palate closes the nasopharynx, and the food enters the pharynx, the larynx is elevated closing the epiglottis and respiration stops. The upper pharyngeal constrictor contracts, initiating sequential contractions of the other pharyngeal constrictors. These contraction waves are involuntary and push the food towards the oesophagus. Peristalsis in the oesophagus is started as the pharyngeal wave passes through the upper oesophageal sphincter (Fig. 22-2). When the propulsive wave reaches the lower oesophageal sphincter (LES), the relaxed muscle wall preceding the bolus momentarily relaxes the LES, and the food passes the cardia to enter the stomach. Vagal stimulation relaxes both sphincters (see achalasia, below).

The upper third of the oesophagus is composed of striated muscle, the middle third contains mixed smooth and striated muscle, and the lower third contains only smooth muscle.

Swallowing is controlled by brainstem neurons. They form a swallowing centre (Fig. 22-2). The vagus nerve contains both somatic motor neurons (originate in the nucleus ambiguus) that form motor endplates on striated muscle fibres, and visceral, preganglionic motor neurons (from the dorsal motor vagal nucleus to the myenteric plexus). The swallowing reflex coordinate motor signals from both oesophageal striated and smooth muscles as well as signals to the upper and lower oesophageal sphincters.

Sympathetic stimulation contracts the LES mediated by noradrenaline acting on a-receptors. When a swallow is initiated via touch receptors in the pharynx, or when the lower oesophagus is distended by a bolus, it will relax the LES by reflexes in inhibitory vagal fibres joining the enteric nervous system. VIP and NO act as transmitters.

4. Gastric and intestinal motility

In the stomach, digestion continues (*salivary amylase*) and the stomach regulates emptying of its content into the duodenum. The fundus has a *high compliance*, so food can accumulate without much increase in gastric pressure. Vagal fibres releasing VIP to inhibitory neurons of the myenteric plexus mediate this receptive relaxation. The body of the stomach mixes and grinds the food with gastric juice - also by *retropulsion* (backward or oral movement) - and then propels the content toward the antrum and pyloric region for regulated emptying. The distal stomach reduces solids to a fluid consistently composed of particles less than 2 mm. Here is a *forceful peristalsis* (ie, propagating contractions), so the pyloric sphincter opens and the chyme is ejected into the duodenum (Fig. 22-3).

Fig. 22-3: Intestinal smooth muscle potentials (left) and contractions (right).

Along the greater curvature of the stomach is a region of rapid spontaneous depolarization, which is called the *gastric pacemaker* establishing the maximum rate of gastric contractions. The gastric smooth muscle wall generates two types of electrical activity. *Slow waves* (basic electrical rhythm) are *slow depolarisation's* occurring at a frequency of three in the stomach, up to 18 in the duodenum and 8 per min in the terminal ileum. The slow waves are oscillations of the resting membrane potential (Fig. 22-3). Voltage-gated (potential sensitive) Ca^{2+}-channels open at a certain threshold of depolarization, causing a Ca^{2+}-influx to the smooth muscle cell resulting in the so-called spikes and contractions. *Spikes* are *periodic fast waves of depolarisation* that always initiate gastric contractions, elicited by the rise in cytosolic [Ca^{2+}]. These contractions last up till 3 s, because the Ca^{2+} -channels open slowly and remain open longer than the Na^{+} -channels. Spikes are elicited by vagal signals, by acetylcholine (muscarinic receptors), by stretch, by myenteric signals and by gastrin (Fig. 22-3). Adrenaline and noradrenaline relax smooth muscle by hyperpolarization through α*-adrenergic receptors.* Relaxation occurs when intracellular Ca^{2+} is returned to the extracellular fluid and to the endoplasmic reticulum.

The small intestine is about 8 m long and commonly divided into three

segments: the duodenum, jejunum and ileum. The intestinal contents must be moved in a manner that brings them into contact with the mucosa of the intestine, and propels the contents along this tubular organ. Several pacemaker regions in the small intestine control the slow waves. The pacemaker rate is highest in the duodenum (about 18 each minute), and decreases down to 8 waves each min in the terminal ileum.

During fasting, a migrating sequence of events called the *migrating motor complex* occurs each 80-90 min. The complex consists of an 80-90 min long quiet period (I) followed by a period of irregular propulsive contractions (II), culminating in a *peristaltic rush* (III) to begin in the stomach, accompanied by increased gastric, pancreatic and biliary secretion. The migrating motor complex is the "*intestinal housekeeper*", which cleanses the digestive tract of non-absorbable substances, and provides an effective emptying of the tract all the way.

During the fed state, *segmentation* serves to mix chyme with enzyme-containing digestive fluid, and brings the mixture into contact with the mucosal surface for absorption. Segmentation divides the small intestinal content into many segments by localised circular smooth muscle contractions with only a small propulsive effect (Fig. 22-3).

Propulsive motility is accomplished by *peristalsis.* Peristalsis is a propagating contraction of successive sections of circular smooth muscle preceded by a dilatation (Fig. 22-3). The dilatated intestinal wall is drawn over its content in this reflex mechanism, which has been called the *law of the gut*. Peristaltic contractions usually travel along a small length of the small intestine, except for the *peristaltic rush* related to the migrating motor complex.

The *ileocoecal sphincter* prevents retrograde flow of colonic matter. The sphincter regulates emptying of ileum five hours after a meal. The emptying of ileum is stimulated by *gastrin,* possibly via the *gastro-ileal reflex,* but a distended colon inhibits the emptying. The gastro-ileal reflex is an increased motility of the terminal ileum caused by elevated gastric activity. On the other hand, distension of the terminal ileum decreases gastric motility. The ileocoecal sphincter is normally passed by *one litre* of faecal matters per day.

5. Vomiting

The feeling of nausea, and an array of sympathetic and parasympathetic responses initiate vomiting or emesis. *Sympathetic responses* include sweating, pallor, increased respiration and heart rate and dilatation of pupils. *Parasympathetic responses* include profuse salivation, pronounced motility of the oesophagus, stomach, and duodenum, relaxation of the oesophageal sphincters. Duodenal contents can be forced into the stomach by anti-peristalsis (Fig. 22-4). During the expulsion of gastric contents, the person *takes a deep breath*, the pylorus is *closed*, the glottis is *closed* so respiration stops, and the stomach is *squeezed* between the diaphragm and the abdominal muscles, causing rapid emptying (Fig. 22-4). Vomiting is co-ordinated by the *vomiting centre* in the medulla.

Fig. 22-4: Vomiting co-ordinated by the vomiting centre.

Vomiting is stimulated in certain areas of the brain (hypothalamus) and the cerebellum through sensory stimuli or injury. Vomiting is also provoked by

certain labyrinthine signals, and from the chemoreceptive *trigger zone* located on the floor of the 4th ventricle close to *area postrema*.

During deep anaesthesia the vomiting and swallowing mechanisms are *paralysed*. Any patient must abstain from food and water for at least six hours before deep anaesthesia is administered. Otherwise, the patient may vomit into the pharynx, and suck his own vomit into the trachea. Over the years, many patients have choked to death due to this mechanism. The survivors develop *aspiration pneumonia*. Such events are clearly malpractice.

The swallowing mechanism is also cut-off by injury of the 5th, 9th, or 10th cranial nerve, by *poliomyelitis*, by *myasthenia gravis* and by *botulism* (Chapter 33).

An acute loss of H^+ from the extracellular fluid (ECF) by vomiting creates a *metabolic alkalosis* (high pH with high Base Excess, see Chapter 17).

6. Colonic motility and defaecation

Colonic transit is measured in *days*. Mixing occurs in the ascending colon, because peristalsis is followed by anti-peristalsis. *Slow waves* of contraction move the content in the oral direction to delay propulsion and increase absorption of water and electrolytes. Colonic segmentation is a mixing of the content by regular segments called *haustrae*. Prominent haustration along the length of the colon is characteristic for the X-ray image of the normal colon. The colon provides an optimal environment for bacterial growth. *Peristaltic rushes* in the colon occur several times per day. They often start in the transverse colon as a *tight ring*, continuing as a *long contraction wave*. Gastro-colic and duodeno-colic reflexes assisted by gastrin and by cholecystokinin (CCK) promote peristaltic rushes**.**

Defaecation is a complex act involving both voluntary and reflex actions in colon, rectum, anal sphincters and many striated muscles (diaphragm, abdominal and pelvic muscles). Defaecation is a temporal release of anal continence brought about by a reflex. The rectum is usually empty, and its wall has a rich sensory supply. Distension of the recto-sigmoid region with faecal matter releases awareness of the urge to defaecate, an *intrinsic defaecation reflex*, and a *strong, spinal reflex*. There is a reflex contraction of the descending colon and the recto-sigmoideum.

The smooth internal *anal sphincter muscle* maintains a tonic contraction during continence, due to its sympathetic fibres from the *lumbar medulla* (through hypogastric nerves and the inferior mesenteric ganglion). The muscle relaxes due to its parasympathetic, cholinergic fibres in the pelvic splancnic nerves (S_2-S_4). The strong spinal reflex produces relaxation of the smooth muscles of the internal anal sphincter (Fig. 22-5) and contraction of the striated muscles of the external anal sphincter (innervated by somatic fibres in the pudendal nerve) inhibiting the reflex and causing *receptive relaxation*. This is the last decision - before defaecation.

Fig. 22-5: Defaecation reflexes.

The *levator ani muscle* contributes to the closure of anus, because contractions increase the angle between the rectum and the anus.

Destruction of the lower sacral medulla (the *defaecation centre*) destroys the spinal reflex and thus the normal defecation. Higher spinal lesions

destroy the voluntary control, whereas the defaecation reflexes persist. An acceptable status is obtainable in paraplegics by mechanical release of the reflex (manual expansion of the external sphincter) once daily following a meal.

7. Gastrointestinal hormones

Gastrointestinal hormones are peptides secreted by the gastrointestinal mucosa, and controlling all gastrointestinal functions together with other hormones and transmitters. As an example insulin works together with acetylcholine and parasympathomimetics to *stimulate* secretion and motility, whereas catecholamines, sympatomimetics and parasympatolytics, such as atropine, *inhibit* gastrointestinal secretion and motility.

Peptide hormone families are groups of regulatory peptides that exhibit *sequence homology* (ie, they possess a common amino acid sequence). The *gastrin-family* and the *secretin-glucagon family* are the most important.

7a. The gastrin family

consists of gastrin and cholecystokinin (CCK) in three different forms (CCK-8, CCK-22, and CCK-33). Gastrin and CCK release *pancreatic glucagon* from the islet cells. There are two major forms of gastrin in the plasma, *normal gastrin* or *G-17* and *big gastrin* or *G-34*. They are 17 and 34 amino acid polypeptides, respectively. Gastrin is produced by G-cells of the gastric antrum and duodenum. The duodenal Brunner glands secrete half of the G-34.

Gastrin is the strongest stimulator of gastric acid secretion. Gastrin also imposes *tropic* (growth-stimulating) actions on the parietal cells, the mucosa of the small and large intestine and possibly the pancreas. Gastrin stimulates the *pepsin secretion* from peptic cells, and the *glucagon secretion* from the a-cells of the pancreatic islets.

Gastrin is derived from *parietal* or *oxyntic* cells in the stomach. When stimulating gastric acidity, gastrin relaxes the gastric muscles, thus retarding the passage of chyme into the duodenum.

Feeding induces the secretion of gastrin to the interstitial fluid and then to the blood. Neural signals pass through the vagal nerve to the gastrin-secreting *G-cells* of the gastric antrum and duodenum (Fig. 22-6). The afferent input begins with the smell and taste of food, and is reinforced by vago-vagal reflexes elicited by oesophageal and gastric distension. *Digested protein* (polypeptides and amino acids) act directly on G-cells.

Fig. 22-6: Gastric HCl secretion following feeding. GRP: Gastrin Releasing Peptide. NANC: Non-adrenergic, Non-cholinergic postganglionic neurons.

Vagal, cholinergic preganglionic fibres transfer signals to the G-cells via *non-adrenergic, non-cholinergic* (NANC) postganglionic neurons. These enteric neurons liberate *gastrin-releasing peptide* (GRP) to the gastrin producing G-cells. The gastrin released reaches the parietal cells through the blood and increases the HCl secretion. GRP thus releases gastrin and hereby stimulates the secretion of gastric acid. - GRP consists of 27 amino acid moieties and is also released from neurons in the brain.

An indirect vagal route to the G-cells is via *postganglionic cholinergic enteric neurons* to *somatostatin cells* that are located close to the G-cells

(Fig. 22-6). When these enteric neurons release acetylcholine, the response of the somatostatin cells is inhibition of somatostatin release. Somatostatin inhibits G-cell secretion by paracrine action. The result of both vagal inputs to the G-cells is *gastrin release* (Fig. 22-6). An elevated [H^+] in the duodenal lumen inhibits gastrin release.

Cholecystokinin, CCK, according to its function and structure, belongs to the *gastrin family*. Cholecystokinin empties the gall bladder as the name implies, and stimulates pancreatic secretion of an enzyme rich juice. However, CCK has a higher affinity for receptors stimulating gallbladder contraction and pancreatic enzyme secretion. CCK has a maximal effect only in the presence of secretin (potentiation) and normal vagal influence.

Both gastrin and CCK release glucagon from the a-cells of the pancreatic islets.

CCK is cleaved from *pre-pro*-CCK in the duodenum, upper jejunum (I-cells) and in the *brain*. CCK molecules consist of a group of peptides. CCK-8, CCK-22 and CCK 33 are the dominant forms in the blood.

The most important stimulus for CCK liberation is amino acids and fatty acids, which reach the duodenal mucosa. Bile is ejected into the duodenum, where fat is emulgated to ease its absorption. CCK also acts as an *enterogastrone* - an intestinal hormone that *inhibits* gastric activity and emptying. This leaves more time for the bile to emulgate fat.

7 b. The secretin-glucagon family

Secretin exhibits sequence homology with pancreatic glucagon, vasoactive intestinal peptide (VIP), growth hormone-releasing hormone (GHRH) and gastric inhibitory polypeptide (GIP). A family of five genes code for these five hormones.

Secretin is secreted by S-cells in the mucosa of the upper small intestine, when acid chyme (pH below 4.5) arrives to the first part of the duodenum. Fatty acids from fat digestion also contribute to secretin release.

Secretin stimulates the secretion of bicarbonate and water by pancreatic duct cells, and of bicarbonate-rich aqueous bile. Secretin potentiates the action of CCK including an *enterogastrone effect* (gastric inhibiting effect). Secretin antagonises gastrin - and potentiates CCK. Secretin is an enterogastrone that is released by H^+ to stimulate pancreatic juice secretion.

Gastric inhibitory polypeptide (GIP or Glucose-dependent Insulin releasing peptide) works as the two names imply: GIP inhibits the gastric mucosa and releases insulin from the a-cells of the pancreatic islets.

Glucagon is actually two different molecules: *Intestinal* glucagon (*glicentin*) and *pancreatic* glucagon. Both are hepatic insulin-antagonists. Glucagon stimulate glycogenolysis, gluconeogenesis (urea genesis-glycogenic amino acids), and ketogenesis.

The function of other peptide hormones is given in Box 22-2.

Box 22-2: Effects of some gastrointestinal hormones and transmitters.

	Duokrinin stimulates duodenal secretion.

Endogenous (*enkephalins*) and exogenous opiates inhibit ganglionary transmission.

Enterokrinin stimulates secretion in the small intestine.

Gastrin releasing peptide (GRP) and **bombesin** release gastrin from G-cells.

Glicentin (intestinal glucagon) stimulates insulin secretion as other incretins.

Motilin stimulates gastrointestinal motility.

Neuropeptide Y and neurotensin stimulate neurotransmission.

Nitric oxide (NO) is a possible neurotransmitter between the preganglionic and the NANC postganglionic neurons.

Pancreatic Polypeptide (PP) from the PP-cells inhibits pancreatic and biliary secretion, which delay the absorption of nutrients. PP is released by meals.

Pancreotonin: Inhibits the pancreatic exocrine secretion.

Somatostatin (Growth hormone-inhibiting hormone, GHIH; 14 amino acid moieties) is a strong, universal inhibitor - both blood-born and paracrine.

Substance P (11 amino acid residues) stimulates smooth muscle contraction and thus the gastrointestinal motility.

Vasoactive intestinal peptide (VIP; 28 amino acid residues; vessel wall and brain neurons) is a vasodilatator in line with adenosine, ATP, and NO. The increased bloodflow increases intestinal secretion. VIP is also involved in penile erection and in bronchiolar dilatation.

Villikrinin: Stimulates the rhythmic movement of villi in the intestine.

Traditionally, the important peptides are also divided into two *functional* groups: *Enterogastrones inhibit* gastric motility and secretion. When gastric acid, fats, and hyperosmolar solutions have entered and distended the duodenum, GIP and other enterogastrones (somatostatin, CCK, and secretin) are released and suppress gastric acid secretion and motility of the stomach

Incretins stimulate insulin secretion. Incretins are liberated to the blood as gastric chyme enters the duodenum - and before the glucose of the chyme can be absorbed. Incretins increase *insulin secretion* from the b-cells of the pancreatic islets much earlier and to a greater extent, than when the blood [glucose] is elevated by intravenous infusion. Incretins are GIP, glicentin, and glucagon-like peptides: GLP-1 and -2.

8. Saliva

Saliva is a watery solution of electrolytes (bicarbonate and K^+) and organic substances, which is a mixture of secretions from three pairs of glands. The *parotid* is the largest and serous (watery saliva), the *sublingual* is mucous (viscous, containing mucin), and the *submandibular* salivary gland is build of mucous acini surrounded by serous *half moons*. The *primary* saliva is produced in the acini, but *secondary* processes in the salivary ducts (secretion and reabsorption) are involved in the final saliva production. Salivary glands have a high bloodflow and produce up to one l of saliva daily. The maximal secretion rate is one ml of saliva per g salivary tissue per min (ie, 60 times that of pancreas).

Salivary *mucin* (a glycoprotein) and *water* lubricate food, dissolve particles, and salivary enzymes initiate digestion. Ptyalin or a-amylase cleaves a-1-4 glycoside bindings in starch. *Salivary buffers* maintain the pH-optimum (6.8) of amylase during the first period in the stomach. The saliva dilutes injurious agents.

Saliva cleans the mouth and pharynx (prevents *caries*), and ease swallowing. Salivary *lysozyme* lyses bacterial cell walls. The salivary *epidermal growth factor* promotes the healing of wounds. Animals instinctively lick their wounds. Saliva contains immuno-defensive secretory globulin A (IgA), amino acids, urea, and blood-type antigens in secreting persons. Saliva may inactivate human immunoactive virus (HIV). The most common infection of the salivary glands is *acute parotitis* caused by the mumps virus.

The virus causing infectious mononucleosis is probably transferred with saliva by "deep kissing". *Infectious mononucleosis* is a disease characterised by lympadenopathy, lympho-cytosis and duration longer than an ordinary tonsillitis. The condition is dangerous, because spontaneous rupture of the spleen occurs.

Salivary secretion is controlled by the autonomic nervous system, and minimally influenced by hormones. Unconditioned reflexes (taste-, olfactory- and mechano-receptors) control salivation as well as *conditioned reflexes* (the thought of food). These signals reach the *brain stem salivary centres*, which activate the parasympathetic nerves to the salivary glands. The *primary* salivary secretion into the acini resembles an ultrafiltrate of plasma, but the final saliva is hypotonic.

Parasympathetic, cholinergic fibres, originating in the *salivary nuclei* of the brain stem, synapse with postganglionic neurons close to the secretory cells. These neurons transmit signals to the *cholinergic, muscarinic receptors* (Fig. 22-7). Parasympathetic activity can release maximal salivary secretion and bloodflow resulting in a amylase-rich saliva with mucin (glycoproteins). *Atropine* blocks the muscarinic, cholinergic receptors (during anaesthesia where the mouth becomes dry). The rise in bloodflow is atropine-resistant and caused by the vasodilatating VIP, which is released from peptidergic nerve terminals that also contain acetylcholine. *β_1-adrenergic agonists* and VIP elevate cAMP in the acinar cells, an effect potentiating the secretory effect of acetylcholine. The vascular smooth muscle relaxation by VIP is probably also mediated via cAMP.

Fig. 22-7: Salivary enzymes, ions and mucin production from two

acinar cells. Solid and dashed arrows indicate active and passive transport, respectively. Circles are carrier molecules, whereas tubes symbolise transport channels. - To the left is shown receptors and second messengers.

1. Neural or humoral (acetylcholine) stimulation of cholinergic, muscarinic receptors on the basolateral membrane of acinar cells leads to a rise in intracellular [Ca^{2+}].

2. This rise triggers luminal Cl^- - and basolateral K^+ -channels. Hereby, K^+ is transferred to ISF and Cl^- to the acinar lumen in a balanced relationship (Fig. 22-7). Therefore, Cl^- flows down its electrochemical potential gradient into the lumen of the acinus. K^+ flows down its gradient to the ISF through activated channels. These ion flows create a negative electric field in the lumen.

3. The initial fall in intracellular [K^+] increases the driving force of the electroneutral Na^+-K^+-$2Cl^-$ co-transporter to transport two Cl^- into the cell together with Na^+ and K^+. Thus the electrochemical potential of Cl^- and K^+ is greater in the cell, than in the interstitial fluid (ISF) and in the saliva.

4. The negative field provides an electric force that drives a passive Na^+ flux into the acinar lumen through leaky tight junctions. Osmotic water transport through leaky junctions and trans-cellularly through water channels in the cell membranes follow the NaCl flux into the lumen. The trans-cellular Cl^- transport is coupled to the paracellular Na^+ transport. The net result is an isosmotic NaCl transport produced by a secondary active Cl^-- secretion.

5. The basolateral membranes of acinar cells contain a Na^+-K^+-pump that provides the energy for the primary salivary secretion (Fig. 22-7). The rise in intracellular [Na^+] from 2., activates the Na^+-K^+-pump, whereby [Na^+] is kept almost constant. Ouabain inhibits salivary secretion, because it blocks the pump.

Sympathetic nerve signals, and circulating catecholamines via b-adrenergic receptors, inhibit the bloodflow and the secretion of serous saliva (b_1-receptors in Fig. 22-7). A small, transient, mucous secretion with a high [K^+] and [bicarbonate], and a low [Na^+] is produced, because of the low secretion rate. Noradrenaline (NA) stimulates both a_1-adrenergic and b_1-adrenergic receptors. Binding of NA or b-adrenergic agonists elevates intracellular cAMP, which correlates with a small increase in primary salivary secretion. This explains why the mouth becomes dry during events, where the sympathetic system dominates (anxiety, excitement etc).

The salivary ducts are almost watertight. Therefore, the final salivary flow is dependent upon the primary salivary secretion rate in the acini.

The duct systems, in particular the small-striated ducts with a substantial O_2 consumption reabsorb large amounts of Na^+ and Cl^-, whereas bicarbonate and K^+ are secreted. Saliva becomes more and more hypotonic at low

secretion rates, because the Na^+ and Cl^- reabsorption dominate.

1. The reabsorption of Na^+ and the secretion of K^+ are processes stimulated by the mineralo-corticoid, aldosterone. Aldosterone stimulates Na^+-influx through the luminal Na^+-H^+-exchanger (Fig. 22-8). Na^+ enters the cell in exchange with H^+. The resulting intracellular rise in $[Na^+]$ activates the basolateral Na^+-K^+-pump. Thus, Na^+ is reabsorbed trans-cellularly from the salivary duct. The pump maintains the electrochemical potential gradients of Na^+ and K^+.

2. The Cl^- follows passively, and is partly exchanged with bicarbonate along the duct system through a luminal Cl^- - bicarbonate exchanger (Fig. 22-8). The secretion of bicarbonate is so great that its concentration in the final saliva exceeds that in plasma.

3. At the basolateral membrane Cl^- leaves the cell via an electrogenic Cl^- channel, while Na^+ is pumped out.

4. K^+, taken up by the Na^+-K^+-pump, leaves the cell through K^+-channels in the basolateral membrane, recycling K^+ to balance the Cl^- efflux.

5. Some of the K^+ leaves the cell by luminal H^+-K^+-exchange. At low secretion rates the H^+-K^+-exchanger (antiport) in the luminal membrane transfers sufficient K^+ for the $[K^+]$ in the final saliva to exceed the concentration in plasma. The net result is K^+-secretion from blood to the duct lumen.

The final salivary $[Na^+]$ and $[Cl^-]$ increase with increasing salivary secretion rate, because the high flow provides less time for reabsorption in the duct system. Bicarbonate may be secreted even without Cl^--reabsorption. At low salivary secretion rates the final saliva becomes hypotonic down toward half of the osmolarity of plasma.

Fig. 22-8: Secretion from salivary duct cells.

The aldosterone effects described above (increased Na^+ reabsorption and increased K^+ secretion) are similar to those in the distal, renal tubules and in the sweat glands.

9.Gastric secretion

The stomach is divided into three main regions: the fundus, corpus and pyloric antrum. The gastric mucosa is highly invaginated and is mainly composed of gastric glands, with mucous neck cells, parietal cells secreting HCl, and peptic (chief) cells secreting pepsinogen. The parietal cells also secrete the peptide intrinsic factor, which is necessary for absorption of vitamin-B_{12}. G-cells in the mucosa produce the hormone gastrin (Fig. 22-6). The gastric secretions include hydrochloric acid (HCl), pepsin and basic mucus, which contains mucin (glycoproteins) and salts.

Efferent signals from the dorsal motor nuclei of the vagi stimulate gastric motility and HCl production. Acetylcholine is released from the short

postganglionic vagal fibres and directly stimulates parietal cells to secrete HCl. The parietal cells contain muscarinic receptors on the basolateral membrane. Vagal fibres work together with intrinsic, peptidergic neurons containing vasoactive intestinal peptide (VIP) and gastrin releasing peptide (GRP). VIP controls the bloodflow of the gastric mucosa; GRP releases the important gastrin from the antral G cells and the peptic cells secrete pepsinogen.

The secretion related to a meal occurs in three phases (cephalic, gastric and intestinal).

The gastric juice is hyperosmotic (325 mOsmol/l), contains 10 mM of K^+ and is low in Na^+ at moderate and high secretion rates; the $[H^+]$ is 170 mM and the $[Cl^-]$ is 180 mM. Gastric juice has an approximate pH of 1, forming a million-fold gradient of H^+ across the gastric mucosa to the blood. The HCl activates pepsinogen, maintains the optimal pH for pepsin activity and denatures proteins and microbes.

The peptic cells, located in the base of the gastric gland, produce pepsinogen. Pepsinogen is stored in granules of the peptic cell. Pepsinogen secretion is stimulated by cholinergic, muscarinic substances and by b-adrenergic agents, but peptic cells have no histamine receptors. Exocytosis releases pepsinogen into the gastric juice, where it is cleaved into pepsin, if HCl is present. Pepsin is the major hydrolytic enzyme in the stomach, but it is only active in the acidic gastric juice.

Fig. 22-9: Secretion of parietal and non- parietal cell juice.

Adult humans produce up to two l of gastric juice daily. The gastric juice is produced from two different sources: The *parietal cell juice* with 170 mM [HCl], 10 mM $[K^+]$, and a low $[Na^+]$. A juice with an ionic composition similar to that of plasma is produced from other cells - the *non parietal juice*. Each of the two secretion products has almost a constant composition.

Increased secretion of gastric juice means increased secretion of parietal cell juice. This explains why the [HCl] increases more and more in the mixed product, whereas $[Na^+]$ falls with increasing secretion rate.

Fatty chyme entering the duodenum delays gastric emptying by negative feedback through duodenal reflexes and by the release of gut inhibiting hormones (so-called enterogastrones: somatostatin, VIP, gastric inhibitory peptide, GIP, neurotensin and secretin). These inhibitors not only inhibit gastric motility; they also inhibit the gastrin release from the antral G cells, and also the HCl production from the parietal cells. Mucus contains mucin (glycoproteins) and electrolytes with bicarbonate that protect the gastric mucosa from adversive effects.

Stimulation of the parietal cells with acetylcholine, histamine and gastrin has two consequences for their content of second messengers (Fig. 22-10, right). The cellular $[Ca^{2+}]$ and [cAMP] is elevated.

Fig. 22-10: HCl secretion from parietal cell in the stomach (left). Secretory receptors on the parietal cell are also shown (right).

1. These second messengers activate luminal Cl^-- and K^+-channels. Cl^- and K^+ pass into the lumen, whereby their cellular concentrations

decrease (Fig. 22-10 left). The luminal [K] activates the K -H - pump. In addition, more pumps are inserted into the luminal membrane from cellular tubulo-vesicles.

2. The fall in cellular [Cl^-], and a rise -see below - in cellular [bicarbonate], stimulates the basolateral Cl^--bicarbonate exchanger, whereby the cellular [bicarbonate] is reduced. The fall in cellular [H^+] and [bicarbonate] stimulates formation of H^+ and bicarbonate, under the influence of carbo-anhydrase (*). The H^+ and bicarbonate are derived from metabolic carbon dioxide from the blood. Bicarbonate diffuses from the interstitial fluid space (ISF) into the blood. Every time the gastric juice receives one H^+, the blood will receive one HCO_3^-. This explains why the pH of the gastric venous blood increases after a meal - the alkaline tide.

3. Cellular [H^+] is a substrate for the luminal gastric proton pump (the K^+-H^+-pump), already activated by K^+. The net result is H^+-secretion to the lumen in a balanced relationship to Cl^--secretion. The surface of the gastric mucosa is always electrically negative with respect to the serosa. H^+ moves against a large concentration gradient into the gastric lumen. The intracellular [H^+] of the parietal cells is 10^{-7} mol/l, so with a [H^+] of 10^{-1} mol/l in the gastric juice, a million-fold concentration gradient is present across the luminal membrane. Accordingly, energy is required for the transport of both ions. The HCl secretion requires ATP.

4. The cellular concentration of cations is maintained by the basolateral Na^+-K^+-pump.

   The parietal cells contain more mitochondrial mass per volume unit than any other cells in the body, indicating a rich oxidative metabolism.

Histamine, acetylcholine and gastrin stimulate acid secretion. We have two types of histamine receptors in the human body: H_1 receptors (blocked by diphenhydramine) and H_2 receptors. Only H_2 receptors are located on the parietal cells.

1. The H_2 receptors make histamine a potent stimulant of HCl secretion. When histamine is bound to the H_2 receptor it activates adenylcyclase, an enzyme generating cAMP from ATP. This increase in intracellular [cAMP] is specific for histamine. The cAMP binds to and activates cAMP-dependent protein kinase (consisting of a regulatory and an active catalytic subunit). The cAMP binding releases the active catalytic subunit, which phosphorylate a variety of target proteins.

   H_2 receptor antagonists (cimetidine and ranitidine) prevent histamine from binding to the H_2 receptors of the basolateral membrane of the parietal cells, which reduces acid secretion. Synthetic analogues of prostaglandin E can inhibit both the cAMP and the Ca^{2+} release mechanisms, thus promoting ulcer healing (see later).

2. Acetylcholine (ACh) is released by vagal stimulation that leads to a stimulation of acid secretion. This secretion is inhibited by atropine. Thus the parietal cells contain muscarinic, cholinergic receptors (M_3).

3. Gastrin is the most potent stimulant of acid secretion in humans. Gastrin receptors were previously supposed not to be present on human parietal cells. Gastrin from G-cells was thought to release histamine from the granules of the mast cells in the gastric glands (Fig. 22-10). This is probably not the case. A direct gastrin effect on human gastrin receptors occurs, and an additional indirect effect via histamine increases the HCl secretion markedly (H_2 receptors).

   However, the three-receptor hypothesis is still under debate.

   Gastrin and acetylcholine release inositol-triphosphate (IP_3), which is produced with diacylglycerol (DAG) by a membrane phospholipase. The target system for IP_3 is a Ca^{2+}-channel protein located in the endoplasmic reticulum. Ca^{2+} is released from the reticulum, and Ca^{2+} also enters the cell through the basolateral membrane.

   Combined stimulation of all three receptors results in maximal gastric secretion (potentiation).

10. Intestinal digestion and absorption

Almost all of the dietary nutrients, water and electrolytes that enter the upper small intestine are absorbed. The small intestine, with its epithelial folds, villi, and microvilli, has an internal surface area of 200 m^2.

10a. Carbohydrates

Carbohydrates are the most important energy-containing components of the diet. The energetic value of most carbohydrates is 17.5 kJ per g, so that a daily diet of 400 g carbohydrates covers 7 000 kJ, which is 56% of the usable energy in a diet of 12 500 kJ daily. The formation of metabolic water on a mixed diet is 0.032 g of water per J.

Fig. 22-11: Absorption of carbohydrates by the enterocyte.

The common sources of digestible carbohydrates are starches (amylose), table sugar, fruits and milk. Plant and animal starch (amylopectin and glycogen) are branched molecules of glucose monomers. Indigestible carbohydrates are present in vegetables, fruits and grains (cellulose, hemicellulose, pectin) and in legumes (raffinose). Indigestible carbohydrates are also referred to as dietary fibres.

Digestion of starches to simple hexoses occurs in two phases: The luminal phase begins in the mouth with the action of salivary amylase (ptyalin), but most of this phase occurs in the upper small intestine as pancreatic a-amylase reach the chyme. The starch polymer is reduced to maltose, maltriose and a-limit dextran or dextrins (Fig. 22-11). The three substrates are pushed through the intestine and are now ready for the brush-border phase. Some of the substrate molecules get into contact with the brush-borders of the absorbing mucosal cell via the unstirred water layer. Enterocytes carry disaccharidases and trisaccharidases (oligosaccharidases)

on their surface that cleave these substrates to glucose, G.

Milk sugar (lactose) and cane sugar (sucrose) only require a brush-border phase of digestion, since they are disaccharides. Sucrose is reduced to glucose and fructose (G-F), and lactose to glucose and galactose (G-Ga) by the action of disaccharidases (sucrase and lactase).

Glucose in the intestinal lumen is absorbed by active transport.

1. The mechanism of active glucose transport is a carrier-mediated, Na^+ - glucose cotransport. As the luminal [glucose] falls below the fasting blood [glucose], active glucose transport becomes essential and sequesters all remaining luminal glucose into the blood. Glucose and Na^+ bind to apical membrane transport proteins (a glucose-transporter, GLUT 5). The two substances are deposited in the cytoplasm, because of conformational changes in GLUT 5, whereby the affinity of GLUT 5 for glucose-Na^+ changes from high to low. Glucose accumulates inside the cell to a level that exceeds blood [glucose].
2. Glucose therefore diffuses down its concentration gradient, through a specific uniport carrier in the basolateral membrane, out into the interstitial space and into the blood (Fig. 22-11). The basolateral uniport carrier for glucose is highly specific (glucose only), and does not depend upon Na^+. Galactose is also actively transported by the luminal glucose carrier system, and is a competitive inhibitor of glucose transport. Phlorrhizin blocks the glucose absorption, when its glucose moiety binds to the transporter instead of glucose.
3. Cytoplasmic Na^+ is actively pumped out through the basolateral membrane by the Na^+-K^+-pump. The low intracellular [Na^+] creates the Na^+ gradient and energises the transport of hexoses over the luminal enterocyte membrane.
4. Fructose has no effect on the absorption of glucose and galactose. Fructose is not actively transported by the Enterocytes, but is absorbed by a carrier-mediated, facilitated diffusion system, where energy is not required.

10 b. Proteins

The typical Western diet contains *100 g* of protein, which is equivalent to an energy input of 1700 kJ daily, although an adult needs only less than one g pr kg of body weight. This luxury combustion is an inappropriate use of global resources. Moreover, a high protein intake implies a long-term risk of uric acid accumulation from purine degradation (Chapter 20). Meats, fish, eggs, and diary products are high in proteins and expensive. Vegetable proteins are not as expensive as animal proteins.

Residents of areas with carbohydrate dominated nutrition and protein hunger develops diseases of protein deficiency, such as *Kwashiorkor* (Chapter 20).

Digestion of dietary proteins begins in the stomach, with the action of the gastric enzyme pepsin (pH optimum is 1), which cleaves proteins to proteoses, peptones and polypeptides. Pepsin is produced from pepsinogen in the presence of HCl. Pepsinogen is secreted by the gastric chief cells. The

digestion is continued in the intestine by proteolytic enzymes of the pancreas. Enteropeptidase converts trypsinogen to trypsin. Trypsin acts auto-catalytically to activate trypsinogen, and also convert chymo-trypsinogen, pro-carboxy-peptidases A/B, and pro-elastase to their active form. When the chyme is pushed into the duodenum, the pancreatic juice neutralises the chyme and the activity of pepsin is stopped. The proteolysis in the small intestine plays the major role, because the digestion and absorption of dietary protein is not impaired by total absence of pepsin.

Cytosolic peptidases from the enterocytes and brush border peptidases from the brush borders of the villous cells then cleave the small peptides into single amino acids (Enteropeptidase, amino-polypeptidase and di-peptidases). The end products of protein digestion by pancreatic proteases and brush border peptidases are di- and tri-peptides and amino acids. The cytosolic peptidases are abundant and particularly active against di- and tri-peptides.

Hydrolytic digestive products such as tripeptides, dipeptides and amino acids can be absorbed intact across the intestinal mucosa and into the blood. Two transport routes are dominant:

1. A peptide transporter, with high affinity for di- and tri-peptides, is absorbing the small peptides (Fig. 22-12). The system is stereospecific and prefers peptides of physiologic L-amino acids. This peptide transport across the brush border membrane is a secondary active process powered by the electrochemical potential difference of Na^+ across the membrane. The total amount of each amino acid that enters the enterocytes in the form of small peptides is considerably greater than the amount that enters as single amino acids.

2. The absorption of single amino acids from the intestinal lumen is an active process that involves a Na^+-dependent, carrier-mediated cotransport system similar to that for glucose. Competitive inhibition, saturation kinetics, Na+ dependency, and expenditure of metabolic energy in this case also characterise active transport.

   Selective carrier systems appear to be present for certain groups of amino acids: neutral, acidic, imino and basic groups. The neutral brush border (NBB) system transports most of the neutral amino acids. The imino acid system handles proline and hydroxyproline.

Fig. 22-12: Absorption of peptides and single amino acids by the enterocyte.

Basic amino acids and phenylalanine are absorbed primarily through facilitated diffusion from the gut lumen to the blood.

The basolateral membrane is more permeable to amino acids than is the brush border membrane. Therefore diffusion is more important for the basolateral transport, especially for amino acids with hydrophobic side chains.

The amino acids are carried in the blood to the liver via the portal vein.

Half of the amino acids absorbed in the intestine are from the diet, the remaining part is from digestive secretions and from desquamated mucosal cells.

Only 1 % of the dietary protein is excreted in the faeces, the remaining faecal protein is derived from micro-organisms and desquamated cells.

The reabsorption of amino acids (and glucose) in the renal tubules bares many similarities to the active absorption mechanism in the intestine.

A rare genetic disease involves defective intestinal absorption of neutral amino acids and a similar defective renal reabsorption. This condition is called Hartnups disease, which is caused by defects in the NBB transport system of the brush border coated epithelial cells of the jejunum and the proximal renal tubules.

10 c. Lipids

The typical Western diet contains *100 g* of lipids (3900 kJ) daily. Most of the dietary lipids consumed are triglycerides (only 2-4% is made up of phospholipids, cholesterol, cholesterol esters etc). Lipids would comprise just above 30% (ie, 100 g = 3900 kJ) of a standard diet of 12 500 kJ daily. An optimal diet should contain only 20% lipids, such as the lipids of fish oil and olive oil.

Absorption of excess lipids results in accumulation (obesity). The consequences of long term obesity are described in relation to diabetes mellitus in Chapter 27.

Essential dietary fatty acids are poly-unsaturated and cannot be synthesized in the body (linoleic acid, linolenic acid and arachidonic acid).

Dietary triglycerides are broken down into simpler molecules, to facilite absorption. A small fraction of the triglycerides is digested in the mouth and stomach by salivary, lingual lipase.

Most dietary triglycerides (TG) are digested in the small intestine. However, two problems must be solved before digestion can occur. Triglycerides are insoluble in water, and the chyme in the intestine is an emulsion of large fat particles in water. All the lipase proteins by contrast are water-soluble. It follows that, triglycerides must be dissolved in the aqueous phase before they can be digested.

The lipolytic activity requires the emulsifying action of bile salts in order to dissolve triglycerides in water. Pancreatic lipase binds to the surface of the small emulsion particles.

1. Simple bile micelles are aggregates of bile salt monomers that form spherical structures with a diameter of 5 nm, and the micelles have a negative charge. Following a meal, bile micelles are formed above a certain concentration of bile salts, called the critical micellar concentration. The lipophilic, hydrophobic, apolar end of the bile acids faces inward creating a hydrophobic core (Fig. 22-13). The hydrophilic polar end of the bile salts (hydroxyl-, carboxyl- and amino- groups) points outward, so that they are mixed with the polar water molecules. The simple lipids must pass a diffusion barrier - an unstirred water layer, which is the water layer immediately adjacent to the mucosa, where the intestinal flow rate is essentially zero. This water layer contains the water-soluble lipases and cholesterol esterases.

Fig. 22-13: Absorption of lipids by the enterocyte (2-MG is 2-monoglyceride).

2. Mixed micelles. Simple lipid molecules (cholesterol, phospholipids, fatty acids, 2-monoglycerides or 2-MG, fat-soluble vitamins and lyso-lecithin) diffuse into the lipophilic core of the simple bile micelles and form a mixed micelle (Fig. 22-13). A solution of micelles is water-clear and stable.

The mixed micelles carry the major part of all the lipids that are absorbed by the intestinal microvilli. When the lipids of the mixed micelle have diffused into the enterocyte, emulsifying more hydrolysed lipids recycles the empty bile micelle. Neither bile salt micelles nor bile salt molecules diffuse into the enterocyte (Fig. 22-13).

The fatty acids with a short chain (up to 12 C-atoms) are more hydrophilic than the rest. They can diffuse directly to the portal blood as fatty acids. Once fatty acids enter the enterocyte, they are primarily activated to acetyl coenzyme A by a process that requires ATP and acetyl coenzyme A synthetase. Acetyl coenzyme A enters one of two pathways: the 2-MG and the a-glycerol phosphate pathways. Both bring about the resynthesis of triglycerides (TG) in the enterocyte.

In the enterocyte the lipids are reformed to triglycerides, cholesterol, phospholipids etc. The reformed triglycerides, cholesterol, phospholipids, fatty acids, esters and fat-soluble vitamins reach the endoplasmic reticulum, where they are packed in another lipid-carrying particle: the chylomicron.

The centre of the chylomicron is a cholesterol ester (E in Fig. 22-13). Chylomicrons are packed into vesicles in the Golgi-system. These vesicles reach the basolateral membrane, and their contents pass through this membrane by exocytosis. Thus the chylomicrons reach the lymphatic channel of the villus (the central lacteal). The lymph delivers the chylomicrons to the blood through the thoracic duct. Plasma is milky (lipaemic) following a fatty meal.

All of the dietary lipid is normally absorbed in the intestine. Faecal fat derives from bacterial lipids and lipids of desquamated mucosal cells. - Disorders such as gallstones, pancreatitis, Crohn's disease, and liver disease can lead to fat malabsorption (steatorrhoea or fat-diarrhoea).

Lipids are mainly absorbed through the enterocyte and transported by the lymph, which reaches the blood via the thoracic duct. Lipids thus reach the liver through the hepatic artery, with the exception of short-chain fatty acids that enter the portal blood directly. Other nutrients are absorbed directly to the blood and reach the liver through the portal vein.

Fat-soluble vitamins, such as vitamin A, D and K, are absorbed in the chylomicrons along with lipid nutrients (Fig. 22-13). In contrast, the water-soluble vitamins, such as vitamin- B and -C, cross the mucosa by diffusion and by association to specific membrane transporter proteins. Vitamin B_{12} (cyanocobalamine) is the largest of the vitamins, and its absorption in the terminal ileum utilises a specific transport mucoprotein called intrinsic factor.

10.d Fluids and electrolytes

The intestinal content is isosmolar with plasma, and the water is absorbed from the lumen to the blood by passive osmosis. The membranes of the intestinal mucosal cells and even the tight junctions are highly permeable to

water. Hereby, active transport of Na and Cl from the lumen to the small interstitial space builds up a forceful osmotic gradient, drawing water the same way by a passive process. In the small interstitial space water creates a hydrostatic overpressure. Since the capillary and lymph endothelial membranes are no barriers for Na^+, Cl^- and water, a bulk flow of fluid from the interstitial space passes into the blood- and lymph vessels. The intestinal mucosa possesses elevations called villi, and pitted areas called crypts. The villous cells have a typical brush border responsible for net absorption of ions and water, whereas the crypt cells contain secretory mechanisms causing net secretion.

The villous cells absorb Na^+ through the luminal brush border membrane by three mechanisms:

1. An inward diffusion gradient through a Na^+-channel,
2. A Na^+-H^+-exchange, and
3. A Na^+ -solute coupled cotransport (the solute being glucose, galactose, bile salts, water-soluble vitamins and amino acids).

Fig. 22-14: Ion transport processes in jejunal enterocyte.

Ad 1.: The [Na^+] is kept low (14 mM) in the cell, whereas [Na^+] is 140 mM in the intestinal lumen. This concentration gradient work together with an electrical gradient, since the cytosol of the cell is -40 mV with the intestinal content as a reference (Fig. 22-14). Thus Na^+ can easily pass the luminal brush border membrane passively. The intestinal mucosa has ion permeable tight junctions - it is leaky. This paracellular transport is so great that the net absorption of Na^+ and Cl^- through the cells only amounts to 10% of the total transport through the mucosa.

Ad 2.: The transport of Na^+ into the enterocyte (Fig. 22-14) is through a co-exchange protein (Na^+/H^+). Part of the energy released by Na^+ moving down its gradient is used to extrude H^+ into the intestinal lumen. Here H^+ reacts with bicarbonate from bile and pancreatic juice to produce CO_2 and water, thus reducing the pH of the intestinal fluid.

Ad 3.: Na^+ -solute coupled cotransport.

The basolateral membrane of the enterocyte contains a Na^+-K^+-pump, which maintains the inward directed Na^+-gradient. The pump is energised by the hydrolysis of ATP, which provides the driving force for Na^+ entry. Thus an active process pumps Na+ out in the small interstitial space and K^+ is pumped into the cell. The basolateral membrane also contains many K^+-channel proteins, so K^+ will leak back to the interstitial space almost as soon as it has entered the cell. The K^+ is absorbed by diffusion - a daily net total of 80 mmol.

A Na^+-K^+-2 Cl^- co-transporter located on the basolateral membrane (Fig.

22-15) maintains the Cl gradient, with an elevated intracellular [Cl]. This transporter drags Cl^- from the interstitial fluid (ISF).

Fig. 22-15: Net Cl^--secretion by crypt cells of the small intestine.

The transporter system uses the electrochemical Na^+ gradient to transport K^+ and Cl^- into the cell (Fig. 22-15). The crypt cells hereby can secrete Cl^- through the luminal membrane via an electrogenic channel. The Cl^- secretion produces a net luminal electronegativity, which drags Na^+ across the tight junctions resulting in net secretion (Fig. 22-15). Water (about 2 l daily) is secreted by passive osmosis.

A dramatic rise in Cl^- and water secretion - caused by gut inflammation with cholera - can lead to secretory diarrhoea.

Fluid absorption in the colon is determined by the absorption of NaCl. The Na^+ transport involves 1. Electrogenic Na^+ transfers via Na^+ channels, and 2. Na^+-co-exchange as in the small intestine (Fig. 22-14). Both transport processes are driven by the Na^+ gradient maintained by the basolateral Na^+-K^+-pump. (The Na^+-solute coupled co-transporter is not present in the human colon). The colonic Na^+-K^+- pump is more sensitive to aldosterone than that in the small intestine. Aldosterone is a steroid hormone. Steroids bind directly to cytosolic receptors and do not need second messengers. The colonic Na^+-K^+-pump activity accumulates K^+ in the enterocyte, and this gradient drives the K^+ secretion across the luminal K^+ channel. The Cl^- absorption is accomplished by diffusion along a Cl^--gradient, and by a luminal Cl^--bicarbonate exchanger producing bicarbonate secretion. We have a bicarbonate-chloride-shift just as in the red cells. Since electrolyte absorption exceeds secretion, there is a net water absorption in the healthy colon (1-1.5 l daily and with a colonic salvage capacity of 4 500 ml).

Nutrient malabsorption of the small intestine increases the fluid volume delivered to the colon and can provide an osmotic effect in the colon with diarrhoea. Up till 4 600 ml of fluid normally passes the ileocoecal valve without causing diarrhoea.

In conditions such as cholera, the excess fluid from the ileum exceeds the colonic salvage, leading to life-threatening diarrhoea. The cholera toxin can enhance the Cl^--secretion drastically and cause secretory diarrhoea with large quantities of Cl^- and water.

In inflammatory diseases of the colon, the colonic salvage capacity is markedly reduced, resulting in colonic diarrhoea.

10.e Iron absorption

Two-third of the iron content of the body (3-4 g) is stored in the haeme group of haemoglobin. The ability to transport O_2 depends on the presence of haeme. Haeme gives the red cell its characteristic red colour. Only haemoglobin with iron in the ferrous state binds O_2, whereas the dark red methaemoglobin with the iron in ferric state cannot bind O_2. Soluble ferritin forms an intracellular store (25% of total). Essential, but minor amounts of iron, is bound in myoglobin and in the electron-transporting enzymes of the

mitochondria in all respiring cells. Haemosiderin is an insoluble degradation product of ferritin that aggregates into cytoplasmic granules. Haemosiderin is a normal microscopic finding in the spleen, bone marrow and the Kupffers cells of the liver.

1. Ascorbate in the food reduces Fe^{3+} to Fe^{2+}, and forms a soluble complex with iron, thereby effectively promoting the iron absorption. We normally ingest about 20 mg iron daily, and less than 1 mg is absorbed in healthy adults, because iron form insoluble salts and complexes in the gastrointestinal secretions.
2. Iron is transported from the lumen of the upper jejunum, across the mucosa, and into the plasma by an iron-binding protein called gut transferrin.
3. Receptor proteins in the brush border membrane bind the transferrin-iron complex, and the complex is taken up into the cell by receptor-mediated endocytosis (Fig. 22-16).

Fig. 22-16: Iron absorption through an enterocyte.

4. There is a free pool of iron in the cytosol. Iron exists in one of two states in the cytosol: The ferrous state (Fe^{2+}) or the ferric state (Fe^{3+}). The Fe^{2+} ions, after absorption into the mucosal cell, are oxidised to Fe^{3+} (Fig. 22-16).
5. When intracellular iron is available in excess, it is bound to apoferritin, an ubiquitous iron-binding protein, and stored within the mucosal cells as ferritin. The synthesis of apoferritin is stimulated by iron. This translational mechanism protects against excessive absorption.
6. At the basolateral membrane the Fe^{3+} are reduced to Fe^{2+} and passes from the interstitial space to the blood. Here Fe^{2+} are again oxidised to Fe^{3+} and binds to plasma transferrin. Cellular iron stores are mobilised by autophagocytosis of enterocyte ferritin, when body stores of iron are deficient.

Normally, serum-iron is 12-36 mM, which is about one-third of the total iron-binding capacity in the plasma of adults. This means that one-third of the circulating plasma transferrin is saturated with iron.

In iron deficiency the serum-iron is falling, whereas the iron binding capacity increases. The red cell count, haematocrit and the haemoglobin concentration fall in continued deficiency, as does the concentration of iron containing cellular enzymes. Latent (or untreated) iron deficiency anaemia is found in 25-33% of all fertile females.

Increase of the total iron content takes place by enhanced intestinal iron absorption or by blood transfusions.

Ferritin is further saturated with iron to form Haemosiderin in the liver and elsewhere, when abnormal amounts are ingested over months. Extreme accumulation of excess iron in cells throughout the body (heart, lungs, pancreas, kidneys, glands and skin) finally damages vital organs and is called *haemochromatosis*.

When blood-containing products are ingested, proteolytic enzymes release

the haeme groups from the haemoglobin in the intestinal lumen. Haem is absorbed by facilitated transport. Approximately 20% of the haem iron ingested are absorbed. Blood containing products are effective in iron deficiency anaemia.

Pathophysiology

The following is a short description of classical gastrointestinal disorders, such as:

1. Achalasia, 2. Gastro-oesophageal reflux, 3. Gastritis, 4. Peptic ulcer disease, 5. Gastric tumours, 6. Gastrointestinal bleeding, 7. Coeliac disease, 8. Crohns disease and ulcerative colitis, 9. Diarrhoea, 10. Acute abdomen, 11. Colon irritabile, diverticulosis and constipation, 12. Megacolon, 13. Colonic cancer, 14. Dry mouth, and 15. Carbohydrate malabsorption.

1. Achalasia

Achalasia is a disease characterised by *lack of peristalsis* in oesophagus and relaxation failure of the lower oesophageal sphincter (LOS or american LES) in response to swallowing (Fig. 14-2). Vomiting and weight loss is major symptoms. There is no *receptive relaxation*, because the myenteric plexus does not work. The aetiology is unknown.

There is *absence of ganglion cells* in the myenteric plexus of the oesophageal wall and the LOS. The peptidergic neurons in the LOS normally secrete VIP (Vasoactive Intestinal Peptide), which relaxes the LOS, but these neurons are lost in achalasia.

The food gets stuck because of the lack of peristalsis, the oesophagus dilates and the patient regurgitates. Intermittent dysphagia during meals is typical. Many patients leave the table, provoke vomiting and are relieved. Vomiting is a classical vagal reflex phenomenon relaxing LOS.

Fig. 22-17: Oesophageal disorders

The diagnosis is confirmed by chest X-ray in particular following a barium swallow, and oesophagoscopy is necessary to exclude malignancy in the region.

A pneumatic bag is placed in the LOS opening and pressurised until LOS is sufficiently dilatated. Surgical division of the LOS muscle is performed by laparoscopy.

American trypanosomiasis (*Chagas´ disease* in Latin America) produces achalasia by microbial destruction of the ganglion cells.

2. Gastro-oesophageal reflux disease

Gastroesophageal reflux with oesophagitis is caused by incomplete closure of the LOS. Gastric contents with acid reaction then reflux into the oesophagus causing inflammation, erosion and bleeding.

This disorder is also called *reflux oesophagitis*. It results from regurgitation of gastric contents (with HCl and pepsin) into the lower oesophagus causing long lasting damage of its mucosa. The wall becomes hyperaemic, and white patches are seen on the epithelium (leucoplakias). The dysphagia most often presents as *heartburn*. As dysphagia progress it is likely that an oesophageal stricture is developing. If the squamous epithelium of the lower

oesophagus is replaced by columnar epithelium, as a response to long lasting injury, there is an increased risk of transformation of the epithelium into an adenocarcinoma.

The most important barrier to the reflux is the LOS. Normally, LOS contracts as soon as the food has passed into the stomach, and the oesophagus is cleared by secondary peristalsis.

Gastro-oesophageal reflux disease is usually treated with H_2-receptor antagonists, who inhibit the gastric acid production, or with *proton pump inhibitors*, which inhibit the gastric proton pump and thus effectively reduce gastric acidity. Major complications such as *strictures* usually need surgery.

3. Gastritis

Gastritis occurs as at least two typical manifestations: Acute, erosive gastritis and chronic, non-erosive gastritis.

Focal inflammatory lesions of the mucosa characterise acute gastritis. Sometimes the erosions extend into the deeper layers of the wall (beyond the lamina propria) to form acute ulcers (Fig. 22-18). Acute gastritis is produced by alcohol, drugs (corticosteroids, ASA and NSAIDs) or infections with Helicobacter pylori or virus. After severe stress the gastritis may develop into a life-threatening condition with stress ulcers and haemorrhage. The stress conditions are severe burns, trauma, shock, and sepsis.

Chronic gastritis is a long-lasting inflammation of the gastric wall. The superficial layers are infiltrated with lymphocytes and plasma cells. Atrophia develops with loss of both parietal and chief cells. Helicobacter pylori are the chief cause of chronic gastritis in the antrum. The loss of parietal cells leads to *achlorhydria* (absent HCl production), and to deficiency of *intrinsic* factor.

Autoimmune gastritis is a pangastritis, where autoantibodies to parietal cells can be demonstrated in the blood. Vitamin B_{12} is not absorbed in the ileum in the absence of intrinsic factor, so the result is pernicious anaemia (Chapter 8).

Fig. 22-18: Peptic ulcers extend beyond the lamina propria, whereas erosions are superficial.

4. Peptic ulcer disease

Peptic ulcer disease is a mucusal ulcer in an acid- producing zone in the distal stomach or the proximal duodenum.

The normal stomach produces enough mucus and alkaline juice to protect the gastric and duodenal mucosa against HCl. The mucine molecules swell and form a non-stirred layer covering the mucosa. In duodenum the pancreatic bicarbonate creates a pH of 7.5 at the luminal membrane of the mucosa.

Epidemiological occurrence can be explained on the prevalence of *Helicobacter pylori* infection of the stomach and the colonisation of the upper gastrointestinal tract with this bacteria. Helicobacter pylori infection destroys the protective system, and at the same time provokes excess acid secretion.

The patient, whose pain complains typically occur a few hours following a meal or awaken the patient at night, points out Epigastric pains.

Bleeding from ulcers can be *fatal*. Upper gastrointestinal tract bleeding implies a significant loss of blood into the lumen of the foregut. Haematemesis and melaena demonstrate such a bleeding. *Haematemesis* is defined as vomiting of whole blood or blood clots. *Melaena* is defined as passage of dark tarry stools (coal-black, shiny, sticky, and foul smelling).

Risk factors for peptic ulcer disease are *drugs* (ASA, NSAIDs and corticoids), *hyperparathyroidism* (the high Ca^{2+} level stimulates gastric acid secretion), and *gastrin-producing tumours* of the pancreas (Zollinger-Ellisons syndrome). Other contributing factors are increased *pepsinogen* from the chief cells, increased parietal cell mass, reduced somatostatin secretion from the antral D cells, and damage of the mucosa. Acetylsalicylic acid and other non-steroid anti-inflammatory drugs deplete the gastric mucosa for prostaglandins, which leads to mucosal damage. Strong alcoholic beverages also damage the gastric mucosal barrier and stimulate acid secretion. Caffein stimulates gastric acid secretion.

Genetic factors must be considered, since persons who do not secrete blood group 0 antigen into the saliva and gastric juice, have an increased risk of developing duodenal ulcers.

The diagnosis is confirmed with endoscopy and biopsy or with double-contrast barium technique.

The following five therapeutic strategies are used in the treatment of peptic ulcer disease:

1. Eradication of Helicobacter pylori with *antibiotics* is the treatment of choice for most cases of peptic ulcer disease, since it seems to cure the patient. *Clarithromycin* is a macrolide that binds to and prevents translocation on Helicobacter pylori- ribosomes, which is an effective basic therapy of peptic ulcers.
2. Inhibition of the *gastric proton pum***p** in the luminal membrane of the parietal cells. Omeprazole is a *proton pump inhibitor*, which relieves symptoms and cure most duodenal ulcers within four weeks - often in combination with antibiotics. Omeprazole and similar antagonists to the gastric proton pump are especially effective in treatment of *persistent* HCl-secretion caused by the Zollinger-Ellison syndrome.
3. Histamine acts through H_2 receptors on the basolateral membrane of the parietal cells. The second messengers for histamine is cAMP. All other cells contain H_1 receptors. Accordingly, *H_2 receptor antagonists* (cimetidine, ranitidine, famotidine, and nizatidine) inhibit acid secretion because they fit the H_2 receptors specifically. The H_2 receptor antagonists prevent histamine from binding to the H_2 receptors on the basolateral membrane of the parietal cells.
4. *Prostaglandin E_1 analogues*, such as misoprostol, inhibits gastric acid secretion by unspecific inhibition of the second messenger, cAMP, in the parietal cell and elsewhere. Prostaglandin E_1 analogues hereby promote ulcer healing.
5. *Surgical management* is rarely used unless complications occur.

Highly selective vagotomy, in which only the nerve fibres to the parietal cells were cut was previously used, but this is not an alternative to chemical vagotomy (procedure 2., 3., 4.).

All treatment procedures, which work by inhibition of gastric acid secretion, have a common drawback. To the extent that gastric acid secretion is reduced there is no inhibition of the *gastrin release* from the antral G cells. Accordingly, the blood [gastrin] increases, and during treatment of the patients this concentration is constantly increased. The high gastrin level counteracts the expected effect on the acid production. Since gastrin is a trophical hormone for the gastric mucosa, long-term treatment with acid suppression might result in *mucosal hypertrophy* with a further rise in acid production and in cellular modifications. These complications are probably related to the rather high ulcer recurrence rate of most treatment procedures. Obviously, the only rational strategy is to eliminate the cause of the peptic ulcer disease.

5. Gastrointestinal tumours

The *leiomyoma* is the most frequent benign gastric tumour. This is a tumour of smooth muscle cells. Leiomyoma are usually discovered at autopsies or by chance, as they do not produce symptoms except when they ulcerate and bleed.

Carcinoma of the stomach is frequently located in the antrum and is almost always adenocarcinoma.

Risk factors for gastric cancer are Helicobacter pylori colonisation with chronic gastritis, atrophia and metaplasia. Dietary factors include spiced, salted or smoked food (with benzpyren). Nitrosamines are probably carcinogenic in man, and they are produced in food and water with a high nitrate content.

One third of the general population have blood group A, but 50% of all patients with gastric cancer belong to blood group A.

Enterochromaffin cells of the intestinal wall form carcinoid tumours. The tumour secretes serotonin, bradykinin, histamine, tachykinins and prostaglandins.

Somatostatin is an almost universal hormone-inhibitor. A somatostatin analogue, octreotide, inhibits the secretion of many gut hormones including those outlined above. Often the typical signs of carcinoid tumour, facial flushing and diarrhoea are totally alleviated with octreotide treatment.

6. Gastrointestinal bleeding

Acute gastrointestinal bleeding occurs in the form of haematemesis or dramatic vomiting of blood.

A bleeding peptic ulcer causes most cases. Less frequent is bleeding oesophageal varicose veins, and gastric carcinoma.

The danger is bleeding shock, with tachycardia, falling blood pressure and pallor in a cold sweating patient. Urgent and adequate blood transfusion is life saving.

Ulcers, infections, tumours, polyps, and varicose veins throughout the gastrointestinal tract cause chronic gastrointestinal bleeding. These patients present with iron deficiency anaemia (Chapter 8).

The patients are first examined with gastroscopy, often followed by Colonoscopy or enteroscopy.

7. Coeliac disease

Gluten-sensitive enteropathy or coeliac disease (*sprue*) describes a condition where the duodenal and jejunal mucosa is more or less destroyed by hypersensitivity towards gluten (see Chapter 32).

8. Crohns disease and ulcerative colitis

These two disorders may be different manifestations of a single disease, *non-specific inflammatory bowel disease* (see Chapter 32).

9. Diarrhoea

This term is usually used for an *increased stool frequency* and implies a larger than normal stool weight (Fig. 22-19).

One pathophysiological differentiation of diarrhoea is the following:

1. Zollinger-Ellisons syndrome with tremendous gastric secretion can cause diarrhoea.

Fig. 22-19: Diarrhoea of different origin.

2. Bacterial or Secretory diarrhoea is caused by increased Cl and reduced Na^+ - reabsorption. Enterotoxins from bacteria on the microvillus surface affect the toxin receptors, which increases the cAMP level in the cell. This in turn activates the chloride- channel and inhibits the NaCl reabsorption process.
3. Inflammatory diarrhoea is caused by mucosal destruction with outflow of fluid and blood such as in ulcerative colitis.
4. Osmotic active substances in the gut lumen cause osmotic diarrhoea. These substances are normal nutrients in case of malabsorption, or non-absorbable substances taken for some reason or other.
5. Diarrhoea following ileal resection. Bile acids are normally reabsorbed in the terminal ileum. Following ileal resection the bile acids enter the colon. Bile acids are toxic to the colonic mucosa and stimulate colonic secretion of large volumes,

10. The acute abdomen

Acute appendicitis is the dominant cause of *acute abdomen*. Mechanical obstruction of the orifice of the appendix by a faecolith is demonstrated in less than half the operated cases. Secretions dilatate the obstructed appendix, until the mucosa ulcerates and the wall is invaded by intestinal bacteria. In many cases only generalized inflammation is found, and in 10% of all removed appendices, the microscopy is normal.

The patient typically experiences periumbilical or diffuse pain, which moves towards the right iliac fossa within hours. The patient is subfebrile and there is nausea and vomiting. The examinator finds a tender right iliac fossa with defence musculaire (guarding), showing local peritonitis. Rectal exploration often reveals tenderness to the right.

Perforation of an inflamed appendix can cause several severe complications: Periappendiceal or hepatic abscesses, fistulae, generalized peritonitis, and

septicaemia with septic shock.

Appendectomy is performed as early as possible by open surgery or by laparoscopy.

A history of more than 48 hours of abdominal pain, with a solid mass in the right fossa iliaca indicates disaster. Perforation is most likely present with formation of a periappendiceal abscess. Here, the patient is preferably treated with antibiotics for some days, and appendectomy is delayed (French: a fraud) until the danger of generalized spread to the peritoneal cavity is minimal.

Acute peritonitis is frequently caused by perforation and presented as a sudden, severe abdominal pain. High fever develops rapidly with nausea, vomiting and paralytic ileus. As the bacterial infection spreads to affect the peritoneum in general, the condition becomes serious and septic shock may develop.

Spontaneous peritonitis with ascites in adults is caused by hepatic, alcoholic cirrhosis with portal hypertension (see Chapter 23).

11. Colon irritabile (irritable bowel), diverticulosis and constipation

These are disorders of *slow* colonic motility.

The patient with *irritabile bowel syndrome* complains of abdominal pain (diffuse or localised to the left iliac fossa), which is relieved by defecation or flatulence. There are often frequent small-volume stools, but the patient feels that the emptying is incomplete. The abdomen is distended. This is a condition with painful spasms causing constipation alternating with mucous diarrhoea. The condition is related to stress and sedentary life style, and is relieved by daily exercise.

Fig. 22-20: Frequent colonic disorders

Diverticulosis or *diverticular disease* is a condition with herniation of the mucosa through the muscular layers of the colon, caused by increased intraluminal pressure. The diverticules are recognized following a barium enema, and if they are inflamed the condition is called *diverticulitis*. Persons with disturbed stool-habits are likely to develop increased intraluminal pressure during defaecation, and they may develop hernias at weak spots in the gut wall. The incidence is high in inactive persons and low in vegetarians or in persons with a high dietary fibre content.

Mild clinical cases can be treated with light daily exercise such as walking in a hilly environment. Emergency cases may need surgery.

Constipation is frequently caused by a low fibre intake in sedentary persons. They often exhibit *irregular* defaecation habits, and irrational use of laxatives. Such habits suppress the natural reflexes.

The condition is improved by a high-fibre diet or by daily walking. Suppositories may be necessary, but long-term use of laxatives is contraindicated.

12. Megacolon

Megacolon covers several disorders, where the colon is dilatated.

Congenital megacolon or *Hirschprungs disease* is colonic dilatation resulting from congenital absence of ganglion cells in the myenteric plexus

at the region, where the colon passes into rectum. Migration of cells from the neural crest is disturbed.

The cause is mutation of a gene localised on chromosome 10. The a-ganglionic segment is permanently contracted and stenotic, so the intestinal content is accumulated proximal to stenosis. The markedly distended colon gives rise to the term *megacolon*. Large amounts of faecal matters accumulate, because peristalsis and mass movements are impossible.

The diagnosis is confirmed by a transmural rectal biopsy showing absent ganglion cells. Surgical removal of the segment cures most of the young patients.

Fig. 22-21: Hirschprungs disease with a-ganglionosis and megacolon.

Acquired Megacolon usually occurs in adults with Parkinsonism, diabetic neuropathy, Chagas disease (Chapter 33) or any other disorder that affect the innervation of the smooth muscles.

13. Colonic cancer

Colonic cancer is related to slow passage of faecal material with *carcinogens* through the colon. *Carcinogens* are chemicals, whose end-products bind to DNA and damage it.

Sedentary persons have a high frequency (morbidity) of *constipation* and a high mortality of *colon cancer*, but not of rectal cancer. The colon cancer is clearly related to an *inactive life style*, and regular exercise reduces morbidity and mortality.

Prolonged accumulation of faecal content with carcinogens in the colon increases the exposure time of the mucosa and may be of importance. *High-fibre diet* and *daily walking* reduce the exposure time. There is a firm correlation between colonic cancer and the activity level of persons in industrial societies. The same is true for groups of persons living on a low fibre diet with a high content of meat and animal fat.

Usually the recto-sigmoid area is involved, a location where the faecal content is moved to and fro for varying periods (Fig. 22-22).

Patients with chronic gastrointestinal bleeding usually present with iron deficiency anaemia. Measurements for faecal occult blood are easy to perform and of value as a mass population screening for large bowel malignancy.

Fig. 22-22: Colon cancer in the ascending colon (polypoid) and in the sigmoid (constricting cancer). – A rectal cancer tumour is shown in the upper rectum.

A correlation between *rectal cancer* and exposure time for carcinogens is not to be expected, because the faecal content passes this part of the tract without delay. A correlation has also been disproved in large population groups.

14.Dry mouth

Patients with a rare autoimmune disorder (the *Sjögren syndrome*) suffer from dry mouth (*xerostomia*), dry eyes (*xerophtalmia*) and rheumatoid arthritis.

In patients lacking functional salivary glands, xerostomia, infections of the

buccal mucosa, and dental caries are prevalent.

In most cases of xerostomia the condition is therapy-resistant and unexplained. Some cases are caused by dehydration or by antidepressants.

15. Carbohydrate malabsorption

The most common chronic disorder in humans is lactose malabsorption or *hypolactasia* (lactose-induced diarrhoea or lactose intolerance), which is due to a *genetically deficiency of lactase* in the brush-border of the duodeno-jejunal enterocytes (see Chapter 31).

Self-Assessment

Multiple Choice Questions

I. Each of the following statements has True/False options:

A. The receptive relaxation response of the stomach decreases Gastroesophageal reflux.

B. The intrinsic innervation of the digestive, secretory epithelium responds to parasympathetic input with decreased secretion.

C. The sympathetic nerve fibres to the gut act presynaptically to inhibit acetylcholine release in the myenteric ganglia and activate a-receptors. Hereby, sphincter muscles are contracted, blood vessels are constricted, and secretion is inhibited.

D. Relaxation of the lower oesophageal sphincter is not caused by increased vagal inhibitory fibre discharge.

E. Oesophageal reflex activity is controlled by primary peristalsis that are co-ordinated by a swallowing centre in the solitary tract nucleus, vagal nuclei, and reticular formation. Local distension stimulates the secondary peristalsis.

II. Each of the following statements has False/True options:

A. Gastrin originates in the antral and duodenal mucosa, where it is released from G-cells.

B. Secretin is a hormone that is released from the duodenum in response to HCl.

C. Pancreozymin (CCK) contracts the sphincter of Oddi.

D. GIP stimulates insulin secretion.

E. GRP is involved in vagal gastrin secretion.

III. The following five statements have True/False options.

A. The major source of cholesterol is food intake.

B. A sweat test resulting in a Na^+ -concentration above 60 mM in the

sweat, is strongly indicative of cystic fibrosis.

C. G-cells in the pancreatic islets produce large amounts of a certain hormone, but they are named after their G-protein systems, which amplify a signal, by production of second messengers.

D. Glucagon stimulate glycogenolysis, gluconeogenesis, ureagenesis and ketogenesis.

E. VIP controls the bloodflow of the gastric mucosa, and GRP releases gastrin from the antral G-cells.

IV. Each of the following five statements have False/True options:

A: The basic electrical rhythm is an electrical event that always causes contractions in the digestive system.

B: The basic electrical rhythm determines the maximal rate of peristaltic contractions.

C: Slow waves in the colon cannot result in anti-peristalsis.

D: The major role of the human colon is to reabsorb water and electrolytes.

E: The only entirely voluntary motor process of the motility patterns in the digestive tract is chewing.

V. Each of the following five statements have False/True options:

A. Hot and acidic liquids are buffered by saliva in the mouth, and the salivary epidermal growth factor promotes the healing of wounds.

B. The parotid secretion is watery and serves to solubilize food, so it can be tasted.

C. Salivary buffers maintain the activity of amylase during the first period in the stomach.

D. Saliva has bactericide effects due to lysozyme.

E. AIDS is transferred via saliva.

VI. Each of the following five statements have False/True options:

A. Acetylcholine, gastrin and histamine stimulate gastric acid secretion.

B. H_2 blockers bind to histamine receptors at the basolateral membrane.

C. The parietal cells increase their O_2 consumption, acid secretion, intracellular [cAMP] and [Ca^{2+}], when stimulated by histamine.

D. The H^+-K^+-ATPase is responsible for gastric acid secretion.

E. Gastrin and acetylcholine does not release IP_3.

Case History A

A resting male patient, age 54 years, body weights 76 kg, is suspected of Zollinger-Ellison syndrome and examined in the morning after fasting overnight. The throat is sprayed with lignocaine and a gastroscope is introduced into the pharynx under direct vision and passed down the oesophagus into the stomach and duodenum. No ulcers, tumours or bleeding is found. A biopsy of the mucosa shows an overgrowth of parietal cells. A sample of gastric juice is aspirated. Following stimulation by an injection of pentagastrin, gastric juice is aspirated via a nasogastric tube for one hour. The hydrogen ion concentration in the aspirate is 150 mM, and the volume is 350 ml.

Due to lung complications the blood gasses of the patient are measured in the morning (P_{aCO2} 40 mmHg, pH_a 7.40 , Base Excess zero, actual [bicarbonate] 24 mM) and just after completion of the aspiration (P_{aCO2} 40 mmHg, pHa 7.48, Base Excess 7 mM, actual

[Bicarbonate] 30 mM). The next morning blood gases were normalised.

1. *Calculate the gastric acid secretion rate of the patient, and compare the result with a normal value of 30 mmol per hour.*

2. *Describe the acid-base status of the patient just following the aspiration.*

3. *Explain the normalisation of the acid-base status the following morning.*

4. *Suggest a better diagnostic tool for the Zollinger-Ellison syndrome.*

Case History B

A nervous, smoking male, age 36 years, is admitted to hospital with severe hunger Epigastric pain reduced by eating, acid hiccups, diarrhoea, and steatorrhoea (ie, fatty stools). He has a stressful work, and over the last months he has frequently used drugs containing acetyl salicylic acid for headache, and used whisky on the rocks. Radiological examination of the stomach and duodenum suggests the presence of an ulcer in the duodenal bulb. This is confirmed by endoscopy. Gastric juice is removed by aspiration. The basal rate of HCl secretion is found to be 5 times normal. Histological examination of the gastric mucosa reveals a higher density of parietal cells and gastric glands than normal, but no hyperplasia of antral G cells.

The serum [gastrin] of the patient is 10 times higher than normal, and does not increase following a test meal.

One dose of the proton pump blocker, Omeprazole, reduces the HCl secretion rate of the patient to normal for 24 hours.

1. *Present a likely explanation for the development of the patient's duodenal ulcer.*

2. *Why does the patient have elevated serum [gastrin]?*

3. *Explain why a test meal did not induce a rise in serum [gastrin]?*

4. *Explain the mechanism for the patient's steatorrhoea and diarrhoea?*

5. *Why is one dose of omeprazole effective for such a length of time?*

6. *Transportation of one mol of H^+ from the cytosol of the parietal cell to the gastric lumen costs at least an oxidation of 30 mmol of glucose. Calculate the free energy necessary for the active transport of one mol of H^+.*

Case History C

A 35-year old male computer expert visits his general practitioner complaining of exhaustion. For weeks he has suffered from constipation, fatty stools and abdominal pain. He is loosing weight and gets out of breath when he is stair climbing.The patient looks pale and emaciated. Palpation of the abdomen reveals a soft mass in the right iliac fossa. Haematological tests show that the blood haemoglobin is 5.2 mM, the red cell count is $(3.1*10^{12})\ l^{-1}$ and the mean cell volume is 68 fl. Endoscopy with a duodenal biopsy show a normal mucosa with long intact villi. Colonoscopy shows patchy reddening of the mucosa and biopsies show granulomas in the lamina propria. A barium examination reveals narrowing of the terminal ileum.

1. *What is the haematological diagnosis?*

2. *What is wrong with the intestine of the patient?*

3. *What is the therapy of this condition?*

4. *Describe the complications of this chronic condition.*

5. *Describe two disorders which may mimic the condition of this patient.*

Try to solve the problems before looking up the answers.

Highlights

- *Epithelial and glandular cells of the gastrointestinal tract produce important digestive secretions that contain electrolytes, enzymes and hormones. The control of gastrointestinal secretion is effected ny neurons and by hormones.*
- *Saliva is a hypotonic fluid with high bicarbonate and potassium concentrations, and an α-amylase that cleaves α-1-4-glycoside bindings in starch.*
- *Saliva cleans the mouth and pharynx (prevents caries), and ease swallowing. Salivary lysozyme lyses bacterial cell walls. The salivary epidermal growth factor promotes the healing of wounds.*
- *Swallowing is a reflex controlled by brainstem neurons forming a swallowing centre.*

- *The swallowing and vomiting mechanisms are blocked by deep anaesthesia and by injury of the 5.th, 9.th or 10.th cranial nerve.*
- *Gastric motility mixes food with gastric juice and subdivides solids to form a fluid composed of small particles.*
- *Gastric glands and mucosa secrete gastrin (G-cells), HCl (parietal cells), pepsinogen (peptic cells), and mucus (mucous neck cells). Mucus and bicarbonate protect the gastric mucosa from adverse HCl effects.*
- *Segmentation mixes the content of the small intestine.*
- *The migrating motor complex is the "intestinal housekeeper", which cleanses the gastrointestinal tract.*
- *Vagal, cholinergic preganglionic fibres transfer signals to the gastrin-producing G-cells in the mucosa via non-adrenergic, noncholinergic (NANC) postganglionic neurons. These enteric neurons liberate gastrin-releasing peptide (GRP) to the G-cells.*
- *The ileocoecal sphincter prevents retrograde flow of colonic matter. The sphincter regulates emptying of ileum some five hours after a meal. The emptying of ileum is stimulated by gastrin, possibly via the gastroileal reflex, but a distended colon inhibits the emptying. The ileocoecal sphincter is normally passed by one litre of faecal matters daily.*
- *In the ascending colon, peristalsis is followed by antiperistalsis, which allow time for absorption of water and electrolytes.*
- *Gluten-sensitive enteropathy or coeliac disease describes a condition where the duodenal and jejunal mucosa is more or less destroyed by hypersensitivity to wards gluten. Gluten is found in barley, rye, wheat and oats.*
- *Acetylsalicylic acid and other non-steroid anti-inflammatory drugs deplete the gastric mucosa for prostaglandins, which leads to mucosal damage. Strong alcoholic beverages also damage the gastric mucosal barrier and stimulate acid secretion. Caffein stimulates gastric acid secretion.*
- *Peptic ulcer disease is a mucosal ulcer in an acid-producing zone in the distal stomach or the proximal duodenum.*
- *All treatment procedures, which work by inhibition of gastric acid secretion in peptic ulcer disease, have a common drawback. To the extent that gastric acid secretion is reduced there is no inhibition of the gastrin release from the antral G cells. Accordingly, the blood [gastrin] increases, and during treatment of the patients this concentration is constantly increased. The high gastrin level counteracts the expected effect on the acid production.*
- *Eradication of Helicobacter pylori with antibiotics is the treatment of choice for most cases of peptic ulcer disease, since it seems to cure the*

patient. Clarithromycin is a macrolide that binds to and prevents translocation on Helicobacter pylori- ribosomes, which is an effective basic therapy of peptic ulcers.

- *Inhibition of the gastric proton pump in the luminal membrane of the parietal cells. Omeprazole is a proton pump inhibitor, which relieves symptoms and cure most duodenal ulcers within four weeks - often in combination with antibiotics. Omeprazole and similar antagonists to the gastric proton pump are especially effective in treatment of persistent HCl-secretion caused by the Zollinger-Ellison syndrome.*

- *Histamine acts through H_2 receptors on the basolateral membrane of the parietal cells. The second messengers for histamine is cAMP. H_2 receptor antagonists (cimetidine, ranitidine, famotidine, and nizatidine) inhibit acid secretion because they fit the H_2 receptors specifically. The H_2 receptor antagonists prevent histamine from binding to the H_2 receptors.*

- *Crohns disease is a chronic infection or inflammation of the gut with a particular prevalence for the terminal ileum, but it can be located all the way along the tract.*

- *Ulcerative colitis is always confined to the colon. Ulcerative colitis is a mucosal inflammation with haemorrhage and rectal bleeding.*

Further Reading

Calver, A., J. Collier and P. Vallance. "Nitric oxide and cardiovascular control." *Experimental Physiology* 78: 303-326, 1993.

Furness, J.B. et al. "Roles of peptides in the enteric nervous system." *Trends Neurosci* 15:66, 1992.

Return to top

Return to content

Chapter 23.

Hepatic And Pancreatic Function And Disorders

Study Objectives

- To *define* concepts such as ascites, a basic hepatic unit, bile acids, biliary cirrhosis, cirrhosis, hepatitis, icterus, ileus, micelles and pancreatitis.
- To *describe* the normal hepatic and pancreatic function and related clinical tests.
- To *explain* the formation of hepatic bile and pancreatic juice, the entero-hepatic bile acid cycle, the function of the bile bladder, the consequences of insufficient bile secretion, the stimulus-secretion coupling in the acinar cells of the pancreas. To explain the control of bile secretion and of exocrine pancreatic secretion. To explain the pathophysiology of common hepatic and pancreatic disorders.
- To *use* the concepts in problem solving and case histories.

Principles

- *The liver is responsible for the key elements of intermediary metabolism, regulating the metabolism of carbohydrates, lipids and proteins.*
- *Pancreas is the classical mixed gland with both exocrine and endocrine elements.*

Definitions

- **Ascites** refers to abnormal accumulation of transudate in the peritoneal cavity.
- **Bile acids** are detergents synthesized from cholesterol in the liver. Cholic acid and cheno-deoxy-cholic acids are conjugated with glycine and taurine, whereby they become water-soluble.
- **Biliary cirrhosis** is a progressive cholestasis with cirrhosis caused by destruction of bile ducts and bile ductules. Antibodies to mitochondria are found in the blood and the aetiology probably includes immunological phenomena.
- **Cholelithiasis** or *gallstone disease* is defined as a condition with gallstones within the lumen of the gallbladder, whether symptoms occur or not.
- **Cholestasis** refers to intra- or extra-hepatic obstruction of the bile flow.
- **Cirrhosis** refers to destruction of the normal hepatic lobular structure by fibrous septa, necrotic hepatocytes and regenerative nodules of hepatocytes.
- **Hepatitis** is either infectious or toxic. Virus, bacteria or protozoa cause infectious hepatitis. Liver toxins, haemolytic toxins, metabolic toxins and drugs cause toxic hepatitis.
- **Icterus** refers to yellow coloration of skin, blood plasma, mucous membranes and tissues. The threshold for visible jaundice (*icterus*) is a [bilirubin] in blood plasma *above 18 mg* l^{-1}

or *30 mM* in most people.

- **Ileus** means intestinal obstruction either due to blockage of the intestinal lumen or to damage or death of smooth muscles that results in paralysis (paralytic or adynamic ileus). The intestine proximal to the lesion is dilatated by fluid and chyme.

- **The liver lobule** is the basic hepatic unit consisting of the hepatic triad: Centrally located is the *central vein* with columns of hepatic cells arranged in radials. Branches of the *hepatic artery* and the *portal vein* are located on the periphery of the lobule. Blood from these vessels perfuse the sinusoids between the hepatocytes.

- **Micelles** are molecular vehicles, consisting of amphipathic bile acids (bile acid micelles) and often aggregates of lipids (mixed micelles). Amphipathic molecules have a hydrophilic and a hydrophobic surface.

- **Pancreatitis** is an inflammatory disease with interstitial oedema in mild cases, and necrosis of the acinar cells of the pancreas in severe cases. The inflammation is not seriously affecting the pancreatic islets.

Essentials

This paragraph deals with: *1.* *Bile*, *2.* *bile acid production*, *3.* *bile pigment production*, *and 4.* *pancreatic exocrine secretion*.

1. Bile

The *basic hepatic unit* is the *liver lobule*. The liver cells (hepatocytes) are arranged into walls of cells, which are separated by highly porous capillaries (Fig. 8-11) called *venous sinusoids*. The portal vein brings blood from the intestine, and the hepatic artery brings arterialised blood from the heart. The venous sinusoids are lined with fenestrated endothelial cells and with specialised, reticuloendothelial *Kupffer-cells*. The large demand of O_2 and nutrients is satisfied from this pool of mixed blood by the hepatocytes; then the blood drains into the *central veins* and leaves the liver through the *hepatic vein*. The hepatocytes form bile and secrete it into *bile canaliculi*, which converge to form a *ductal system*, where bile flows in the opposite direction of the blood. Hereby, cleared blood passes "new" bile. The many bile ducts converge to form the *hepatic duct*. Secretin stimulates the secretion of bicarbonate from the bile ductal system.

Bile has a golden colour and is nearly isotonic with blood plasma. The bile contains NaCl and bicarbonate in concentrations similar to those of plasma, but the bile contains more Ca^{2+} (bound to bile acids) than plasma. We normally produce 0.5-1 litre of hepatic bile per day with bile salts, lecithin, cholesterol and 1.5 g of bile pigments.

The normal gall bladder can concentrate the hepatic bile by a factor up to 5. The bile salts (Cholic acid and Deoxy-cholic acid) are made from cholesterol, which is also abundant in bile. The formation of *mixed micelles* (special fatty aggregates) containing cholesterol, phospholipid, and bile salts, provide concentrations of both phospholipid and cholesterol far exceeding their normal solubility in water. *Cholecystokinin* (CCK) is released by the duodenal mucosa in response to contact with fat and essential amino acids. CCK reaches the gallbladder wall via the blood, and it causes contractions of the gall bladder and relaxation of the sphincter of Oddi. *Gastrin* has a small CCK-effect, and VIP/acetylcholine inhibits gallbladder contractions.

Red blood cells have a life span of 120 days and are continuously being degraded, and the haeme released is taken up from the blood by the hepatocytes to produce bilirubin. Bilirubin is conjugated to glucuronic acid by a transferase in the liver cell to form the golden yellow *bilirubin mono-* and *di-glucuronide.* These conjugates are much more water-soluble than bilirubin, and are thus easily excreted by the bile capillaries (Fig 23-1).

The liver contains a *large store* of vitamin B_{12}. Only 0.1% of the store is lost daily in the bile, because most of its content is reabsorbed in the terminal ileum. Even if absorption totally ceases the hepatic vitamin B_{12} store lasts for *5-6 years.* In the absence of vitamin B_{12} the maturation of erythrocytes is retarded, and *pernicious anaemia* develops (Chapter 8).

2. Bile acid production

Bile acids are detergents synthesized from cholesterol in the liver. Cholic acid and cheno-deoxy-cholic acid are conjugated with glycine and taurine, whereby they become water-soluble. Bile acids have lipophilic and lipophobic terminals, which increase the lipid solubility of the intestinal chyme by micelle formation. Phospholipids and cholesterol expand *simple micelles* into effective *mixed micelles*. In the terminal ileum intestinal bacteria change the two major bile acids into deoxy-cholic acid and litho-cholic acid. The bile acid production is reduced without the return of bile acids from the terminal ileum and steatorrhoea develops.

Cholate and *desoxycholate* are *fat-soluble agents*. They have *fat-soluble hydrocarbon rings* that enable them to mix with fats and *several charged groups* that enable them to mix with water. Large fat droplets in the duodenal chyme become dispersed, forming smaller fat particles - a process called *emulsification*. The bile salts contribute to emulsification.

These smaller fat particles are efficiently digested by the *water-soluble pancreatic lipases*, forming glycerides and fatty acids in the *micelles* (Fig. 22-13).

The micelle contents are readily absorbed by the enterocytes. The co-lipase helps the lipase to eliminate the inhibitory bile salts from the surface, so that the *lipase* is fixed to the *lipids.*

The pancreatic lipase cleaves the ester linkage of tri-acyl-glycerol at the 1- and 3-position, releasing two fatty acids and 2-monoglyceride (2 MG), or occasionally a free glycerol molecule (Fig. 22-13). Free glycerol is readily absorbed. A protein - *fatty acid binding protein* (FABP; 12 kDa) - is present in the cytosol of the enterocytes. FABP binds fatty acids in order to *re-esterify the fatty acids* and to protect the cell from adverse effects of *cytotoxic fatty acids*. Once the fatty acids are formed, the fatty acids and 2-monoglyceride participate in the emulsification process, but the fatty acids still require bile salts for complete water solubility. Micelles are passed through the aqueous bowel lumen to reach the absorbing mucosa (Fig. 22-13). A large concentration of *bile salts* helps the *lipid laden micelles* to get access to the absorbing surface. Then lipids diffuse easily out of the *lipophilic micellar core* and into the *lipid layer* of the apical membrane of the mucosal cell.

Lipids do not adequately form micelles, if there is no bile present.

3. Bile pigment production and excretion

Mature red cells are continuously being degraded in the macrophages of the liver and in the reticulo-endothelial system (RES) of the spleen and bone marrow. The haeme is converted to biliverdin, which is reduced to produce *bilirubin* (Fig. 23-1). About 10-15% of the bilirubin arises from the breakdown of immature red cells and cytochrome. Bilirubin released to the blood from these tissue macrophages is bound to albumin during its transport, so the concentration of neurotoxic, free bilirubin is normally low. Bilirubin is taken up at the hepatocyte cell membranes after dissociation from the albumin. Within the hepatocyte bilirubin is transferred to the endoplasmic reticulum, which contains a *transferase* that conjugates bilirubin with glucuronic acid. Conjugated bilirubin (bilirubin di- and mono-glucuronide) is water-soluble, so it diffuses easily through the cytoplasm and is actively secreted into the bile canaliculi. From here conjugated bilirubin is excreted into the intestine with the bile. Bacterial enzymes in the terminal ileum and the colon hydrolyse the large molecule. The free bilirubin is then reduced to *urobilinogen*.

Most of the urobilinogen is excreted in the faeces; the remainder is absorbed in the terminal ileum, returned to the liver via enterohepatic blood, and again excreted as bile into the

intestine (the enterohepatic bile pigment circuit). A small amount of the urobilinogen is excreted in the normal urine.

Fig. 23-1: Normal bilirubin metabolism. The free bilirubin is fat-soluble and toxic. The conjugated bilirubin is water-soluble bilirubin-glucuronide and non toxic.

Part of the bilirubin is broken down to colourless substances, hepatocytes produce urobilinogen, and colonic bacteria stercobilinogen. Both substances can be oxidised to *yellow* urinary *urobilin* and *brown* faecal *stercobilin* (Fig. 23-1). The renal excretion of urobilin and stercobilinogen is increased in cases with intrahepatic icterus (ie, hepatitis and other damages of hepatocytes).

The terminal ileum is essential for life.

1. The terminal ileum of humans absorbs conjugated bile salts efficiently by an active Na^+ -dependent *co-transporter* that is similar to the glucose/Na^+ - and amino acid/Na^+ -co-transporter in the duodenum-jejunum (Fig. 23-2).

Fig. 23-2: Absorption of bile acids by the terminal ileum.

2. Bile acids also cross the brush border by *diffusion* in unconjugated form.
3. In the cytosol the bile acids are probably bound to macromolecules, and they traverse the basolateral membrane by
4. facilitated transport and diffusion into the portal blood.

The absorbed bile salts reach the liver, where they are conjugated and reprocessed, and the hepatocytes clear the portal blood from bile acids in a single passage. The reabsorbed bile acids are essential stimuli for the liberation of bile with new (15%) and reprocessed (85%) bile acids, but when entering the liver they *inhibit* the synthesis of new bile acids.

The total bile acid pool in the body is about 3 g, and this pool can be recycled up to 12 times per day.

By contrast, the intestine reabsorbs only a small part of the bile pigments - the *enterohepatic bile pigment circuit* (Fig. 23-1).

4. Pancreatic exocrine secretion

PAN KREAS is Greek and means *all meat*. Pancreas is the classical mixed gland with both endocrine and exocrine elements. The exocrine pancreas is an *abdominal* “salivary gland.” The endocrine pancreas is described in Chapter 29.

Secretions from the zymogen containing acinar cells collect in the acinar duct and travel through a network of converging ducts to the *main pancreatic duct*, which run into the common bile duct entering the duodenum at the duodenal papilla (Vateri), where the sphincter of Oddi is located.

The *exocrine* glandular tissue consists of *acinar cells* producing a primary secretion, with an ionic composition similar to that of plasma, and *duct cells* forming the secondary secretion by modification of the primary secretion.

The organic components, secreted by acinar cells, are the major enzymes necessary for digestion of dietary nutrients. The acinar cells also secrete mucus and ions.

Fig. 23-3: Secretion of enzymes from pancreatic acinar cells. ISF means interstitial fluid.

1. Upon stimulation of the duodenal mucosa with acid chyme containing peptides and long chain fatty acids, *cholecystokinin* (CCK) is released to the blood, whereby it can reach the pancreatic acinar cells. They carry specific receptors that bind the *gastrin-family* (gastrin and CCK competing for the same receptor) as well as the neurotransmitter,

acetylcholine (ACh in Fig. 23-3). Receptor-ligand binding activates IP_3, thus elevating $[Ca^{2+}]$ in the cells. Ca^{2+} triggers exocytosis of the enzymes from the zymogen granules, utilising either a Ca^{2+}-calmodulin complex or a Ca^{2+}-phosphatidyl serine-dependent *proteinkinase C*. - The secretin family potentiates the action of CCK.

2. The rise in $[Ca^{2+}]$ opens a *luminal* Cl^-- and a *basolateral* K^+*-channel*, whereby these ions are leaving the cell in a balanced relationship and produce a negative electric field in the acinar lumen. A small amount of bicarbonate also leaks through the anion channel.
3. The fall in intracellular $[Cl^-]$ and $[K^+]$ activates a basolateral Na^+-K^+-*2* Cl^- *co-exchanger* through which NaCl enters the cell from the ISF.
4. The negative electric field in the acinar lumen provides a force that drives a passive Na^+- and water transport as an isotonic solution into the acinar lumen through leaky tight junctions.
5. The secretory energy is from the basolateral Na^+-K^+-pump, which maintains the intracellular ion composition.

As the primary juice leaves the acini and proceeds down the pancreatic ducts, it is supplied isotonically with water and *electrolytes* (mainly bicarbonate salts) from the duct cells (Fig. 23-4).

Fig. 23-4: Secretion from pancreatic duct cells.

Upon stimulation of the duodenal mucosa with acid chyme, *secretin* is secreted to the blood and transported to the pancreatic duct cells. This induces an important rise in cellular [cAMP].

The rise in [cAMP] activates luminal Cl^-- and basolateral K^+-channels, so these ions leave the cell in a balanced relationship. This triggers a luminal Cl^-/HCO_3^-*co-exchanger* through which the cell eliminates the bicarbonate produced by carbon dioxide from the blood. Cellular carboanhydrase is essential for the bicarbonate production. A certain luminal $[Cl^-]$ is necessary for recycling. The net result is a secretion of bicarbonate.

This net secretion of bicarbonate induces a (lumen negative) transepithelial potential difference (-6 mV), which constitutes the driving force for the paracellular transport of Na^+ and K^+. The net secretion of salt drags water trans-epithelially in isosmotic proportion.

The fall in cellular pH upon secretion of bicarbonate activates a basolateral Na^+/H^+-*coexchanger*, whereby the cells eliminate H^+ to the blood.

The secretory energy is from the basolateral Na^+-K^+-pump. The whole system is analogous to the formation of saliva.

The pancreas (weight 100 g) of adult humans is capable of elaborating approximately 1.5 l of pancreatic juice daily, and its pH increases with increasing secretion rate. The maximal secretion rate is one ml/g of tissue each hour (ie, 60 times less than that of the salivary glands).The pancreatic juice is a clear fluid, isosmolar with plasma. The basic reaction is due to bicarbonate, and the [bicarbonate] can approach the $[H^+]$ in gastric juice (150 mM).

Fig. 23-5: Concentrations of ions (mM) in pancreatic juice as a function of the secretory flow rate (left). – The control of pancreatic secretion is shown to the right.

With increasing secretion rate, the [bicarbonate] in the final pancreatic juice increases at the expense of $[Cl^-]$, whereas the $[Na^+]$ and $[K^+]$ remain relatively constant (Fig. 23-5). Pancreatic juice (pH 8) thus buffers the *extremely acid* gastric juice and protects the duodenal

mucosa against *erosion*. Buffering of gastric juice also optimises the activity of pancreatic digestive enzymes in the duodenum.

The pancreatic secretion is regulated by two intestinal hormone families (Fig. 23-5, right): The *secretin*- (secretin and VIP) and the *gastrin*-family (gastrin and CCK), as well as by the autonomic nervous system.

Signals in *cholinergic, vagal fibres* stimulate both pancreatic secretions via acetylcholine-receptors (Fig. 23-3), whereas noradrenergic, sympathetic stimuli inhibit secretion via a-receptors. The secretion is also stimulated by signals in *peptidergic nerve fibres*. The free radical gas *nitric oxide* (NO) stimulates the exocrine pancreatic secretion, and simultaneously inhibits the non-adrenergic, non-cholinergic intestinal activity (Fig. 22-6).

Related to the meal there are *three* phases of pancreatic secretion (cephalic, gastric and intestinal).

1. The *cephalic phase* is elicited before food reaches the stomach. Olfactory signals (via the limbic system) as well as visual and tactile signals (via the thalamic relay station) are processed in the brain, and vagal signals reach the antral mucosa. Here gastrin is released from G-cells. Gastrin induces the secretion of a low volume of pancreatic juice with a high enzyme content.

2. The *gastric phase* is elicited by the presence of food in the stomach. Gastric distension and peptides reaching the antral mucosa trigger the release of more gastrin from the G-cells. Hereby, the secretion of a small volume of pancreatic juice rich in enzymes is continued.

3. The *intestinal phase* is elicited by duodenal and jejunal mechanisms. When chyme enters the duodenum both secretin and CCK is released for different reasons. *Secretin* is secreted by S-cells in the mucosa of the upper small intestine, when acid chyme (pH below 4.5) arrives to the first part of the duodenum. This is an appropriate arrangement, because secretin stimulates both the secretion of bicarbonate and water by pancreatic duct cells (Fig. 23-4), and of bicarbonate-rich bile by small biliary ductules. Secretin inhibits gastric secretion. Secretin inhibits both the gastrin release by the antral G-cells, and the gastrin effect on the parietal cells.

   CCK from the duodenal I-cells stimulates gallbladder contraction as its name implies, and stimulates pancreatic acinar secretion of an enzyme-rich fluid (Fig. 23-3). The most important stimulus for CCK liberation is when an acid chyme with amino acids, peptides and long chain fatty acids reach the duodenal mucosa. This is essential. CCK contracts the gallbladder and stimulates the pancreatic secretion of an enzyme rich juice. Bile is ejected into the duodenum, where fat is emulgated to ease absorption. CCK also acts as an *enterogastrone* - an intestinal hormone that inhibits gastric activity and emptying. This leaves more time for the bile to emulgate fat and for the digestible enzymes to work.

Pancreatic α-amylase does not pose any danger to pancreatic tissue. Pancreatic a-amylase - like the salivary a-amylase - cleaves the large dietary carbohydrate molecules at the internal 1,4-glycosidic bonds, but cannot hydrolyse terminal 1,4-bonds or 1,6-bonds. The end-products are oligo- and di-saccharides like maltose (two glucose), maltriose (three glucose) and branched oligosaccharides known as a-limit dextrins. Other enzymes such as maltase and lactase secreted by the intestinal mucosa digest these end-products into monosaccharides (glucose, fructose and galactose). The carbohydrate absorption is shown in Fig. 22-11.

The protein digestion is continued in the duodenum and jejunum, where the protein breakdown products are attacked by the proteolytic enzymes of the pancreas (trypsin, trypsinogen, chymo-trypsinogen, pro-carboxy-peptidase, and pro-elastase). The pancreatic proteases are secreted as

inactive proenzymes, and they are crucially important. The proenzymes are normally not activated before they arrive in the intestinal lumen. Trypsin catalyses its own activation (autocatalysis), and also activates chymotrypsinogen and the pro-carboxypeptidases in the trypsin cascade.

Duodenal enterokinase cleaves trypsinogen to trypsin, and hereby activates the trypsin cascade. When the chyme is pushed into the duodenum, the pancreatic juice neutralises the chyme and the pepsin activity is stopped. The peptides and amino acids are absorbed as shown in Fig. 22-12.

Normally, trypsin inhibitor inhibits the trypsin cascade from the pancreas, but cases of acute pancreatitis cannot inhibit the trypsin cascade, so autodigestion occurs.

Enzymes for the breakdown of fats are pancreatic lipase, phospholipase A, and lecithinase.

Pancreatic lipase and co-lipase cleave triglycerides into free glycerol and fatty acids or to mono-glycerides (MG) and fatty acids. Free glycerol is readily absorbed. The lipolytic activity requires the emulsifying action of bile salts in order to solubilize triglycerides in water. Once liberated fatty acids and mono-glycerides participate into bile salt micelle formation. Micelles pass by diffusion through the unstirred water layer of the intestinal lumen to reach the absorbing mucosa (Fig. 22-13).

Patophysiology

This paragraph deals with *1.Jaundice, 2.Gallstones, and 3. Hepatitis, 4. Liver cirrhosis, 5. liver cancer, 6. Pancreatitis, 7. Cystic fibrosis, 8. Carcinoma of the pancreas, 9. Endocrine pancreatic tumours*.

1. Jaundice (icterus)

Bilirubin has a molecular weight of 588 g per mol. The normal [bilirubin] in blood plasma is up to 17 mg l^{-1} or 29 mmol l^{-1} (mM). The threshold for visible jaundice (icterus) is a [bilirubin] in blood plasma *above 18 mg l^{-1} or 30 mM* in most people.

Three types of icterus can be distinguished:

1.a. Prehepatic or haemolytic icterus is caused by excessive destruction of mature or immature red cells. Haemolytic anaemia causes haemolytic jaundice. Increased destruction of red cells (haemolysis) increases the bilirubin production (normally 35 * 6 = 210 mg daily) to the extent that the hepatocytes cannot conjugate the bilirubin as rapidly as it is formed (the key hole enzyme is *glucuronyl transferase*). The neurotoxic free bilirubin in blood plasma rises much above normal, and large quantities of urobilinogen is excreted in the urine.

1.b. Intrahepatic icterusis caused by poor hepatocyte function. Damages of the hepatocytes by infections, tumours, or toxic agents impair the uptake, transport and conjugation of bilirubin. Absence of glucuronyl transferase or inhibition of the enzyme by steroids *block* conjugation of bilirubin.

1.c. Posthepatic icterus is caused by *cholestasis* due to gallstones or pancreatic tumours. Gallstones or tumour masses obstruct the bile ducts, which is causing extrahepatic cholestasis with impaired excretion of conjugated bilirubin to the intestine. Hereby, *conjugated bilirubin reflux* to the blood. Most of the bilirubin in plasma is therefore conjugated and some of it strongly bound to plasma albumin.

Hepatic Failure results from destruction of liver cells or impairment of hepatocyte function. Liver failure causes *severe jaundice*, hepatic encephalopathy, the hepatorenal syndrome, pulmonary veno-arterial shunts, and low coagulability of the blood.

2.Gallstones

Cholelithiasis or *gallstone disease* is defined as a condition with gallstones within the lumen

of the gallbladder, wether symptoms occur or not.

More than 70% are *cholesterol stones;* the remainder is a so-called brown *pigment gallstone* composed of Ca^{2+} salts of bilirubin, carbonate, cholesterol and phosphate (Fig.23-6).

Fig. 23-6: The two common types of gallstones: Cholesterol stones and brown pigment stones.

Approximately half of the *pigment stones* are *radiopaque*.

The incidence of cholesterol stones among females is three times higher than among males. The cause is both genetic and environmental. Oestrogens stimulate hepatic secretion of cholesterol and reduce the formation of bile acids. Diets high in cholesterol increase the incidence of gallstones.

Bile can be supersaturated with cholesterol. When this happens, crystals can precipitate out of bile. Cholesterol crystals and Ca^{2+} can aggregate and develop into *gallstones* in the common bile duct. Mixed gallstones and gallstones made of bile pigment and other bile substances are also found.

Excessive removal of water in the gallbladder can be pathogenic. Enlarged gallstones can obstruct the common bile duct thus causing bile with bilirubin to flow back into the liver and leak into the blood plasma (*jaundice* or *icterus*).

Most gallstones are asymptomatic, but those occluding the biliary tract cause a severe pain called *biliary colic.*

3. Hepatitis

Hepatitis is either infectious or toxic. Virus, bacteria, or protozoa cause infectious hepatitis.

Virus molecules cause most of all global hepatitis (see Chapter 33).

Bacterial and rickettsial hepatitis is caused by leptospirosis interrogans icterohaemorrhagia (Weil-hepatitis), rickettsia (typhus-hepatitis), streptococci, pneumococci etc - sometimes by ascending infection after biliary tract disorders.

Protozoan hepatitis is caused by the species of plasmodium in malaria-hepatitis, trypanosoma in trypanosomiasis, and toxoplasma gondii in toxoplasmosis-hepatitis. The clinical course is often like infectious mononucleosis.

Toxic hepatitis is caused by liver toxins, haemolytic toxins, drugs (isoniazid, methyl-dopa, nitrofurantoin, oxyphenisatin) and metabolic toxins (ileus, ulcerative colitis, pregnancy hyperemesis, thyrotoxicosis).

4. Hepatic cirrhosis

Cirrhosis is the end result after *necrosis of the hepatocytes*, with destruction of the normal lobular structure by fibrous septa and regenerative nodules of hepatocytes. The clinical picture includes liver failure and signs of portal hypertension such as oesophageal varicose veins and ascites. The terminal stage is *hepatic coma*.

Two pathological types are considered:

a. Micronodular cirrhosis is characterized by nodules less than 3 mm in diameter. This disorder was previously termed Laennec's cirrhosis (after a French pathologist). The cause is alcohol abuse (*alcoholic cirrhosis*) or biliary tract disease (*biliary cirrhosis*).

b. Macronodular cirrhosis is characterized by larger nodules sometimes including normal lobules. The cause is acute and chronic hepatic infection (hepatitis B virus, hepatitis C virus, hepatitis D virus) often in carriers.

Alcoholic cirrhosis

This type of cirrhosis follows years of alcohol abuse with fatty liver and alcoholic hepatitis.

Fig. 23-7: Alcohol metabolism and fat accumulation.

Ethanol is metabolised by the liver hepatocytes to acetaldehyde and acetate. At rest 5% of the molecules are excreted unchanged in the urine, sweat and expired air. The major route of ethanol oxidation is via alcohol-dehydrogenase an NAD-dependent enzyme (Fig. 23-7). A minor pathway is microsomal ethanol oxidising system (MEOS) in the smooth endoplasmic reticulum using NADP as a cofactor. Both enzyme systems are easily saturated, so a fixed quantity is metabolised per time unit (7 g per hour as an average). Both pathways result in an increased NADH/NAD ratio, whereby the fatty acid synthesis is increased. This leads to *fat accumulation* with fatty liver and alcoholic cirrhosis. The altered redox-potential and the accumulation of acetaldehyde, causes *centrilobular necrosis* of the micronodular type.

Biliary cirrhosis

The primary type is a progressive cholestasis with cirrhosis caused by destruction of bile ducts and bile ductules. Antibodies to mitochondria are found in the blood and the aetiology probably includes immunological phenomena. There is a high total cholesterol (especially HDL), pruritus, xanthomas, osteoporosis, steatorrhoea, and portal hypertension. The condition is associated with many autoimmune diseases such as rheumatoid arthritis, scleroderma, renal tubular acidosis, and membranous glomerulonephritis (Chapter 32).

The secondary type of biliary cirrhosis result from maintained extrahepatic cholestasis.

Cirrhosis related to hepatitis

The most frequent cause is viral hepatitis in particular infection with hepatitis B virus (also hepatitis D and C virus).Viral markers are examined in the blood such as HBsAg or its antibodies, or IgM anti-HCV (Chapter 33).

5. Liver cancer

Seemingly healthy carriers of HBV and HCV are at risk of developing *hepatocellular carcinoma.* The *α-foetoprotein* is raised in the blood plasma. Other risk factors for hepatocellular carcinoma are alcoholic damage, haemochromatosis, aflatoxin from peanuts, androgens, and oestrogens.

Metastases from breast cancer, bronchial cancer and gut cancer to the liver are much more frequent than primary hepatic tumours.

6. Pancreatitis

Pancreatitis is an inflammatory disease with interstitial oedema in mild cases, and necrosis of the acinar cells of the pancreas in severe cases. The incidence of pancreatitis in alcoholics is high, whereas infection as such does not account for many cases. Both the first acute case of pancreatitis and the chronic cases are linked to alcohol abuse.

a. *Acute pancreatitis* is characterised by *foci* of necrotic fat cells besides the acinar cell necrosis and the infiltration of polymorpho-nuclear leucocytes. The injured acinar cells release digestive enzymes (amylase, lipase, and protease) into the blood stream and into the peritoneal fluid causing ascites.

Severe epigastric pain is referred to the back between the shoulder blades. The patient is critically ill and develops a shock condition, which may end in terminal renal failure. The diagnosis relies on a 500% rise in serum amylase concentration together with the demonstration of amylase in the abundant peritoneal fluid (*ascites*).

Premature activation of the pancreatic digestive enzymes causes the pancreas to digest itself. The essential enzymes are trypsin, phospholipase A and elastase. The normal balance between *trypsin* and *trypsin inhibitor* is destroyed. Protease inhibitors such as *α -antitrypsin, C -*

esterase and *trypsin inhibitor* bind to the proteases and reduce their enzyme activity until the inhibitors are digested by the enzymes. Phospholipase A destruct cell membranes and converts lecithin to the cytotoxic lysolecithin. Vessel walls are broken down and haemorrhage can be fatal. Capillary destruction causes anoxia and necrosis.

Alcohol (and acid chyme) stimulates the secretion of *secretin* from the duodenum and thus the secretion of an enzyme-rich pancreatic fluid. Simultaneously, alcohol closes the *sphincter of Oddi*, and pancreatic secretions from the obstructed duct are filtered into periductal tissues.

The treatment is symptomatic with intravenous nutrition and analgesia. Adequate therapy of shock and respiratory insufficiency must be instituted (transfusions, endotracheal intubation and oxygen if necessary).

b. *Chronic pancreatitis* is a progressive destruction of pancreatic acini, followed by irreversible fibrosis and possibly calcification of the pancreatic tissues. Calcifying pancreatitis is strongly linked to alcohol abuse.

The patient has episodes of epigastric pain, and the anorexia results in severe malnutrition. Most patients with calcifying pancreatitis develop *steatorrhoea* (a copious fatty faeces) and *diabetes.*

Steatorrhoea is treated with low-fat diet and the enzyme, pancreatin, to each meal. The diabetic condition is frequently characterized by a simultaneous lack of pancreatic insulin and glucagon, whereby the daily, external insulin requirement is increased. Alcoholics with pancreatitis must stop drinking alcohol.

Pancreatic failure is caused by *pancreatitis* (inflammation not seriously affecting the Islets), *blockage* of the pancreatic duct, pancreatic *carcinomas* and *surgical removal* of the pancreatic head. Loss of pancreatic juice means lack of pancreatic lipase, pancreatic amylase, trypsin, chymotrypsin, carboxy-poly-peptidase, and elastase. Lack of these enzymes means that half of the fat entering the small intestine pass unabsorbed to the faeces, and one third of the starches and proteins. Steatorrhoea is found.

The major metabolic disorder caused by loss of pancreatic endocrine secretion is diabetes mellitus. Removal of pancreas is compatible with *survival* as long as both the exocrine and endocrine vital substances are supplied artificially.

7. Pancreatic cystic fibrosis (mucoviscidosis)

This is a recessive genetic defect with dysfunction of exocrine glands (see Chapter 31).

8. Carcinoma of the pancreas

This is almost exclusively adenocarcinoma originating from the duct cells of the pancreatic head. Its occurrence is related to alcohol abuse.

The diagnosis is made by CT or by ultrasound technique. Surgical removal of pancreatic adenocarcinoma is frequently unsuccessful.

Fig. 23-8: Endocrine pancreatic tumours from the pancreatic islet cells.

9. Endocrine pancreatic tumours

1. *Insulinomas* are islet cell tumours of *β-cells,* which release sufficient insulin into the blood to induce serious hypoglycaemia. The patient must eat extensively in order to survive, so obesity is almost unavoidable.
2. *Somatostatinomas are* islet cell tumours of D- or d-cells that secrete somatostatin to the blood. Somatostatin causes diabetes, steatorrhoea, gallstones and hypo-chlorhydria.
3. *Glucagonomas* are islet cell tumours of *α-cells* that release large amounts of glucagon into

the blood stream. This causes diabetes, anaemia and a typical erythematosus rash.

4. *Vipomas produce* large quantities of VIP, and the diagnosis is made by a *high plasma VIP*. VIP increases the intestinal secretion and causes watery diarrhoea with loss of K^+ and H^+. Localisation and removal of the vipoma is efficient.

5. *Gastrinomas* (Zollinger-Ellisons syndrome) consist of G-cells in the pancreatic islets. The G-cells produce large amounts of *gastrin* causing extensive gastric HCl secretion and peptic ulcers. The serum gastrin is high. The patient has diarrhoea, because of the high H^+-concentration in the intestinal lumen. Omeprazole inhibits the *proton pump*, whereby the gastric HCl secretion is blocked.

6. *PP-producing tumours.* The concentration of pancreatic polypeptide (PP) in plasma is increased. PP is released by most pancreatic islet cell tumours, and its plasma concentration is always measured in screenings of islet cell tumours. PP stimulates the gastrointestinal enzyme secretion and inhibits smooth muscle contraction. Localisation and ablation of the tumour is the ideal therapy, but *steroids* and **o***ctreotide* are of help.

Equations

- For calculation of hepatic bloodflow see the Fick Principle in Chapter 10.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

A. With increasing rate of pancreatic secretion its [Cl^-] will increase.

B. Carboanhydrase is an important enzyme for the secretion of pancreatic bicarbonate.

C. Duodenal acidification stimulates the pancreatic secretion.

D. Duodenal chyme with a pH of 7 inhibits pancreatic bicarbonate secretion.

E. Duodenal enterokinase cleaves trypsinogen to trypsin.

II. Each of the following five statements have False/True options:

A. An elevated concentration of cAMP in the intestinal mucosal cells inhibits Na^+ absorption.

B. The intestinal Na^+ absorption is parallel to the Cl^- absorption.

C. Increased intracellular [Ca^{2+}] increases intestinal Na^+ absorption.

D. The basolateral Na^+-K^+-pump (ATPase) maintains an essential electrochemical gradient with a high intracellular [Na^+].

E. The intestinal Na^+ absorption is secondary to the water transport across the mucosal cells.

III. Each of the following five statements have True/False options:

A: Bile acids are essential for solubilizing cholesterol and phospholipids by formation of

micelle aggregates.

B: Bilirubin binds to cytoplasmic proteins within the hepatocyte.

C: The primary bile acids are deconjugated and dehydroxylated to form the secondary bile acids.

D: Intrinsic factor-cobalamin complexes are inactivated by pancreatic proteases.

E: Cholate and desoxycholate have water-soluble hydrocarbon rings.

IV. Each of the following five statements have False/True options:

A**:** Hexoses and amino acids require Na^+ for active transport into the enterocyte.

B: A person with lactase deficiency cannot digest lactose, so undigested lactose from a milky diet would enter the colon.

C: The Na^+-K^+-pump is essential for intestinal Na^+ absorption.

D: All lipase proteins are lipid soluble.

E: Cytosolic peptidases from the enterocytes and brush border peptidases cannot cleave small peptides into single amino acids.

Case History A

A male patient with coecal cancer still maintains his colon function. Approximately 1.5 l of intestinal fluid passes the ileocoecal valve in 24 hours, and only 150 ml is found in the daily faeces. The intestinal fluid has a [Na^+] and [K^+] of 120 and 4 mM, respectively. The 75% water in the faecal volume of 150 ml contains 20 mM Na^+ and 5 mM of K^+.

1. Calculate the water absorption in the colon and rectum.

2. Calculate the net Na^+-and K^+-absorption in the colon.

3. Calculate the loss of Na^+ and K^+ with faeces.

4. Removal of the coecum and the ascending colon with the tumour necessitates an ileostomy. Calculate the loss of Na^+ and K^+ with an unchanged fluid flux through the terminal ileum.

5. Are dietary measures important for the ileostomy patient?

Case History B

A female, age 42 years, is admitted to hospital due to fatigue. She describes a serious gastrointestinal infection for which she was cured some years ago. Suspicion of vitamin B_{12} deficiency reveals a seriously low vitamin B_{12} concentration and plasma antibodies against her parietal cells in the gastric mucosa.

Assume that the absorption of vitamin B_{12} totally ceased at the time where her parietal cells were destroyed by autoimmune disease. Assume further that she had a normal liver store of 5 mg vitamin B_{12}, and that she has lost 1 permille daily of the hepatic store in the bile.

1. Calculate the half-time period necessary to reduce the hepatic vitamin B_{12} store by 50%.

2. Calculate the number of years it takes to empty the hepatic vitamin B_{12} store down to 0.5 mg (manifest pernicious anaemia).

Case History C

An alcoholic male with hepatic insufficiency is brought to the intensive care unit of a hospital in hepatic coma.

Normally, ammonia is formed in the gastrointestinal tract as a product of protein digestion and bacterial action. The liver usually removes a major portion by converting ammonia into urea. Hereby, the toxic ammonia is eliminated.

The impaired liver function of this patient has lead to development of collateral venous shunts with oesophageal varicosities. Large quantities of blood from the gut, with a high [NH_4^+], are transported directly into the systemic veins and the brain of the patient. His blood [NH_4^+] is drastically increased, and his blood [glucose] is 2 mM (hypoglycaemia).

1. What is hepatic coma? What is causing the unconsciousness?

2. Explain his condition in terms of abnormal glucose metabolism.

Try to solve the problems before looking up the answers.

Highlights

- *The liver controls the intermediary metabolism of carbohydrates, lipids, and proteins. The liver produces important plasma proteins including coagulation factors, angiotensinogen, trypsin-inhibitor etc.*
- *The liver is a vital glucose exchanger for the hypothalamic glucostat.*
- *The hepatocytes secrete bile (CCK-stimulated) into the bile capillaries and the bile finally enters the small intestine. Bile acids facilitate fat digestion and absorption. Bile acids emulsify lipids, so they are easily accessible for the action of lipid-digesting enzymes.*
- *Bile acids form micelles, which can diffuse to the brush-border membrane for absorption.*
- *Bile acids and vitamin B12 are absorbed in the terminal ileum and returned to the liver in the portal vein.*
- *Hepatocytes clear the blood for bile acids and they are recycled several times daily (enterohepatic recycling).*
- *The bile duct cells and the pancreatic duct cells secrete a bicarbonate-rich fluid (secretin-stimulated).*
- *The liver has an important excretory function, because the hepatocytes excrete bile pigments, and deactivate hormones, toxins and drugs by hydroxylation, proteolysis and hydrogenation. Lipophilic drugs are converted into water-soluble drugs that are easily excreted in bile or urine.*

- *The hepatic reticuloendothelial system normally stimulates repair of tissue damages through influence on T- and B-lymphocytes. Intrahepatic disease impairs the normal immune response to infection elsewhere.*
- *The reticuloendothelial system of the liver eliminates microbes and other antigens transferred to the liver with the blood from the intestines. Antigens are phagocytized by macrophages attached to the endothelium (Kupffer-cells). The macrophages produce collagenase, hydrolases, and interleukins and tumour necrosis factor that degrade the antigens without formation of antibodies.*
- *The normal hepatic store of vitamin B_{12} is sufficient for 3-6 years, and there is also an important store of other vitamins (A, D, and K). Coagulation factors are stored as well as iron in ferritin.*
- *Approximately half of the total lymph produced in the body is liver lymph, although the liver is only 1.5 kg of the total body weight. Chylomicrons filled with lipids reach the blood via the liver lymph and liver.*
- *The primary oxidation of alcohol (ethanol) occurs in the hepatocytes. Since the alcohol-dehydrogenase activity has a maximum capacity, the elimination rate is constant (0.0025 permille per min).*
- *A pressure rise in the hepatic veins from zero (normally) to 5 mmHg (0.7 kPa) result in hepatic stasis with ascites. Hepatic stasis leads to hepatic failure with jaundice and fatty stools (steatorrhoea).*
- *The pancreas (weight 100 g) of adult humans is capable of elaborating approximately 1.5 l of pancreatic juice daily, and its pH increases with increasing secretion rate. The maximal secretion rate is one ml per g of tissue each hour (ie, 60 times less than that of the salivary glands).*
- *The pancreatic juice is a clear fluid, isosmolar with plasma. The basic reaction is due to bicarbonate, and the [bicarbonate] can approach the $[H^+]$ in gastric juice (150 mM).*
- *Pancreatic acinar cells produce enzymes for digestion of carbohydrates, proteins and fats. CCK stimulates the enzyme secretion.*
- *Acute biliary tract disease is diagnosed by biliary pain and confirmed by ultrasonographic or computer tomographic evidence of a distended/inflamed gall-bladder. The disorders are subdivided into inflammatory (acute cholecystitis) or obstructive (eg, stone in the common bile duct).*

Further Reading

Baillière´s Clinical Gastroenterology. Quarterly reviews of Gastroenterology. London: Bailliere Tindall.

Gastroenterology. Monthly journal published by the Am. Gastroenterological Association, WB Saunders Co, The Curtis Center Suite 300, Independence Square West, Philadelphia , Pa 19106-3399, USA.

Hug, M., C. Pahl, and I. Novak. "Effect of ATP, carbachol and other agonists on intracellular calcium activity and membrane voltage of pancreatic ducts." *Pflügers Arch* 426: 412-

18, 1994.

Return to top

Return to Content

Chapter 24.

Body Fluids And Regulation

Study Objectives

- To *define* the concepts: Dehydration, hyponatraemia, intracellular fluid volume (ICV), extracellular fluid volume (ECV), interstitial fluid (ISF), overhydration, oxidation water, radioactivity, specific activity, and total body water.

- To *describe* the daily water balance, the K^+- and Na^+-balance, sweat secretion, the ionic composition in blood plasma, the water content of fat- and muscle- tissue and the daily water transfer across the gastro-intestinal mucosa. To describe the osmotic pressure in the body fluids, the measurement of fluid compartments by indicator dilution, the measurement of total body-K^+ and -Na^+ and the related dynamic pools.

- To *draw* models of the body fluid compartments.

- To *explain* the influence of age, sex and weight on the size of the total body water and its phases. To explain disorders with increased or reduced extracellular fluid volume and shock.

- To *apply* and *use* the above concepts in problem solving and in case histories.

Principles

- *The law of conservation of matter states that mass or energy can neither be created nor destroyed (the principle of mass balance). The principle is here used to measure physiological fluid compartments and the body content of ions.*

Definitions

- **Concentration:** The concentration of a solute is the amount of solute in a given fluid volume.

- **Dehydration** is a clinical condition with an abnormal reduction of one or more of the major fluid compartments (ie, total body water with shrinkage of blood volume or ISF).

- **Dextrans** are polysaccharides of high molecular weight.

- **Intracellular fluid volume** (ICV) refers to the volume of fluid inside all cells. This volume normally contains 26-28 litre (l) out of the total 42 l of water in a 70-kg person. - One litre of water equals one kg of water.

- **Extracellular fluid volume** (ECV) refers to the interstitial and the plasma volume. The ECV contains the remaining water (14-16 kg) with most of the water in tissue fluid (ISF) and about 3 kg of water in plasma. - Interstitial fluid (ISF) is the tissue fluid between the cells in the extravascular space.

- **Hyperkalaemia** refers to a clinical condition with plasma-[K^+] above 5 mM (mmol/l of plasma).
- **Hypokalaemia** refers to a clinical condition with plasma-[K^+] below 3.5 mM.
- **Hypernatraemia** refers to a clinical condition with plasma-[Na^+] above 145 mM.
- **Hyponatraemia** refers to a clinical condition with plasma-[Na^+] below 135 mM.
- **Oedema** refers to a clinical condition with an abnormal accumulation of tissue fluid or interstitial fluid.
- **Osmolality** is a measure of the osmotic active particles in one kg of water. Plasma-osmolality is given in Osmol per kg of water. Water occupies 93-94% of plasma in healthy persons. Plasma osmolality is normally maintained constant by the antidiuretic hormone feedback system.
- **Overhydration** refers to a clinical condition with an abnormal increase in total body water resulting in an increased ECV and thus salt accumulation.
- **Oxidation water** or *metabolic water* (oxidative phosphorylation) refers to the daily water production by combustion of food - normally 300-400 g of water daily in an adult.
- **Radioactivity** is measured as the number of radioactive disintegrations per s (in **Becquerel** or Bq per l). *One disintegration per s* equals *one Bq*.
- **Total body water** is destributed between two compartments separated by the cell membrane: The intracellular and the extracellular fluid.

Essentials

This paragraph deals with *1. The three major fluid compartments*, *2. Water balance*, *3. Body potassium*, *4. Body sodium*, *5. The indicator dilution principle*, *6. The renin-angiotensin-aldosterone cascade*, *7. Output contol*, *8. Regulation of renal water excretion*, *and 9. Regulation of renal sodium excretion*.

Read first about the nephron (paragraph 1 of Chapter 25).

1. The three major fluid compartments

The three major body fluid compartments are the intracellular fluid volume (ICV), the interstitial fluid volume (ISV) and the vascular space (Chapter 1, Fig.1-4). *Water permeable* membranes separate the three compartments, so that they contain almost the same number of osmotically active particles per kg. The three compartments have the same concentration expressed as mOsmol per kg of water or the same freeze-point depression. They are said to be *isosmolal*, because they have the *same osmolality*.

The so-called *lean body mass*, which means a body stripped of fat, contains 0.69 parts of water (69%) of the total body weight in all persons. - Such high values are observed in the newborn and in extremely fit athletes with minimal body fat. Babies have a tenfold higher water turnover per kg of body weight than adults do.

As an average females have a low body water percentage compared to males. Such differences show *sex dependency*, but the important factor is the relative content of body fat, since fat tissue contains significantly less water (only 10%) than muscle and other tissues (70%). This is

why the relative water content depends upon the relative fat content.

The average for most healthy persons is 60% of the body weight. Sedentary, overweight persons contain only 50-55 % water dependent on the body fat content.

The relative content of *body fat* rises with increasing age and body weight, and the relative mass of muscle tissue becomes less. Consequently, the body water fraction falls with increasing body weight and age. Aging implies loss of cells, but the ECV is remarkably constant through life and under disease conditions.

Each body (weight 70 kg) contains 4 mol of both sodium and potassium (ie, the total ion pool). A minor fraction of the potassium is radioactive. The calcium and magnesium content is 25 and 1 mol, respectively.

In the renal tubule cells the epithelium is a single layer of cells, joined by junctional complexes near their luminal border (Fig. 25-7). Solutes can traverse the epithelium through transcellular or paracellular pathways. Virtually every cell membrane in the body contains the Na^+- K^+-pump, which maintains the *low* intracellular Na^+-concentration and develops the negative, intracellular voltage. In the renal tubule cells the Na^+- K^+-pump, is located in the basolateral membrane. Read more about the nephron in Chapter 25 and about hormonal control later in paragraph 8 and 11 of this Chapter.

Unfortunately, the simple laws of dilute solutions are unprecise at physiological concentrations. Rough estimates are based on the assumptions that extracellular sodium is associated with monovalent anions and that deviations in osmolality are twice the deviation in plasma sodium concentration.

ICV: The dominating intracellular solute is potassium (K^+), balanced by phosphate and anionic protein, whilst the dominating extracellular solute is NaCl. All compartments have almost the same osmolality 300 mOsmol* kg^{-1} of water. The thin cell membrane - or the endothelial barrier between ISF and plasma in the vascular phase - cannot carry any important hydrostatic gradient. Water passes freely between the extra- and intra-cellular compartment, as osmotic forces govern its distribution and the membranes are water permeable.

Fig. 24-1: The daily water transfer across the gastrointestinal barrier in a healthy standard person.

The ICV comprises 26-28 kg out of the total 42-kg of water in a 70-kg person (Fig. 1-4).

ECV: The ECV compartment comprises the remaining water (14-16 kg) with most of the water in *tissue fluid (interstitial fluid* or *ISF)* and 3 kg of water in plasma (Chapter 1, Fig. 1-4). The size of the ECV compartment is proportional to the total body Na^+. Changes in plasma osmolality indicate problems in water balance.

A $[Na^+]$ in ECV of 150 mmol per kg of plasma water corresponds to a total osmolality of 300 mOsmol per kg.

Alterations in plasma-$[Na^+]$ (osmolality) will be followed by similar changes of the ECV osmolality, because the permeability of of the capillary barrier for Na^+ and water is almost equal.

The daily water transfer across the gastrointestinal tract amounts to approximately 9 l in each direction (Fig. 24-1).

2. Water balance

A healthy person on a mixed diet in a temperate climate receives 1000 ml with the food and drinks 1200 ml daily. Balance is maintained as long as the water loss is the same (Fig. 24-2).

Fig. 24-2: The daily water balance in a 70-kg healthy person on a mixed diet. The apparent imbalance between input (2200 ml) and output (2500 ml) is covered by 300 ml of metabolic water.

Water is lost in the urine (1500 ml), in the stools (100 ml), in sweat and evaporation from the respiratory tract (900 ml) as a typical example.

The total loss of water is 2500 ml, and this corresponds perfectly to the intake plus a normal production of 300 ml of *metabolic water* per 24 hours (Fig. 24-2).

3. Body potassium

The daily dietary intake of potassium varies with the amount of fruit and vegetables consumed (75-150 mmol K^+daily).

More than 90% of the body potassium is located intracellularly. Only a few percent of the K^+ in the body pool are found outside the cells and subject to control (Fig. 24-3). The main renal K^+-reabsorption is passive and paracellular through *tight junctions* of the proximal tubules. Moreover K^+-excretion can vary over a wide range from almost complete reabsorption of filtered K^+ to urinary excretion rates in excess of filtered load (ie, net secretion of K^+).

The Na^+-K^+-pump located in the cell membrane, maintains the *high* intracellular [K^+] and the *low* intracellular [Na^+]. The energy of the terminal phosphate bond of ATP is used to actively extrude Na^+ and pump K^+ into the cell. The membrane also contains many K^+ - and Cl^- - channels, through which the two ions leak out of the cell.

In myocardial cells, as in skeletal muscle and nerve cells, K^+ plays a major role in determining the resting membrane potential (RMP), and K^+ is important for optimal operation of enzymatic processes.

Under normal conditions, the RMP of the myocardial cell is determined by the dynamic balance between the membrane conductance to K^+ and to Na^+. As $[K^+]_{out}$ is reduced during hypokalaemia, the membrane depolarises causing voltage-dependent inactivation of K^+-channels and activation of Na^+-channels, allowing Na^+ to make a proportionally larger contribution to the RMP.

Fig. 24-3: The total body K^+-pool in a healthy person comprises 4000 mmol with more than 90% intracellularly. The normal ECG and the ECG of a patient with hyperkalaemia is shown to the right.

The K^+ -permeability is around 50 times larger than the Na^+ -permeability, so the RMP of normal myocardial cells (typically: -90 mV) almost equals the equilibrium potential for K^+ (-94 mV).

The excretion of K^+ by overload is almost entirely determined by the *extent* of distal tubular secretion in the *principal cells*. Any rise in serum [K^+] immediately results in a marked rise in K^+-secretion. This transport mechanism is controlled by *aldosterone* and by K^+. Aldosterone stimulates the secretion of K^+ and H^+ by the principal cells of the renal distal tubules and collecting ducts (Fig. 25-11). This is why chronic acidosis decreases and chronic alkalosis increases K^+-secretion. – Actually, acute acidosis may reduce K^+-secretion.

Of the consumed K^+, 75-150 mmol is daily absorbed in the intestine. Since 90% is excreted

renally in a healthy person, there must be a minimum in a typical volume of 1500 ml of daily urine with a concentration of (75/1.5) = 50 mM. Normal urinary [K^+] is at least 30 mM. A high urinary [K^+] is indicative of a high *total body* K^+ or a high intake of K^+.

The normal *excretion fraction* (Chapter 25) for K^+ is 0.10 (10% or 90 mmol of the 900 mmol in the daily filtrate) corresponding to the daily intake (Fig. 24-4). A K^+ -poor diet leads to hypokalaemia with less than 20 mmol K^+ in the daily urine. A K^+ -rich diet triggers a large secretion and a high excretion in the urine (Box 25-1). A low urinary [K^+] is indicative of a low total body K^+ or of *extracellular acidosis* with transfer of K^+ from the cells in exchange of H^+. A low [K^+] in the distal tubule cells reduces the K^+ -excretion**.**

The *normal plasma-[K^+]* level is dependent upon the exchange with the cells, the renal excretion rate, and the extrarenal losses through the gastrointestinal tract or through sweat.

Measurement of total and exchangeable body potassium

Our natural body potassium is ^{39}K, but we also contain traces of naturally occurring radioactivity (0.00012 or 0.012% is ^{40}K with a half-life of 1.3×10^9 years). When using this natural tracer, injection of radioactive tracer is avoided.

The person to be examined is placed in a sensitive whole body counter, and the *total activity* of the tracer ^{40}K in the body (S Bq) is measured.

Specific activity (SA) is the concentration of radioactive tracer in a fluid volume divided by the concentration of naturally occurring, non-radioactive mother-substance. The concentration of mother-substance is traditionally measured in mmol per l (mM). SA is equal to *radioactivity* (A) per non-radioactive mass unit, m (ie, A/m in Bq/mol). Following even distribution, the SA for a certain substance must be the same all over the body. SA is preferably measured in plasma (with scintillation counters or similar equipment).

Specific activity (SA) is here the number of Bq ^{40}K per mol of mother substance (^{39}K) in the whole body. We can calcu
when SA is known to be 0.012% or a fraction of 0.00012. The *total body potassium* of a healthy person is 4000 mmol. The SA of ^{40}K implies a $^{40}K/^{39}K$ ratio of 0.48 mmol/4000 mmol (=0.00012).

An *exchangeable ion pool* in our body is the dynamic part of the total specific ion content. The remaining content is fixed as insoluble salts in the bones. The dynamic character implies the use of a dilution principle to measure such a pool.

In order to measure the *exchangeable* body potassium pool, a radioactive tracer is injected, such as ^{42}K with a physical half-life of 12 hours (12.4 hours) and urine is collected. The first urine sample is from the first 12 hours, and the second sample is covering 12 - 24 hours. The total tracer dose given must be adjusted for by the loss of tracer in the urine and by the radioactive decay during the first 12 hours mixing period. The two urine samples obtained are examined for tracer and for natural potassium. The tracer is assumed to distribute just as natural potassium after 12 - 24 hours. When the tracer is distributed evenly in the exchangeable body potassium, its SA must be the same in urine, plasma or elsewhere in the body. The exchangeable body potassium is calculated by Eq. 24-2 .

The specific activity for the tracer (SA Bq per mol) is known from the plasma measurements. In this way we measure the exchangeable body potassium. The normal values are *41 mmol* ^{39}K per kg body weight for females, and *46 mmol* per kg for males.

4. Body sodium (^{23}Na)

The *exchangeable* body sodium is easy to measure using the dilution principle and a minimum of equipment.

Our natural non-radioactive body sodium is ^{23}Na. We administer the radioactive tracer, ^{24}Na, with a physical half-life of 15 hours. We have to use a total period of 30 hours to secure even distribution in the ECV.

The total tracer dose given, must be adjusted for by the loss of tracer in the urine, and the radioactive decay of ^{24}Na (see the decay law in Chapter 1). The exchangeable body sodium is calculated by Eq. 24-2.

We know the specific activity for the tracer (SA Bq/mol) from the plasma measurements; therefore calculation of the exchangeable body ^{23}Na is easy.

The normal value for exchangeable body sodium is 40 mmol/kg of body weight. In a patient with a body weight of 75 kg the exchangeable sodium is $(75 \times 40) = 3000$ *mmol*. The non-exchangeable sodium is fixed in the bones.

The *total body sodium* is measured following discrete radiation with a method called *neutron activation analysis*. The whole body of the patient is exposed to radiation with neutrons. A small fraction of the natural ^{23}Na now becomes radioactive sodium (^{24}Na) by uptake of an extra neutron.

A sensitive *whole body counter* records the radiation from ^{24}Na. Now we can calculate the *total body sodium.*

Normally, the total body sodium is 1000 mmol larger than the *exchangeable* sodium due to the fixed sodium content of the bones (1000 + 3000 mmol = 4000 mmol ^{23}Na).

Fig. 24-4: Body fluid electrolytes. Water permeable membranes separate the three compartments, which contain almost the same number of osmotically active particles per kg.

The sum columns of electrolyte equivalents in muscle cells are essentially higher than the extracellular sum columns of equivalents, because cells contain proteins, Ca^{2+}, Mg^{2+} and other molecules with several charges per particle (Fig. 24-4).

The above columns show the ionic composition per kg of water, so we have 150 mmol of Na per kg of plasma water. Normally, one litre of plasma has a weight of 1.040 kg and contains 10% of dry material. Consequently, one litre of plasma contains 0.940 l of water, and the rest consists of plasma proteins and small ions. Thus the fraction of water in plasma (F_{water}) is typically 0.94.

5. The indicator dilution principle

Mass conservation is always the underlying principle. The amount of indicator n mol distributes in V litres of distribution volume.

We measure the concentration C_p in mM, following even distribution, and calculate V:

$$V = n/C_p.$$

Errors: Uneven distribution of indicator introduces a systematic error. - A non-representative concentration of indicator in the plasma makes it insufficient to correct for plasma proteins alone. - Loss of indicator to other compartments is inevitable. - Elimination or synthesis of indicator in the body occurs as frequent errors. - The indicator may be toxic or in other ways change the size of the compartment to be measured.

Total body water, ECV, plasma volume, and the elimination rate constant are measured as follows:

5 a. Total body water

Total water is measured by the help of the dilution principle. *Tritium marked water* is a good tracer. The equilibrium period is 3-6 hours. n mol of indicator divided by C_p mmol of indicator per l is equal to the distribution volume (V) for the indicator.

Healthy adolescents and children have normal values around 60% of the body weight assuming one l of water to be equal to one kg. Adult males and females with a sedentary life style and larger fat fractions contain only 50% of water.

5 b. The extracellular fluid volume (ECV)

is measured by administration of a *priming dose* of inulin intravenously. Then inulin is infused to maintain a steady state with constancy of the plasma concentration of inulin (C_p).

The patient then urinates, and the infusion is stopped with collection of a plasma sample. For the next 10 hours the patient collects his urine, which makes it possible to measure all the body inulin present at the end of the infusion (n mol) assuming all inulin excreted.

Dividing n with C_p gives the volume of distribution (V) after correcting for the difference in protein concentration between plasma and ISF (Eq. 24-1).

Chromium-ethylene-diamine-tetra-acetate (^{51}Cr-EDTA) is a chelate with a structure that cannot enter into cells. The chelate molecule contains radioactive Cr, making it easy to measure. The ^{51}Cr-EDTA distributes and eliminates itself in the extracellular fluid volume (ECV) just as inulin and is therefore used to measure ECV. – For clearance measurements, we inject a single dose intravenously, and draw blood samples every hour for 5 hours. The clearance of ^{51}Cr-EDTA is independent of C_p and a good estimate of GFR just like the *inulin clearance*. Since the indicator is cleared from the ECV only, it is possible to measure its size. Such methods - including renal lithium reabsorption - are important during renal function studies. Normal values for ECV are approximately 20% of the body weight or 14-17 kg.

Chronically ill patients with debilitating diseases often maintain their ECV remarkably well in spite of marked reductions in the cell mass of their body.

5 c. The plasma volume

Also here, the dilution principle is used. The indicator for plasma volume can be Evans Blue (T_{1824}) that binds to circulating plasma albumin. A small dose of albumin, marked with radioactive iodine, is also a good indicator (iodine ^{131} has a physical half-life of 8 days).

The indicator concentration in plasma (C_p) is measured every 10-min for an hour after the administration, and the log of C_p is plotted with time. Extrapolation to the time zero determines the maximum concentration of indicator in plasma. This corrects for the biological loss, while the indicator distributes itself in the plasma phase. The tracer dose divided by C_p at time zero provides us with the *intravascular* plasma volume. Normal values for the plasma volume are close to 5% of the body weight.

In *diabetics* and *hypertensive* patients the tracer is lost more readily through their leaky capillaries to the interstitial fluid than in healthy persons (increased transcapillary escape).

6. The renin-angiotensin-aldosterone cascade

Macula densa is described in paragraph 9 of Chapter 25.

The most likely intrarenal trigger of the renin-angiotensin-aldosterone cascade is the **falling** NaCl concentration of the reduced fluid flow at the macula densa in the distal renal tubules (Fig. 24-5).

The NaCl concentration at the macula densa falls, when we lose extracellular fluid, move into the upright position and when the blood pressure falls.

Renin is a proteinase that separates the decapeptide, angiotensin I, from the liver globulin, angiotensinogen.

When angiotensin I passes the lungs or the kidneys, a dipeptide is separated from the decapeptide by angiotensin converting enzyme (ACE). This process produces the octapeptide, angiotensin II.

Angiotensin II has multiple actions that minimize renal fluid and sodium losses and maintain arterial blood pressure.

1. Angiotensin II stimulates the aldosterone secretion by the adrenal cortex, and through this hormone it stimulates Na^+-reabsorption and K^+-(H^+)-secretion in the distal tubules (Fig. 24-5). - Angiotensin II is in itself a potent stimulator of tubular Na^+-reabsorption.
2. Angiotensin II inhibits further renin release by negative feedback.
3. Angiotensin II constricts arterioles all over the body including a strong constriction of the efferent and to some extent also the afferent arteriole. Hereby, the renal bloodflow (RBF) and to a lesser extent the glomerular filtration rate (GFR) is reduced.
4. Angiotensin II inhibits the absolute proximal tubular reabsorption – contributing to the reduction of GFR.
5. Angiotensin II enhances sympathetic nervous activity.

Fig. 24-5: The renin-angiotensin-aldosterone cascade.

Sympathetic stimulation of the renal nerves stimulates renin secretion directly via b-adrenergic receptors on the JG cells just as falling blood pressure in the preglomerular arterioles. - b-blocking drugs and angiotensin II inhibit the renin secretion (Fig 24-5).

The combined effects from the whole renin cascade is extracellular fluid homeostasis.

In contrast, exposure to stress and painful stimuli triggers the combined sympatho-adrenergic system with release of catecholamines, gluco- and mineralo-corticoids, and ACTH from the hypophysis. ACTH stimulates further the secretion of the glucocorticoid, cortisol, from the adrenal cortex.

7. Output control

The body uses *output control*, when it is overloaded with water or with sodium.

The most important osmotically active solute in ECV is NaCl, because it only passes into cells in small amounts. Urea, glucose and other molecules with modest concentration gradients are without importance, because they distribute almost evenly in the fluid compartments.

Healthy persons use two primary control systems: 1) The osmolality (osmol per kg of water) or ion concentration controls our elimination of water. 2) The change of blood volume (ECV) or pressure controls sodium excretion - not osmolality.

Only when the arterial blood pressure falls *drastically* the body will drop its protection of normal concentration. In such a disease state large amounts of ADH molecules are released in an attempt to improve the volume and blood pressure.

8. Regulation of renal water excretion

The primary control of the renal water excretion is *osmolality control* (Fig. 24-6). Since 2/3 of the body water normally is located within the cells, this is also an intracellular volume control.

Following *water deprivation* even an increase in plasma osmolality of only one per cent stimulates both the hypothalamic osmoreceptors and similar (angiotensin-II-sensitive) thirst receptors. Thirst may increase the water intake of the individual and thus increase the ECV, with negative feedback to the thirst receptors.

Activation of the hypothalamic osmoreceptors and thirst receptors increases the hypothalamic neurosecretion to the neurohypophysis and releases antidiuretic hormone (ADH or vasopressin). Hyperosmolality elicits a linear increase in plasma ADH, which causes water retention (Fig. 24-6) until isosmolality is reached.

ADH increases the reabsorption of water from the fluid in the renal cortical and medullary collecting ducts. ADH binds to receptors on the basolateral surface of the tubule cells, where they liberate and accumulate cAMP. This messenger passes through intermediary steps across the cell to the luminal membrane, where the number of water channels (aquaporin 2) are increased. The luminal cell membrane is thus rendered water-permeable, which increases the renal water retention. The increased water reabsorption leads to a small, concentrated urine volume (antidiuresis), and a net gain of water that returns ECF osmolality towards normal. Initially, osmolality control overrides blood volume control.

Fig. 24-6: Primary osmolality control of the renal water excretion. ADH and thirst systems maintain osmolality and ICV within narrow limits.

Water overload decreases ECF osmolality and has the reverse effect, because the hypothalamic osmoreceptors suppress the ADH release, and the renal water excretion is increased already after 30 min (Fig. 24-6). When a person rapidly drinks one litre of water, the intestine absorbs water. Ions diffuse into the intestinal lumen and the blood osmolality falls causing a block of the ADH secretion (Fig. 24-6).

Pure water is distributed evenly in all three body fluid compartments – just like intravenous infusion of one litre of 5% glucose in water.

Intake of one l of isotonic saline implies ECV expansion, without dilution of body fluids. This expansion will not increase the urine volume much, so the increased ECV can be sustained for many hours. An intravenous infusion of one l of large dextran molecules (macrodex) stays mainly in the vascular space.

9. Regulation of renal sodium excretion

In healthy persons, changes of *blood volume* (or ECV) or *blood pressure* control sodium excretion (Fig. 24-7). The dominating cation of the ECV is Na^+ . The sodium intake is balanced by the sodium excretion as long as the thirst and other homeostatic systems are functional.

During conditions where sodium intake exceeds renal sodium excretion, total body sodium and ECV increase. Conversely, total body sodium and ECV decrease, when sodium intake is lower than renal sodium excretion. This is because volume-pressor-receptors detect the size of the circulating blood volume (ECV) or pressure, and effector mechanisms adjust the renal sodium excretion accordingly.

The volume-pressor-receptors are widely distributed. *Low-pressure receptors* are found in the

pulmonary vessels and in the atria. An increased blood volume can also increase the arterial blood pressure and stimulate the well-known *high-pressure baroreceptors* in the carotid sinus and the aortic arch. Increased arterial pressure reduces sympathetic tone – also in the kidneys, whereas decreasing arterial pressure enhances sympathetic tone and renal salt retention. Arterial pressure receptors are also located in the renal *preglomerular* arterioles. Both stimuli in Fig. 24-7 release renin from macula densa, whereby angiotensin II and aldosterone is secreted (both sodium retaining hormones).

A decrease in circulating blood volume leads to a decrease in NaCl delivery to the macula densa and release of the renin cascade. Conversely, an increase in circulating blood volume with increased NaCl delivery to the macula densa suppresses renin release and increases sodium excretion (Fig. 24-7).

Fig. 24-7: Primary blood volume-pressure control of the renal Na^+ -excretion. The effective circulating blood volume is protected – also during shock (Na^+ -retention) and during hypertension (natriuresis).

Increased salt intake increases blood volume and leads to natriuresis, possibly augmented by release of ANP (see below), nitric oxide and other factors. The excretion of Na^+ depends upon several effector mechanisms out of which three are classical:

The *first* factor is the *glomerular filtration rate* (GFR), which is responsible for the size of the filtered flux of Na^+ across the glomerular barrier in the kidneys. Renal prostaglandins, generated in response to angiotensin II, are involved in maintaining the filtered flux of Na^+.

The *second* factor is the *renin-angiotensin-aldosterone cascade* (Fig. 24-5).

The *third* factor consists of *peptides* with natriuretic effects. The most well-known peptide is called *atrial natriuretic peptide* (ANP) and originates from granules of the atrial myocytes. A low circulating blood volume with low atrial pressure increases renal sympathetic tone, reduces the stimulus of the low-pressure receptors in the atrial wall and thus the ANP secretion. Hereby, the natriuresis is reduced. - Renal natriuretic peptide or urodilatin from the distal tubule cells is related to ANP. Urodilatin has been isolated from human urine and contains four amino acids more than ANP.

An increase in effective circulating blood volume, increases atrial pressure, reduces sympathetic tone and releases ANP and urodilatin leading to increased natriuresis.

The main purpose of these mechanisms is to maintain an effective circulating blood volume by an increase or a decrease of the renal excretion of Na^+. Initially, osmolality control is dominating. Finally, after a dangerous reduction in blood volume, volume-pressure receptors override the hypothalamic osmoreceptors and stimulate the ADH release and thirst. In the terminal phase, the body protects effective circulating blood volume at the expense of ECF osmolality.

Pathophysiology

This paragraph deals with 1. Dehydration, 2. Overhydration, 3. Hyponatraemia, 4. Hypernatraemia, 5. Hypokalaemia, and 6. Hyperkalaemia.

1. Dehydration

Dehydration is an *abnormal reduction of the major fluid volumes (total body water* with shrinkage of ECV). When we lose more than 5% of the total body water it has clinical

consequences. The condition is life threatening if the patient loses 20 %.

Accidents and surgery with a period of water deprivation, imply a rise in ECF osmolality and thus stimulation of both thirst and the hypothalamic osmoreceptors, whereby ADH is released. - Symptoms and signs of dehydration are *thirst*, dry *mucous membranes*, and decreased skin elasticity or *turgor* due to loss of ISF.

Loss of effective circulating blood volume implies a low blood pressure in both the venous and the arterial system. Loss of more than one litre of ECV causes *postural hypotension* with dizziness, confusion and cerebral failure. Empty veins and cold skin characterise the peripheral venoconstriction. Finally, there is extreme tachycardia, which turns into terminal bradycardia and an arterial blood pressure that approach zero.

Loss of salt and water frequently develops into *hypo-osmolal dehydration* (Fig. 24-8). This is because the thirst forces the patient to drink (salt free) water. Water dragged into the cells further reduces the hyposmolal ECV (Fig. 24-8). The small ECV elicits a hyperaldosteronism, which is called *secondary*, because it is not initiated as primary hypercorticism in the adrenal cortex. A precise compensation of the water loss results in *pure hyponatraemia*, where water eventually is drawn from ECV into the cells. The low [Na^+] around the swelling cells reduces the potential gradient across the cell membranes with increased neuromuscular irritability (muscular twitching) and cardiac arrhythmias.

Isosmolal dehydration is a proportional loss of water and solutes. There is no concentration gradient over the cell membranes, and the loss is mainly from ECV (Fig. 24-8).

Fig. 24-8: Dehydration (hyperosmolal, isosmolal and hyposmolal).

Hyperosmolal dehydration occurs in persons deprived of water. The hypero
water from ICV and dehydrates the cells (Fig. 24-8). This is intracellular dehydration.

The hyperosmolality liberates ADH to restrict the water loss. The patient excretes a very small urine volume.

Persons deprived of water at sea may drink seawater. Sea water is hypertonic saline and the victims die faster. When hypertonic saline reaches the ECV it aggravates the intracellular dehydration simultaneously with an extracellular overhydration. Intracellular dehydration leads to respiratory arrest and death of thirst.

2. Overhydration

Overhydration is an *abnormal increase of total body water* - in particular ECV, and thus salt accumulation. The increase in the interstitial fluid volume is called oedema. Overhydration frequently occurs among patients in fluid therapy (ie, overhydration of iatrogenous origin).

Increased salt intake by mouth is compensated by increased salt excretion by normal kidneys.

However, a large saline infusion (0.9% NaCl) will expand ECV and total body water (isosmolal overhydration in Fig. 24-9). Inappropriately large infusions of saline lead to iatrogenous hyperosmolal overhydration, if they lose more water than salt (Fig. 24-9).

Hyperosmolality drags water from the cells, so that the patient develops *intracel dehydration* with hallucinations, loss of consciousness and eventually respiratory arrest.

The patient with *hyposmolal overhydration* is typically in fluid treatment and develops muscle cramps and disorientation. The skin turgor is normal. A low serum - [Na^+] confirms the diagnosis. The water overload in ECV is dragged into the cells in hyposmolal overhydration

until osmolality balance (Fig. 24-9).

In the brain and the muscles this *intracellular overhydration* causes headache, disorientation, increased spinal pressure, coma and muscle cramps. Both *hyposmolal* and hyperosmolal *intracellular overhydration conditions* are characterised by cerebral symptoms and signs.

Fig. 24-9: Overhydration (hyperosmolal, isosmolal, and hyposmolal).

Acute renal failure with decreased GFR reduces the flux of filtered NaCl (*first factor*) and thus the Na^+ -excretion.

Oedema is a clinical condition where the interstitial fluid volume (ISF) is abnormally large.

A *voluminous ISF* is usually due to *increased* hydrostatic venous pressure (heart insufficiency), or a *reduced* colloid osmotic pressure (hypoproteinaemia) as predicted from Starlings law for transcapillary transport.

Reduced protein synthesis (liver disease) and *abnor*
(proteinuria) causes hypoproteinaemia. Thus protein-losing kidneys are involved.

Capillary damage (allergy, burns, inflammation etc) with increased capillary permeability *causes* local oedema. – Obstruction to lymphatic drainage can also cause oedema (scarring after radiation therapy, elephantiasis etc).

Cardiac insufficiency with increased venous pressure and oedema formation increases sympathetic tone and thus releases the *renin-angiotensin-aldosterone cascade* (Fig. 24-5) causing Na^+ -retention.

Hepatic cirrhosis activates the cascade in a similar way - possibly including the release of nitric *oxide*.

Hypoalbuminaemia reduces the colloid osmotic pressure of plasma, whereby water is distributed from the vascular space to the ISF. The fall in effective circulatory volume activates the renin cascade and leads to Na^+ -retention.

NSAIDs can activate the renin-angiotensin-aldosterone cascade, and the increased aldosterone leads to Na^+ -retention and overhydration.

Angiotensin II-receptor antagonists and *ACE-inhibitors* are utilized clinically to block the effects of angiotensin II in congestive heart failure, diabetes mellitus and hypertension. Blockade of the cascade reduces both preglomerular and postglomerular resistances.

The *supine position* at bed rest increases venous return. This implies an increased cardiac output (Starlings law), a reduced ANF secretion from the atrial walls and a reduced renin-angiotensin-aldosterone cascade. This is why bed rest is beneficial for disorders with salt accumulation.

3. Hyponatraemia

Hyponatraemia (ie, plasma-[Na^+] below 135 mM) is associated with dehydration, overhydration or normohydration (ie, a normal ECV and total body sodium content).

Hyponatraemia with *reduced* ECV (ie, salt-deficient hyponatraemia) is caused by a salt loss in excess of the high water loss (ie, hyposmolal dehydration in Fig 24-10). This is seen in any type of hypoadrenalism including the rare primary hypoadrenalism (Addison's disease).

In Addison's disease the entire adrenal cortex is destroyed by autoimmune reactions (80%) or

by malignancy or infection. All three types of hormones are insufficiently produced (mineralocorticoids, glucocorticoids and sex hormones). The lack of aldosterone leads to Na^+ -excretion and K^+-retention with hyponatraemia combined with hyperkalaemia resulting in dehydration and hypotension.

Hyponatraemia is developed in the following way (Fig. 24-10):

1. The first step is the salt loss in excess of the water loss.

2. Since the ECF-[Na^+] is low, the ADH secretion is suppressed, and the water excretion is increased. Hereby, both the ISF and the vascular spaces are reduced often by more than 10%.

3. This is an adequate stimulus for the *volume-pressure receptors,* which override the osmoreceptors, whenever the effective circulatory volume is threatened.

Fig. 24-10: The three body fluid compartments in a patient with salt-deficient hyponatraemia.

The volume-pressure receptors stimulate both thirst and the release of ADH. The effective circulating volume is protected at the expense of osmolality! Still the blood pressure is falling, which impairs cerebral perfusion, causing confusion, headache and coma.

The hyponatraemia implies a reduced resting membrane potential and thus a low threshold for neuromuscular stimulation resulting in muscle cramps.

The large renal loss is seen with osmotic diuresis (hyperglycaemia and uraemia), excessive use of diuretics, renal tubular reabsorption defects, adreno-cortical insufficiency as aldosterone-antagonist-intoxication or other types of hypoaldosteronism.

The extra-renal loss is often large from excessive sweating, diarrhoea, haemorrhage, vomiting, loss with ascites or bronchial secretion, and transudation from cutaneous defects. Normal kidneys normally compensate extra-renal loss. The urinary excretion of salt and water falls in response to volume depletion, so the urine is concentrated - but with less than 10 mM Na^+.

Normal sweat is a *hypotonic* solution, because Na^+ is reabsorbed in the duct system. The [Na^+] can increase up to 80 mM with increasing sweat flow - due to the limited time for the aldosterone-controlled Na^+-reabsorption.

Increased salt intake by mouth or intravenously is required as a supplement to the treatment directed at the primary cause.

Low plasma- [Na^+] in a chronically salt-deficient patient suggests a high aldosterone secretion from the adrenal zona glomerulosa. Further administration of aldosterone therefore may not have any effect.

Hyponatraemia with *increased* ECV (water-excess hyponatraemia) is often caused by cardiac, hepatic, and renal insufficiency or by hypoalbuminaemia - see hyposmolal overhydration (Fig. 24-9).

Hyponatraemia with *normal* ECV is often caused by stress (surgery, psychogenic polydipsia), abnormally high ADH release (in the syndrome of inappropriate antidiuretic hormone secretion, and in vagal neuropathy), increased sensitivity to ADH by drugs such as chlorpropamide and tolbutamide, or by intake of ADH-like substances (oxytocin).

Pseudo-hyponatraemia is characterised by a spuriously low plasma value measured

conventionally in the total volume of plasma, which includes an extra volume in cases with hyperlipidaemia or hyperproteinaemia etc. Plasma osmolality or plasma-Na^+ measured with ion selective electrodes is the choise and the direct read value is normal. This is because Na^+ is confined to the aqueous phase.

Treatment of *artefactual hyponatraemia* (taking blood from an extremity into which isotonic glucose is infused) is also unnecessary.

4. Hypernatraemia

The *normal* plasma-[Na^+] is *135-145 mM*, and values above 170 mM are rare. Excessive infusion of saline (0.9% NaCl or 154 mM) can lead to hypernatraemia. Such alarmingly high levels create an emergency situation, where glucose infusion is indicated initially in order to reduce the high level slowly. The increased plasma osmolality elicits a strong desire to drink.

The cause is sometimes *water deficit* due to pituitary diabetes insipidus, or to nephrogenic diabetes insipidus, where ingestion of nephrotoxic drugs have made the renal collecting ducts resistant to ADH. – Osmotic diuresis also causes water deficit with hypernatraemia just as excessive loss of water through the skin or lungs.

Primary hyperaldosteronism (Conns disease) and all types of secondary hyperaldosteronism also lead to hypernatraemia combined with hypokalaemia and enlarged blood volume.

Cerebral failure and convulsions are alarming signs, but there are no specific symptoms and signs of hypernatraemia.

Polyuria, polydipsia and thirst suggest diabetes. Diabetes mellitus is easy to diagnose, and diabetes insipidus shows a low urinary osmolality. Pituitary diabetes insipidus is treated with an analogue of ADH (desmopressin, with a low pressor-effect).

5. Hypokalaemia

The normal potassium ion concentration in blood plasma is 3.5-5 mM. Hypokalaemia is caused by renal or extra-renal K^+ -loss or by restricted intake.

Long standing use of diuretics without KCl compensation is a frequent cause of hypokalaemia.

Hyperaldosteronism (increased aldosterone secretion) is another cause.

Vomit fluid only contains 5-10 mM of K^+. Still, prolonged vomiting develops into hypokalaemia, because the Na^+-loss stimulates the aldosterone secretion, which increases $\mathbf{K^+}$-excretion in the kidneys.

Profuse diarrhoea causes marked hypokalaemia, also because the diarrhoea fluid contains up to 50 mM of $\mathbf{K^+}$.

Hypokalaemia is seen in cardiac patients receiving *digoxin* treatment. Digoxin toxicity is imminent, because digoxin firmly binds to myocardial cells in hypokalaemia. Treatment must be directed towards the underlying cause. Infusion of potassium -rich fluid is dangerous, because of the marginal distance to hyperkalaemia.

The reduced extracellular K^+ hyperpolarises the cell membrane (increases the negativity of the voltage across the membrane). This reduces the excitability of neurons and muscle cells. Thus, hypokalaemia can result in muscle weakness and paresis. Hypokalaemia is associated with an increased frequency of cardiac arrhythmias with atrial and ventricular ectopic beats in

particular in patients with cardiac disease . - Hypokalaemia inhibits release of adrenaline, aldosterone and insulin.

6. Hyperkalaemia

Acute hyperkalaemia (ie, plasma-[K^+] above 5 mM) is a normal condition following severe exercise, and normal kidneys easily eliminate K^+ .

In disease states the causes are *insufficient* renal excretion or *increased* release from damaged body cells as during long lasting hunger, exercise or in severe burns. A plasma- [K^+] above 7 mM is life threatening due to asystolic cardiac arrest.

Long term intake of b-blocking drugs, which inhibit the Na^+-K^+-pump, leads to hyperkalaemia that is accentuated by exercise.

Hyperkalaemia reduces the size of the resting membrane potential (reduces the negativity of the voltage), whereby the threshold for firing is approached in neurons and striated muscle cells. The increased excitability in hyperkalaemia results in muscle contractions, cramps followed by muscle weakness. Hyperkalaemia leads to decreased cardiac excitability, hypotension, bradycardia and eventual asystole. The ECG is characterised by increased duration of the QRS-complexes and *tented* T-waves due to abrupt Ca^{2+}-influx, contraction, and abrupt Ca^{2+} -binding (Fig. 24-3). Cardiac arrest occurs as ventricular fibrillation (the heart can never produce smooth tetanus) or as asystole.

Insulin is used to drive K^+ back into the cells - either by insulin infusion or by glucose infusion in order to release more insulin from the pancreatic islets. Usually, a combined glucose-insulin drop is applied.

Other hormones (adrenaline, aldosterone) also stimulate the Na^+-K^+-pump and thereby increase cellular K^+ -influx (Fig. 24-3)

Equations

- *The indicator dilution method*: The indicator n mmol distributes in V litres of distribu volume. We measure the concentration C_p in mM, following even distribution, and calculate the volume, V:

  Eq. 24-1: $V = n/C_p$. (litre = mmol/(mmol/l)

- When the tracer is radioactive potassium and thus distributed evenly in the *exchangeable potassium pool*, its specific Activity (SA) must be the same in urine, plasma or elsewhere in the pool.

  Eq. 24-2: Exchangeable body potassium =

  (Injected - eliminated)/SA. (Mol = Bq/(Bq per mol)

We know the specific activity for the tracer (SA Bq per mol) from the plasma measurements. In this way we measure the exchangeable body potassium. The normal values are *41 mmol* ^{39}K per kg body weight for females, and *46 mmol* per kg for males.

- The following *concentrations* are found in normal plasma:

[Na^+] 135-145, [K^+] 3.5-5, [Cl^-] 96-106, [bicarbonate] 24, and total-[Ca^{2+}] 2.5 mM.

- The concentration of low molecular ions in the *ultrafiltrate* is affected by the Donnan

effect (normally 5% for monovalent ions), and by the fractional content of water in plasma (0.94 normally):

Eq. 24-3: [Low molecular ions] = Plasma conc. * Donnan factor/ 0.94.

$[Na^+]$ = 141× 0.95/0.94 = 143 mmol/l of ultrafiltrate. Based on the Donnan effect alone, this result is less than 141. The Donnan effect on monovalent cations is simply more than compensated by the protein volume effect or fractional content of water in plasma (0.94).

$[Cl^-]$ = 103×1.05/0.94 = 115 mmol/l of ultrafiltrate. Based on the Donnan effect this result should be greater than 103 and the protein volume effect contribute further. Such an ultrafiltrate is present in the kidneys and in ISF.

- The *extracellular fluid volume* (ECV) can be measured if all inulin molecules are collected in the urine over 10-15 hours after the inulin infusion stopped.

  Eq. 24-4: ECV = Amount of inulin excreted/(C_p/0.94).

  The inulin distribution volume is more or less identical to the ECV.

- Concentration of molecules in the filtrate are calculated as follows:

  Eq. 24-5: $C_{filtr} = C_p \times F_{free}/0.94$ (mmol per l of ultrafiltrate).

This value depends upon the fractional content of water in plasma (F_{water} = 0.94 l of water per l of plasma) and of the fraction of free, unbound molecules (F_{free}). For uncharged, free molecules like inulin F_{free} is 1, and for protein-bound molecules F_{free} is lower than 1

Self-Assessment

Multiple Choice Questions:

Each of the following statements has True/False options:

A. Hyponatraemia with normal ECV is often caused by stress, abnormally high ADH release, increased sensitivity to ADH by drugs, or by intake of ADH-like substances.

B. The total water content of a healthy person is 60%, and an extremely obese adult contains relatively more water.

C. Hyponatraemia is defined as a plasma-$[Na^+]$ below 145 mM.

D. A plasma- $[K^+]$ above 4.5 mM is life-threatening.

E. An infusion of one l of 5% glucose is distributed evenly into all three compartments just as pure water. An infusion of one l of saline remains mainly in the ECV, whereas an infusion of one l of macrodex stays mainly in the vascular space.

Case History A

A healthy male with a body weight of 70 kg has a normal extracellular osmolality (300 mOsmol kg^{-1}), and a normal ICV/ECV of 28/14 kg or l of water.

One day he is the victim of severe burns and he suffers a water loss of 2.5 l of water (the salt loss is covered).

1. Calculate the new ECV osmolality following the water loss.

2. Does this hyperosmolality have consequences?

3. Following total restitution of the water compartments the patient undergoes surgery with skin grafts. During the long procedure he receives sufficient water by glucose infusion, but he looses 900 mOsmol NaCl. Calculate the new osmolality.

4. Is it dangerous for a healthy individual to lose 6 kg of water without solutes?

5. Is it dangerous for a healthy individual to lose 6 kg of water as an isosmolal fluid from the ECV?

Case History B

A female patient (age 22 years; weight 71 kg) is in hospital suspect of potassium imbalance. She has taken diuretics for 2 years. She is tired and sleepy; her legs are paretic. The ECG shows prolongation of the Q-T interval, depression of the S-T segments and flattening of the T-waves. Her blood pH is 7.57 and the serum K^+ -concentration is 2.9 mM. One morning she receives an intravenous injection of a solution containing the radioactive isotope of potassium (555,000 Becquerel, Bq, of $^{42}K^+$ with a physical half-life of 12 hours). Following the injection her urine is collected in two periods (0-12 and 12 -24 hours). The first urine collection contained 40 mmol K^+ ($^{39}K^+$) and 4144 Bq $^{42}K^+$. The second urine specimen contained 40 mmol K^+ and 2220 Bq $^{42}K^+$. Both urine specimens were analysed for radioactivity exactly 24 hours after the injection, where the specific activity of her plasma was 55.5 Bq/mmol. The $^{42}K^+$, retained after the first 12 hours, distribute in her body just like all other exchangeable K^+. The body contains traces (0.012% of the total) of naturally occurring radioactivity (^{40}K) with a half-life of 1.3×10^9 years.

1. Calculate the exchangeable K^+ pool of her body after the 12-hour distribution period. - Is the result normal?

2. Calculate the elimination rate constant (k) for exchangeable K^+ in her body, and the biological half-life for this K^+ in hours. Calculate the ratio between the physical and the biological half-life of K^+.

3. What is the cause of her disease?

4. Describe the actions of diuretics.

5. Describe a method for measurement of her total body potassium.

Case History C

This case requires knowledge of the renal function (Chapter 25).

Two groups of substances are evenly distributed in the ECV of a healthy 25-year-old man. His weight is 70 kg, and his extracellular volume (ECV) is 14 L. Both groups of substances disappear solely by excretion through the kidneys. His GFR is 120 ml/min, and his renal plasma flow (RPF) is 700 ml/min.

1. Inulin is representative for one family of substances. Inulin is only ultrafiltered in the kidneys. What fraction (k_1) of the total amount of inulin in the body is maximally excreted in the urine per min?

2. The other substances are not only ultrafiltered, but they are also undergoing tubular secretion to such an extent that they totally disappear from the blood during the first passage. What is the elimination rate constant (k_2) for these substances?

Try to solve the problems before looking up the answers.

Highlights

- *Water permeable membranes separate the three body fluid compartments, so that they contain almost the same number of osmotically active particles (expressed as mOsmol per kg of water or the same freeze-point depression). The three compartments are the intracellular fluid volume (ICV), the interstitial fluid volume (ISV) and the vascular space.*
- *The sum columns of electrolyte equivalents in muscle cells are essentially higher than the extracellular sum columns, because cells contain proteins, Ca^{2+}, Mg^{2+} and other molecules with several charges per particle.*
- *Females contain less water as an average compared to males. Such differences show sex dependency, but the important factor is the fraction of body fat, since fat tissue contains significantly less water than other tissues (only 10%). Sedentary, overweight persons contain 50-55 % water dependent on the body fat content, and regardless of sex.*
- *Primary hyperaldosteronism (Conns hypercorticism disease) and all types of secondary hyperaldosteronism also lead to hypernatraemia combined with hypokalaemia and enlarged blood volume. Cerebral failure and convulsions are alarming signs, but there are no specific symptoms and signs of hypernatraemia.*
- *Polyuria, polydipsia and thirst suggest diabetes insipidus and low urinary osmolality is a clear indication. Pituitary diabetes insipidus is treated with an analogue of ADH (desmopressin, with a low pressor-effect).*
- *Regulation of K^+-balance: The daily intake of K^+ is matched by the renal K^+-excretion and our daily urine contains 2-5 g of K^+.*
- *Acid-base balance. The pH of the ICV and the ECV is maintained within narrow limits (many metabolic processes are sensitive to pH). The acid-base balance is accomplished by co-operative action of the kidneys and the lungs.*
- *Hypokalaemia reduces the excitability of neurons, muscle cells and the myocardial syncytium. Thus, hypokalaemia can result in muscle weakness, paresis, and cardiac arrhythmias with ectopic beats and cardiac arrest in diastole.*
- *Hyperkalaemia increases the excitability of neurons, muscle cells and the myocardium.*
- *Acute hyperkalaemia is a normal condition following severe exercise, and normal kidneys easily eliminate this. In disease states the causes of hyperkalaemia are insufficient renal excretion or increased release from the body cells as during long lasting hunger.*
- *A plasma- $[K^+]$ above 7 mM is life threatening due to ventricular fibrillation or cardiac arrest in systole. Tented T-waves and increased QRS-complexes characterise the ECG.*

Further Reading

Astrup, P, P. Bie, and H.C. Engell**.** 'Salt and water in culture and medicine.' *Munksgaard*, Copenhagen, 1993**.**

Knox, F. G. "Physiology of potassium balance." *Am.J. Physiol*. 275 (*Adv. Physiol. Educ*. 20): S142-S147, 1998.

Return to chapter 24

Return to Content

Chapter 25.

Renal Physiology And Disease

Study Objectives

- To *define* the concepts: Nephron, glomerular filtration, tubular secretion and reabsorption, renal lobulus, renal plasma clearance, osmolar clearance, tubular passage fraction, reabsorption fraction, excretion fraction, filtration fraction, plasma extraction fraction, proximal and distal system, glomerular propulsion pressure, net filtration pressure, renal threshold, and the maximal transfer (T_{max}) for tubular secretion and reabsorption.

- To *describe* the renal circulation and measurement of renal bloodflow, a superficial and a juxtamedullary nephron, the juxtaglomerular apparatus, and the concentrating mechanism of the kidney.

- To *calculate* the relation between half-life, elimination rate constant, clearance and distribution volume of a substance treated in the kidneys.

- To *explain* the normal renal function including the control functions, use of endogenous creatinine clearance as a renal test, the renal treatment of the filtration- reabsorption- and secretion-families of substances, the glomerular filtration rate (GFR), the angiotensin-renin-aldosterone cascade, the tubulo-glomerular feedback, the proximal and distal transport processes, and micturition. To explain the pathophysiology of common renal disorders including renal oedema.

- To *use* the above concepts in problem solving and in case histories.

Principles

- *The glomerulus and the proximal tubule are responsible for filtration of plasma and for major reabsorption of water and solutes. Glomerular filtration is due to a hydrostatic/colloid osmotic pressure gradient.*

- *Tubular reabsorption is the movement of water and solute from the tubular lumen to the tubule cells and often further on to the peritubular capillary network.*

- *Tubular secretion represents the net addition of solute to the tubular fluid in the lumen.*

- *All substances treated by the kidneys can be divided into three groups or families, namely the filtration group, the reabsorption group and the secretion group.*

Definitions

- **Anuria** refers to a total stop of urine production frequently caused by circulatory failure with anoxic damage of the tubular system.

- (Renal plasma) **Clearance** is a cleaning index for blood plasma passing the kidneys. The efficacy of this cleaning process is directly proportional to the excretion rate for the substance, and inversely proportional to its plasma concentration.

- **Diuresis** is an *increased* urine flow (ie, volume of urine produced per time unit).
- **Excretion fraction** (EF) for a substance is the fraction of its glomerular filtration rate, which passes to and is *excreted* in the urine.
- from the total amount of substance delivered to the kidney during one passage of the arterial blood plasma.
- **Free water clearance** is the difference between urine flow and osmolar clearance (see below). The free water clearance is an indicator of the excretion of solute-free water by the kidneys. *Excess water* is excreted compared to solutes, when free-water clearance is positive. *Excess solutes* are excreted compared to water, when free-water clearance is negative. – Free water clearance is an estimate of the renal capacity for excretion of solute-free water.
- **Glomerular filtration** is due to a hydrostatic/colloid osmotic pressure gradient - the Starling forces.
- **Glomerular filtration fraction** (GFF) is the fraction of the plasma flowing to the kidneys that is ultrafiltered (GFR/RPF). GFF is normally 0.20 or 1/5. - The GFF is reduced during acute glomerulonephritis.
- **Glomerulonephritis** is an autoimmune injury of the glomeruli of both kidneys.
- **Glomerular filtration rate** (GFR) is the volume of glomerular filtrate produced per min.
- **Glomerular propulsion pressure** in the blood of the glomerular capillaries is the hydrostatic minus the colloid osmotic pressure of the blood (ie, 2-3 kPa in a healthy resting person).
- **Glomerulo-tubular balance** refers to the simultaneous increase in NaCl and water reabsorption in the proximal tubules as a result of an increase in GFR and filtration rate of NaCl. An almost constant fraction of salt and water is thus reabsorbed regardless of the size of GFR.
- **Nephron**: A nephron consists of a glomerulus, a proximal tubule forming several coils (pars convoluta) before ending in a straight segment (pars recta), the thin part of the Henle loop and a distal tubule also with a pars recta and a pars convoluta.
- **The nephrotic syndrome** refers to a serious increase in the permeability of the glomerular barrier to albumin, resulting in a marked loss of albumin in the urine. The albuminuria (more than 3 g per day) causes hypoalbuminaemia and generalized oedema.
- **Net ultrafiltration pressure** is the pressure gradient governing the glomerular filtration - the net result of the so-called *Starling forces* (see Fig. 25-7).
- **Osmolar clearance** is the plasma volume cleared of osmoles (solutes) each minute. – Osmolar clearance is also defined as the fictive urine flow that would have rendered the urine isosmolar with plasma. - Osmolar clearance is the difference between the urine flow and the free water clearance, and osmolar clearance estimates the renal capacity to excrete solutes.

- **Osmolarity** is the amount of osmotically active particles dissolved in a litre of solution.
- **Proximal tubule** consists of the proximal convoluted tubule and pars recta.
- **Renal threshold for glucose** is the blood glucose concentration at which the glucose can be first detected in the urine (appearance threshold) or at which the reabsorption capacities of all tubules are saturated (saturation threshold).
- **Renal ultrafiltrate** is also compared to *plasma water*, because it is composed like plasma minus proteins. The fraction of one litre of plasma that is pure water is typically 0.94. Thus, the concentration of many substances in the ultrafiltrate, C_{filtr}, is equal to $C_p/0.94$.
- **Single effect gradient** is a transepithelial concentration gradient between the tubular fluid and the medullary interstitial fluid established at each level of the thick ascending limb by active NaCl reabsorption.
- $\mathbf{T_{max}}$ refers to the maximal net transfer rate of substance by tubular secretion or reabsorption.
- **Tubular passage fraction.** The fraction of the amount ultrafiltered of substance passing a cross section of the nephron is the *passage fraction*. The passage fraction for inulin does not vary at all throughout the nephron. The passage fraction for inulin is *one* and remains so.
- **Tubular reabsorption fraction.** The *reabsorption fraction* is the reverse of the passage fraction (1 minus the passage fraction).
- **Tubular reabsorption** (active or passive) is the net movement of water and solute from the tubular lumen to the tubule cells and often further on into the peritubular capillary network.
- **Tubular secretion** (active or passive) represents the net addition of solute to the tubular lumen.
- **Tubulo-glomerular feedback** (TGF) controls the glomerular capillary pressure and the proximal tubular pressure – thus stabilising delivery of solute and volume to the distal nephron. The macula densa-TGF mechanism responds to disturbances in distal tubular fluid flow passing the macula densa. - Renal autoregulation is caused by myogenic feedback and by the macula densa-TGF mechanism.

Essentials

This paragraph deals with *1. The nephron, 2. Clearance and three clearance families, 3. Ultrafiltration and the inulin family, 4. Tubular reabsorption and the glucose family, 5. Tubular secretion and the PAH family, 6. Water and solute shunting by vasa recta, 7. Concentration or dilution of urine, 8. Renal bloodflow, 9. Macula densa-tubulo-glomerular feedback, 10. Non-ionic diffusion, 11. Tests for proximal and distal tubular function, 12. Stix testing with dipstics, and 13. Diuretics.*

1. The nephron

The kidneys transport substances by three vectorial processes. Vectorial processes are characterized by their direction and size only (Fig. 25-1).

Fig. 25-1: Renal transport. Black arrows indicate three vectorial transporting processes in a nephron: 1. Glomerular ultrafiltration is caused by a hydrostatic/colloid osmotic

pressure gradient (the Starling forces), 2. Tubular reabsorption is the net movement of water and solute from the tubular lumen to the tubule cells and to the peritubular capillaries, and 3. Tubular secretion represents the net addition of solute to the tubular fluid.

The final excretion rate of the substance s in the urine is called net-flux, J_s, in Fig. 25-1.

1a. Nephron anatomy

The functional unit is the *nephron*. Each human kidney contains *1 million units* at birth. Each nephron consists of a glomerulus (ie, many glomerular capillaries in a Bowman's capsule), a proximal tubule forming several coils (pars convoluta) before ending in a straight segment (pars recta), the thin part of the Henle loop and a distal tubule also with a pars recta and a pars convoluta. The distal tubule ends in a *collecting duct* together with tubules from several other nephrons.

The kidney (average normal weight 150 g) consists of a cortex and a medulla. The medulla is composed of renal pyramids, the base of which originates at the corticomedullary junction. Each pyramid consists of an inner zone (the papilla) and an outer zone. The outer zone is divided into the outer medullary ray and the inner ray. The rays consist of collecting ducts and thick ascending limbs of the nephron.

A *kidney lobulus* is a medullary ray with adjacent cortical tissue. A kidney lobule is a pyramid with adjacent cortical tissue.

The *loop of Henle* is a regulating unit. Actually, the *Henle loop* consists of the proximal pars recta, the thin Henle loop and the distal pars recta, which ends at the level of macula densa.

The thin descending limb contains a water channel (called aquaporin 1) in both the luminal and the basolateral membrane. The last segment of the thick ascending limb is called the macula densa. The juxtaglomerular (JG) apparatus include the macula densa and granular cells of the afferent and efferent arterioles. Granular cells are modified smooth muscle cells that produce and release renin.

The distal tubule is convoluted from the macula densa of the JG apparatus (Fig. 25-2). The illustration shows a collecting duct, which receives urine from many neph
collecting ducts join to empty through the duct of Bellini into a renal cup or calyx in the renal pelvis.

The *superficial nephron* (represented on the left side of Fig. 25-2 A) does not reach the inner zone of the medulla, because its loop of Henle is short. These small, cortical nephrons have a smaller blood flow and glomerular filtration rate (GFR) than the deep, *juxtamedullary nephrons* (which are located close to the medulla and comprise 15% of all nephrons). The total inner surface area of all the glomerular capillaries is approximately *50-100* m^2. Mesangial and endothelial cells in the glomerulus secrete prostaglandins and exhibit phagocytosis. Many vasoconstrictors contract the mesangial cells, reduce the gomerular filtration coefficient (K_f – see later) and thus also GFR.

The *proximal tubules* have an inner area of 25 m^2 due to characteristic microvilli or brush borders (containing carboanhydrase).

Fig. 25-2: A: A superficial and a deep, juxtamedullary nephron leading to the same collecting duct. B: A juxtamedullary nephron with related blood vessels.

The juxtamedullary nephron has a *long, U-shaped* Henle loop. The bottom of this loop extends towards the tip of the papilla (apex papillae) at the outlet of the collecting duct (Fig. 25-2). The juxtamedullary nephrons have large corpuscles with relatively large bloodflow. These nephrons also receive blood through afferent arterioles with large diameters, and return blood through efferent arterioles with small diameters. When the blood has passed the juxtamedullary

glomeruli it continues to a primary capillary network and to the vasa recta in the medulla. The blood collects in vena arcuata, vena interlobaris and finally into vena renalis.

1b. The glomerular barrier

The filtration barrier of the glomerulus consists of capillary endothelium, basement membrane and the epithelial layer of Bowmans capsule consisting of podocytes with foot processes. The holes or fenestrae of the endothelium have a radius of approximately 40 nm (covered by a thin diaphragm) and are permeable to peptides and small protein molecules. The basement membrane consists of a network of fibrils permeable to water and small solutes. The podocytes cover the basement membrane with foot processes separated by gaps called *split-pores* through which the filtrate is retarded, because each split is covered by a membrane.

All small ions and molecules with an effective radius below 1.8 nm (water, ions, glucose, inulin etc) filtrate freely. Substances with a radius of 1.8-4.2 nm are less filterable, and substances with a radius above 4.2 nm cannot cross the barrier.

All channels of the glomerular barrier carry *negatively charged* molecules that facilitate the passage of positively charged molecules (eg, polycationic dextrans, Fig.25-3). Dextran macromolecules can be electrically neutral or they have negative (anionic) or positive (cationic) charges.

Fig. 25-3: Filtration of dextran molecules across the glomerular barrier. The barrier contains glycoproteins with negative charges. Positive charged dextran molecules are attracted by the negative charges and filter easily.

Positive charged molecules with an effective radius of 3 nm filter easier than negative charged molecules of the same size. These molecules can act as effective osmotic diuretics.

Immunological or inflammatory damages of the glomerular barrier reduce the negative charge of the barrier. Hereby, negative protein molecules leave the plasma easier and proteinuria occurs in a number of glomerular disorders.

1c. Pregnancy and age

The glomeruli grow and the size and weight of the kidneys increase during *pregnancy*, accompanied by increases in both renal bloodflow and filtration rate.

The number of glomeruli and their tubules *decrease* with age. Drugs that are excreted by renal mechanisms can easily cause toxic accumulation in the elderly with poor kidney function.

2. Clearance

In 1926 Poul Brandt Rehberg, an associate of August Krogh, found the muscle metabolite *creatinine* extremely concentrated in human urine (C_U mg per ml) compared to plasma (C_P mg per ml). He also measured the urine flow (urine production per min).

Thus, the concentration index, C_U/C_P, is large for creatinine. Multiplying this index with the urine flow yields a result greater than
1). Brandt Rehberg used this concept (later termed clearance) as his measure of *renal filtration rate*. The work with these matters developed into the idea of a *filtration-reabsorption* type of kidney. Rehberg was the first to realise that the reabsorption in the proximal tubules controls the filtration. A few years later Rehberg´s *renal filtration rate* was called *creatinine clearance* and used as a measure of the *glomerular filtration rate* (GFR).

The *renal plasma clearance* is a *cleaning index* for blood plasma passing the kidneys. The efficacy of this cleaning process is directly proportional to the excretion rate for the substance and inversely proportional to its plasma concentration (Eq. 25-1).

Clearance is the ratio between excretion rate and plasma concentration for the substance. Renal clearance can also be thought of as the volume of arterial plasma completely cleared of

the substance in the kidneys within one min, or the *number of ml* arterial plasma containing the same amount of substance as contained in the urine flow per minute (Eq. 25-1).

2a. Glomerular filtration rate

The glomerular filtration rate, GFR, is the volume of glomerular filtrate produced per min.

In healthy adults the GFR is remarkably constant about 180 l each day or 125 ml per min due to intrarenal control mechanisms. In many diseases the renal bloodflow, RBF, and GFR will fall, whereby the ability to eliminate waste products and to regulate body fluid volume and composition will decline. The degree of impaired renal function is shown by the measured GFR.

GFR is routinely measured as the *endogenous* creatinine clearance.

The endogenous creatinine production is from the creatine metabolism in muscles and proportional to the muscle mass. In a 70-kg person creatinine is produced at a constant rate of *1.2 mg per min* (1730 mg daily). This production is remarkably constant from day to day, only slightly affected by a normal protein intake, and equal to the rate of creatinine excretion. Both the serum creatinine and the renal creatinine excretion fluctuate throughout the day. Therefore, it is necessary to collect the urine for 6-24 hours and measure the creatinine excretion rate (ie, the urine flow rate multiplied by the creatinine concentration in the urine). A single venous blood sample analysed for creatinine in plasma is all that is needed to provide the *endogenous* creatinine clearance (Eq. 25-1).

Theoretically, two small errors disturb the picture, but both are overestimates.

At the normally low plasma concentrations of creatinine, a modest tubular secretion of creatinine from the blood is detectable resulting in up to 15% overestimation of the creatinine excretion flux. Most laboratories measure creatinine in *serum* instead of plasma, which results in an overestimation of plasma creatinine.

Thus, calculation of a fraction with both an overestimated nominator and denominator results in a value close to that of GFR in almost all situations, where the renal function is near normal.

With progressive renal failure the plasma creatinine rises, and the creatinine secretion increases the nominator in the clearance expression even more, so the measured clearance will overestimate GFR. Still, the clearance provides a fair clinical estimate of the renal filtration capacity (GFR).

In most cases a normal creatinine clearance (above 70 ml plasma per min at any age) is comparable with the normal range for serum creatinine (around 0.09 mM in Fig. 25-4). The serum creatinine concentration is inversely proportional to the creatinine clearance, and also a good estimate of GFR. Renal failure is almost always *irreversible,* when the serum creatinine is above 0.7 mM.

Fig. 25-4: Creatinine clearance versus serum creatinine. – A low serum creatinine indicates normal kidney function, but not always (see false negative concentrations). – An elevated serum creatinine indicates kidney failure, but not always (see false positive concentrations).

Serum [creatinine] and serum [urea] depend upon both *protein turnover* and *kidney function*. The serum [creatinine] and [urea] are large after intake of meals extremely rich in (fried) meat, although the kidney function is normal (false positive concentrations in Fig. 25-4). In some materials up to 15% of measured serum creatinine concentrations are normal, although the kidney function fails (false negative values in Fig. 25-4). Long-term hospitalisation often leads to muscular atrophy, which reduces creatinine production and excretion. The serum creatinine concentration is maintained normal because of a similar fall in kidney function (GFR).

Half the osmolality of normal urine is due to *urea,* and the other half is mainly due to *NaCl*.

The osmolarity of urine varies tremendously (from 50 to 1400 mOsmol per l).

Physiological changes of the renal bloodflow often parallel changes of GFR. A *reduced* GFR implies a smaller tubular Na^+-reabsorption and thus a smaller O_2 *demand*. When kidneys are perfused by anoxic blood the tubular reabsorption is blocked first, and then the GFR is reduced. As tubular Na^+ -reabsorption is the main oxidative energy demanding activity, a high GFR is correlated to high *oxygen consumption* in the normal kidney.

The size of GFR is determined by the factors shown in Fig. 25-7. The resistance of the glomerular barrier is extremely small in healthy human kidneys.

2b. Inulin

Inulin is the ideal indicator for determination of GFR, because of the following three relations:

1. Inulin is a polyfructose (from Jewish artichokes) without effect on GFR. Inulin has a spherical configuration and a molecular weight of 5000. Inulin filters *freely* through the glomerular barrier. Inulin is uncharged and not bound to proteins in plasma. Inulin crosses freely most capillaries and yet does not traverse the cell membrane (distribution volume is ECV). Since one litre of plasma contains around 0.94 l of water, the ultrafiltrate concentration of inulin is $C_p/0.94$.

2 All ultrafiltered inulin molecules pass to the urine. In other words, they are neither reabsorbed nor secreted in the tubules. Inulin is an exogenous substance - not synthesised or broken down in the body.

3. Inulin is non-toxic and easy to measure.

Thus, under steady-state conditions, the rate of inulin leaving the Bowman's capsulesmust be exactly equal to the rate of inulin arriving in the final urine. The main idea is to measure the *amount* of inulin excreted in the urine during a timeperiod were the plasma [inulin] is maintained constant by constant infusion of inulin. After one hour the subject urinates, and the urine volume and inulin concentration in the urine and plasma is measured. The amount of inulin filtered through the glomerular barrier per min is: $(GFR \times C_p/0.94)$.

All inulin molecules remain in the preurine until the subject urinates. Thus, the amountexcreted is equal to the amount filtered and Eq. 25-4 is developed (see later).

Since the *inulin clearance* is 180 l per 24 hours for young, healthy males or 125 ml per min, the GFR must be $(125 \times 0.94) = 118$ ml per min. The inulin clearance is 10% lower for young females than for young males due to the difference in average body weight and body surface area.

The normal values for both sexes decrease with age to *70 ml* per min after the age of *70*.

Inulin clearance is a precise experimental measure and the ideal standard, but inulin must be infused intravenously, and the method is not necessary in clinical routine.

If the clearance of a substance has the *same value* as the inulin clearance for the person, then the substance is only subject to *ultrafiltration.* Theoretically, reabsorption might balance tubular secretion and give the same result.

If the clearance of a substance is *greater* than the inulin clearance, then clearly this substance is being added to the urine as it flows along the tubules; in other words, it is being *secreted.*

Similarly, if the clearance of a substance is *less* than the inulin clearance, it means that the substance is being *reabsorbed* at a higher rate than any possible secretion.

The extracellular fluid volume (ECV) can be measured with inulin as inulin does not pass the cell membrane (see Chapter 24 and Eq. 24-4). The elimination of inulin is exponential - ie, the

fraction (k) of the remaining amount in the body that disappears per time unit is constant (see Chapter 1). Since the filtration family of substances is eliminated from the blood solely by filtration, the elimination depends only on GFR, and the distribution volume is that of inulin (ECV). Thus, the elimination rate constant ($k = 0.69/T_{½}$) for the inulin family is roughly equal to $(GFR*C_p)/(ECV*C_p)$.

2c. The three clearance-families

All substances treated by the kidneys can be divided into three groups or *families*, namely the filtration-, the reabsorption-, and the secretion- family.

The kidney treats the *filtration family* of substances (see later) just like inulin.

The filtration rate (J_{filtr}) for inulin equals the excretion rate (J_{excr}), and both increase in direct proportion to the rise in C_p (Fig. 25-5). The clearance is the slope of the curve, and it is obviously a constant value that is independent of C_p.

Fig. 25-5: The straight line shows a direct relationship between the filtration rate and the concentration for the inulin family of substances in plasma.

The reabsorption or glucose family contains many vital substances (see later). For the reabsorption family of compounds, the excretion flux is equal to the filtration flux minus the reabsorption flux. The maximal reabsorption flux (T_{max}) is reached above a certain threshold. Above this saturation threshold the clearance for the reabsorption family is equal to (the inulin clearance - T_{max}/C_p), according to the mathematical argument in Fig. 25-8.

The secretion or PAH family comprises endogenous substances and drugs (see later). Foreign substances are often distributed in the ECV, but some of them are also entering cells (ICV). At low concentrations their elimination rate constant (k) is roughly equal to renal plasma flow (RPF) divided by ECV: ($RPF*C_p/ECV*C_p$) = RPF/ECV. Thus, k equals RPF/ECV or 1/20 min^{-1} in most healthy persons. The k value corresponds to a half-life of 14 min ($T_{½} = \ln 2/k$).

2d. Excretion rate and clearance.

Excretion rate curves for inulin can be changed into clearance by a simple mathematical procedure:

Differentiating the excretion flux curve for the inulin family with respect to C_p produce the renal plasma clearance curves for these substances. Let us assume that the curves are from a resting person in steady state with a normal inulin clearance (the slope of the line in Fig. 25-6,A).

For the *inulin family* the excretion flux equals (urine flow × C_u), and by division with C_p we have the inulin clearance.

Fig. 25-6: A, B, and C are the filtration-reabsorption- and secre substances, respectively. - D shows the clearance curves.

For all substances belonging to the inulin family the excretion flux curves are linear, so the *rate of change* (which is the clearance) must be constant in a given condition (Fig. 25-6A).

The results of the three excretion fluxes are plotted with C_p as the dependent variable (x-axis of Fig. 25-6, ABC).

The excretion flux curves for the three families of substances, when differentiated (dJ_{excr}/dC_p), provide us with the three possible clearance curves (Fig. 25-6, D).

For the *reabsorption family*, the clearance is zero at first, because the excretion is zero (Fig.

25-6 D). The clearance increases, and finally it approaches the inulin clearance. Therefore, the clearance is steadily increasing towards inulin clearance with increasing C_p.

For the *secretion family*, the clearance must also be equal to the excretion flux divided by C_p. When the [PAH] increases, more and more PAH is eliminated by filtration, and the secretory elimination is relatively suppressed (so-called *auto-suppression*). The clearance for the secretion family is falling with increasing C_p, and approaches that of inulin (Fig. 25-6 D).

Box 25-1: Composition of urine

Component	Concentration	Daily renal excretion	Finding/Disease
Water		500-2500 ml	<500 ml/Nephropathy, shock >2500 ml/Diabetes
Potassium	60-70 mM	90 mmol daily	<20 mmol daily/Low diet >150 mmol daily/Rich diet
Sodium	50-120 mM	150 mmol daily	
Protein	20 mg*l-1	30-150 mg daily	Microalbuminuria/Diabetes
Proteinuria/Nephropathy			
Glucose	zero	Negligible	Glucosuria/Diabetes mellitus Glucosuria/Proximal defect
Urea	200-400mM	500 mmol daily	High excretion/Uraemia
Creatinine	0.1	1500-2000 mg daily	High excretion/Large m. mass Low excretion/Muscul. atrophy
Osmolality	>600 mOsmol*kg-1		Acceptable conc. capacity

The composition of urine in Box 25-1 is the basis for simple diagnostics. Anuria or oliguria (<500 ml daily) indicates the presence of hypotension or renal disease. Polyuria (>2500 ml of urine daily) is the sign of diabetes – both diabetes mellitus and diabetes insipidus. Microalbuminuria (ie, 50-150 mg per l) indicates glomerular barrier disorder such as diabetic glomerular disease. Glucosuria with hyperglycaemia is the sign of diabetes mellitus, and without hyperglycaemia it is a sign of a proximal reabsorption defect. High urea excretion is seen in uraemia, and high creatinine excretion indicates a large muscle mass in a healthy person. A low creatinine excretion is the sign of muscular atrophy or ageing.

3. Ultrafiltation and the inulin family

In a healthy person at rest almost 25% of *cardiac output* passes the two kidneys (1200 ml each min). The blood reaches the first part of the nephron through the afferent arteriole to the glomerular capillaries. In the *glomerular capillaries* the hydrostatic pressure is approximately 60 mmHg at the start and 55 mmHg at the end (Fig. 25-7).

The inulin or filtration family consists of inulin, radioactive indicators(^{51}Cr-EDTA, ^{57}Co-marked B_{12}, ^{14}C-marked inulin, ^{3}H-marked inulin, iothalamate marked with ^{125}I or ^{131}I),

mannitol, raffinose, sucrose, thiocyanate, and thiosulfate. These substances are more or less evenly distributed in the ECV.

3a. The Starling forces

The pressures governing the glomerular ultrafiltration rate (GFR) are called the *Starling forces* (see equation in Fig. 25-7). Normally, filtration continues throughout the entire length of the glomerular capillaries in humans, because the net ultrafiltration pressure (P_{net}) is positive also at the efferent arteriole. The average values for determinants of GFR are given in the first equation of Fig. 25-7. The hydrostatic pressure gradient is an important determinant of GFR. The *glomerular filtration coefficient* is called K_f. The K_f is equal to the filtration surface area divided by the resistance of the glomerular barrier and thus a constant for a given barrier (Fig. 25-7). The value of K_f (also called the *reciprocal glomerular hydrodynamic resistance*) is reduced in diabetes, glomerulonephritis and hypertension. Vasoactive substances constrict or dilatate the glomerular mesangial cells and change the value of K_f.

In other conditions, the forces opposing filtration become equal to the forces favouring filtration at some point along the glomerular capillaries. This is called *filtration equilibrium.*

The hydrostatic pressure in *Bowmans space* below the glomerular barrier is about 15 mmHg or 2 kPa (P_{Bow} in Fig. 25-7). This pressure is almost equal to the proximal tubule pressure, since there is no measurable pressure fall along this segment.

Fig. 25-7: Net ultrafiltration pressures in afferent and efferent end of glomerular capillaries. The Starling forces determine the final ultrafiltration pressure (P_{net}) across the glomerular barrier.

There is almost no colloid-osmotic pressure in Bowmans space, but an oncotic pressure of approximately 25 mmHg in the incoming plasma, mainly due to proteins, which are up-concentrated, when fluid leaves the the plasma for Bowmans space. Hereby, the protein-oncotic pressure (p_{gc}) may increase from 25 to 35 mmHg at the end of the glomerular capillary (Fig. 25-7). The higher the renal plasma flow (RPF), the lower is the rise in p_{gc}.

A selective increase in the resistance of the afferent arteriole reduces both the RPF and the glomerular hydrostatic pressure (P_{gc}), but GFR decreases more than RPF, so the filtration fraction (= GFR/RPF) falls. In contrast, a rise in the resistance of the efferent arteriole reduces RPF but increases P_{gc} (Fig. 25-7). Instantly, GFR increases slightly, but GFR eventually decreases due to the rise in p_{gc}. As RPF falls more than GFR the filtration fraction increases.

A combined increase in both the afferent and the efferent arteriolar resistance (as caused by most vasoconstrictors) may also reduce RPF more than GFR, and increase the filtration fraction.

3b. The net ultrafiltration pressure

The *net ultrafiltration pressure* (P_{net}) varies from 20 to 5 mmHg through the glomerular capillaries, and provides the force for ultrafiltration of a fat- and protein- free fluid across the glomerular barrier into Bowmans space and flow through the renal tubules (Fig. 25-7).

The ultrafiltrate is isosmolar with plasma, almost protein free, and contains low molecular substances in almost the same concentration as in plasma water.

The proximal tubular reabsorption takes place through para- and trans-cellular pathways. In the

peritubular capillaries, the Starling forces are seemingly adequate for capillary uptake of interstitial fluid (Fig. 25-7).

The hydrostatic *net pressure* in the proximal tubules – and with it the GFR - is remarkably well maintained in spite of changes in proximal reabsorption of salt and water.

An acute defect in the proximal reabsorption mechanism results in an initial rise in proximal hydrostatic pressure and the GFR is reduced. Due to *autoregulation* (see paragraph 9), the proximal hydrostatic pressure is rapidly normalised at a new steady state.

Sympathetic stimulation increases both the proximal reabsorption rate and the peritubular capillary uptake (Fig. 25-7). Hereby, the hydrostatic pressure falls in the proximal tubules and Bowman's capsule so GFR may increase. In reverse, angiotensin II secretion inhibits the proximal reabsorption rate, increases the proximal pressure and may reduce GFR.

The *total distal flow resistance* below the proximal tubules (ie, in the distal system) is large and important. The distal resistance has two major components namely a high resistance in the Henle loop and an even higher resistance in the remaining distal system including the collecting ducts.

The *resistance of the glomerular barrier* is calculated in Fig. 25-7 to be extremely small.

Normally, there is hardly any hydrodynamic resistance to glomerular ultrafiltration.

4. Tubular reabsorption and the glucose family

The *reabsorption or glucose family* contains vital substances such as glucose, amino acids, albumin, acetoacetates, ascorbic acid, beta-hydroxybutyrate, carboxylate, vitamins, lactate, pyruvate, Na^+, Cl^-, HCO_3^-, phosphate, sulphate and urea.

4a. Tubular handling of glucose

T_{max} is the maximum transfer or net reabsorption flux (J_{reabs}) for glucose (mol.wt. 180 g per mol) in the proximal tubules. The optimal value for this *glucose transporter* is 300 mg/min or 300/180 = 1.7 mmol/min for healthy, young subjects with a body weight of 70 kg.

For the reabsorption family of substances, the excretion is zero at first since the entire filtered load is reabsorbed (all glucose is reabsorbed, see Fig. 25-8). The excretion flux increases then linearly with increasing filtration flux.

Fig. 25-8: Renal Glucose rates as a function of the plasma concentration (C_p).

The *appearance threshold* is the blood plasma [glucose] at which the glucose can be first detected in the urine (normally 8.3 mM or 150 mg%). This occurs when most but not all nephrons are saturated (Fig. 25-8).

The actual *saturation threshold,* the point where all nephrons are glucose-saturated, is much higher (normally above 13.3 mM). The concentration difference (13.3 - 8.3 = 5 mM) represents a similar reabsorption rate difference (1.7 - 1.0 = 0.7 mmol/min at normal GFR) called *splay*. The reabsorption capacity for glucose in the proximal tubule cells becomes saturated at these high blood concentrations (Fig. 25-8).

4b. Urea transport

The water reabsorption in the proximal tubules increases the urea concentration in the fluid. Since urea is uncharged and diffuses easily, it will diffuse passively to the peritubular capillary blood. The passage fraction at the outlet of the proximal tubule is around 0.5 (50% of the

filtered load).

Urea is thus reabsorbed in the proximal tubules and also in the inner medullary collecting ducts and secreted in the thin descending and ascending limb of the Henle loop (see later).

The kidney reuses urea by recirculation in the intra-renal urea recycling circuit: Inner medullary collecting ducts – medullary interstitium – loop of Henle – distal tubules – collecting ducts.

The net reabsorption flux is around 50% of the filtration flux at normal urine flow. The normal urea concentration in plasma is 5mM, and the excretion flux for urea is proportional to this urea concentration.

4c. Proximal tubular reabsorption

Healthy proximal tubules reabsorb approximately 70% of the filtered water, Na^+, Cl^-, K^+ and other substances. The tubular passage fraction for these substances at the outlet of the proximal tubule is 0.3 (30%). The reabsorption of fluid is isosmotic. Almost all filtered glucose, peptides and amino acids are also reabsorbed by the proximal tubules. The Cl- reabsorption is passive.

This ion follows the secondary active reabsorption of Na^+ in order to maintain electrical neutrality. Reabsorption of water is passive as a result of the osmotic force created by the reabsorption of NaCl. All reabsorption processes are linked to the function of the basolateral Na^+-K^+-pump. The extremely high water permeability of the proximal tubule is essential for its nearly isosmotic volume reabsorption. The active reabsorption of solutes makes the fluid slightly dilute and the interstitial fluid slightly hypertonic. If inulin and PAH molecules are present their concentration in the fluid will rise (PAH also because of proximal secretion). The actively reabsorbed solutes have lower permeabilities (higher reflection coefficients) than NaCl.

In the *first half of the proximal tubule*, Na^+- is reabsorbed with carbonic acid and organicmolecules belonging to the reabsorption family. - The proximal and distal reabsorption ofbicarbonate is already described in Chapter 17.

Fig. 25-9: Reabsorption of NaCl in the early and the late part of the proximal tubule. CA stands for carboanhydrase in the brush borders of the cell.

The reabsorption family of substances (X) enters the tubule cells by *specific symporter proteins* coupled to the Na^+ -reabsorption (1.in Fig. 25-9). This is secondary active transport showing saturation kinetics. Na^+ -reabsorption is also coupled to H^+ -secretion from the cell by the function of the *Na^+ -H^+-antiporter protein* (2. in Fig. 25-9). This H^+ -secretion is linked to bicarbonate reabsorption in the upper part of the proximal tubules. The driving force for the Na^+ -entry is the Na^+ -K^+-pump located in the basolateral membrane, which extrudes the Na^+ to the intercellular space and the blood (3. in Fig. 25-9). Glucose is a typical example. The luminal membrane contains a sodium-glucose-cotransporter (SGLT 2). A genetic defect in this protein produces familial renal glucosuria – just as a genetic defect in a similar intestinal protein (SGLT 1) produces glucose-galactose malabsorption. - The passage of glucose across the basolateral membrane is by carrier-mediated (facilitated) diffusion.

In the *second half of the proximal tubule*, Na^+ is reabsorbed together with Cl^- across the cell membrane or through paracellular routes (Fig. 25-9, below). In this segment the tubular fluid contains a high concentration of Cl^- and a minimum of organic molecules. Na^+ crosses the

luminal membrane by the operation of *Na^+-H^+-antiporters* and *Cl^- -anion antiporters*. In the tubular lumen the secreted H^+ and anion form a H^+-anion complex. The accumulation of a lipid-soluble H^+-anion-complex establishes a concentration gradient that allows *H^+-anion-complex* recycling (Fig. 25-9). Transfer of the Cl^- -ion from the tubular fluid to the blood causes the tubular fluid to become positively charged relative to the blood.

4d. Reabsorption in the thick ascending limb

The Na^+-K^+-pump maintains a low intracellular Na^+ , which drives the simultaneous, electroneutral *reabsorption* of 1 Na^+, 1 K^+, and 2 Cl^- by the *luminal Na^+-K^+-$2Cl^-$-symporter*. The Cl^--channels are only located in the basolateral membrane, so accumulated Cl^- reaches the ISF. The K^+-channels are located in all membranes and K^+ recirculates (Fig. 25-10). Paracellular reabsorption of positive ions by diffusion is augmented by the positive charge of the tubular fluid (Fig. 25-10).

The secondary active re*absorption of Na+ (and Cl-) is the basis for the transepithelial single effect gradient a*t each transverse level of the thick ascending limb (see later).

Fig. 25-10: Reabsorption of NaCl in the thick ascending limb of the Henle loop. There is a luminal Na^+ K^+-$2Cl^-$-symporter and a basolateral Na^+ K^+-pump. This mechanism is essential for development of medullary hypertonicity by NaCl and thus for counter current mutiplication (see later).

The electrochemical energy for the function of the basolateral Na^+ K^+-pump is provided by its Na^+-K^+-*ATPase*. The pump throws Na^+ into the peritubular fluid. The K^+ and Cl^- ions leak out passively. The thick ascending limb is impermeable to water in the absence of ADH, and reabsorbs Na^+ actively.

Loop diuretics, which abolish the entire osmolar gradient in the outer renal medulla, inhibit the *luminal Na^+ K^+-$2Cl^-$-symporter* of the thick ascending limb.

4e. Reabsorption in the distal tubule and collecting duct

The distal tubule is divided into an early and a late segment, since the early segment reabsorbs NaCl and is impermeable to water (as the thick ascending limb), whereas the late segment functions more like the collecting duct. In the early segment, the NaCl transfer is mediated by [a] *NaCl-symporter* (Fig. 25-11). Na^+ leaves the cell through the basolateral Na^+-K^+-pump, and Cl^- leaves the cell by diffusion across the basolateral Cl^--channels. Only a small fraction of the glomerular filtrate reaches the distal tubules. Thiazide diuretics inhibit the NaCl-symporter.

Fig. 25-11: Cellular transport processes in the distal tubule and collecting duct.

The late segment is composed of two cell types just as the collecting ducts. The light *principal cells* reabsorb Na^+ and secrete K^+. The Na^+-K^+-pump in the basolateral membrane draws Na^+ out into the ISF and K^+ into the principal cells (Fig. 25-11). These cells have special ion channels in the luminal membrane, which is permeable to Na^+, but also to K^+. The Na^+-uptake depolarises the luminal membrane (-70 mV) and makes the lumen electronegative (-12 mV) compared to the interstitial fluid (reference potential zero). K^+ rapidly diffuses into the tubular fluid. This secretion of K^+ into the tubular fluid from the principal cell is thus linked to the

Na^+-reabsorption. The amount of Na^+ reabsorbed in the distal tubule system is much less than in the proximal, but it can be increased by the adrenocortical hormone, aldosterone.

Aldosterone is a mineralocorticoid, which promotes the reabsorption of Na^+ (and thus Cl^-) and the secretion of K^+ (and H^+) in principal cells. Aldosterone enters the cell from the blood and binds to an intracellular receptor to form a complex. The complex increases the formation of membrane proteins including the Na^+-K^+-pump and the luminal Na^+-channels. This is the essential control mechanism for $[K^+]$ in the ECV. Secretion mainly occurs when the $[K^+]$ in the ECV is higher than normal.

Aldosterone also promotes the reabsorption of Na^+ (and thus Cl^-) and the secretion of K^+ (and H^+) in the collecting ducts of sweat and salivary glands just as in the principal cells of the distal tubules of the kidney. Aldosterone-an

intercalated cells secrete H^+ across the luminal membrane and reabsorb K^+.

Intercalated cells are mitochondrial-rich and most active in persons with a low K^+-pool. The H^+ -secretion by the H^+-pump is precisely determined by the $[H^+]$ in the ECV.

The *collecting duct* contains principal and intercalated cells just as the late distal segment, but the intercalated cell disappears in the inner medullary collecting ducts.

The luminal membrane of the principal cells in the collecting ducts can be regulated from nearly water-impermeable (in the absence of antidiuretic hormone, ADH) to water-permeable (in the presence of ADH). The hormone increases the water-permeability by insertion of water-channels called *aquaporin 2*. The water-channels are stored in cytoplasmic vesicles that fuse with the luminal membrane. The basolateral membrane of the principal cell contains other aquaporins and they remain water-permeable even in the absence of ADH. Mutations in the genes for these channel proteins cause *nephrogenic diabetes insipidus*.

5. Tubular secretion and the PAH family

Substances secreted like PAH constitute the secretion or PAH family. The filtration flux (J_{filtr}) as usual increases in direct proportion to the rise in C_p (Fig. 25-12). Dividing the excretion flux for PAH with C_p provides us with the *PAH clearance*. The clearance is the slope of the excretion flux curve (Fig. 25-12). The secretion flux approaches a maximum (T_{max}). Most of the PAH molecules are free, but 10-20% are bound to plasma proteins.

Fig. 25-12: Renal PAH net rates (fluxes or *J*) as a function of plasma concentration, C_p.

Organic acids and bases secreted in the proximal tubules include endogenous substances and drugs. The *endogenous* substances include adrenaline, bile salts, cAMP, creatinine, dopamine, hippurates, noradrenaline, organic acids and bases, oxalate, prostaglandins, steroids and urate. The *drugs* comprise acetazolamine, amiloride, atropine, bumetanide, chlorothiazide, cimetidine, diodrast, furosemide, hydrochlorothiazide, morphine, nitrofurantoin, para-aminohippuric acid (PAH), penicillin, phenol red, probenecid, sulphonamides, and acetylsalicylic acid. The secretion is often competitive. All these substances have varying but high affinity to an *organic acid-base secretory system* in the proximal tubule cells showing saturation kinetics with a T_{max}. The organic cation secretion is analogous to the anion secretion.

5a. Tubular handling of PAH

T_{max} is the maximum secretion rate for PAH in the tubules (Fig. 25-13). Normally, the T_{max} is 0.40 mmol per min (80 mg/min) for PAH.

At low PAH concentrations in the plasma (Fig. 25-13), the slope of the excretion rate curve is high (the clearance for PAH is high). Here the PAH clearance is an accept
minimal renal plasma flow (see effective RPF later), because the blood is almost cleared by one transit.

The secretion flux is maximal, when the plasma-[PAH] is high enough to achieve saturation. The weak organic acids and bases mentioned above are similarly secreted into the proximal tubule, and have secretory T_{max} -values just like PAH (Fig. 25-14). In humans of average size (with an average body surface area of 1.7 m^2), the T_{max} for diodrast and phenol red average 57 and 36 mg/min, respectively.

5b. Tubular handling of urate

The *active reabsorption* of urate ions is accomplished in the proximal tubules by an electroneutral Na^+-cotransport. The tubular reabsorptive capacity is normally far greater than the amount delivered in the glomerular filtrate. Above a critical concentration in the ECV of about 0.42 mM, the urate precipitates in the form of uric acid crystals, provided the environment is acid. Precipitation in the joints is termed *gout* (arthritis urica), often affecting several joints. Urate ions are accumulated in the ECV of gout patients, and often also in patients with uraemia. High doses of probenecid compete with urate for the proximal reabsorption mechanism. Use of this drug to patients with acute gout increases the excretion of urate in the urine.

The *active* secretion of urate ions occurs from the blood plasma to the tubular fluid by the *organic acid-base secretory system*, which has a low capacity for urate.

Thus, the renal tubules have a capacity of both actively reabsorbing urate ions and actively secreting them.

5c. Tubular handling of creatinine

Essentially all creatinine in the glomerular filtrate passes on and is excreted in the urine. The molecule is larger than that of urea, and none of it is reabsorbed. Contrary, creatinine is secreted into the proximal tubules, so that the creatinine concentration in the urine increases more than 100-fold.

5d. The secretion mechanism

The molecules of the secretion family leave the blood plasma of the peritubular capillaries and binds to basolateral receptors with symporters on the tubule cell (Fig. 25-13). These channels are driven by energy from the basolateral Na^+-K^+-pump transporting the molecules against their chemical gradient across the basolateral membrane. Inside the cell the molecules accumulate until they can diffuse towards the luminal membrane. Here, an antiporter transfers the ions into the tubular fluid. All these molecules compete for transport, so intake of the drug probenecid can reduce the penicillin secretion loss.

Fig. 25-13: Secretion of organic anions across the proximal tubules

The luminal membrane contains *specific* receptor proteins for *nutritive mono- and di-carboxylates*. These receptor functions are also coupled to Na^+ -transfer.

6. Water and solute shunting by vasa recta

The normal perfusion of the renal medulla is typically 5-10% of RBF. This bloodflow is larger than the fluid flow through the loop of Henle. Both the vasa recta and the closely located loops of Henle (from juxtamedullary nephrons) consists of two parallel limbs with counter-current fluid flow in the medulla.

Vasa recta are designed as a counter current bloodflow and act as *water-solute shunts* that protect the medullary hyperosmotic gradient. The endothelial lining of vasa recta is highly permeable for small molecules (water, urea, NaCl, oxygen and carbon dioxide). Vasa recta also serve as a nutritive source to the medulla.

Vasa recta receive blood from the efferent arterioles and consequently have an elevated colloid osmotic pressure and reduced hydrostatic pressure (Fig. 25-14). The net force in these vascular loops favours *net fluid reabsorption*.

Let us consider the situation with a hyperosmotic medullary gradient and ADH present, so a concentrated urine is produced. The blood in the descending limb of vasa recta is first passed on in the direction of increasing medul
supply water to the hyperosmolar, interstitial fluid by passive osmosis, and passively reabsorb solutes (NaCl and urea) by diffusion. Hereby, the interstitium is temporarily diluted and the blood is concentrated. In the ascending portion the blood passes regions with falling osmolarity, and the blood gradually absorbs water osmotically and delivers solutes to the interstitium by diffusion. The flow in the ascending vasa recta is larger than in the descending limb, because water from the Henle loop is also reabsorbed.

Fig. 25-14: A: Passive counter-current exchange occurs in vasa recta, with diffusion of solutes along black arrows. Passive osmotic flux of water from the blood to the hyperosmolar interstitium occurs along stippled, blue arrows. – B: The active counter-current multiplier in the thick ascending limb with a single effect at each horizontal level.

The gross effect of the passive counter-current exchange in the vasa recta is that of a *water shunt* passing the medullary tissue, whereas solutes recycle and thus are maintained in medulla. Water is shunted from limb to limb without disturbing the inner medulla. The passive counter-current exchange and low bloodflow through the vasa recta curtail the medullary hyperosmotic gradient (Fig. 25-14). The meagreness of the medullary blood flow, reduced by ADH, contribute to the maintenance of the medullary hyperosmotic gradient, but reduce the nutritive supply to the inner medulla.

7. Concentration or dilution of urine

The thin ascending limb of Henle is impermeable for water, but highly permeable for NaCl and less so for urea. The thick ascending limb is also impermeable for water and also for urea. The water permeability of the cortical and medullary collecting ducts increase with increasing concentrations of antidiuretic hormone (ADH) in the peritubular blood.

Concentration of urine. Initially, the osmolarity of the tubular fluid, the vasa recta blood, and the interstitial fluid is 300 mOsmol * l^{-1}. The ascending limb of the Henle loop is impermeable to water and actively transports NaCl from the preurine into the surrounding interstitium. Thus solute and fluid is separated and the tubular fluid becomes diluted. At each horizontal level of the thick ascending limb, a hyperosmotic gradient (a *single effect*) of typically 200 mOsmol * l^{-1} is established (Fig. 25-14B). Energy is necessary to establish the hyperosmotic gradient.

The energy is from Skou´s basolateral Na^+-K^+-pump, working in conjunction with the Na^+-K^+-$2Cl^-$symporter of the thick ascending limb (Fig. 25-10).

The total osmolarity in the inner medullary interstitial tissue can be as high as *1400 mOsmol* per l, when the urine is maximally concentrated.

The renal cortex fluid is isotonic with the plasma. When the isotonic fluid from the proximal tubules passes down through the hypertonic medulla in the *descending thin limb* of the Henle loop, water moves out into the medullary interstitium by osmosis, making the tubular fluid concentrated. This is because the epithelial cells of the thin descending limb are highly permeable to water but less so to solutes (NaCl and urea). Water is reabsorbed and returned to the body via vasa recta and the renal veins. At the bend of the loop the fluid has an osmolarity equal to that of the surrounding medullary interstitial fluid. However, the tubular fluid has a greater concentration of NaCl and a smaller concentration of urea than the surroundings.

In contrast to the thin and thick ascending limb, most cell membranes including those of the proximal tubules and the thin descending limb of the Henle loop, are water-permeable under all circumstances. This is because these cell membranes contain water-channel proteins called *aquaporins*.

As new fluid enters the descending limb of the Henle loop, the hyperosmotic fluid in the bottom of the loop is pushed into the ascending limb, where NaCl is separated from water.

The osmolarity of the isosmotic tubular fluid running into the thin descending loop of the outer medulla is 300 $mOsmol*l^{-1}$ and the output to the distal tubule is 100 $mOsmol*l^{-1}$ (Fig. 25-14, B). At the bottom of the Henle loop the osmolarity can increase to at least 1300 $mOsmol*l^{-1}$. In a steady state with continuous fluid flow the total osmotic gradient along the entire system is thus (1300 - 100) = *1200*. The gradient along the entire system is *6 multiples* of the 200 $mOsmol*l^{-1}$ single effect gradient. The thick ascending limb is a *counter-current multiplier* with a high multiplication capacity.

The NaCl is reabsorbed repeatedly in the thick ascending limb of the Henle loop. The passive counter-current exchange in the vasa recta and the active counter-current NaCl reabsorption in the thick ascending limb combine into a *solute-water separator*, when ADH is present.

Another component in the maintenance of the medullary hyperosmotic gradient is addition of *urea* to the tubular fluid in the thin segment of the Henle loop. Urea is then trapped in the lumen, because all nephron segments, from the thick ascending limb through the outer medullary collecting duct, are impermeable to urea.

As the tubular fluid flows through the distal tubules, cortical collecting ducts and outer medullary collecting ducts, its urea concentration rises progressively, because these segments are essentially urea-impermeable whether or not ADH is present. In the presence of ADH, water is reabsorbed but urea is not and the osmolarity of the fluid increases. The maximal osmolarity in the *cortical* collecting duct is up to 300 $mOsmol*l^{-1}$, which is equal to the surrounding interstitial fluid.

The distal fluid contains much urea and less NaCl. In reverse, the inner medullary collecting duct cells have urea-transporters that are ADH-sensitive. Thus large amounts of urea are reabsorbed at low urine flows, and the inner medullary interstitial fluid is loaded with urea that diffuses back to the tubular fluid through the thin descending and ascending limb in this urea recycling process. Urea covers 700 and NaCl also 700 $mOsmol*l^{-1}$ out of the total 1400.

Without passive urea recycling, the medullary interstitial osmolarity contributed by NaCl would have to double and thus the energy demand. Without the medullary hypertonic gradient we would be unable to produce concentrated urine when water depleted.

A high osmolarity in the medullary interstitium enhances passive water reabsorption when ADH is present. ADH increases the concentration of solutes in the collecting ducts, and reduces the loss of water. A hyperosmotic concentration – moving from 300 up to *1400* mOsmol*l^{-1} in the inner medulla - has established a large concentration gradient between the tubular and the interstitial fluid.

In man, the maximal urine osmolarity – when ADH is high - is 1400 mOsmol*l^{-1}, which in a daily urine volume of *500 ml* corresponds to a daily solute loss of up to 700 mOsmol. The small urine volume contains high concentrations of urea and nonreabsorbed or secreted solutes.

Dilution of urine (large urine flow)

In the absence of ADH, the distal tubules, cortical collecting ducts and outer medullary collecting ducts are impermeable to water. The osmolarity of the passing tubular fluid is reduced (towards 100 mOsmol*l^{-1}) when we need a diluted urine. The medullary collecting duct reabsorbs NaCl (actively) and is slightly permeable to water and urea in the absence of ADH. The final urine – with small concentrations of NaCl and urea - has an osmolarity of 50-150 mOsmol*l^{-1}, with a volume of up to 10% of the daily GFR.

When ADH is absent, the fluid leaving the distal tubules remains hypotonic. *Large amounts of hypotonic urine* would then flow into the renal pelvis (with an osmolarity down towards 50 mOsmol*$^{l-1}$). A daily solute loss of 700 mOsmol, under these conditions, implies a daily water loss of at least *14 l*.

8. Renal bloodflow (RBF)

The Fick's principle (mass balance principle) is used to measure the renal plasma clearance at low plasma [PAH] , since at low concentrations - the blood is almost cleared by one transit. Thus, the renal plasma clearance for PAH is almost equal to the renal plasma flow RPF in Eq. 25-5. The law of mass balance states that the infusion rate of PAH is equal to its excretion rate at steady state.

Only one passage through the kidneys *effectively* eliminates PAH from the venous blood plasma at low [PAH]. A methodological short cut is to measure the [PAH] in the medial cubital vein only, instead of the true arterial [PAH] by arterial catheterisation. PAH clearance is an acceptable approximation called the *effective* renal plasma flow (ERPF). In a healthy, resting person the ERPF is 600-700 ml of plasma per min and lower than the RPF. The ERPF principle avoids complex invasive procedures such as catheterisations.

The T_{max} for PAH is also a valuable measure of the secreting tubular mass, because the proximal tubule cells are saturated with PAH at high plasma-[PAH].

The RBF falls drastically, when the mean arterial pressure is below 9.3 kPa (70 mmHg). The *medullary* bloodflow is always small in both absolute and relative terms. Any severe RBF reduction as in shock, easily leads to ischaemic damage of the medullary tissues resulting in papillary necrosis and ultimately to failure of renal function.

During such pathophysiological conditions, prostaglandins (PGE_2 and PGI_2) are secreted from

the mesangial and endothelial cells due to sympathetic stimulation. These prostaglandins dilatate the afferent and efferent glomerular arterioles and dampen the renal ischaemia caused by sympatho-adrenergic vasoconstriction.

Both RBF and GFR show *autoregulation* following acute changes in the perfusion pressure within the physiological pressure range (Fig. 25-15). The renal autoregulation is mediated by myogenic feedback and by the macula densa-tubulo-glomerular feedback mechanism. *Myogenic feedback* is an intrinsic property of the smooth muscle cells of the afferent and efferent arterioles. The myogenic response allows preglomerular arterioles to sense changes in vessel wall tension (T) and respond with appropriate adjustments in arteriolar tone. Stretching of the cells by a rise in arterial transmural pressure (DP) elicits smooth muscle contraction in interlobular arteries and afferent arterioles (Fig. 25-15). During sleep the mean arterial pressure decreases 1-2 kPa, which would lower P_{gc} and GFR without autoregulation.

Autoregulation with maintained RBF and GFR means that also the filtered load and the sodium excretion is maintained during sleep and variations in daily activities. The *macula densa- TGF mechanism* is described below.

When the renal perfusion pressure rises, the cortical bloodflow is effectively autoregulated. However, during certain circumstances the papillary bloodflow may increase due to release of NO, prostaglandins, kinins or other factors. The increased medullary bloodflow increases the interstitial hydrostatic pressure and thus the resistance towards Na^+-reabsorption, whereby the Na^+-excretion increases.

Sympathetic vasoconstriction reduces the renal perfusion pressure and thus the resting RBF.

Increased renal sympathetic tone releases renin and enhances Na^+-reabsorption in the proximal and distal tubules via nerve fibres. At maximum exercise RBF falls to half the resting level. - RBF also drops during emotional stress and during haemorrhage.

Fig. 25-15: Pressure-flow relations in the kidney. The RBF curve shows autoregulation, and GFR follows the bloodflow.

Noradrenaline/dopamine from adrenergic fibres and circulating adrenaline from the adrenal medulla, constrict the afferent and efferent glomerular arterioles, when the hormones are bound to a_1-adrenergic receptors. This constriction decreases both RBF and GFR. Sympathetic stimulation releases renin from the granular JG-cells of the arterioles via b_1-adrenergic receptors. Activation of the adrenergic fibres enhanches the Na^+-reabsorption along the whole nephron.

The normal 300-g's of kidney tissue receive a total bloodflow (RBF) of 1200 ml per min, which is 20-25% of the cardiac output at rest. Thus, on an average, RBF is 400 ml of blood per min and per 100-g kidney tissue. These units are actually called *Flow Units* (FU) or perfusion coefficients. The renal blood flow per weight unit is higher than any other major organ in the body. The renal cortex receives 90% of the total RBF, and only 5-10% reaches the outer medulla. The blood supply is at a minimum in the inner medulla, and the oxygen tensions falls off sharply in the papillary tissue. The medullary bloodflow can be reduced towards 1% by vasopressin.

The *counter current exchange* of oxygen in vasa recta is a disadvantage to the renal papillae because their cells are last fed with oxygen by the blood. The inner cells meet their energy requirements primarily by anaerobic breakdown of glucose by glycolysis. The amount of energy obtained here is only 1/10 of the oxidative breakdown of 1 mol of glucose (2 888 kJ

free energy).

The *cortical bloodflow* is much larger than the medullary bloodflow. Here, 1/5 of the whole plasma stream passes the glomerular barrier by ultrafiltration and becomes *preurine*. Fortunately, we obtain the greater part of the energy required for cortical tubular transport by oxidative metabolism.

9. Macula densa-tubulo-glomerular feed-back (TGF)

The macula densa-TGF mechanism responds to disturbances in distal tubular fluid flow passing the macula densa.

The JG-apparatus includes 1) the renin-producing granular cells of the afferent and efferent arterioles, 2) the macula densa of the thick ascending limb, and 3) the extraglomerular mesangial cells connecting the afferent and the efferent arteriole (Fig. 25-16).

Renin is described in paragraph 6 of Chapter 24.

Fig. 25-16: The juxtaglomerular apparatus with renin secretion.

Regulation of renal sodium excretion is described in paragraph 9 of Chapter 24.

The TGF mechanism thus includes the renin-angiotensin II-aldosterone cascade (Fig. 24-5). Prostaglandins, adenosine and NO can modulate the response. These renin responses are part of the autoregulation to maintain RBF and GFR normal.

10. Non-ionic diffusion

Non-ionic diffusion is a passive tubular reabsorption of weak organic acids and bases, which are lipid-soluble in the undissociated or non-ionised state. In this state these compounds penetrate the lipid membrane of the tubule cell by diffusion. The tubule cells, however, are practically impermeable to the dissociated form of these compounds. Therefore, the *ionic* form of the weak acid or base is fixed in the tubular fluid and *favoured* for urinary excretion.

A weak organic acid is mainly undissociated at low urinary pH, whereas an organic base is more dissociated. In acid urine the reabsorption rate of weak organic acids is *increased*, whereas the reabsorption rate of weak organic bases is *reduced*. In alkaline urine the opposite situation prevails.

Examples of weak acids showing this phenomenon are phenobarbital and procain (both with pK just below 7), NH_4^+, acetylsalicylic acid, and many other therapeutics. Weak bases are the doping substance, amphetamine, and many therapeutics.

In rare cases of poisoning with weak bases, the patients are treated with infusions of ammonium chloride solutions or amino acid-HCl solutions, which acidifies the urine (see Chapter 17). In cases of poisoning with weak acids, some patients receive infusions of bicarbonate solutions, whereby alkalisation of the urine is instituted.

11. Tests for proximal and distal tubular function

Several proximal tests are available.

1. About 30 g of plasma albumin passes through the glomerular barrier each day. Fortunately, most of this albumin is absorbed through the brush border of the proximal tubules by pinocytosis. Inside the cell the protein molecule is digested into amino acids, which are then absorbed by *facilitated diffusion* through the basolateral membrane. Proteins derived from proximal tubule cells, such as *$ß_2$-microglobulin*, are reabsorbed by the proximal tubules. If this protein is demonstrated by urine electrophoresis, a proximal reabsorption

defect is present. This is also the case, when *generalized aminoaciduria* is present.

2. *Glucosuria in the absence of hyperglycaemia* indicates a proximal reabsorption defect of glucose, since all glucose is reabsorbed before the fluid reaches the end of the proximal tubules in the normal state.

3. *The lithium clearance*. The lithium ion, Li^+, is filtered freely across the glomerular barrier, and its concentration in the ultrafiltrate is equal to that in plasma water. Lithium carbonate is used in the treatment of manic phases (catecholamine over-reaction) of *manic depressive psychosis*. A plasma concentration of 0.5-1 mM provides enough Li^+ to block membrane receptors on the neurons involved for catecholamine binding.

Fig. 25-17: Lithium clearance used as a measure of the proximal reabsorption capacity in the nephron.

Li^+ is reabsorbed isosmotically in the proximal tubules together with water and Na^+ (Fig. 25-17). The amount of Li^+ that leaves the proximal tubules (pars recta) is equal to its excretion rate in the final urine. This is because there is practically no reabsorption or secretion of Li^+ distal to this location. Accordingly, a large lithium clearance depicts a low proximal lithium reabsorption, and thus a poor proximal tubular function at a given GFR. Normally, the passage fraction of Li^+ is 0.25-0.3 at the end of the proximal tubules and almost the same fraction passes into the urine.

4 *Hypokalaemia* combined with *normal or increased renal K^+-excretion* suggests a defective proximal K^+-reabsorption (see Chapter 24 or Box 25-1).

5 Secretion across the proximal tubules (PAH clearance).

Tests of distal tubular function**:**

1. Renal concentrating capacity is easily estimated as osmolalities in morning plasma and urine. Normal plasma osmolality ranges over 275-290 mOsmol per kg, and a urine osmolality above 600 mOsmol per kg suggests an acceptable renal concentrating capacity (more accurate is a standardized water deprivation test).
2. Inability to lower urine pH below 5.3 despite a metabolic acidosis is indicative of distal renal tubular acidosis (ie, a bicarbonate reabsorption defect). This is a rare inherited condition with failure of bicarbonate reabsorption in the distal tubules and the collecting ducts. The metabolic acidosis is instituted by the oral intake of 100 mg ammonium chloride per kg and confirmed by a pHa less than 7.35 with a negative base excess and [bicarbonate] below 21 mM.
3. NaCl reabsorption in the early part of the distal tubule dilutes the tubular fluid, because this segment is impermeable to water (Fig. 25-11). Thiazide diuretics inhibit the Na+-Cl- symporter protein that causes a measurable increase in NaCl excretion and in diuresis (Fig. 25-11).

12. Stix testing with dipstics

Routine stix testing for blood, glucose, protein etc. is necessary for the clinical evaluation of renal patients. *Reagent strips* for red blood cells are extremely sensitive. Even a trivial bleeding from a small capillary results in a positive answer indicating the presence of a few red cells. In such cases microscopy is necessary. Microscopy of fresh urine reveals red cells in cases of bleeding from the urinary tract, and red-cell casts in cases of kidney bleeding as in glomerulonephritis.

Since the concentration threshold in urine for most reagent strips is *150 mg* albumin per litre

(l), there is no reaction to the normal albumin concentration of 20 mg l^{-1}. Even 50-100 mg of protein is often excreted daily due to the upright posture and exercise.

An early sign of diabetic glomerular leakage or nephropathy is microalbuminuria, which is defined as an albumin concentration of 50-150 mg per l of urine, and measured by radioimmunoassay (RIA).

Some laboratories measure the *Tamm-Horsefall glycoprotein*, which is secreted from the cells of the thick ascending limb of Henle, and thus a normal constituent of urine.

Bacteria in the urine produce nitrite from the urinary nitrate, and dipsticks easily demonstrate the nitrite. Urinary tract infection also results in white blood cells in the urine, and more than 10 cells per µl are abnormal.

13. Diuretics

Diuretics are therapeutic agents that increase the production of urine. Diuretics are employed to enhance the excretion of salt and water in cases of cardiac oedema or arterial hypertension. The so-called *natriuretics* inhibit tubular Na^+-reabsorption, but since the secretion of K^+ and H^+ is also increased, the patient must have compensatory treatment. The sites of action for different groups of diuretics are shown in Fig. 25-18.

13 a. *Carboanhydrase inhibitors* (eg, acetazolamide) act on the carboanhydrase (CA) in the brush borders and inside the cells of the proximal tubules. Inhibition of the metallo-enzyme reduces the conversion of filtered bicarbonate to carbon dioxide. As a result, there is a high concentration of bicarbonate and sodium in the tubular fluid of the proximal tubules. Up to half of the bicarbonate normally reabsorbed is eliminated in the urine causing a high urine flow and a metabolic acidosis.

Thus, these inhibitors are diuretics. They are mainly used in the treatment of open-angle glaucoma (ie, an intraocular pressure above 22 mmHg). Acetazolamide promotes the outflow of the aqueous humour and probably diminishes its isosmotic secretion.

Fig. 25-18: Sites of action on the nephron of different groups of diuretics

13 b. Loop diuretics (bumetanide and furosemide) inhibit primarily the reabsorption of NaCl in the *thick ascending limb* of Henle by blocking the luminal Na^+-K^+-$2Cl^-$-symporter. The reabsorption of NaCl, K^+ and divalent cations is reduced, and also the medullary hypertonicity is decreased. Hereby, the distal system receives a much higher rate of NaCl, water in isotonic fluid, and K^+. The overall result is an increased excretion of NaCl, water, K^+ and divalent cations. The patient's plasma- [K^+] should be checked regularly.

13 c. *Thiazide diuretics* (bendroflurazide, hydrochlorothiazide) act on the early part of the distal tubule by inhibiting the (Na^+- Cl^-)-symporter. They increase K^+ excretion by increased tubular flow rate. *Thiazide* and many other diuretics are secreted in the proximal tubules. This secretion inhibits the secretion of uric acid, so thiazide is contraindicated by *gout*.

13 d. *Potassium-sparing diuretics* (eg, amiloride) inhibit Na^+-reabsorption by inhibition of sensitive Na^+-channels in the principal cells of the distal tubules and collecting ducts. Hereby, they reduce the negative charge in the lumen and thus the K^+-secretion. Amiloride causes natriuresis and reduces urinary H^+- and K^+-excretion

13 e. *Aldosterone-antagonists* (eg, spironolactone) compete with aldosterone for receptor sites on principal cells. As aldosterone promotes Na^+-reabsorption and H^+/ K^+ -secretion, aldosterone-antagonists cause a natriuresis and reduce urinary H^+ - and K^+ -excretion. Aldosterone-antagonists are weak potassium-sparing diuretics, mainly used to reduce K^+ -excretion caused by thiazide or loop diuretics.

13 f. *Angiotensin-converting-enzyme (ACE)-inhibitors* (captopril, enapril and lisinopril) reversibly inhibit the production of angiotensin II, reduce systemic blood pressure, renal vascular resistance and K^+ -secretion. ACE-inhibitors promote NaCl and water excretion. ACE-inhibitors increase RBF without much increase in GFR, because of a decrease in both afferent and efferent arteriolar resistance. The development of diabetic nephropathy can be markedly delayed by early reduction of blood pressure with ACE-inhibitors and by careful diabetic management.

13 g. *Osmotically* active *diuretics* are substances such as mannitol and dextrans. These substances retard the normal passive reabsorption of water in the proximal tubules. Osmotic therapy with mannitol is used in the treatment of cerebral oedema.

Mannitol is a hexahydric alcohol related to mannose and an isomer of sorbitol. Mannitol passes freely through the glomerular barrier and has hardly any reabsorption in the renal tubules. Its presence in the tubular fluid increases flow according to the concentration of osmotically active particles, which inhibit reabsorption of water. The high flow of tubular fluid means that the excretion of Na^+ is great - despite the rather low Na^+ concentration. Mannitol may help to flush out tubular debris in shock with acute renal failure, but the results are controversial.

Dextrans (ie, polysaccharides) have a powerful osmotic and diuretic effect. - The larger, molecules (macrodex) are seldom used as volume expanders during shock because of allergic reactions.

Pathophysiology

This paragraph deals with *1. Glomerulonephritis*, *2. Renal insufficiency*, *3. Acute tubular necrosis*, *4. Diabetic nephropathy*, *5. Nephrotic syndrome*, *6. Urinary tract infection*, *7. Tubulo-interstitial nephritis*, *8. Gouty nephropathy*, *9. Renal hypertension*, *10. Urinary tract obstruction*, *and 11. Tumours of the kidney*.

The severity and cause of kidney disease is evaluated by measurement of the GFR.

1. Glomerulonephritis

Glomerulonephritis is an immunologically mediated injury of the glomeruli of both kidneys.

The majority of patients suffer from postinfectious glomerulonephritis or immune complex nephritis. This is a disorder, where circulating antigen-antibody complexes are deposited in the glomeruli or free antigen is bound to antibodies trapped in the capillary network. Typically, the antigen is derived from Lancefield group Aß- haemolytic streptococci, but also other bacteria, viruses, parasites (malaria), and drugs may be the origin. A few patients produce antibodies against their own antigens (eg, host DNA in systemic lupus erythematosus, malignant tumour antigen, or anti-glomerular basement antibody, anti-GBM).

The inflammation is an abnormal immune reaction often caused by repeated streptococcal tonsillitis. An insoluble antigen-antibody complex precipitates in the basement membrane of the glomerular capillaries. The cells of the glomeruli proliferate, and disease will of course

reduce GFR and to some extent, the RBF (measured as PAH clearance). Thus the infection depresses the glomerular filtration fraction (GFF = GFR/RPF).

The acute postinfectious glomerulonephritis occurs typically in a child, who has suffered from streptococcal tonsillitis a few weeks before.

Haematuria, proteinuria, and oliguria characterise acute nephritis with salt-water retention causing oedemas and hypertension. Pulmonary oedema and hypertensive encephalopathy with fits is life threatening.

Uraemia is a clinical syndrome dominated by retention of non-protein nitrogen (eg, urea, uric acid, NH_4^+ creatinine and creatine). Uraemic patients generally exhibit hyperkalaemia (plasma- $[K^+]$ above 5.5 mM) and metabolic acidosis (pH below 7.35 and a negative base excess). This is due to the inadequate secretion of K^+, NH_4^+ and H^+. In complete renal shutdown, the patient dies within 1-2 weeks without dialysis.

Dialysis is mandatory with severe uraemia. When serum creatinine rises above 0.7 mM, renal insufficiency is usually terminal (Fig. 25-4).

Recording of blood pressure and fluid balance with weighing is important in order to prevent hypertension and pulmonary oedema to develop into a life-threatening condition.

Fig. 25-19: Post-streptococcal glomerulonephritis.

The parietal and visceral epithelial cells of the glomeruli grow and proliferate, just as the mesangial cells (Fig. 25-19). This proliferation and the damage of the basement membrane with accumulation of insoluble complexes all impair the glomerular barrier and reduce the glomerular filtration rate (GFR). Production of cytokines and autocoids enhance the inflammation. Capillary injuries with reduction of the lumen also reduce the renal bloodflow (RBF) to some extent (Fig. 25-19).

Children with poststreptococcal glomerulonephritis are treated with a course of penicillin - often with an excellent prognosis.

Glomerulonephritis as a part of systemic lupus erythematosus (SLE) is frequent in female lupus patients - in particular during pregnancy, where hypertension may precipitate glomerular injuries. Oestrogens accelerate progression of SLE, and there is a genetic predisposition. In SLE there is hyperactivity of the B-cell system, which may involve any organ, but typically affects the kidneys, joints, serosal membranes and the skin (Chapter 32). The B-cell system releases many antibodies to host antigens both in and outside the cell nuclei (single- and double-stranded DNA, RNA, plasma proteins, cell surface antigens, and nucleoproteins). Lymphocytotoxic antibodies are also liberated, which may explain the inhibition of the T-cell system. The most important autoantibodies are those against nuclear antigens. Accumulation of immune complexes with double-stranded DNA probably causes the glomerular lesions as well as vasculitis and synovitis.

Fig. 25-20: Anti-GBM glomerulonephritis with anti-GBM of the IgG type. Complement is shown as a small circle.

Anti-GBM glomerulonephritis is a seldom disorder, where the patient produces antibodies (IgG type) against his own basement membrane. The antibody is known as anti-GBM or anti-Glomerular Basement Membrane antibody. The antigen is localised both in the glomerular basement membrane and in the basement membrane of the alveolar capillaries. The histological picture is characterized by proliferation of both parietal epithelial cells, and

mesangial cells (Fig. 25-20).

The capillary basement membrane is disrupted, and there is red cells and fibrin in Bowmans space. The diagnosis is confirmed by identification of circulating anti-GBM (Y-shape in Fig. 25-20). Glomerulonephritis with pulmonary haemorrhage is termed Goodpastures syndrome. The recurrent haemoptyses can be life threatening.

2. Renal Insufficiency

Renal insufficiency is a clinical condition, where the glomerular filtration rate is inadequate to clear the blood of nitrogenous substances classified as non-protein nitrogen (urea, uric acid, creatinine, and creatine). The retention of nonprotein nitrogen in the plasma water is called azotemia, and the clinical syndrome is called uraemia. The number of filtrating nephrons falls below 1/3 of normal, as determined by measurement of a GFR below 40 ml/min.

Acute renal insufficiency accompanies extremely severe states of circulatory shock (prerenal cause). The prerenal causes are hypovolaemia with hypotension or impaired cardiac pump function or the combination.

Also a large group of renal causes to failure occurs (Box 25-2). Finally, the postrenal causes are all types of urinary tract obstruction.

Acute renal failure is a serious disorder, which leads to progressive uraemia and chronic renal insufficiency.

Box 25-2. Causes of renal failure	
Prerenal Causes: Cardiogenic and hypovolaemic shock	
Renal Causes: ACE-inhibitors and NSAID´s impair renal autoregulation	
	Fulminant hypertension.
	Renal artery stenosis and embolism
	Vasculitis in glomerular capillaries
	Renal vein thrombosis
	Toxic tubular damage (organic solvents, myoglobin, aminoglycosides, and X-ray contrast).
Postrenal Causes: Urinary tract obstruction is caused by obstructions of the lumen, the wall and by pressure from outside	
	***Lumen*:** Tumours, calculus and blood clots within the lumen of the renal pelvis, ureter, and bladder
	***Wall*:** Strictures of the ureter, the ureterovesical region, urethra, and pinhole meatus.
	Congenital disorders such as megaureter, bladder neck obstruction, and urethral valve.
	Neuromuscular dysfunction in the urinary tract
	Pressure: Compression by tumours, aortic aneurysm, retroperitoneal fibrosis or gland enlargement, retrocaval ureter, prostate hypertrophy, phimosis, and diverticulitis.

Two complications to *chronic* renal failure must be considered:

1. Renal osteodystrophy develops in patients with severe renal failure. The kidneys fail in producing sufficient 1,25-dihydroxy-cholecalciferol. This is active vitamin D or a potent steroid hormone. The active vitamin D metabolite stimulates the Ca^{2+}-transport across the cell and mitochondrial membranes.

   Lack of active vitamin D has the following two effects:

a. Poor gut absorption of dietary Ca^{2+}, so that plasma [Ca^{2+}] falls.

b.The PTH release is stimulated, because the normal inhibitory effect of active vitamin D is lost.

   After some time a secondary hyperparathyroidism develops with increased resorption of calcium from bone and increased proximal tubular reabsorption of calcium in an attempt to correct the low serum calcium. The calcium release from bone results in osteomalacia and in osteoporosis. Osteomalacia or soft bones is the result of demineralisation of the osteoid matrix usually caused by insufficient active vitamin D. Osteoporosis or thin bones is characterized by a reduction in all components of the bones.

2. Normochromic, normocytic anaemia. When normal kidneys are perfused with hypoxaemic blood, the peritubular interstitial cells produce large amounts of the glycoprotein hormone, erythropoietin, with strong effect on erythrogenesis.

   Chronic renal failure leads to erythropoietin deficiency, and thus to anaemia, which is of the normochromic, normocytic type.

Haemodialysis

The aim of haemodialysis is to eliminate nitrogenous wastes in patients with renal failure, and maintain normal electrolyte concentrations, serum glucose and normal ECV. In other words, the haemodialyzer or artificial kidney mimics the normal renal excretion of waste products (Fig. 25-21)

Fig. 25-21: An artificial kidney (dialyser) with an area of 1 m^2 and a membrane thickness of 10 µm.

Blood from the patient is pumped through a container with series of semi-permeable membranes separating the blood from dialysate (Fig. 25-21).

Dialysate is a mixture of purified water with salts, and glucose in a composition comparable to normal fasting plasma apart from proteins. Bicarbonate or acetate buffer is present at a concentration about 35 mM.

Haemodialysis is performed with a bloodflow of 200-300 ml per min. The patient is often connected to the dialyzer by an arteriovenous shunt made by plastic cannulae between the radial artery and an adjacent vein. The arterial blood flows into the artificial kidney and after dialysis the blood is returned to the venous system (Fig. 25-21). Dialysate is pumped through the container at a rate of 500 ml each min.

A plastic shunt connects the two cannulae on the forearm between dialysis sessions, and the large arterial bloodflow is sufficient to avoid coagulation in the plast shunt. Also dual-lumen venous catheters placed centrally are in use.

If the sodium concentration of the dialysate is too high, the patient complains of thirst and the arterial pressure starts to rise. Low dialysate calcium may result eventually in secondary hyperparathyroidism, whereas a high dialysate calcium concentration causes hypercalcaemia.

An adult patient with acute renal failure (so-called shock kidney) requires 4 -5 hours dialysis 3 times a week.

Renal Transplantation

Fit patients with chronic renal failure are offered renal transplantation. Rejection of the transplant is due to complement-fixing antibodies in the blood, or later caused by cellular or humoral immunity. Rejection years after the transplantation is frequently caused by ischaemic damages of the kidney. Donation of a kidney leaves the donor with one kidney only.

Immediately after the removal, the GFR of the patient falls to half its original value, because half the functioning nephrons have been removed.

Soon, most individuals will increase their GFR towards normal values by compensatory work hypertrophia of the remaining kidney. The hypertrophia-factor is not known. Each remaining nephron must filter and excrete more osmotically active particles than before.

3. Acute Tubular Necrosis

This disorder has haemodynamic or toxic causes.

Cardiogenic and hypovolaemic shock cause acute renal failures just as renal vasoconstriction. Renal ischaemia leads to hypoxic damage, in particular damage of the renal medulla, which is especially susceptible to ischaemia, because of the normally relatively poor oxygenation. Ischaemic tubular damage also reduces the GFR further, because of reflex spasms of the afferent arterioles, and due to tubular blockage with accumulation of filtrate in the early part of the proximal tubules, and hypoxic damage of the proximal tubular reabsorption capacity.

Loss of appetite and energy, nausea and vomiting, nocturia and polyuria characterise the condition. Only when the GFR is severely depressed there is oliguria. Even a GFR of only 1 ml each min, as a contrast to the normal 125 ml per min, may result in a daily urine flow of 1440 ml (1*1440 min daily), if there is a total loss of tubular reabsorption and no luminal obstruction. This urine flow is normal, but unfortunately based on an almost total loss of glomerular and tubular function. Sufficient regeneration of the tubular epithelium allows clinical recovery.

Sometimes also the renal cortex is necrotic, and following healing of the injuries, the result is scarring with glomerulosclerosis. This condition is also found following radiation nephritis.

4. Diabetic nephropathy

Diabetic nephropathy includes *glomerulosclerosis*, with thickening of the basement membrane and damage of the glomerular filter by disruption of the protein cross-linkages and glomerular hyperfiltration. Excess NO production reduces the afferent arteriolar resistance and increases the glomerular capillary pressure. The earliest evidence of glomerular damage may occur 5-15 years following diagnosis in the form of *microalbuminuria*. The patient later develops intermittent albuminuria followed by persistent albuminuria. *Diabetic nephropathy* includes hypertension, persistent albuminuria, and a decline in GFR. One third of all insulin-dependent diabetics develop nephropathy. The mortality rate is high. The metabolic disturbance in diabetics causes hypertension and leaky renal glomeruli, but the mechanism remains uncertain.

Ascending infections result in interstitial lesions and diabetes typically show hypertrophy and hyalinization of afferent and efferent arterioles. Obstruction of the renal bloodflow (ischaemia) leads to hypoxic damage of the renal tissue. The tenuous bloodflow to the renal papillae via the vasa recta explains why renal papillary necrosis is frequent in diabetics.

Treatment with ACE- inhibitors reduce urinary albumin excretion. Prophylactic therapy also postpones the development of diabetic nephropathy and hypertension with persistent microalbuminuria. The effectiveness of this treatment suggests that relative oversecretion of angiotensin may be involved in the pathogenesis of diabetic nephropathy.

5. Nephrotic syndrome

The nephrotic syndrome refers to a serious increase in the permeability of the glomerular barrier to albumin, resulting in a marked loss of albumin in the urine. The albuminuria (more than 3 g per day) causes hypoalbuminaemia and generalized oedema.

The number and size of pores in the glomerular barrier increase due to disruption of protein-linkages. Negatively charged glycoproteins in the glomerular barrier repel negatively charged proteins. The amount of negatively charged glycoproteins is reduced in glomerular disease.

Oedema is visible in the face - especially around the eyes.

A serious but rare complication may develop when a large volume of fluid accumulates in the abdominal cavity as ascites.

6. Urinary Tract Infection

Urination (micturition) is controlled by the micturition reflex. Stretch or contraction of the smooth muscles in the bladder wall is sensed by mechanoreceptors and signalled via the pelvic nerve to the sacral spinal cord. Increased parasympathetic tone (via pelvic nerves and muscarinic receptors) cause sustained bladder contraction. Normally, contraction of the bladder muscles by micturition almost completely empties the bladder.

Recurrent infections of the urinary tract are frequent among females. Faecal bacteria are transferred to the periurethral region, and finally to the bladder via the short female urethra. Bladder urine is normally sterile owing to bladder mucosal factors and other local defence mechanisms. Bacteria adhere to the bladder epithelium and multiplicate, when defence mechanisms function insufficiently. Prolonged bladder catheterisation predisposes to bladder infection, and even a few days can be critical.

The diagnosis bladder infection is based on more than 100 000 bacteria per ml of clean-catch mid-stream urine. Quite a few patients with significant bacteriuria do not develop nitrite enough to be shown by dipstick tests.

Typical symptoms are frequent micturition (polyuria), painful voiding (dysuria), suprapubic pain and smelly urine perhaps with haematuria.

Echerichia coli and other coliform bacteria cause the majority of urinary tract infections; these infections are treated successfully with antibiotics (amoxyllin, trimethoprim etc) either as a single shot or for longer periods.

7. Tubulo-Interstitial Nephritis

Bacterial pyelonephritis typically causes interstitial inflammation of the kidneys, but the interstitial inflammation is more often caused by a hypersensitivity reaction to drugs (antibiotics, phenacetin and non-steroid anti-inflammatory drugs, NSAIDs).

Pyelonephritis begins in the renal pelvis, and then progresses into the renal medullary tissue.

The essential function of the medulla is to concentrate the urine during water depletion. Therefore, in patients with pyelonephritis, the ability to concentrate the urine is abolished/decreased (isosthenuria/hyposthenuria). The ability to dilute the urine deteriorates

also. Thus, in isosthenuria the urine is always isotonic with the plasma.

The patient with acute nephritis has fever, skin rashes and acute renal failure with eosinophiluria and eosinophilia. First of all the offending drug must be withdrawn, and the renal failure may require dialysis.

Chronic tubulo-interstitial nephritis is caused by pyelonephritis, NSAIDs, diabetes mellitus, hyperuricaemia, irradiation damage etc. The major problem is that long lasting consumption of large amounts of analgesics leads to terminal renal failure. Nephrotoxic analgesics must be abandoned.

The patient presents with uraemia, albuminuria, polyuria, haematuria, anaemia, and most often a history of analgesic abuse. Papillary necrosis can be present with papillary tissue passed in the urine or obstructing the ureter or urethra. In patients with tubular damage of the renal medulla, the ability to concentrate the urine is abolished together with the ability to dilute the urine. Thus, the urine is always isotonic with the plasma (isosthenuria). The result is polyuria and salt wasting. As the inflammation progresses to the cortex also the glomerular filtration deteriorates with accumulation of non-protein nitrogen in the plasma water (azotaemia), and the clinical syndrome uraemia.

An isolated damage of the Na^+ -reabsorption (salt-losing nephritis) is a condition in which the disease processes are mainly due to dysfunction in the renal medulla. There is a marked loss of Na^+ in the urine and seriously low ECV and blood volume (hypovolaemia with threat of imminent shock). Thus the patient must have a high salt intake to prevent shock and keep alive.

8. Gouty Nephropathy

Acute hyperuraemic nephropathy occurs in patients, where the condition leads to rapid destruction of cell nuclei (at the start of treatment for malignant disorders or obesity). Large quantities of nucleoproteins are released, and the production of uric acid is increased. The urate concentration increases in the extracellular volume (ECV). Above a critical concentration of 420 mM, the urate precipitates in the form of uric acid crystals, provided the fluid is acid. This concentration threshold defines hyperuricaemia.

Precipitation in the joints with pain is termed gout (arthritis urica), and precipitation of uric acid crystals also occurs in the tubules, the collecting ducts and the urinary tract. Normally, urate ions are actively reabsorbed in the proximal tubules by a Na^+-cotransport. Urate ions can also be actively secreted from the blood to the tubular fluid.

Allopurinol is prescribed during radiotherapy or cytotoxic therapy. Acute cases are also treated with allopurinol and forced alkaline diuresis.

Uric acid stones are found in patients with hyperuricaemia, and in patients secreting sufficient urate without hyperuricaemia. Calcium stones may be formed around a nucleus of uric acid crystals.

9. Renal Hypertension

Bilateral renal disease such as *chronic glomerulonephritis* is a frequent cause of hypertension (Chapter 12), whereas unilateral renal disease, such as renal artery stenosis, is a fairly seldom cause of hypertension. Stenosis (narrowing of the lumen) of one renal artery leads to renal hypotension with excess renin production (see below) and systemic (secondary) hypertension.

Exposure to fluid loss, reduced glomerular propulsion pressure, and increased sympathetic activity releases renin from the juxtaglomerular cells in the afferent glomerular arteriole, so the renin-angiotensin-aldosterone cascade is triggered (Fig. 24-5).

Angiotensin II stimulates the aldosterone liberation from zona glomerulosa of the adrenal cortex, and thus stimulates Na^+ -reabsorption and K^+ -secretion in the distal tubules. The result is salt and water retention with increase in blood volume and blood pressure. Angiotensin II also constricts arterioles, with an especially strong effect on the efferent renal arteriole. This reduces the renal bloodflow further and also the proximal reabsorption. The development of hypertension in high renin states is mainly due to salt-retention and systemic vasoconstriction.

Stenosis of one renal artery does not always lead to increased erythrogenesis. Stenosis of the renal artery implies a small renal bloodflow, a small glomerular filtration and a small NaCl-reabsorption with a related small oxygen consumption on the affected side. As long as the renal oxygenation is sufficient, the erythropoietin production is normal.

Severe renal artery stenosis implies renal ischaemia and hypoxia, which is probably always consequential with complications. A hypoxic kidney has a low creatinine and PAH clearance.

A long-term increase in sodium intake results in changes of the kidney function. Surprisingly, the changes are similar in hypertensive and normotensive humans! Most people increase their ECV and GFR without changing the absolute reabsorption rate of Na^+ and water in the proximal tubules. Therefore, the rise in filtration rate of Na^+ and water will reach the loop of Henle and the distal tubule. The arterial blood pressure and heart rate is unaffected by the amount of sodium in the diet. The plasma concentrations of active renin (Fig. 24-7), angiotensin II and aldosterone decrease with increasing Na^+ intake, but atrial natriuretic factor (ANF) and cyclic GMP increase. Arginine vasopressin (ADH) in plasma does not change.

The reason why this increase in NaCl load to the loop of Henle is not counterbalanced by the TGF-system is due to resetting of the TGF-mechanism, so a contraction is avoided in spite of the increased salt load.These homeostatic reactions are all appropriate physiological responses in both healthy and hypertensive humans.

A rare cause of renal hypertension is due to Liddles syndrome. This is an autosomal dominant defect characterised by severe hypertension, hypokalaemia and metabolic alkalosis. The syndrome is similar to primary hyperaldosteronism, but the renin-aldosterone concentration in plasma is not increased. Liddles syndrome is caused by mutation of the gene for the amiloride-sensitive Na^+-channel (Fig. 25-11), whereby the channel is wide open. The Na^+-entry depolarises the membrane and favours secretion of K^+ and H^+.

10. Urinary Tract Obstruction

Obstruction of the urinary tract may occur at any location, and cause dilatation of the above structures. The obstruction is localised within the lumen (stone, sloughed papilla, or tumour), within the wall (neuromuscular dysfunction, stricture, congenital urethral valve, or pin hole meatus), or pressure from the outside obstruct the tract (eg, tumours, diverticulitis, aortic aneurysm, prostatic obstruction, retrocaval ureter).

Stretching of the renal calyces as they collect urine promotes their pacemaker activity and initiate a peristaltic contraction along the smooth muscle syncytium of the urinary tract.

Obstruction of the urinary tract for weeks may lead to irreversible damage of the renal function

in particular when combined with infection. Obstruction of the upper urinary tract with backpressure damage of the kidney is especially dangerous.

Kidney stone disease (nephrolithiasis) attacks only a few percent of the Western population at any time. Most stones in male patients are composed of calcium complexed with oxalate and phosphate, whereas magnesium ammonium phosphate/acetate stones are more common in females. Only a few percent of all renal stones are composed of uric acid crystals or cysteine (mainly in children). Calcium-containing and cysteine stones are radiopaque, whereas stones of pure uric acid are radiolucent.

In the presence of infection with urea-splitting bacteria, urea is hydrolysed to form the strong base ammonium hydroxide:

$$CO(NH_2)_2 + H_2O \rightarrow 2\,NH_3 + CO_2\,;\; NH_3 + H_2O \rightarrow NH_4^+ + OH^-.$$

Alkaline urine favours stone formation by crystallization in the supersaturated fluid. *Magnesium ammonium phosphate* stones are also termed *mixed* infection stones.

Obstruction or spasm of the ureter causes reflex constriction around the stone with ureteric or renal *colic pain*. The pain is an excruciating flank pain, with radiation to the iliac fossa and the genitals. The wall of the ureter is innervated with sensory nerve fibres running in the pelvic nerves. Renal colic is considered to be one of the most severe pain experience known.

Excretion urography and plain X-ray examination are important in the diagnosis of renal stone disease.

Percutaneous nephrolithotomy, pyelolithotomy or ureterolithotomy can avoid many cutting operations. Also shock-wave disintegration is in use (lithotripsy).

Nephrocalcinosis refers to diffuse renal calcification that is detectable on a plain abdominal X-ray. Patients with hypercalcaemia (eg, primary hyperparathyroidism, hypervitaminosis D, and sarcoidosis) or with hyperoxaluria precipitate calcium oxalate and calcium phosphate in the renal parenchyma. Patients with renal tubular acidosis fail to acidify their urine, which favour precipitation of calcium oxalate and phosphate.

Abdominal radiography

A plain X-ray can identify calcification at any site including the renal system.

Intravenous pyelography

An *organic iodine-containing* contrast substance is injected slowly. Serial X-rays are taken, while compression bands are applied to the abdomen in order to obstruct ureteral emptying. Hereby, the upper renal tract is distended by the excreted contrast medium. Following removal of the compression bands, the rate of excretion of contrast is studied with films before and after voiding.

11. Tumours of the Kidney

Benign and malignant tumours occur in the kidney.

Benign renal fibroma, cortical adenomas or simple cysts seldom cause symptoms and signs. Those of no clinical importance are found incidentally at autopsy. *Juxtaglomerular cell tumours* are seldom. They produce large amounts of renin, which causes hypertension.

Haemangiomas may bleed following trauma and cause fatal blood loss.

Malignant renal tumours are nephroblastoma and renal cell carcinoma.

Nephroblastoma (Wilms´ tumour) is the most frequent intraabdominal tumour in both girls and boys. It usually presents within the first three years of life. A *large abdominal mass* is found sometimes with signs of intestinal obstruction. The tumour grows rapidly and spread to the lungs. The diagnosis is confirmed with excretion urography, arteriography or scanning.

Radiotherapy and chemotherapy, combined with nephrectomy have improved the long-term survival rate.

Renal cell carcinoma (hypernephroma) accounts for more than 90% of all the malignant renal tumours in adults - in particular smokers. There is a strong association with a rare autosomal dominant inherited disease called *Von Hippel-Lindau´ syndrome* (haemangioblastomas in the cerebellum and the retina). The genetic locus is on *chromosome 3p*.The tumour arises from proximal tubular epithelium, and lies within the kidney, but the prognosis is worse, if the tumour penetrates the renal capsule. The tumour is often protruding and the neoplastic cells have an unusually clear cytoplasm.

Renal cell carcinoma is a likely source of ectopic hormone production. Increased production of erythropoietin leads to erythrocytosis and polycythaemia. Release of a parathyroid-hormone-like substance leads to hyperparathyroidism and hypercalcaemia. Release of abnormal quantities of renin triggers the renin-angiotensin-aldosterone cascade and leads to systemic hypertension.

Metastases to distant regions are frequently found in the lungs and in the bones (osteolytic metastases). Solitary tumours are treated by partial or total nephrectomy or with interferon.

Equations

- The **plasma clearance** is defined as follows:

  Eq. 25-1: Clearance = $(C_u \times V^{\circ}_u) / C_p$ [(mg/ml)×(ml/min)/(mg/ml)= ml/min].

Clearance can also be thought of as the volume of arterial plasma containing the same amount of substance as contained in the urine flow per minute.

- **Excretion fraction (EF).** EF for a substance is the fraction of its glomerular filtration flux, which passes to and is *excreted* in the urine.

  $EF = J_{excr}/J_{filtr}$

  Since $J_{excr} = (C_u \times V^{\circ}_u)$ and $J_{filtr} = (GFR \times C_{filtr})$ it follows that:

  Eq. 25-2: $EF = (C_u \times V^{\circ}_u) / (GFR \times C_{filtr})$

$\mathbf{C_{filtr}}$ is the concentration of the substance in the ultrafiltrate. The excretion fraction for inulin is one (1). Substances with an EF *above one* are subject to *net secretion*. Substances with an EF *below one* are subject to *net reabsorption.*

- **Extraction fraction** (E). E for a substance is the fraction *extracted* by glomerular filtration from the total substance delivery to the kidney via renal blood plasma.

  Eq. 25-3: $E = J_{filtr}/J_{total} = (C_a - C_{vr})/C_a$.

Substances with an E of *one* are cleared totally from the plasma during their first passage of the kidneys. Inulin has an extraction fraction of 1/5. PAH has an extraction fraction of 0.9.

- **Inulin clearance.** The flux of inulin filtered through the glomerular barrier per min is:

($GFR \times C_p/0.94$). All inulin molecules remain in the preurine and is excreted in the final urine.

Thus, the amount excreted is equal to the amount filtered:

$$GFR \times C_p/0{,}94 = (C_u \times V°_u) \text{ mmol/min}$$

Eq. 25-4: $GFR = ((C_u \times V°_u) / C_p) \times 0.94 = CLEARANCE_{inulin} \times 0.94$.

- The **Fick's principle** (mass balance principle) is used to measure the renal plasma clearance at low plasma [PAH], since at low concentrations the blood is almost cleared (90%) by one transit. Thus the renal plasma clearance is equal to the effective renal plasma flow (ERPF):

  Eq. 25-5: $ERPF = J_{excr}/Cp$; $RPF = ERPF/E_{PAH}$

- The law of mass balance states that the delivery of PAH to the kidney is equal to its excretion rate at steady state. The *Effective Renal Blood Flow* (ERBF) is calculated by the help of a total body haematocrit (normally 0.45). If ERPF is measured to be 600 ml plasma per min, we can calculate ERBF: 600/(1 - 0.45) = 1090 ml whole blood per min at rest. This is 20-25 % of cardiac output. The true RBF is 10% higher than the measured ERBF (ie, 1200 compared to 1090 ml whole blood).

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A: The B-cell system releases antibodies to host antigens.

B: The glomerular barrier facilitates the passage of negatively charged polyanionic macromolecules.

C: Thiazide diuretics may have serious side effects such as hypercholesterolaemia, hyperglycaemia (eg, glucose intolerance), hyperuricaemia, hypokalaemia, and impotence.

D: Loop diuretics inhibit the reabsorption of NaCl in the thick ascending limb of Henle – and proximal pars recta - by blocking the cotransport process in the luminal entry membrane.

E: Aldosterone antagonists, such as spironolactone, act on the aldosterone receptors on the late distal tubule cell and inhibit the K^+-excretion.

Case History A

A male office worker, 58 years of age, body weight 70 kg, suffers from insulin-dependent diabetes mellitus. The disorder is complicated with arterial hypertension, hypercholesterolaemia, albuminuria and open-angle glaucoma. The patient is in anti-hypertensive therapy with a ß-adrenergic antagonist. The open-angle glaucoma is treated with acetazolamide (a carboanhydrase-inhibitor used as a diuretic to reduce the intra-ocular pressure).

Scanning of the kidneys show a normal picture with an estimated normal kidney weight of 300 g. During renal catheterisation, a renal arteriovenous oxygen content difference is measured to 15 ml per l of blood, and the renal bloodflow is 1.2 l (normal). – The first 3 questions necessitate pharmacological knowledge.

1. Is it recommendable to treat hypertensive complications to diabetes with ß-blockers?

2. Describe the effects of carboanhydrase-inhibitor- treatment.

3. Are thiazide diuretics without risks when prescribed to diabetics?

4. Calculate the renal oxygen uptake. Calculate the renal oxygen uptake in percentage of the total oxygen uptake of 250 ml per min.

5. Calculate the kidney weight in percentage of the total body weight.

6. Is the renal bloodflow redundant compared to the renal oxygen consumption?

Case History B

A female patient (weight 57-kg) of 23 years, with an inherited defect in renal tubular function, has a lowered tubular threshold for glucose reabsorption. The patient has a blood- [glucose] of 1000 mg per litre, and just above this level glucose appears in the urine (her appearance threshold). The diuresis is 1.5 ml per min, the plasma -[creatinine] is 0.09 mM, and the urine [creatinine] is 6 mM. The normal blood-glucose level is 5-6 mM.

1. Is the above blood -[glucose] normal?

2. Calculate the creatinine clearance?

3. Calculate the glucose reabsorption at this glucose level and compare it to the normal maximal capacity: 1.78 mmol min^{-1}.

4. Is the appearance threshold defined above equal to the saturation threshold?

Case History C

A 14-year old girl has a history of previous upper respiratory tract infections, and is now treated for another sore throat (ie, tonsillitis and high fever) with ampicillin for 10 days. Two weeks later she returns to her general practitioner (GP) complaining of tender knee joints from playing handball. There is abdominal pain.

The girl is obviously ill and has a higher blood pressure than normally (145/90 mmHg or 19.3/12.7 kPa). The tonsillitis is cured and there is no fever. The upper abdomen is tender. A freshly passes urine sample is examined with a combined quantitative stick test. There is found haematuria and albuminuria (300 mg l^{-1}).

1. What is the cause of the arthritis?

2. What are the causes of the haematuria and albuminuria?

3. Does the GP admit the girl to a hospital?

Case History D

During her working hours a 24-year old nurse delivered an arterial sample for blood gas tensions. She had no symptoms or signs of disease, but doubted that an arterial sample could be taken without causing pain. The sample was taken from a radial artery with a fine needle following local anaesthesia and she experienced no pain. The arterial blood gas values were: CO_2 partial pressure 24 mmHg, O_2 partial pressure 102 mmHg, pH_a 7.36, and Base Excess -8 mM. The nurse had been starving for 24 hours.

1. What was the explanation of her acid-base disturbance?

2. What was the rational treatment?

Case History E

A young female (body weight 56 kg) with an inulin clearance of 125 ml of plasma per min is tested with para-amino-hippuric acid (PAH). The free fraction of PAH in the plasma is 0.80, and the rest binds to plasma proteins.

Her urine is collected in a period and the excretion flux of PAH is measured to 100 mg each min. The average concentration of PAH in plasma from the renal arterial and venous blood is 0.2 and 0.02 g per l, respectively. The haematocrit is 43%.

1. Calculate the clearance for PAH.

2. Calculate the tubular secretion flux for PAH at the blood plasma concentration concerned.

3. Calculate the renal blood flow (RBF).

   The patient collects the urine in a second period, where the average concentration of PAH in plasma from the arterial blood is 1 g per l. The maximal tubular secretion rate for PAH is defined as T_{max} for PAH and is 80 mg per min.

4. Calculate the excretion flux for PAH in the urine.

5. Calculate the new clearance for PAH.

Try to solve the problems before looking up the answers

Highlights

- *Creatinine clearance provides a fair clinical estimate of the renal filtration capacity.*
- *The renal control of body fluid osmolality maintains the normal cell volume (ICV) by changes of renal water excretion.*
- *Normally, we excrete 1500 (range: 1200-1800) ml of water and 2-5 g of Na^+ (= 5-12 g NaCl) daily.*
- *Renal excretion of waste products. Urea from amino acids is excreted with about 30 g or half a mol of urea per day. The daily renal excretion of uric acid, creatinine, hormone metabolites and haemoglobin derivatives matches their daily production.*
- *The daily renal excretion of metabolic intermediates and foreign molecules (drugs, toxins, chemicals, and pesticides) is carefully matched to the intake or production.*
- *Secretion of hormones: The kidney secretes erythropoietin, renin, kinins, prostaglandins and 1,25-dihydroxy-cholecalciferol.*
- *Acute Tubular Necrosis has haemodynamic or toxic causes. Cardiogenic and hypovolaemic shock cause acute renal failures just as renal vasoconstriction. Renal ischaemia leads to hypoxic damage, in particular damage of the renal medulla. Ischaemic tubular damage also reduces the GFR further, because of reflex spasms of the afferent arterioles, and due to tubular blockage with accumulation of filtrate in the early part of the proximal tubules.*
- *Bacterial pyelonephritis typically causes interstitial inflammation of the kidneys, but the*

interstitial inflammation is more often caused by a hypersensitivity reaction to drugs (antibiotics, phenacetin and non-steroid anti-inflammatory drugs, NSAIDs).

- *Diabetic nephropathy includes hypertension, albuminuria and low GFR with glomerulosclerosis (thickening of the basement membrane and damage of the glomerular filter by disruption of the protein cross-linkages). The earliest evidence may be microalbuminuria. The patient later develops intermittent albuminuria followed by persistent albuminuria.*

- *Nephroblastoma (Wilms´ tumour) is the most frequent intraabdominal tumour in both girls and boys. A large abdominal mass is found sometimes with signs of intestinal obstruction. The tumour grows rapidly and spread to the lungs. The diagnosis is confirmed with excretion urography and arteriography.*

- *Renal cell carcinoma (hypernephroma) accounts for more than 90% of all the malignant renal tumours in adults (smokers). There is a strong association with a rare autosomal dominant inherited disease called Von Hippel-Lindau syndrome (haemangioblastomas in the cerebellum and the retina). The genetic locus is on chromosome 3p.*

Further Reading

Nephron. Monthly journal published by the International Society of Neprology. S Karger AG, Allschwilerstrasse 10, PO Box CH-4009 Basel, Switzerland.

Rehberg, P. Brandt. "Studies on kidney function: I. The rate of filtration and reabsorption in the human kidney." *Biochem. J.* 20: 447, 1926.

Schafer, JA. Renal water and ion transport systems. *Am. J. Physiol*. 275 (*Adv. Physiol. Educ*. 20): S119-S131, 1998.

Return to chapter 25

Return to Content

Chapter 26

The Hypothalamo-Pituitary System

Study Objectives

- To *define* autocrine, endocrine, neurocrine and paracrine transmission, endocrine feedback, endocytosis, exocytosis, hormones, hormone receptors, membrane receptors, neurohormones, neurotransmitters, phagocytosis, transcription, and tropic hormones.
- To *describe* the structural relations between the brain, hypothalamus and hypophysis. To describe secretion mechanisms, second messengers, hormonal sensitivity, melanotropin secretion and function, and stimulation-secretion coupling. To describe the structure and function of hypothalamo-hypophyseal hormones, of tropic hormones from the adenohypophysis and of pro-opio-melano-corticotropin. To describe acromegaly, Cushings syndrome, gigantism, dwarf growth, panhypopituitarism, and hyperpituitarism.
- To *explain* the secretion and function of growth hormone, somatomedins and somatostatin. To explain the secretion and function of gonadotropins, prolactin, relaxin, oxytocin, vasopressin. To explain hormonal function tests.
- To *use* the above concepts in problem solving and case histories

Principles

- *The endocrine and nervous systems co-ordinate the functions of the other organ systems as regulators of the function of the whole body.*
- *The endocrine system exerts its influence through blood-borne substances (hormones) produced in glands without secretory ducts (endocrine glands).*
- *The endocrine glands comprise the hypothalamo-hypophyseal axis, which regulates the function of the thyroid, parathyroid, adrenal, and reproductive glands. Other important hormones are the growth factors, cytokines and gastrointestinal hormones.*

Definitions

- **Autocrine transmission** refers to liberation and diffusion of signal molecules inside a cell to control functions in the cell of origin.
- **Calmodulin** is a specific binding protein for Ca^{2+} inside cells. Different calcium-calmodulin complexes activate or inhibit the activities of *calcium-dependent enzymes*.
- **Calsequestrin** is a specific binding protein for Ca^{2+} inside the sarcoplasmic reticulum of muscle cells. Calsequestrin is a buffer for cytosolic Ca^{2+}-concentration.
- **Catecholamines** are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The important catecholamines in humans are

adrenaline, noradrenaline and dopamine.

- **Cushing's disease** is *hypercorticism* (increased glucocorticoid production) caused by a pituitary basophilic adenoma.
- **Cushing's syndrome** refers to the consequences of increased plasma glucocorticoid concentration from any source.
- **Cytokines** are secreted polypeptides that affect the functions of other cells.
- **Domain** is a segment of a protein molecule with a functional role independent of the rest.
- **Endocrine feedback** is a system, whereby the first hormone, liberated to the blood stream, controls the secretion and liberation of the second. The second hormone acts by feedback and modulates the secretion of the first.
- **Endocrine transmission** refers to transport of hormones along the blood stream to a distant target organ.
- **Endocytosis** or *pinocytosis* refers to transport of molecules or material into the cell in vesicles of cell membrane. In some cases the coating is made by a surface protein called **clathrin**. Endocytosis requires metabolic energy (ATP).
- **Exocytosis** is a process whereby the contents of intracellular vesicles (hormones, transmitters) are released to the external environment.
- **Feedback systems** which are negative contain at least one step of inhibition. The total effect is to minimise any external change introduced to the system. Almost all hormone systems maintain homeostasis by negative feedback.
- **Feedback systems** which are positive are systems, where an external change leads to increased secretion of hormone 1, which also leads to a secondary rise in hormone 2's concentration. This is an auto-accelerating phenomenon and a rarity.
- **Hormones** are *messenger* or *signal* molecules. Classical hormones are conveyed by the blood (endocrine substances) and their target cells are equipped with *receptors* that recognise each hormone. Hormone molecules form a large *signal family* together with neurotransmitters, and local diffusive (autocrine- paracrine) substances.
- **Hormone receptors** are proteins, to which hormones bind, they are present in cell *membranes, cytoplasm* and *nucleus,* and serve two functions. Firstly, they are required for selectivity. Secondly, they are connected to an effector mechanism in the cell. In response to hormone binding, the receptor conformation is changed, and this activates a specific enzyme system that serves as an amplifier.
- **Membrane receptors** are surface glycoproteins (just like immunoglobulins), that bind the water-soluble hormones (catecholamines and peptides). Some receptors have an amino acid sequence similar to a sequence within the hormone.
- **Neurocrine** (neurosecretory) **transmission** refers to transport of a *neurohormone* first from the cell body of a neuron along its axon, and then with the blood to its target cells.
- **Neurotransmitters** are signal molecules functioning in axonal transfer or between neurons.

- **Phagocytosis** refers to transport into cells of bacteria and large foreign bodies. The cells a leucocytes and cells of the reticuloendothelial system that can destroy noxious substances.
- **Paracrine transmission** is a release and diffusion of signal molecules with regulatory action on neighbour cells.
- **Radio-immuno-assays** (RIA) refers to any method for detecting or quantitating antigens or antibodies utilising radiolabeled reactants. RIA utilises the competitive binding between a hormone and its induced antibody. RIA can be used to detect small quantities (high sensitivity), even in complex mixtures (high specificity).
- **Transcription** (copying) is defined in Chapter 31 together with other genetic concepts.
- **Tropic hormones** regulate the growth and hormone secretion from other cells. The five classical hormones from the adenohypophysis are tropic hormones.

Essentials

This paragraph deals with *1.Hormones in general, 2. Hormone receptors, 3. Monoamines, amino acids and peptides, 4. Radio-immuno-assays, 5. Hormone function tests, 6. Clinical application of hormones, 7. The hypothalamo-hypophyseal system, 8. The adenohypophysis, 9. The neurohypophysis.*

1. Hormones in general

The scientists of the past used extirpation, substitution and transplantation to obtain the classical part of our present knowledge on endocrinology (ie, the discipline covering the internal secretion of signal molecules to the blood). Removal of the pancreas produced diabetes in animals. Removal of the pituitary gland, followed a few days later by removal of the pancreas, produced no symptoms of diabetes. This is because the adenohypophysis produces a hormone that is antagonistic to the effect of insulin in glucose metabolism. The hormone is human growth hormone. Houssay received a Nobel Prize for this work in 1947.

Hormones are *messenger* or *signal* molecules. Classical endocrine hormones are secreted into the blood and transported to their distant target cells, which are equipped with *receptors* that recognise each hormone. The hormones co-ordinate the activities of different cells in order to maintain homeostasis and to secure growth and reproduction. Hormone molecules form a large *signal family* together with neurotransmitters, autocrine and paracrine acting substances.

Paracrine and *autocrine* signal molecules are secreted and diffuse into the interstitial fluid surrounding the cells and their actions are restricted either to nearby cells (paracrine) or to the cell of origin (autocrine).

Neurotransmitters (acetylcholine, adenosine, amines, amino acids, ATP, peptides) exert a type of paracrine action, since they are released in the synaptic region.

All reactions in the cell linking stimulation and secretion together are termed *stimulation-secretion couplings*. Stimulation-secretion coupling involves depolarization of the cell membrane or opening of Ca^{2+} -channels, so that Ca^{2+} can diffuse into the cell and combine with its Ca^{2+}-binding proteins. A rise in intracellular concentration [Ca^{2+}] is necessary for exocytosis.

Elimination of hormones takes place by metabolic processes such as the inactivation of peptide hormones by proteolytic enzymes, or the transformation of hormones in the liver. Hormones are also eliminated by excretion in the urine or bile. In the liver hormones are coupled to glucuronic acid or sulphate, but these hormones are in part reabsorbed in the *entero-hepatic-*

circuit.

Protein binding protects small hormone molecules (such as the thyroid hormone) from elimination. Protein binding also eases the transportation of the lipid-soluble steroids, and main
is maintained.

Hormones can be divided into three chemical categories:

Peptides and proteins include neuropeptides, pituitary and gastrointestinal hormones.

Steroids consist of adrenal and gonadal steroids and vitamin D, which is converted to a hormone. Steroids are lipid soluble (lipophilic).

Monoamines (modified amino acids) comprise catecholamines, histamine, serotonin, and melatonin. Catecholamines (dopamine, noradrenaline and adrenaline) are derived from tyrosine - and serotonin/melatonin from tryptophan - by a series of enzymatic conversions. Monoamines and amino acid hormones are water soluble just as peptides. Thyroid hormones are iodinated derivatives of tyrosine, and thyroid hormones are lipophilic.

The water-soluble hormones are packed in the Golgi complex in secretory granules that migrate to the cell surface.

Exocytosis of the granule contents to the interstitial fluid (ISF) and diffusion through fenestrae to the capillary blood is a common method. The secretory cells are first stimulated by chemical or electrical signals.

Synthesis of protein or peptide hormones takes place as outlined in Chapter 31. Transcription of the hormone gene results in a specific mRNA determining the synthesis of a single hormone. However, a single gene may dictate the synthesis of different peptides in different cells. As the signal protein is cut off, the prohormone is formed and transported to the Golgi apparatus and stored in granules. The hormone specific amino acid sequence is contained in the prohormone.

An *endocrine feedback system* is a system whereby the first hormone controls the secretion and liberation of the second. The second hormone acts by feedback to modulate the secretion of the first.

A *negative feedback system* contains at least one step of inhibition. The total effect is to minimise any external change introduced to the system. Almost all hormone systems maintain homeostasis by negative feedback.

A *positive feedback system* exagerates any primary change initiated. - This is an auto-accelerating phenomenon and a rarity.

The most important example in humans is the *steep* rise in blood [oestradiol] in the middle of the menstrual cycle. High [oestradiol], when maintained for longer than 35 hours, stimulates by positive feedback, the luteinizing hormone (LH) and follicle stimulating hormone (FSH) secretion from the adenohypophysis, which further stimulate oestradiol secretion etc.

By contrast, moderate plasma [oestradiol] levels, which are present during the other parts of the cycle, provide negative instead of positive feedback. *Long feedback systems* act on the hypothalamo-pituitary system from remote target organs.

Short feedback systems use a short distance feedback, such as the influence of the hypophysis back to the hypothalamus. *Auto-feedback* refers to the action of a liberated hormone that was secreted on the cell from where it came thereby modulating its own secretion.

2. Hormone receptors

These are proteins, to which hormones bind. They are present in cell membranes, cytoplasm

and nucleus, and serve two functions. Firstly, they are required for selectivity. Secondly, they are connected to an effector mechanism in the cell (Fig. 26-1).

In response to hormone binding the receptor conformation is changed, and this activates a specific enzyme system that serves as an amplifier. In the cytosol, multiple second messengers have evolved to serve such purposes, whereas in the nucleus, the hormone-receptor complex binds to DNA and regulates gene expression (Fig. 26-1). The effector domain of the membrane receptor is directly coupled to the regulatory portion of the effector enzymes (such as adenylcyclase = adenylate cyclase). These effector enzymes control ion fluxes, membrane transport systems, the production of cyclic nucleotides, and the breakdown of phospholipids. Inactive kinases are activated by the use of ATP.

This phosphorylation is critical for for transformation of information and for cell viability (synthesis, transport and metabolism of vital molecules). Many hormones initiate a series of reactions when bound to membrane receptors. One family of coupling molecules, called G-proteins, links some of the receptors to nearby effector molecules (see Chapter 1). Other receptors make use of another system.

Fig. 26-1: Target cells activation by hormones acting at Membrane, Cytoplasmic, and Nuclear receptors.

Steroids and thyroid hormones are lipophilic and therefore pass easily through the cell membrane by diffusion. Steroids bind to specific cytosol-receptor proteins that are then translocated into the cell nucleus where they reversibly bind to DNA (Fig. 26-1). Some unbound receptor proteins may even exist in the nucleus. The binding of the steroid-receptor complex to the specific gene modulates mRNA transcription.

Tri-iodo-thyronine (T_3) binds to nuclear receptor proteins, which then attaches to a thyroid response unit in the gene in a manner similar to that of steroid receptors. The result is increased mRNA formation (Fig. 26-1).

Steroids and thyroid hormones frequently work in conjunction with each other (potentiate amplification of gene expression).

Cell membrane and intracellular receptors can change their affinity and number. A specific ligand for a receptor is able to modulate the total number of this receptor. In concentration of the ligand (hormone, neurotransmitter, drug) often reduces the number of receptors (down-regulation), and other hormones recruit their own receptors at low concentrations (up-regulation). Maximal effects of hormones are generally observed at receptor occupancy of less than 50%.

The myoepithelial cells (myometrium and breast) contain oxytocin receptors. Their number is up regulated by estrogens and down regulated by progesterone. The cardiac muscle contains nor-adrenergic receptors (b_1). Both affinity and number of receptors is increased by thyroid hormone stimulation (T_3/T_4).

Internalisation is the transport of hormone-receptor complex into the cell by an endocytotic vesicle. This is a means of terminating the action of the hormone. After destruction of the hormone by lysosomes, the receptor returns to the surface and is reused.

3. Monoamines, amino acids and peptides

Such water-soluble hormones (first messengers) bind to hormone receptors on the lipid-rich plasma membrane. Peptide hormone and catecholamine receptors are membrane receptors with a binding domain located extracellularly and an effector domain intracellularly (Fig. 26-1).

The second messengers involved are cyclic adenosine monophosphate (cAMP), cyclic guanosine monophosphate (cGMP), inositol trisphosphate (IP_3), Ca^{2+}, diacylglycerol (DAG)

etc. The Ca^{2+}-ion is an important second messenger. The Ca^{2+}-influx to the cytosol is controlled by hormone receptor binding, neural stimuli or modified by other second messengers.

Sutherland discovered cAMP and demonstrated its role as a second messenger in mediating body functions (Nobel Prize 1971).

Increased activity of the sympathetic nervous system including release of adrenaline triggers fight-or-flight reactions. In the heart, adrenaline molecules diffuse to the myocardial cells, where they bind to membrane b-receptors. A stimulatory signal is hereby transmitted to an associated enzyme called adenylcyclase. This enzyme catalyses the conversion of ATP to cAMP. The importance of cAMP is that it activates protein kinase A, which, among many other functions, phosphorylates the Ca^{2+}-channel protein. This activation is correlated with an increase in the magnitude of the Ca^{2+}-influx, the force of contraction, and the heart rate.

The parasympathetic system counteracts the sympathetic by slowing the heart rate and decreasing the force of contraction. Acetylcholine is bound to another set of specific membrane receptors located on the heart cell membrane. Acetylcholine reduces the Ca^{2+}-influx that was increased by adrenaline.

Most hormones have a blood concentration of approximately 10^{-10} mol per l. One molecule bound to a cell receptor releases 10 000 times more cAMP in the cell. Hence, cAMP works as an amplifier of the hormone signal.

Phosphodiesterase (PDE) destroys cAMP. PDE enhances hydrolysis of cAMP to the inactive 5'- AMP by a highly exergonic process.

Inhibitors of the PDE (theophylline and caffeine) act synergistically with hormones that use cAMP as a second messenger.

cAMP stimulates catabolic processes such as lipolysis, glycogenolysis (glucagon), gluconeogenesis, and ketogenesis. The cAMP also stimulates amylase liberation in the saliva by the parotid gland, the HCl secretion by the parietal cells, the insulin release by the b-cells in pancreas, and the increased ion permeability of many cell membranes.

When the glucose concentration increases in the arterial blood and close to the b-cells of the pancreatic islets of Langerhans, it triggers an increase in Ca^{2+} -influx to the cell.

The initial surge in insulin secretion is caused by calmodulin-dependent *protein kinases*.

The high cytosolic [Ca^{2+}] activates the membrane phospholipase A_2 and C. Phospholipase A2 releases arachidonic acid (AA) which stimulates insulin secretion. Phospholipase C catalyses the formation of IP_3 and DAG. The IP_3 releases more Ca^{2+} from the endoplasmic reticulum, and DAG activates protein kinase C.

The decrease in insulin secretion after the initial surge and its subsequent increase can be explained by the action of protein kinase C.

Initially, the active protein kinase C stimulates the Ca^{2+}-pump in the plasma membrane, reduces cytosolic [Ca^{2+}] and thus reduces the initial calmodulin-dependent insulin secretion. Later, protein kinase C stimulates the formation of cAMP and amplifies the induction of calmodulin-dependent protein kinase thereby causing a gradual increase in insulin secretion. Prolonged glucose stimulation probably leads to down-regulation of protein kinase C. An abnormally prolonged glucose stimulation may render b-cells glucose blind and thus spoil their function.

Insulin secretion is not only stimulated by glucose, but also potentated by acetylcholine via

phospholipase C and by glucagon via activation of adenylcyclase. b-Agonists stimulate b-receptors on the glucagon producing a-cells, whereas a-agonists inhibit insulin secretion via a_2-receptors on the b-cells. Acetylcholine and glucagon react by activating protein kinase C and cAMP dependent protein kinase A, respectively. Both mechanisms potentate the Ca^{2+}-triggered insulin secretion.

Transcription in the cell nucleus produces a precursor messenger RNA molecule complementary to part of a DNA. The precursor is processed into messenger RNA and transported through the nuclear membrane into the cytoplasm. Messenger RNA carries the genetic information in triplet codons (Chapter 31). Messenger RNA binds to ribosomes and transfer RNA molecules synthesise peptides (ribosomal translation). Translation produces big precursor molecules (pre-pro-hormones). Precursors have a signal peptide that contains processing information to ensure that the protein enters the rough endoplasmic reticulum. Here enzymes split the precursor into a signal molecule and a prohormone. Finally, peptide hormones undergo post-translational processing (for eg, thyroid stimulating hormone, TSH, and gonadotropins are glycosylated; insulin forms a zinc-complex). The hormones reach the Golgi complex, where they are packed into secretory granules that migrate to the cell surface.

Roger Guillemin synthesized brain peptides that regulate the pituitary secretion in vitro. He received the Nobel Prize in 1977.

4. Radio-immuno-assays (RIA)

RIA refers to any method for detecting or quantitating antigens or antibodies utilising radiolabeled reactants. RIA is used to detect very small quantities of antigens or antibodies, even in complex mixtures.

First a specific antibody is produced towards the antigen (eg, hormone).

In one version of RIA for antigen detection, the antigen is radiolabeled and reacted with a limited amount of specific antibody. The complex containing bound antigen is then separated from free antigen. Unlabeled antigen in a test sample is used to compete with the binding of radiolabeled antigen. The test antigen is quantitated from the extent of inhibition obtained with standards containing defined amounts of the same antigen.

Rosalyn S. Yalow and Saul Berson developed the RIA method. Rosalyn Yalow received the Nobel Prize in 1977.

Recent variations of the RIA technique include immuno-radiometric, chemi-luminescent, and enzyme-linked radioimmuno-sorbent assays. - In the radioreceptor assay a hormone receptor is substituted for the antigen-antibody in RIA.

5. Hormone function tests

The following tests are clinical tools in the diagnosis of hormone disorders:

5.1. The hormone concentration in the blood is commonly used. It can be measured by taking advantage of the new methods described above.

5.2. The secretion flux of T_3 and T_4 from the thyroid gland.

5.3. The metabolic rate or the absorption rate of ^{131}I (radioactive iodine) in the thyroid gland. The physical half-life of ^{131}I is 8 days or 192 hours. The elimination rate constant (k) of a substance is the amount eliminated per unit time divided by the total amount present in the distribution volume, assuming exponential elimination. The variable k is easy to calculate:

$T_{½}$ = ln 2/k = 0.693/k. The value of k for iodine is 0.693/192 or 0.0036 $hours^{-1}$.

5.4. The elimination rate:

Abnormal amounts of catecholamines or VMA (Vanillyl mandelic acid) in a 24-hour urine suggest the presence of a catecholamine-producing tumour (phaeochromocytoma).

5.5. Stimulation test:

Stimulation with ACTH (Adrenocorticotropic hormone) without a substantial rise in plasma -[cortisol] suggests primary, adrenocortical atrophia.

5.6. Suppression test:

Dexamethasone (a cortisol synergist) is administered to a Cushing suspect patient in the evening. The next morning a measurement of plasma [cortisol] shows suppres in normal persons and in patients with a primary, adrenocortical hyperfunc cortisol synergist reduces ACTH secretion and thus cortisol production by negative feedback.

Hypothalamic/pituitary Cushing is never suppressed by cortisol.

5.7. The glucose tolerance test:

A load of glucose normally triggers an increased rate of insulin production.

6. Clinical applications of hormones

Distribution of oestrogens and progesterone in *contraceptives* (P pills) is world-wide. Oestrogens are widely used to relieve postmenopausal discomfort. Now some females with osteoporosis are treated experimentally with calcitonin, because calcitonin inhibits osteoclastic bone resorption.

Insulin is a lifesaver for diabetics, and it is produced and distributed as pure human insulin.

In the affluent areas of the world many women deliver their babies following an oxytocin infusion.

Oestrogens and gonadotropins are used in treatment of sterility and menstrual disturbances.

Huggins received the Nobel Prize in 1966 for the introduction of a new form of cancer therapy in which sex hormones are used to retard their growth. He used androgens for breast cancer and oestrogens for prostate cancer.

7. The hypothalamo-hypophyseal system.

The human pituitary gland consists essentially of two parts both controlled by the hypothalamus.

The *glandular* part is the adenohypophysis or anterior lobe, and the *neural* part is the neurohypophysis or posterior lobe.

The adenohypophysis develops ectodermally from the primitive mouth cavity (Rathkes pouch). Blood-borne signal molecules from the hypothalamus regulate the cells of the adenohypophysis.

The neurohypophysis develops from the neuro-ectoderm in the floor of the third ventricle. The two parts combine to form one body called the adeno-neuro-hypophysis that weighs about 0.5

g.

The infundibular process of the neurohypophysis receives blood from the inferior hypophyseal artery, whose capillary plexus drains into the adenohypophysis (Fig. 26-2). The upper stalk and the median eminence is supplied with blood by the carotid artery to the superior hypophyseal artery, whose primary capillary plexus ends in long portal veins carrying blood to the highly permeable secondary capillary plexus of the adenohypophysis. From this plexus, blood drains into the dural sinus. The adenohypophysis lies outside the blood-brain barrier, and does not receive arterial blood directly. A third capillary plexus, between the neurohypophysis and the median eminence of the hypothalamus, allows *short loop feedback* from the hypophysis to the hypothalamus.

The hypothalamus and the hypophysis connects in the following ways (Fig. 26-2):

1. The neurosecretory axons pass from the cell bodies in the supraoptic and paraventricular nuclei of the hypothalamus to the neurohypophysis. The neurosecretory granules are here stored in the terminals of these axons. The granules are released by exocytosis upon stimulation. The peptides from the granules then enter the capillary plexus of the inferior hypophyseal artery

2. The hypophysiotropic zone in the median eminence of the hypothalamus is connected to the adenohypophysis. Both releasing and inhibiting peptides are synthesized in hypothalamic neuron bodies and transported to the median eminence in granules via axonal transport. At the median eminence the inhibitory and releasing signal molecules are discharged to the capillary plexus of the superior hypophyseal artery. From here they follow the blood through the long portal veins to reach the specific cells in the adenohypophysis. Here, the releasing and inhibiting hormones modulate the output of tropic hormones.

Fig. 26-2: The hypothalamo-pituitary axis. The 5 classical tropic hormones are corticotropin, gonadotropins (FSH, LH), somatotropin, thyrotropin and mammotropin.

Neurosecretory neurons (which have nuclei in the hypothalamus and axons that lead to the median eminence and to the posterior lobe of the hypophysis) and peptidergic neurons (spread in the nervous system and gut) produce and liberate peptides in much the same way.

The secretory granules travel through the axons of the neurosecretory neurons that form the supraoptico-hypophyseal tract, with high velocity (more than 100 mm each hour). This tract runs through the pituitary stalk and end in the neurohypophysis. The transfer is known as axoplasmic transport. The neuro-hy
During transport the pro-hormone splits into its subunits. Oxytocin and vasopression are then released.

The nerve endings of this tract in the neurohypophysis are the storage area for these two neurosecretory hormones; secretion to the blood takes place through fenestrated capillaries. The secretion granules release their content by regulated exocytosis. Exocytosis is triggered when the neurosecretory neuron is depolarised and an action potential is transferred to the terminals.

Even a small rise in the osmolarity of plasma stimulates osmoreceptors, located close to the neurosecretory cells in the hypothalamus.

The osmoreceptors stimulate both production and secretion of vasopressin (ADH) in the neurosecretory cells. The plasma [ADH] will then rise from the basal level that is 2 pmol per l.

The normal secretion flux is 10^{-13}mol ADH per kg body weight per min, and the biological half-life in human plasma is 18 min. Some females increase their plasma [ADH] in the pre-menstrual phase.

The neuroregulatory peptides are endogenous opiates (endorphins and enkephalins), b-lipotropin, neurotensin, substance P, VIP or vasoactive intestinal peptide etc. Many of these peptides are cut off from a big mother molecule: pro-opio-melanocortin (POMC). These peptides may exhibit a permissive effect on other hormones (ACTH, growth hormone) related to behaviour and autonomic responses. During exercise (a Cooper test which lasts 12 min) the plasma [b-endorphin] and [ACTH] increases by 200-300% from the normal resting averages of 1.7 and 2.2 pmol per l, respectively. There is an increase in plasma [ACTH] and its accompanying neuroregulatory peptides during prolonged stress like exhaustive exercise and chronic disease. The endogenous opiates affect stress-adapting behaviour, such as the euphoria observed in the chronically ill.

The pituitary gland normally has a mass of approximately 500 mg, but it increases during pregnancy and decreases with aging.

8. The adenohypophysis

The *five* hormones from the adenohypophysis are tropic hormones - they regulate the growth and hormone secretion of target cells (Fig. 26-2). They include:

1. Thyrotropin or thyroid-stimulating hormone (TSH), which is produced in thyrotropic cells,
2. Gonadotropins (FSH and LH) from gonadotropic cells,
3. Corticotropin (ACTH) from corticotropic cells,
4. Somatotropin (human growth hormone, HGH or GH) from somatotropic cells and
5. Prolactin or mammotropin produced in mammotropic cells.

The five tropic peptide hormones have the following molecular characteristics:

1. One group contains glycoproteins with two peptide chains.

   There is a special, biologically active b-chains for each of the three hormones TSH, FSH and LH, although the inactive a-chain is the same for all of them.

2. Somato-mammotropins are single-chain peptides containing 200 amino acids of almost the same sequence.

   HGH and prolactin (PRL) are probably simple gene duplicates from the same prohormone molecule.

3. POMC peptides are neuroregulatory hormones: ACTH, endogenous opiates, b-endorphin, b-lipoprotein, a-MSH and b-MSH (MSH abbreviates melanocytic stimulating hormone).

   Histamine plays an important role in pituitary hormone secretion. Histamine stimulates the secretion of ACTH, b-endorphin, a-MSH, and PRL. Histamine participates in the release of these hormones during prolonged stress and possibly in the suckling- and oestrogen-induced PRL-release.

   The release of growth hormone (GH) and TSH are predominantly inhibited by histamine. GH is the main stimulator of body growth in humans (Chapter 30).

Histamine increases the secretion of LH in females - mediated by GnRH (gonadotropin releasing hormone). Histamine probably affects the cell bodies in the supraoptic and paraventricular nuclei, stimulating the formation of arginine vasopressin and oxytocin.

9. The neurohypophysis

The neurohypophysis secretes two hormones: vasopressin and oxytocin.

Vasopressin or antidiuretic hormone (ADH) is a vasopressor with a strong antidiuretic effect as the names imply. Vasopressin is normally synthesized as a big pre-prohormone in the ribosomes of neurons in the supraoptic and paraventricular nuclei of the hypothalamus. The pre-prohormone consists of a signal peptide, ADH, neurophysin and a glycopeptide. First, the signal peptide is cut off, and then the precursors are packed in secretion granules in the Golgi complex. The secretion granules travel by axoplasmic transport through the axon of the neurosecretory neurons that form the supraoptic-hypophyseal tract, and then are stored in its terminals in the neurohypophysis. These terminals are located close to the fenestrated capillaries. The smallest rise in the osmolarity of plasma stimulates osmoreceptors located close to the neurosecretory cells of the hypothalamus. The osmoreceptors stimulate both production and release of ADH. Vasopressin is a nonapeptide with a molecular weight of 1084 Da.

ADH has the following effects:

1. ADH eases the renal reabsorption of water in the cortical collecting ducts (and not in the outer medulla but in the inner medulla) - leading to antidiuresis.

2. ADH probably stimulates the active solute reabsorption (NaCl) in the thick ascending limb of the renal Henle loop. Thus, ADH helps maintain the con
the kidney.

3. Vasopressin is a universal vasoconstrictor. Vasopressin reduces the small, medullary bloodflow through vasa recta along the Henle loop.

   ADH acts on the basolateral membrane of the cells, and the result is a rise of [cAMP] in the cytosol. The cAMP diffuses to the luminal side, where it causes vesicular structures to develop and fuse with the luminal membrane. Hereby, the membrane receives a large number of water channels, so the membrane becomes highly water permeable. Water diffuses through the cell to the basolateral membrane and into the interstitial fluid.

Oxytocin

Stimulation of *tactile receptors in the mammary nipple* causes the neurosecretory neurons to release oxytocin through a neuroendocrine reflex.

The latency between the stimulus and milk ejection is due mainly to the transport of oxytocin in the blood from the neurohypophysis to the milk ducts (20-30 s). Oxytocin stimulates the myoepithelial cells in the milk ducts of the lactating breast so that milk is ejected to the baby.

Oxytocin also stimulates the myoepithelial (myometrial) cells of the uterus satisfying the woman sexually during breast-feeding. Oxytocin can perhaps start labour.

Pathophysiology

- This paragraph deals with the five tropical hormones from the adenohypophysis and vasopressin secreted from the neurohypophysis: *1.Pituitary TSH-secreting tumours*, *2. Polycystic ovarian syndrome*, *3. Basophilic pituitary adenoma*, *5. Prolactinomas*, *6. Diabetes insipidus*, *7. Syndrome of inappropriate ADH secretion*, *8. Panhypopituitarism*.

1. Pituitary TSH-secreting tumours

A pituitary TSH-secreting tumour is an extremely rare cause of thyrotoxicosis. Thyrotoxicosis is dealt with in Chapter 28.

2. Polycystic ovarian syndrome

Chaotic LHRH secretion from the hypothalamus to highly sensitive gonadotropic cells in the pituitary increases the LH-level in the blood plasma. Actually, LHRH induces its own receptors. The gonadotropin level is so high in the follicular phase that androgens in excess are produced from the theca cells. These females produce immature or atretic follicles occurring as multiple cysts in enlarged ovaries.

A maintained LH-level can be produced by other causes. The excessive gonadotropin and androgen secretion causes irregular bleedings, subfertility, acne and hirsutism.

3. Basophilic pituitary adenoma

Basophilic pituitary adenoma is the cause of classical *Cushing's disease*. The excessive ACTH secretion induces adrenocortical hypersecretion of cortisol. The hyper-cortisolaemia causes the many symptoms and signs found in *Cushing's syndrome* (Chapter 30).

4. Somatotropic pituitary adenoma

Somatotropic pituitary adenomas produce large amounts of growth hormone leading to *gigantismus* in childhood and to *acromegaly* in adults. These cases of hyperpituitarism are dealt with in Chapter 30. In rare cases the cause is excessive GHRH secretion from the hypothalamus. Some acromegalics also produce excess prolactin in hypertrophic mammotrophs.

Pituitary adenoma cells with TRH receptors also secrete excess GH. TRH is used as a diagnostic test. Other pituitary adenoma cells have somatostatin receptors. Somatostatin and somatostatin agonists inhibit GH secretion, and make the adenomas shrink

5. Prolactinomas

Prolactinomas are microlactinomas (less than 1 cm), which cause anovulatory, irregular bleedings, abnormal milk production (galactorrhoea), and subfertility. The constantly elevated plasma prolactin inhibits the LH-secretion necessary for ovulation. – Dopamine, dopamine agonists and somatostatin analogues inhibit prolactin secretion and can make the prolactinomas shrink.

6. Diabetes insipidus

The true form of *diabetes insipidus* is caused by deficiency of vasopressin (ADH deficiency). There are two types of diabetes insipidus. The primary or idiopathic type, which is due to a genetic defect that blocks the hormone production, and the secondary type, where the hypothalamo-hypophysary system is damaged by disease or surgery.

The renal tubule cells are rarely insensitive to ADH, and this infrequent condition is called *renal* diabetes insipidus or *nephrogenic* diabetes insipidus. This is a sex-linked recessive disorder or it is acquired from renal disorders or hypercalcaemia.

The symptoms and signs are mainly due to the large diuresis (polyuria), nocturia, and a tremendous thirst (polydipsia). A total lack of ADH can result in a diuresis of 25 l daily.

ADH eases the renal reabsorption of water in the cortical collecting ducts and in the inner medulla via a cAMP mechanism which increases the number of water channels in the luminal membrane. ADH stimulates NaCl reabsorption in the thick ascending limb of Henle, and

vasopressin is a universal vasoconstrictor.

Synthetic vasopressin is given intra-nasally as a spray up to 3 times daily.

7. Syndrome of inappropriate ADH secretion

ADH producing tumours in the hypophysis or in the lungs causes the *syndrome of inappropriate ADH secretion*. Water retention, concentrated urine, hyposmolar plasma, and muscle cramps characterise this syndrome.

8. Panhypopituitarism

is typically due to total destruction (lesions or tumour invasion) of all hormones in the hypothalamo-hypophysary system. Lack of GH and somatomedins result in a dwarf without normal sex development (lack of LH and FSH). This dwarf has also hypothyroidism (lack of TSH), and a Cushing-like syndrome (hypercorticism without excess ACTH secretion).

The differentiation of somatotrophs, mammotrophs and thyrotrophs is dependent upon a *protein transcription factor* (Pit-1). Mutation of the *Pit-1 gene* leads to hypoplasia of the adenohypophysis and to insufficient production og GH, prolactin and TSH with hypopituitarism.

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A. High-pressure liquid chromatography is a sensitive analysis used for many hormones and biochemical key molecules.

B. he affinity and number of specific membrane receptors on a given cell is constant.

C. Monoamines and amino acid hormones are water-soluble just as peptides.

D. cAMP inhibits catabolic processes such as lipolysis, glycogenolysis, gluconeogenesis, and ketogenesis.

E. Hyperpituitarism is often caused by microadenomata, which typically cause dwarf growth.

Case History A

The adenohypophysis of a 23 year old woman contains approximately 300 μg of TSH with a molecular weight of 31 000. TSH has a half-life ($T_{½}$) in plasma of 55 min and a concentration of 100 pmol per l of plasma. The haematocrit of the patient is 0.5. The woman secretes TSH to her total blood volume (TBV), which is 4 l.

1. Develop an equation for the calculation of her TSH secretion (J mol/hour). The rate constant k can be used.

2. Calculate the secretion of TSH from her adenohypophysis

3. What fraction of her total TSH store is secreted per 24 hours?

Case History B

A woman (24 years of age; height: 1.70 m; weight: 60 kg) is in hospital due to a tremendous thirst, and she drinks large amounts of water. Since she is producing 10 or more litres of urine each day, the doctors suspect the diagnosis to be diabetes insipidus. The vasopressin concentration in plasma (measured by a RIA method) is 10 fmol per l. Her secretion of

vasopressin is only 5% of the normal flux of 10^{-13} mol per min per kg body weight. The normal plasma [vasopressin] is 2 pmol per l as a mean. The extracellular volume (ECV) is 20% of her body weight.

Vasopressin is injected intravenously at several occasions. A dose of 3 µg vasopressin is the minimum necessary to normalise her diuresis for 4 hours. Before the injection her diuresis is 6 ml of urine per min, but within 25 min her urination is constantly around 0.5 ml/min.

1. Calculate the secretion of vasopressin (in mg/hour) from the neurohypophysis of a normal 60-kg person and of this patient.
2. Calculate the distribution volume for vasopressin, which is 20% higher than ECV.
3. Assume the 3 mg vasopressin injected to be distributed evenly immediately after the intravenous injection. Calculate the rise in vasopressin concentration in the distribution volume.
4. Estimate the relation between this concentration and that of a healthy individual.
5. Does this ratio have implications for the interpretation of her special type of diabetes insipidus?
6. Is it dangerous to lose 10 litres of urine per day?

Try to solve these problems before looking up the answers .

Highlights

- *The hypothalamo-pituitary system controls the function of the adrenal, thyroid, and reproductive glands, as well as regulating growth, lactation, milk secretion and water excretion.*
- *Protein and peptide hormone synthesis starts with processing of a primary gene transcript (code) called a prohormone. The processing includes proteolysis, glycosylation and phosphorylation.*
- *Catecholamines, peptide and protein hormones are stored in secretory granules and discharged by exocytosis.*
- *Catecholamines, peptide and protein hormones are water-soluble and cannot pass the cell membrane. They act on the surface of target cells via membrane receptors. The hormone-receptor-complex activates second messengers in the cell (cAMP, Ca^{2+}, DAG, IP_3) via stimulatory or inhibitory G-proteins.*
- *Thyroid and steroid hormones are lipid-soluble and act through specific nuclear receptors. The hormone-receptor-complex modulates elements in DNA molecules in order to change the expression of target genes.*
- *Peptide hormones produced in the cell bodies of hypothalamic neurons pass down their axons inside secretory granules to be stored in the terminals of the neurohypophysis.*
- *Releasing and inhibiting peptides from the hypothalamus are released in pulses in the adenohypophysis and act via second messengers. They modulate transcription, translation*

and secretion of tropic hormones.

- *Vasopressin or antidiuretic hormone is a vasopressor with a strong antidiuretic effect as the names imply. Vasopressin (ADH) is normally synthesized as a big pre-prohormone in the ribosomes of neurons in the supraoptic and paraventricular nuclei of the hypothalamus. The pre-prohormone consists of a signal peptide, ADH, neurophysin and a glycopeptide.*

- *POMC peptides are neuroregulatory hormones: ACTH, endogenous opiates, β-endorphin, β-lipoprotein, α-MSH and β-MSH.*

- *Stimulation of tactile receptors in the mammary nipple causes the neurosecretory neurons to release oxytocin through a neuroendocrine reflex. Oxytocin stimulates the myoepithelial cells in the milk ducts of the lactating breast. Oxytocin also stimulates the myoepithelial (myometrial) cells of the uterus. Oxytocin can perhaps start labour.*

- *The true form of diabetes insipidus is caused by deficiency of vasopressin (ADH deficiency). There are two types of diabetes insipidus. The primary or idiopathic type, which is due to a genetic defect that blocks the hormone production, and the secondary type, where the hypothalamo-hypophysary system is damaged by disease or surgery.*

- *ADH producing tumours in the hypophysis or in the lungs causes the Syndrome of inappropriate ADH secretion. Water retention, concentrated urine, hyposmolar plasma, and muscle cramps characterise this syndrome.*

- *Panhypopituitarism is due to total destruction (lesions or tumour invasion) of all hormones in the hypothalamo-hypophysary system. Lack of GH and somatomedins result in a dwarf without normal sex development (lack of LH and FSH). This dwarf has also hypothyroidism (lack of TSH), and a Cushing-like hypercorticism (wothout ACTH secretion).*

- *Hyperpituitarism is often caused by prolactin producing microadenomata, which cause abnormal milk production. This leads to disturbance of the menstrual cycle and infertility. Other pituitary adenomas produce large amounts of GH leading to gigantismus in childhood and to acromegaly in adults.*

Further Reading

Nature. Weekly journal published by Macmillan Magazines Ltd, Porters South, 4 Crinan Street, London N1 9XW, UK.

Cell. Bi-weekly journal published by Cell Press, 1050 Massachusetts Av., Cambridge Massachusetts 02138, USA.

Yalow, R. S. “Radioimmunoassay: Historical aspects and general considerations”, in *Handbook of Experiment. Pharmacol*., 1987.

Return to top

Return to Content

nt

Chapter 27.

Blood Glucose And Diabetes

Study Objectives

- To *define* glicentin, gluconeogenesis, glycogenolysis, glycolysis, hyperglycaemia, hypoglycaemia, incretins, insulin antagonists, paracrine secretion, primary and secondary diabetes mellitus.
- To *describe* the structural and functional characteristics of the Langerhans islets with cell types and hormones produced. To describe the incretin effect, insulin effects, the insulin receptor, and the glucose transporter. To describe disorders of the different cell types and their clinical picture. To describe methods for evaluation of the glucose combustion.
- To *draw* oral and intravenous glucose tolerance curves with linear and logarithmic ordinates for blood glucose concentration.
- To *explain* the biosynthesis and the effects of insulin, glucagon, pancreatic polypeptide (PP) and somatostatin. To explain the glucose metabolism and the control of blood glucose in the fed and the fasting state. To explain the consequences and the therapy of high and low blood glucose disorders.
- To *use* the above concepts in problem solving and in case histories.

Principles

- *The physiological principle in treatment of diabetes is to inject a fast-acting insulin three times a day just before meals and a slow-acting insulin at night.*
- *Insulin promotes the storage of energy, the synthesis of glycogen, mRNA and proteins.*
- *Certain major tissues (kidney, brain and intestine) are insensitive to the direct action of insulin.*

Definitions

- **Glicentin** is intestinal glucagon. Glicentin is built up from 69 amino acids in contrast to pancreatic glucagon, which consists of 29 amino acid moieties.
- **Gluconeogenesis** refers to formation of new glucose from glycogenic amino acids, lactate, glycerol and pyruvate.
- **Glycogenolysis** refers to glycogen breakdown to glucose in the liver.
- **Glycolysis** refers to anaerobic breakdown of glycogen.
- **Hyperglycaemia** is a condition, where the blood glucose is above 6.7 mM.
- **Hypoglycaemia** refers to a serious condition, where the blood glucose is below 2 mM.

- **Incretins** are hormones, which strongly potentiates the insulin secretion induced by the rising blood [glucose]. The incretins cause a much larger insulin secretion than the iv. administration of glucose, even at the same rise in blood [glucose]. This extra insulin secretion is called the *incretin effect*.

- **Ketogenesis** refers to accelerated lipolysis with liberation of free fatty acids to the blood. Free fatty acids are broken down to fatty acyl carnitine within the liver cells, and this molecule is converted into acetyl CoA, which in turn reach the mitochondria, where ketone bodies are formed.

- **Paracrine secretion** is a release of signal molecules to neighbour cells.

- **Insulin antagonists** are hormones opposing the effect of insulin: Pancreatic and intestinal (glicentin) glucagon, ACTH, growth hormone.

- **Primary diabetes mellitus** refers to all cases, where the cause is not fully explained.

- **Secondary diabetes mellitus** is caused by hypersecretion of one or more of the many catabolic hormones with hyperglycaemic effect (adrenaline, noradrenaline, glucagon, glucocorticoids and growth hormone) or by total destruction of the pancreas from pancreatitis or carcinoma. The hormone disorders are phaeochromocytoma, glucagonoma, Cushing's syndrome and acromegaly.

- **Somatostatin** (GHIH) is a multipotent hormone inhibitor consisting of a disulphide bridge and 14 amino acid units.

Essentials

This paragraph covers *1. the blood glucose regulation in the fed state*, as well as in *2. the fasting states*. Also *3. The endocrine pancreas*, and *4. Pancreatic exocrine control* is dealt with.

1. Glucose regulation in the fed state

In the *absorptive state* after a balanced meal, nutrients enter the blood and lymph from the gastrointestinal tract (as monosaccharides, triglycerides, and amino acids). All the blood passes directly to the liver, which converts most of the other monosaccharides into *glucose.* Much of the absorbed carbohydrate enters the liver cells, but little of it is oxidised; instead most is stored as glycogen. Absorbed glucose, which did not enter hepatocytes but remained in the blood, is stored as glycogen by muscle cells, or it may enter into adipose tissue. A large fraction is oxidised to CO_2 and water in the various cells of the body. Glucose is the major source of energy during the absorptive state. Homeostatic mechanisms maintain the plasma [glucose] within narrow limits in healthy humans, so that the energy needs during the postabsorptive state can be met by stored fuel.

A high glucose intake results in a high blood [glucose] or extracellular *hyperglycaemia*. Hyperglycaemia increases insulin secretion from the b-cells and inhibits glucagon secretion from the a-cells of the pancreatic islets. These hormones block *hepatic glucose production* by glycogenolysis and gluconeogenesis. *Insulin secretion* dominates over all *insulin-antagonists* (growth hormone, glucagon, cortisol and some catecholamines).

The sight and the smell of a meal triggers *cephalic* insulin secretion. When the meal reaches the intestine, several peptides of the *incretin family* are released; this is the *intestinal* secretion phase. Typical representatives of the *incretin family* are Gastric Inhibitory Peptide (GIP), glicentin (intestinal glucagon), and glucagon-like peptides (GLP-1 and -2). Incretins strongly

potentate the insulin secretion induced by the rising blood [glucose]. The *incretins* cause a much larger insulin secretion than the iv. administration of glucose, even at the same rise in blood [glucose]. This extra insulin secretion is called the *incretin effect*. The insulin released following a meal increases the storage rate of glucose-related energy in the liver, muscles and fat tissues. The storage effect is much larger than when glucose is administered intravenously.

Glucose is absorbed through the luminal membrane of the intestinal cells in *glucose-Na^+ transporter proteins*. The two substances pass through the basolateral membrane via separate routes: Glucose passes in a special glucose-transporter, and Na^+ is transferred by the Na^+-K^+-pump. Glucose transport proteins and insulin receptors are described in Chapter 1.

The filtration flux for glucose (mmol per min) increases proportionally to the concentration in the blood (as for all other filtered substances). Normally, all glucose is reabsorbed in the first part of the proximal renal tubules with a T_{max} of 1.8 mmol per min or 320 mg per min.

In other words, the passage fraction falls from one to zero already halfway through the proximal tubules. The excretion flux for glucose is *zero* in healthy humans.

Glucose appears in the urine of diabetics, who have a blood [glucose] exceeding the *appearance threshold* (10 mM).

Reabsorption of glucose over the luminal membrane of the proximal tubule cell takes place through the *glucose-Na^+ transporter*.

2. Glucose regulation in the fasting state (rest and exercise)

We keep our blood [glucose] surprisingly constant around the fasting level, considering the wide variety of daily activities.

Glucose production (gluconeogenesis and glycogenolysis) must equal *glucose combustion* in the fasting state. Thus, a precise relation must exist between the secretion of insulin and glucagon from the pancreatic islets.

In the fasting state hepatic *glycogenolysis* produces most of the glucose and the remaining glucose is produced by *gluconeogenesis*. Hepatic glucose is produced by *glycogenolysis* (glycogen breakdown to glucose) and by *gluconeogenesis* (glucose formation from glycogenic amino acids, lactate, glycerol, and small amounts of pyruvate. Muscle glycogen cannot deliver glucose to the blood, since muscle tissues lack *glucose-6-phosphatase*.

In the fasting condition a healthy adult has a blood [glucose] of *4.5-5.6* mM. With an average [glucose] of 4 mM in 15 l of extracellular fluid volume (ECV), therefore, the total glucose content in ECV is 60 mmol or 10.8 g (2 teaspoons of sugar). This amount is equal to the glucose eliminated in one hour at rest (60 mmol each hour), but during maximum exercise that same person may use *five* times more glucose.

The CNS and the erythrocytes neither synthesise nor store glucose, which is their primary fuel. Any surplus of glucose is deposited as liver and muscle *glycogen*. The liver cells contain an especially *efficient glucose transporter* (GLUT 2), and its *glucose uptake rate* cannot be increased further by insulin or by other hormones.

Any small *fall* in blood glucose releases more glucagon. During fasting *glycogenolysis, gluconeogenesis, and lipolysis* are dominant. If a normal person does not eat for 24-48 hours the CNS cells revert to combustion of ketone bodies, and a reversible condition that is similar to diabetes develops (*hunger diabetes*).

Exercise

The exercise stress on the hypothalamus increases the activity of the *sympathoadrenergic system*, which includes increased secretion of adrenaline from the adrenal medulla. Sympathoadrenergic activity inhibits the *insulin secre*

activity also increases hepatic glucose production.

We tend to increase glucagon secretion only if the blood [glucose] falls. A slight fall in blood [glucose] can occur both during exercise bouts and during prolonged exercise.

During exercise the blood [glucose] is maintained rather constant by bihormonal interplay between insulin and glucagon.

Generally, an increasing demand of more energy elicits increased glycogenolysis, lipolysis, and increased gluconeogenesis caused by insulin-antagonistic hormones (catecholamines, glucagon, cortisol, and growth hormone).

Adrenaline inhibits the insulin and stimulates the glucagon secretion, so the blood [glucose] increases.

Somatotropin - human growth hormone (GH) - is an *insulin-antagonist,* but together with insulin probably the most important *anabolic hormone.*

Conditions where energy sources are lacking are hypoglycaemia, hunger, fasting state, exhaustion, and stress. These conditions trigger the release of GRH from the hypothalamus, which in turn stimulates the release of GH from the hypophysis. This hormone has a *tropic effect* on the a -cells of the pancreatic islets. GH releases glucagon from these cells, just as sympathetic stimulation from the hypothalamus does.

GH increases blood [glucose] by increasing *hepatic glucose production* (glycogenolysis but not its gluconeogenesis) and by inhibiting the insulin sensitivity of the muscle cells and thus reduces their glucose uptake. GH also has the same effect on fat cells, mobilising fatty acids and glycerol. GH stimulates protein synthesis, mitoses, chondrogenesis, ossification, and phosphate balance, while increasing glycolysis (ie, anaerobic breakdown of glycogen).

Glucocorticoids are insulin-antagonists. They stimulate the *hepatic glucose production* (glycogenolysis and gluconeogenesis) but inhibit the cellular glucose uptake. Glucocorticoids are *permissive* and *potentiating* for catecholamines and glucagon.

Catecholamines (adrenaline & noradrenaline) are *insulin-antagonists.* Adrenaline *stimulates hepatic glucose production* (glycogenolysis). Catecholamines also stimulate lipolysis. The increase in mitochondrial oxygen uptake by T_3/T_4 is potentates by catecholamines.

The glucostat

Glucose sensitive neurons in the hypothalamus (the *glucostatic centre*) react to hypoglycaemia by releasing glucagon from the pancreatic a-cells and catecholamines from the adrenal medulla by action of the sympathetic system.

The glucostatic centre also reacts to hyperglycaemia to release insulin from pan
and to activate hepatic glycogen synthesis by vagal stimuli. Insulin promotes the entry of glucose into tissues, including the neurons of the hypothalamic glucostatic centre (but in no other brain neurons). A balanced blood [glucose] is achieved by sympathetic signals stimulating hepatic glucose production. This balance theory is called the glucostatic theory. In the glucostatic theory the hypothalamus is considered a glucostat and the liver is a unique glucose exchanger, due to the portal system and the hepatic glucose-6-phosphatase. Leptin is dealt with in Chapter 20.

Since the hypophysis hormones ACTH and GH are insulin-antagonists the net effect of the hypophysis, when not balanced by a normal pancreatic insulin secretion, is a reduced glucose tolerance.

3. The endocrine pancreas

The endocrine pancreas or the pancreatic islets are synonyms for the produc
polypeptide hormones: Glucagon, insulin, somatostatin, and pancreatic polypeptide (PP).

The one million islets of Langerhans are discrete structures scattered throughout the pancreas, but which only comprise 1% of its total weight. The islets are arranged along fenestrated capillaries, so that the hormones can pass easily to the portal blood. The islets of Langerhans receive both sympathetic (adrenergic) and parasympathetic (cholinergic) fibres.

The membranes of the islet cells contain gap junctions between neighbour cells, so hormones from one cell can act on its neighbour (paracrine action). Gap junctions allow passage of small molecules from one islet cell to its neighbour. In many pancreatic lobules, the a - b - and d - cells form a paracrine syncytium.

3 a. Glucagon

The a-cells of the pancreatic islets is the source of pancreatic glucagon. Glucagon stimulates adenylcyclase in the hepatocytes. This enzyme activates phosphorylase that breaks down glycogen. Actually, glucagon triggers a glycogenolytic cascade, so those considerable amounts of glucose are released in response to the fall in blood glucose. In addition, glucagon stimulates the hepatic production of glucose (gluconeogenesis) from glycerol, alanine and lactate. Glucagon is a direct antagonist to insulin, being catabolic in its actions (gluconeogenetic, glycogenolytic, lipolytic & ketogenic, and deaminating amino acids).

Intestinal glucagon (glicentin) is built up from 69 amino acids. The glucagon from the a-cells of the pancreatic islets only contains 29 of the 160 amino acid residues in pro-glucagon. Conditions where there is intracellular lack of glucose (hunger, insulin deficiency, protein rich meals, and amino acid infusion) liberate glucagon from the a-cells of the pancreatic islets to the pancreatic vein and then to the portal vein. Glucagon stimulates ketogenesis (formation of ketone bodies). High blood [glucose] and [FFA] inhibit glucagon secretion.

Pancreatic and intestinal (glicentin) glucagon are hepatic insulin-antagonist. Glucagon stimulates hepatic glucose production by glycogenolysis in the hepatocytes and thus increases the blood [glucose].

Glucagon also stimulates gluconeogenesis from glycogenic amino acids in the liver and thus increases urea-genesis. Glucagon stimulates ketogenesis (formation of ketone bodies). In addition to the ketogenic effect, intestinal glucagon is a potent stimulator of insulin secretion - as are other members of the incretin family. Incretins act by increasing cAMP in the b-cells.

3 b. Insulin

Banting shared the Nobel Prize with Macleod in 1923 for their work in identifying the role of insulin in the carbohydrate metabolism. Their research led to the practice of insulin therapy for diabetes.

Pre-proinsulin is the precursor of insulin. When pre-proinsulin reaches the endoplasmic reticulum, enzymes separate the molecule from the signal molecule, to form proinsulin. In the Golgi apparatus enzymes cleave proinsulin to insulin (51 amino acids in two chains: A and B) and the C peptide (Connecting peptide). Insulin and C peptide are wrapped in the same secretion granule. The content of these secretion granules is expelled from the cell by exocytosis.

When the secretory granules release proinsulin to the portal blood and later the extracellular fluid volume (ECV), connecting peptide (C-peptide) and two amino acids breaks off. The split products are carried to the liver, where half of the insulin molecules are degraded and extracted. The degradation products are broken down and eliminated by the kidney. The kidneys only eliminate c-peptide and its rate of production reflects the rate of insulin secretion. Insulin contains 51 amino acid residues in two chains (m.w. 5734).

In healthy persons the blood glucose concentration, B-[glucose], is controlled exactly. The fasting value is within the range of 4-7 mM, with minimum individual variance from day to

day, despite varying life conditions with food and exercise.

The liver is a glucose exchanger, because it absorbs glucose from the intestine, stores glucose as glycogen, and produce glucose from fat and protein residues (gluconeogenesis). The liver releases glucose to the ECV in exact proportion to the peripheral rate of glucose utilisation in the postabsorptive state (180-200 g per day).

The brain metabolism of a healthy standard person requires 100 g of glucose per day. The brain glucose is totally oxidised, ir-regardless of the insulin status.

Each meal elicits a peak of insulin secretion, because of the rise in blood [glucose] (Fig. 27-1). The blood [glucose] increases after a meal. Increasing [glucose] is a strong stimulus to the b-cells of the pancreas. Glucose enters these cells through GLUT 2. The cells empty their granules into the ECV, and the granule dissolve immediately after entering the blood. This sequence of events supplies the blood with insulin, C-peptide and proinsulin in the ratio 19:19:1.

Insulin reduces the blood glucose for the following reasons: insulin increases the cell uptake of glucose (and potassium) in most tissues (adipocytes, heart and other muscle tissue). The exceptions are the brain, kidney and erythrocytes. The uptake capacity for glucose in hepatocytes is so large, that any insulin effect is immaterial.

Insulin promotes the formation of tissue stores from circulating nutrients (the actions are all anabolic). The insulin receptor is a tetrameric protein complex with two a-units extracellularly, and two b-units traversing the membrane of target cells (ie, skeletal muscle fibres, cardiac myocytes and adipocytes). The target cells contain the glucose transporter, GLUT-4. In the absence of insulin, all the GLUT-4 units are located in the intracellular vesicles. Insulin binding to the insulin-receptor activates the tyrosine kinase, which resides in the b-units, and promotes the transport of GLUT-4 vesicles towards the surface of the cell, where they melt together with the membrane. This phenomenon increases the number of glucose-channels through the membrane and thus promotes glucose uptake into target cells. Finally, the insulin-receptor complex is internalised by the cell, insulin is broken down and the insulin-receptor recycles to the cell surface for further use.

Glucose is stored in the muscle cells as glycogen, used in the Krebs cycle or it is broken down to lactate.

In the fat cells glucose is utilised as a substrate for triglycerides synthesis. In the post-prandial phase, lipolysis liberates fatty acids (FFA) from triglyceride together with glycerol. Glycerol and lactate are substrates for hepatic gluconeogenesis.

Fig. 27-1: Effects of insulin on target cells.

Falling blood [glucose] is called hypoglycaemia, which activates the sympathoadrenergic system and deplete the glucose-dependent brain for its only fuel. Accordingly, the hypoglycaemia causes sympathoadrenergic (sweating, hunger, tremor, tachycardia), and cerebral manifestations (anxiety, disorientation, cramps, and unconsciousness). The clinical picture is that of hypoglycaemic shock.

Insulin also increases the rate of glycogen synthesis in the liver and muscles and inhibits the rate of gluconeogenesis. Insulin is a direct antagonist to glucagon being anabolic in its actions (increased glucose entry to cells, increased glycogen and lipid synthesis, decreased protein catabolism and ketogenesis).

Glucose-evoked insulin secretion is the result of a chain of events in the pancreatic b-cell (Fig. 27-2).

Fig. 27-2: Insulin release from pancreatic β-cell.

1. The glucose uptake takes place through a specific transporter protein (GLUT-2). The pancreatic b-cell membrane contains several K^+ channels, of which two are directly involved. This is the K^+-ATP channel and the maxi-K^+ channel (Fig. 27-2).

2. The hyperglycaemia accelerates the glucose uptake and metabolism and thus increases the ATP/ADP ratio.

3. Increased ATP closes the K^+-ATP channels, so the cell depolarises (hypopolarises). During hypopolarisation from the normal resting membrane potential of -70 mV, a threshold is reached at - 50 mV, where the voltage dependent Ca^{2+} channels open.

4. The Ca^{2+} influx triggers exocytosis of insulin and C-peptide containing granules following vesicular fusion with the cell membrane.

5. Normally, the maxi-K+ channel and other K+ channels stop depolarisation. When intracellular [Ca^{2+}] and [K^+] has increased, it opens the maxi-K^+ channel. The K^+ efflux restores the resting membrane potential (- 70 mV) towards the equilibrium potential of K^+ (-100 mV).

Insulin is a vital hormone. Blood from the pancreas passes through the liver, where insulin promotes the produc
substantial amount of the insulin, whereas the C-peptide passes the liver undisturbed. The plasma [C-peptide] is thus a good estimate of insulin secretion. Insulin can now be synthesized from genetically modified micro-organisms.

Insulin is an anabolic hormone.

Insulin reduces the blood [glucose] because it increases glycogen synthesis in the liver and muscles. Insulin increases the uptake of glucose through GLUT 4 (in adipocytes, heart and skeletal muscles). Insulin inhibits the gluconeogenesis from glycogenic amino acids in the liver.

Insulin promotes the storage of energy, the synthesis of glycogen, mRNA and proteins. Insulin thus reduces urea-genesis.

Insulin promotes lipogenesis in the fat stores; however, it inhibits lipolysis. It may be noted that the glycerol portion of the triglyceride molecule is a derivative of glucose.

Insulin increases the synthesis of cholesterol in the liver, in particular the rate of VLDL formation (Very Low Density Lipoprotein). The dangerous cholesterol fraction is LDL (Low Density Lipoprotein).

Insulin increases the GLUT 4 transfer of glucose and K^+ into the muscle cell interior.

Certain major tissues (kidney, brain, and intestine) are insensitive to the direct action of insulin.

3 c. Somatostatin

D-cells or d-cells are the source of somatostatin, a potent and multipotent hormone inhibitor. Somatostatin contains a disulphide bridge and 14 amino acid molecules. Somatostatin produced in the islets inhibits the local secretion of the other islet hormones, while glucagon stimulates the local release of insulin and somatostatin. Somatostatin is also produced in the hypothalamus, where it functions as the Growth Hormone Inhibiting Hormone (GHIH). Pancreatic somatostatin is released in response to high blood [glucose] and [alanine]. Somatostatin inhibits the secretion of the gastrointestinal tract (but not its motility) and functions as a synaptic transmitter in the CNS. Persons with somatostatin-producing tumours

develop diabetes and gallstones.

GHIH is synthesized both in the hypothalamic-pituitary system and in the pancreatic islets. The D-cells of the pancreatic islets of Langerhans produce GHIH, which controls the function of the other islet cells by paracrine action. Somatostatin is a multipotent hormone inhibi Somatostatin inhibits Somatotropin (GH) but also TSH, insulin, and glucagon. Somatostatin blocks the gastrin secretion in the gastric antrum. Somatostatin inhibits the secretion of digestive fluids, but increases gastrointestinal motility.

3 d. Pancreatic polypeptide

The cells responsible for pancreatic polypeptide (PP) secretion are particularly abundant in the head of the pancreas. PP contains 36 amino acid residues in a linear polypeptide. The plasma [PP] increase markedly after a protein rich meal, but it is not released by alanine infusion. The PP secretion is increased by exercise (with high plasma [alanine]), by fasting and by hyperglycaemia. The plasma [PP] is suppressed by glucose infusion. PP inhibits the exocrine pancreas and reduces the gallbladder contractions. This is an appropriate response during exercise and fasting, where any reduction in blood glucose would trigger a PP release.

Meals, rich in protein and fat, release pancreatic polypeptides (from the PP-cells).

Pancreatic polypeptide inhibits both enzyme secretions from the pancreas and the emptying of bile into the small intestine. This leads to a delay in the absorption of nutrients including glucose.

Patients with pancreatic islet cell neoplasm have elevated plasma [PP].

4. Pancreatic exocrine control

Endocrine glandular tissue is localised in the Langerhans islets that produces insulin in the b-cells, glucagon in the a-cells, somatostatin (GHIH) plus gastrin from the d- and G-cells, and pancreatic polypeptide (PP) from the P-cells (Fig. 27-3). Bombesin, galanin, and neuropeptides are present in pancreatic neurons and act as transmitters.

Stimulation of vagal fibres to the pancreas enhances the rate of enzyme secretion into the pancreatic juice. Stimulation of sympathetic fibres reduces bloodflow to the pancreas, and thus inhibits pancreatic secretion. Gastrin enhances enzyme secretion and insulin potentiates the effect, whereas somatostatin inhibits secretion from both acinar and ducts cells (Fig. 27-3: - all).

Fig. 27-3: Liberation of pancreatic islet hormones.

Pathophysiology

Diabetes mellitus (DM)

DM is a collective term for a multitude of metabolic disorders, where lack of insulin (type I diabetes) or insulin resistance (type II) dominates.

Insulin resistance is defined as insufficient sensitivity to insulin.

The diabetic condition is characterized by an abnormal glucose tolerance that is documented by the use of a glucose tolerance test.

Besides hyperglycaemia, the overall phenomena in the diabetic condition are *protein depletion* and *increased lipolysis* with deposition of lipids in the vascular walls of the brain, heart, kidneys, eyes and muscles.

All cases of DM where the cause is not fully explained are termed *primary* DM, whereas *secondary* DM is explainable and sometimes directly curable.

Secondary DM is caused by hypersecretion of one or more of the many catabolic hormones

with hyperglycaemic effect (adrenaline, noradrenaline, glucagon, glucocorticoids and growth-hormone, HG) or by total destruction of the pancreas from pancreatitis or carcinoma. The hormone disorders are phaeochromocytoma, glucagonoma, Cushing's syndrome and acromegaly.

Most of the patients with primary DM can be classified into the two groups already presented above Insulin-Dependent DM (IDDM) or type I diabetes, and Non-Insulin-Dependent DM (NIDDM) or type II diabetes.

This paragraph deals with *1. IDDM*, *2. NIDDM*, *3. Insulin shock (hypoglycaemia)*, *4. Oral and intrveneous glucose tolerance tests*, *5. treatment of diabetes mellitus*, *and* *6. Summary of the diabetic condition.*

1. IDDM or Type I diabetes

The presentation of IDDM is typically a young person with a few week history. This is a serious, life threatening metabolic disease, where continuation of life depends upon insulin treatment. The first treatment with insulin took place in Canada (1922). Until then these patients died within half a year in ketoacidotic coma. Persons, often with hereditary predisposition, are suddenly attacked by autoimmune destruction of *all* b-cells in the pancreatic islets, which results in the complete absence of insulin. This autoimmune destruction occurs more often in populations, where breast feeding is unpopular, and protein-rich cow's milk is used generally.

Lack of insulin leads to extracellular hyperglycaemia, and increased lipolysis.

The classical triad is:

1. Polyuria (osmotic diuresis due to extracellular hyperglycaemia and glucosuria).
2. Polydipsia and thirst due to the loss of salt and water.
3. Weightless due to extracellular fluid volume (ECV) depletion and the breakdown of tissue stores (ie catabolic effect of insulin deficiency) with rapid wasting.

The intracellular lack of glucose activates glycogenolysis in the liver and muscle cells; lipolysis and muscular proteolysis is accelerated. The liberated free fatty acids (FFA) are converted to ketone bodies, whereby a metabolic acidosis (ketoacidosis) is produced.

A patient with a blood (glucose) above 25 mM loses consciousness and to such degrees that contact is impossible and reactions to pain are absent (coma).

Diabetic ketoacidosis is a condition of insulin deficiency causing increased hepatic ketogenesis. This condition of uncontrolled catabolism occurs in IDDM, only.

In a young person, the first sign of IDDM can be coma with diabetic ketoacidosis - a life-threatening condition, if the patient is alone. This is particularly likely, when the patient is under the stress of intercurrent illness such as infection with high fever. Also a recognized IDDM patient can be hit by ketoacidosis during intercurrent illness, where his insulin demand is increased, or the patient may take too little insulin because of lost appetite or for any other reason.

Insulin deficiency has two consequences. First of all, the hepatic glucose release accelerates, and secondly the uptake of glucose by muscle and fat cells in the periphery is reduced. Progressive hyperglycaemia causes osmotic diuresis with loss of salt and water. The abnormally low ECV is termed the dehydrate state, and with falling blood volume also renal bloodflow falls. Insulin deficiency also accelerates the lipolysis (Fig. 27-4).

Fig. 27-4: The development of ketoacidosis during insulin deficiency (FFA = free fatty acids).

Triglycerides are liberated from adipose tissue, and the concentration of free fatty acids (FFA)

in the blood is elevated. FFA are broken down to fatty acyl carnitine within the liver cells, and this molecule is converted to acetyl CoA, which in turn reach the mitochondria, where ketone bodies (aceto-acetate, acetone, b-hydroxybutyrate) are formed (Fig. 27-4).

The breath of the patient smell by acetone, and there is ketosis in the urine. The concentration of ketone bodies in the blood passes *5 mM*, and when pH falls below 7, there is life-threatening or terminal coma. The condition is called an acute metabolic acidosis, characterized by a negative base excess. The patient tries to compensate the metabolic acidosis by hyperventilation, so-called Kussmaull-breathing.

Most patients with recent IDDM have circulating antibodies to islet cells, and tend to develop other organ-specific autoimmune disorders (Addison disease, Hashimotos thyroiditis and pernicious anaemia). Identical twins show a 40% concordance in developing IDDM, so life-style must also play a role. There is an association with HLA-DR4, if also HLA-B8 or HLA-DR3 is present.

2. NIDDM or Type II diabetes

This is a frequent type of DM in particular in populations with a sedentary life style and obesity. The incidence increases with age and development of obesity, which is reflected in the name *maturity-onset diabetes*. The prevalence is high in Afro-Caribbean and Asian population groups. The onset of NIDDM is sometimes triggered by Intercurrent illness or by pregnancy, but not by immunological reactions.

NIDDM is a complex of polygenic disorders. Certain families show an autosomal dominance. the genetic defects differ and many mutations are known. One is the gene on chromosome 7, which code for *glucokinase.* Identical twins show almost absolute concordance in development of NIDDM. The much more frequent *type II diabetes* (maturity-onset) is the result of *insulin resistance* and b-cell defects. Type II diabetes also occurs in younger persons, especially in persons with a *high fat-low muscle mass*. A strong genetic element is always present, but *inactivity* and *stress* (an inactive life style with a low endurance capacity) seems to be involved in the development of type II diabetes. - Lack of exercise predisposes one to obesity, a condition that greatly decreases *insulin sensitivity* of the target cells (adipocytes, heart and skeletal muscle tissues). Reduced glucose combustion creates hyperglycaemia. The hyperglycaemia elicits insulin secretion from defective b-cells in some patients, resulting in raised serum [insulin]. Since insulin is present, the acute complications such as ketonaemia and metabolic acidaemia (often found with IDDM) are rare in these patients. The high serum [insulin] may further down-regulate the activity of their *insulin receptors.* The insulin secreted in NIDDM patients does not increase the uptake of glucose as in normal persons. Many NIDDM patients need much more insulin for a given test effects than IDDM patients and healthy people.

An inactive life style for years, with redundant food intake, seems to be involved in the development of NIDDM in persons with a genetic predisposition. Lack of regular physical activity with development of overweight, increases the incidence if NIDDM. The *impaired glucose tolerance* is demonstrated by a *glucose tolerance test*.

The insulin secretion is abnormal in patients with NIDDM, although they typically possess half of their *β-cell mass* at autopsy. The destroyed b-cells is filled with amyloid material (islet amyloid polypeptide, IAPP). IAPP is a possible antagonist to insulin, and explain some cases of *insulin resistance*.

Many older patients with NIDDM have no symptoms, but a routine examination reveals glucosuria or a raised blood [glucose]. Other patients are tired, have minor genital infections or sugar spots on their underwear. Some patients present with established *late-complications* such as retinopathy (blindness), nephropathy, arteriosclerotic disorders (cerebrovascular insults, myocardial infarction, intermittent claudication, gangrene), susceptibility to infection or

neuropathy.

NIDDM can be caused, theoretically, by *1. β-cell defects* including genetic defects, resulting in abnormal insulin production, or by *2. target cell defects* including receptor failure. The possible defective sites in 1. and 2. have one common denominator. They are all *key proteins* (hormone, receptors and transporters). Muscular activity is required to stimulate the normal production of *key proteins*. NIDDM relates to inactive life style.

The basic problem is therefore possibly a genetically and activity dependent *defect in key protein production* in the cell interior. Actually, a genetic defect has just been demonstrated at certain steps of insulin action in a subset of patients of late-onset NIDDM.

3. Insulin shock (hypoglycaemia)

A high blood [insulin] will cause tissues to store away the available blood glucose rapidly, mainly through muscular GLUT 4, and stop simultaneously the production of new glucose.

A low blood glucose level elicits a large secretion of glucagon to the portal blood. Glucagon is the most important insulin-antagonist. Glucagon increases hepatic glucose production (enhancing glycogenolysis, gluconeogenesis and protein breakdown). Low glucose levels trigger glucagon production, even from denervated, pancreatic a- cells; hence, they must be glucose sensitive.

An increased catecholamine secretion from sympathetic nerve endings (NA) and Ad from the adrenal medulla (elicited from the hypothalamic glucostat via the sympathetic nervous system) helps within minutes to compensate, as catecholamines stimulate glycogenolysis, increase lipolysis and inhibit peripheral glucose uptake. Hours later, cortisol and GH also contribute. An appropriate rise of plasma [cortisol] in response to insulin-induced hyperglycaemia documents an intact CRH-ACTH-adrenal axis, and this is the most widely used stress test.

Adrenergic effects, such as trembling fingers, tachycardia, and muscular stiffness warn the hypoglycaemic patient. The glucose consumption by the heart and brain continues. The lack of glucose in the brain makes the patient uneasy at first; he is then inse
sweat. Later in hypoglycaemia, the patient becomes confused, furious and denies with slurred speech to take glucose.

Blood [glucose] below 2.5 mM elicits hypoglycaemic shock with loss of consciousness (somnolence, sopor or coma), universal cramps and respiratory stop (Fig. 27-5).

Intravenous injection of glucose (50%) is the rational therapy for hypoglycaemic coma. The patients wake up almost immediately, and are then often in a hyperglycaemic state.

Fig. 27-5: Consequences of hypoglycaemic shock.

The b-cell defects are insufficiently described. Type 2 diabetics do not produce sufficient levels of ATP in the pancreatic b-cells to completely block the K^+ -ATP channels (Fig. 27-2). Thus, the b-cell does not hypopolarize adequately in response to hyperglycaemia. Therefore, the voltage dependent Ca^{2+} -channels is insufficiently activated, and intracellular [Ca^{2+}] do not increase enough to trigger the insulin exocytosis needed. Sulfonylurea compounds close the K^+-ATP-channels and thus help to treat type 2 diabetes.

4. Oral and intravenous glucose tolerance tests

The test is performed orally or intravenously.

Oral test. The patient drinks a glucose solution containing 75 g glucose within four minutes (WHO). The blood concentration in venous plasma ([glucose]) is followed over 3 hours by blood sampling.

Normal individuals start from a low fasting [glucose] such as 5.5 mM or less. The blood

[glucose] peaks after one hour and *returns to normal within two hours* (less than 6.7 mM in Fig. 27-5). Persons with *impaired glucose tolerance* start with a fasting level less than 7.8 mM, and after 2 hours the level is 7.8-11.1 mM.

Fig. 27-5: Oral glucose tolerance curves for a normal person, a diabetic, a patient with hyperthyroidism and a patient with myxoedema.

The *diabetic patient* typically starts from a high fasting [glucose], such as values above 7.8 mM, and increase to a very high level. The blood [glucose] is not back to normal within two hours, but stays above 11.1 mM. This test pattern is the clinical WHO criterion of diabetes (Fig. 27-5).

A patient with hyperthyroidism (Graves disease or Morbus Basedowii) has a rapid intestinal absorption and a rapid combustion of glucose. The myxoedematous patient has a slow absorption and a slow combustion of glucose.

Fig. 27-6: Results of i.v. glucose tolerance tests from a normal person and from a diabetic.

Intravenous (i.v.) test. We inject 25 g glucose intravenously over a period of 4 minutes. We then measure the blood [glucose] every 10 min for at least an hour in order to determine the half-life from a semi-logarithmic plot (Fig. 27-6). The metabolic combustion rate for glucose is exponential, so it is easy to calculate the metabolic rate constant (k) expressed in percentages.

Note that the *metabolic rate constant k* here is the amount of glucose combusted divided by the total amount of glucose in a mainly extracellular distribution volume. The half-life ($T_{1/2}$) is equal to 0.693/k.

All glucose combustion rates *above 1.2% per min* are normal.

Fig. 27-7: Intravenous Glucose Tolerance Test

5. Treatment of diabetes mellitus

A normal person with three meals per day will have *three* peak concentrations of glucose and insulin in his blood. It is possible to obtain such a time profile in a diabetic person by the following strategy. Inject a fast-acting insulin three times a day just before meals and a slow-acting insulin at night. This is the *physiologic* principle. The aim of this procedure is to reduce the number of acute and chronic complications for diabetics.

All diabetics are recommended to eat healthy just like anyone else. When diet alone is insufficient to achieve a satisfactory blood glucose, a slim type II diabetic is treated with sulphonylurea compounds. The obese type II diabetic is treated with a biguanide called metformin. Patients, who have been in metabolic acidosis, are usually treated best with insulin.

6. Summary of the diabetic condition

A poorly controlled diabetic condition leads to extracellular hyperglycaemia, glucosuria, metabolic acidosis, polyuria (osmotic diuresis), dehydration and polydipsia. The osmotic diuresis leads to the excretion of Na^+ and water, which results in Na^+ and ECV depletion.

Intracellular lack of glucose activates glycogenolysis in the liver and muscles, and accelerates muscular proteolysis and lipolysis. This liberates free fatty acids, which are converted to ketone bodies.

A patient with hyperglycaemia above 25 mM loses consciousness to such a degree that contact is impossible (ie, coma).

The increased rate of cholesterol production increases the occurrence of atherosclerosis and of diabetic nephropathy. Albuminuria, hypertension and low GFR characterise diabetic

nephropathy.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Peptides and protein hormones are lipophobic.

B. Infertility is a diagnosis used on a couple, which have been unable to conceive during one year of unprotected intercourse.

C. Chronic hypoadrenalism is also called Addison's disease.

D. Nephrogenic diabetes insipidus, is a condition where the renal cells are resistant to ADH.

E. Glycerol and lactate are substrates for hepatic gluconeogenesis.

II. Each of the following five statements have True/False options:

A. Receptors are frequently glycosylated, so one signal molecule linked to a receptor is always enough to elicit a response.

B. The incidence of atherosclerosis is correlated to the total [cholesterol], the [LDL], and the [LDL]/ [HDL] ratio in blood plasma.

C. Oestrogens, exercise, nicotinic acid and alcohol increase the plasma-HDL.

D. Omeprazole stimulates the luminal gastric proton-pump.

E. Prostaglandins are primarily paracrine hormones, which act through receptors linked to G-proteins.

Case History A

A 19 year old male, body weight 80 kg, was suddenly complaining of fever during his work and ordered home to bed. The patient was living alone. Fortunately a colleague visited him the next morning. He had to break the door down and found the patient unconscious. The patient arrived at the hospital in deep coma. A blood sample from the radial artery showed the following results: β-cell antibodies, P_{aCO_2} 27 mmHg, pH 7.21, actual bicarbonate 10.5 mM, O_2 saturation 0.96 and [glucose] 32 mM (5.75 g/L). The Base Excess is -15 mM in the extracellular fluid volume (see Fig. 17-12). The urine contained glucose and ketone bodies.

The patient's breathing was deep and fast, his heart rate was 115 beats/min, and his blood pressure was 90/55 mmHg. The mucous membranes of the mouth were dry and the tonsils were enlarged and infected. The rectal temperature was 39.9 Centigrade.

1. Explain the condition of the patient concerning thermo-balance, carbohydrate metabolism, acid-base balance and fluid balance.

2. Describe a rational treatment of the four homeostatic disturbances mentioned in 1.

3. Following eight hours of treatment the blood pH was 7.41 and P_{aCO_2} 42 mmHg, but the

patient is still hyperventilating. Explain why.

4. Following 24 hours of treatment all blood gas values were normal and the patient was resituated. Why did the patient stop to hyperventilate?

Case History B

A female of 49 years and with a height of 1.52 m is in hospital due to obesity and related problems. Her weight is 74 kg. She has developed skin mycoses and multiple boils. In the morning (fasting state) a blood sample shows a blood [glucose] of 7.4 mM, but glucose is not found in the urine.

At the hospital her total body water is measured following intravenous injection of radioactive water (5.5×10^7 Becquerel or Bq tritium water). Her bladder is emptied at the time of injection. Two hours later the bladder is drained for 95 ml of urine with a concentration of 1 598 400 Bq per l.

At this time the indicator is evenly distributed in the total body water with a concentration of 1 520 700 Bq per l. The amount of indicator lost in the urine is 2/3 of the total loss.

1. What is the principle for estimation of total body water?
2. Calculate the total body water in litres and in fraction of her body weight.
3. Is this a normal result?
4. The patient is obese, but is this a serious overweight?
5. Does she have symptoms and signs of complications?

Case History C

A 23-year old male was saved after 30 days in the ruins of a house following earth quake. There was no food but sufficient water. At the arrival to the hospital the patient was in syncope with frequent, deep respiration, and the expired air smelled of acetone. The skin was dirty with brown pigmentation. The cardiac rate was 85 bpm, and the arterial blood pressure was 11.3/7.3 kPa (85/55 mmHg).

The blood [glucose] was 2.2 mM, and the plasma [FFA] was increased. The serum concentrations of proteins and essential amino acids were reduced. The blood [haemoglobin] was 95 g l^{-1}. There was moderate antidiuresis with ketonuria with signs of water retention and a high nitrogen loss in the urine.

The patient was treated with parenteral administration of glucose, amino acids and electrolytes. Following the glucose intake, the blood [glucose] was increased to 10 mM, and glucosuria occurred. A glucose tolerance test was performed and resulted in a high blood [glucose] level that had not reached the normal level within 2 hours.

1. Describe the energetic events leading to survival.
2. Why did the patient smell of acetone?
3. What happened to the carbohydrate metabolism of the patient?
4. Explain the high nitrogen loss in the urine.

See answers

Highlights

- *Glucose is absorbed through the luminal membrane of the intestinal cells in glucose-Na^+ transporter proteins. The two substances pass through the basolateral membrane via separate routes: Glucose passes in a special glucose-transporter, and Na^+ is transferred by the Na^+-K^+-pump.*
- *Somatotropin - human growth hormone (GH) - is an insulin-antagonist, but together with insulin probably the most important anabolic hormone.*
- *Glucose sensitive neurons in the hypothalamus (the glucostatic centre) react to hypoglycaemia by releasing glucagon from the pancreatic α-cells and catecholamines from the adrenal medulla by action of the sympathetic system.*
- *Since the hypophysis hormones ACTH and GH are insulin-antagonists the net effect of the hypophysis, when not balanced by a normal pancreatic insulin secretion, is a reduced glucose tolerance.*
- *The endocrine pancreas or the pancreatic islets are synonyms for the produc four polypeptide hormones: Glucagon, insulin, somatostatin, and pancreatic polypeptide (PP).*
- *Insulin is synthesized as proinsulin, which is stored in granules close to the cell membrane of the β-cells of the pancreatic islets. When the secretory granules release proinsulin to the portal blood and later the extracellular fluid volume, connecting peptide (C-peptide) and two amino acids breaks off.*
- *A poorly controlled diabetic condition leads to extracellular hyperglycaemia, glucosuria, metabolic acidosis, polyuria (osmotic diuresis), dehydration and polydipsia. The osmotic diuresis leads to the excretion of Na^+ and water, which results in Na^+ and ECV depletion.*
- *Intracellular lack of glucose activates glycogenolysis in the liver and muscles, and accelerates muscular proteolysis and lipolysis. This liberates free fatty acids, which are converted to ketone bodies.*
- *A patient with hyperglycaemia above 25 mM loses consciousness to such a degree that contact is impossible (ie, coma).*
- *The increased rate of cholesterol production increases the occurrence of atherosclerosis and of diabetic nephropathy.*
- *Albuminuria, hypertension and low glomerular filtration rate characterise diabetic nephropathy.*

Further Reading

Almind, K., C. Bjørbæk, H. Vestergaard, T. Hansen, S. Echwald, and O. Pedersen. "Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus." *The Lancet* 342: 828-832, 1993.

Ashcroft, F.M. and S.J.H. Ashcroft. "Insulin." *IRL Press at Oxford Univ*. Press, Oxford 1992.

Banting, F.G. and C. H. Best. "Internal secretion of pancreas." *J. Lab. and Clin. Med.* 7: 251-

326, 1922.

Flatt, P.R. "Nutrient regulation of insulin secretion." *Portland Press Ltd.*, London 1991.

Return to top

Return to Content

Chapter 28.

Thyroid Hormones And Disorders

Study Objectives

- To *define* concepts such as biological and physical half-life, cretinism, Graves disease, hypothyroidism, osteoporosis, PTH, TRH, TSH, thyrotoxicosis and struma.
- To *describe* the synthesis of thyroid hormones, the iodine balance, the iodide trap, the endocytosis of colloid, the transport of thyroid hormone in plasma,
- To *draw* a model of the functioning thyroid follicle.
- To *explain* the thyroid hormone feedback control, the thyroid hormone metabolism, the mechanisms of effect of the thyroid hormones, and the tests of thyroid function. To explain hyperthyroidism, hypothyroidism, myxoedema, struma, iodine deficiency, cretinism, therapy of the disorders, and use of thyroid hormone in doping and obesity.
- To *use* the above concepts in problem solving and case histories

Principles

- *The thyroid gland maintains the metabolic level of almost all cells in the body.*
- *Thyroid hormones are essential for normal neural development, linear bone growth and proper sexual maturation*.

Definitions

- **Biological half-life** refers to the rate of elimination of biologically active substances (hormones) by 50%.
- **Calcitonin** is produced by the parafollicular C-cells of the thyroid. The hormone inhibits bone resorption by blocking the parathyroid hormone (PTH)-receptors on the osteoclasts. Calcitonin is important in bone remodelling and in treatment of osteoporosis.
- **Cretinism** refers to a clinical condition caused by congenital hypothyroidism or infantile iodide deficiency. The clinical picture is a *cretin* or a mentally retarded hypothyroid dwarf.
- **Exophtalmus** refers to bulging eyes - a sign, which is part of the thyroid eye disease.
- **Goitre** *(struma)* is a visible or palpable enlargement of the thyroid gland.
- **Grave's disease** or *Morbus Basedowii* is the combination of *thyrotoxicosis, struma and exophtalmus.*
- **Hypothyroidism** is an abnormally low activity of the thyroid gland with low circulating thyroid hormone levels caused by thyroid disease.

- **Myxoedema** is a severe thyroid gland hypothyroidism in adults with a puffy swollen face due to a hard, non-pitting oedema called myxoedema or tortoise skin.
- **Osteoporosis** (bone waste) is a term used for a marked reduction in all components of the bone mass.
- **Parathyroid hormone** (PTH) refers to a single chain peptide hormone produced by the chief (C)-cells of the parathyroid glands. PTH accelerates osteolysis from bones, reduces the reabsorption of Ca^{2+} and phosphate from the proximal renal tubules and increases the reabsorption of Ca^{2+} from the distal tubules, and stimulates the renal production of biologically active vitamin D.
- **Thyroid releasing hormone** (TRH) is released from the hypothalamus and reaches the adenohypophysis via the portal system. Here, the thyrotropic cells are stimulated to produce TSH.
- **Thyroid stimulating hormone** (TSH) is released from the thyrotropic cells of the adenohypophysis to the systemic blood by which it travels to the thyroid gland.
- **Thyrotoxicosis** *(hyperthyroidism)* is a condition probably caused by *TSH-receptor antibodies*, which bind to the thyroid follicle cells and stimulate the gland to secrete T_3 and T_4. The rise in thyroid hormone concentration will *suppress TSH secretion.*
- **Physical half-life** refers to the rate of 50 % disintegration of radioactive isotopes.

Essentials

This paragraph deals with *1.The thyroid gland, 2. Synthesis and release of thyroid hormones, 3. Control of thyroid gland activity, 4. Metabolism of thyroid hormones, 5. Actions of thyroid hormones, and 6. Calcitonin.*

1. The thyroid gland

The thyroid gland maintains the metabolic level of almost all cells in the body by producing, in its follicular cells, two thyroid hormones: *triiodothyronine* (T_3), and *tetraiodothyronine* (T_4) or *thyroxine*. Iodine (I_2) has an atomic weight of 127 and a molecular weight of 254; T_4 has a molecular weight of 777 Daltons of which 508 is iodide.

Thyroid hormones are essential for normal *neural* development, linear *bone growth,* and proper *sexual maturation*.

Parafollicular cells called C-cells are located close to the follicular cells. *C-cells* produce the polypeptide hormone, *calcitonin.*

2. Synthesis and release of thyroid hormones

Thyroid hormones are synthesised in adults as long as the *dietary* iodine (I_2) supersedes *75 µg daily*. This is an adequate supply to prevent goitre formation. The daily ingestion of iodide is 400-500 mg daily in many areas and the same amount is excreted in the urine in a steady state.

The synthesis in the thyroid gland takes place in the following way:

A. *Dietary iodine* (I_2) is reduced to *iodide* (I^-) in the stomach and gut is rapidly absorbed and circulates as iodide (Fig. 28-1).

Fig. 28-1: The production and secretion of thyroid hormones.

B. *Follicular cells* in the thyroid gland possess an active *iodide trap* that requires and concentrates iodide from the circulating blood (Fig. 28-1). Iodide is transported into the cell against an electrochemical gradient (more than 50 mV) by a Na^+-I^--symport. The iodide pump is linked to a *Na^+-K^+-pump,* which requires energy in the form of oxidative phosphorylation (ATP) and is inhibited by ouabain. The thyroid absorption of iodide is also inhibited by negative ions (such as perchlorate, pertechnetate, thiocyanate and nitrate), because they compete with the iodide at the trap. In the follicular cell, iodide passes down its electrochemical gradient through the apical membrane and into the follicular colloid. Iodide is instantly oxidised – with hydrogen peroxide as oxidant - by a *thyroid peroxidase* to atomic or molecular iodine (I_0 or I_2) at the colloid surface of the apical membrane. Thiouracil and sulfonamides block this peroxidase.

C. The *rough endoplasmic reticulum* synthesises a large storage molecule called *thyroglobulin*. This compound is build up by a long peptide chain with tyrosine units and a carbohydrate unit completed by the Golgi apparatus. Iodide-free thyroglobulin is transported in *vesicles* to the apical membrane, where they fuse with the membrane and finally release thyroglobulin at the apical membrane.

D. At the apical membrane the oxidised iodide is attached to the tyrosine units (L-tyrosine) in thyroglobulin at one or two positions, forming the hormone precursors *mono-iodotyrosine* (MIT), and *di-iodotyrosine* (DIT), respectively. This and the following reactions are dependent on *thyroid peroxidase* in the presence of hydrogen peroxide -both located at the apical membrane. As MIT couples to DIT it produces *tri-iodothyronine* (3,5,3`-$T_3$), whereas two DIT molecules form *tetra-iodothyronine* ($T_4$), or *thyroxine*. These two molecules are the two thyroid hormones. Small amounts of the inactive *reverse* $T_3$ (3,3`,5`- T_3) is also synthesised.

E. Each thyroglobulin molecule contains up to 4 residues of T_4 and zero to one T_3. Thyroglobulin is retrieved back into the follicular cell as *colloid droplets* by *pinocytosis*. Pseudopods engulf a pocket of colloid. These colloid droplets pass towards the basal membrane and fuse with *lysosomes* forming *phagolysosomes.*

F. Lysosomal exopeptidases break the binding between thyroglobulin and T_4 (or T_3). Large quantities of T_4 are released to the capillary blood. Only minor quantities of T_3

are secreted from the thyroid gland.

G. The proteolysis of thyroglobulin also releases MIT and DIT. These molecules are deiodinated by the enzyme deiodinase, whereby iodide can be reused into T_4 or T_3. Normally, only few intact thyroglobulin molecules leave the follicular cells.

H. TSH stimulates almost all processes involved in thyroid hormone synthesis and secretion.

3. Control of thyroid gland activity

The hypothalamic-pituitary-thyroid axis controls the thyroid gland function and growth.

a. The production and release of thyroid hormone is controlled by *thyroid-releasing hormone* (TRH) from the hypothalamus (Fig. 28-2).

TRH reaches the anterior pituitary via the portal system, where the thyrotropic cells are stimulated to produce *thyroid-stimulating hormone* (TSH) or *thyrotropin*. TSH is the only known regulator of thyroid hormone secretion in humans. TSH is released to the systemic blood, by which it travels to the thyroid gland (Fig. 28-2). Here, TSH stimulates the uptake of iodide, and all other processes that promote formation and release of T_4 (and T_3). TSH activates *adenylcyclase* bound to the cell membranes of the follicular cells and increases their cAMP. T_3 has a strong *inhibitory* effect on TRH secretion, as well as on the expression of the

gene for the TRH precursor.

Fig. 28-2: The negative feedback control of the hypothalamic-pituitary-thyroid axis.

b. Almost all *circulating* T_3 is derived from T_4. TSH also stimulates the conversion of T_4 to the more *biologically active* T_3. Most of the circulating thyroid hormones are bound to plasma proteins, whereby the hormone is protected during transport. There is an equilibrium between the pool of protein-bound thyroid hormone and the free, biologically active forms (T_3 and T_4) that can enter the body cells. Thyroid hormones are lipid-soluble and they can easily cross the cellular membrane by diffusion.

c. Inside the cell, T_3 binds to *nuclear receptors* and stimulates cellular metabolism and increases *metabolic rate*.

d. The concentrations of T_3 and T_4 in the blood are recorded by pituitary and hypothalamic receptors. This *negative feedback system* keeps the blood concentrations within normal limits, and there is only a minimal nocturnal increase in TSH secretion and T_4 release.

4. Metabolism of thyroid hormones

In the blood we have only small amounts of *thyroxine-binding globulin* (TBG; approximately 10 mg per l), but the affinity for T_4 is high. The total T_4 is 10^{-7} mol per l equal to 77.7 mg per l of blood serum, because 777 g of T_4 equals one mol. out of the total. Approximately 70% of T_4 and T_3 binds to TBG, and the rest to *thyroxine-binding albumin* (TBA) and to *transthyrenin*. Oestrogens stimulate the synthesis of TBG.

The T_3 hormone is elimina
protein binding. The thyroxine (T_4) molecule has a biological half-life of 7 days, almost equal to the physical half-life of the radioactive isotope ^{131}I (8 days).

T_4 is likely to be a prohormone, which is deiodinised by *monodeiodinase* to the more potent T_3 just before it is used in the cells. Thus T_3 is probably the final active hormone, although it is present only in a very low concentration (10^{-9} mol per l).

Most of the daily T_4 released from the thyroid gland undergoes deiodination, with subsequent deamination and decarboxylation. Some of the hormone molecules are coupled to sulphate and glucuronic acid in the liver and are excreted in the bile. In the intestine most of the coupled molecules are hydrolysed, and the hormones are reabsorbed by the blood, whereby they reach hepar again (the *enterohepatic circuit*).

5. Actions of thyroid hormones

Thyroid hormones are lipid-soluble and pass through cell membranes easily. T_3 binds to specific *nuclear receptor proteins* with an affinity that is tenfold greater than the affinity for T_4. The information alters *DNA transcription* into mRNA, and the information is eventually *translated* into many *effector proteins*. One type of thyroid receptor protein is bound to thyroid regulatory elements in target cell genes.

Important cellular constituents are stimulated by T_3: The mitochondria, the Na^+-K^+-pump, myosin ATPase, adrenergic b-receptors, many enzyme systems and proteins for growth and maturation including CNS development.

Thyroid hormones *stimulate* oxygen consumption in almost all cells.

Thyroid hormones stimulate the rate of 1) hepatic glucose output and peripheral glucose utilisation, 2) hepatic metabolism of fatty acids, cholesterol and triglycerides, 3) the synthesis

of important proteins (the Na^+-K^+-pump, respiratory enzymes, erythropoietin, b-adrenergic receptors, sex hormones, growth factors etc), 4) the absorption of carbohydrates in the intestine and the gut excretion of cholesterol, and 5) the modulation of reproductive function.

The many rate-stimulating effects are summarized in an overall increase in *oxygen consumption*. This slow - but long lasting - *calorigenic* and *thermogenic* effect is confined to the *mitochondria*.

The thyroid hormones and the catecholamines work together in metabolic acceleration.

Thyroid hormones increase cardiac rate and output as well as ventilation.

The high basal metabolic rate raises the core and shell temperature, so that the peripheral vessels dilatate. This vasodilatation forces the cardiac output to increase. A circulatory shock develops, if the rise in cardiac output is insufficient to match the vasodilatation - socalled *high output failure*.

A human body overloaded with thyroid hormones for a prolonged period (*hyperthyroidism*) will suffer from muscle atrophia, bone destruction and hunger damage, due to increased catabolism of cellular proteins and fat. Eventually *hypothyroidism* may develop due to suppression.

6. Calcitonin

is produced by the parafollicular C-cells of the thyroid. Calcitonin inhibits bone resorption by blocking the parathyroid hormone (PTH)-receptors on the osteoclasts. The result is an extremely effective lowering of plasma-[Ca^{2+}] and -[phosphate]. Calcitonin is important in bone remodelling and in treatment of *osteoporosis*.

Calcitonin is a single-chain peptide with a disulphide ring, containing 32 amino acids. Calcitonin is secreted from the thyroid gland in response to hypercalcaemia and it acts to lower plasma [Ca^{2+}], as opposed to the effect of PTH.

Administration of calcitonin leads to a rapid fall in plasma [Ca^{2+}]. *Calcitonin* is the *physiologic antagonist* to PTH and inhibits Ca^{2+} -liberation from bone (ie, inhibits both osteolysis by osteocytes and bone resorption by osteoclasts). But calcitonin reduces plasma phosphate just as PTH.

Calcitonin probably inhibits reabsorption of phospha
calcitonin also inhibits the renal reabsorp
gut absorption of Ca^{2+} and promote phosphate entrance into bone and cause important bone remodelling.

Calcitonin deficiency does not lead to hypercalcaemia, and *excess* calcitonin from tumours does not lead to hypocalcaemia. Therefore, most effects of calcitonin are evidently offset by appropriate regulation through the actions of PTH and vitamin D.

Calcitonin in plasma declines with age and is lower in women than in men. Low levels of calcitonin are involved in accelerated bone loss with age and after menopause (*osteoporosis*).

Calcitonin protects the *female skeleton* from the drain of Ca^{2+} during pregnancy and lactation.

Calcitonin is a *neurotransmitter* in the hypothalamus and in other CNS locations.

Calcitonin is administered to postmenopausal females in attempt to prevent osteoporosis.

Pathophysiology

This paragraph deals with *1. Hyperthyroidism*, *2. Hypothyroidism*, *3. Struma*, *and 4. Thyroid medullary carcinoma*.

1. Hyperthyroidism

The *classical hyperthyroidism* or thyrotoxicosis (Graves thyroiditis, Basedows disease) is a condition characterized by an abnormal rise in basal metabolic rate, struma and eye signs (thyroid eye disease). The eyes of the patient typically bulge (ie, *exophtalmus*). Patients with *thyrotoxicosis* have overwhelmingly high metabolic rates.

Thyroid eye disease (with exophtalmus) is not confined to Graves's hyperthyroidism only. Some exophtalmus patients are euthyroid or hypothyroid. Common to all types of thyroid eye diseases are *specific antibodies* that cause inflammation of the *retro-orbital tissue* with swelling of the extraocular eye muscles, so they cannot move the eyes normally. Proptosis and lid lags are typical signs, and conjunctivitis and scars on the cornea follow due to lack of protective cover. The oedematous retro-orbital tissue may force the eye balls forward and press on the optic nerve to such an extent that vision is impaired or blindness results. The best treatment is to normalise the accompanying thyrotoxicosis. Other therapeutic measures are palliative.

TSH receptor antibody (IgG antibodies) release causes Graves's disease from activated B-cells (Fig. 28-3). A genetic deficiency is involved, which is shown by the 50% concordance in monozygotic twins. Trigger mechanisms are presumed to be bacterial or viral infections producing autoimmune phenomena in genetically deficient individuals. The autoimmune system can produce the following autoantibodies:

1. *TSH-receptor antibodies* to the TSH receptors (antigens) on the surface of the thyroid follicular cells, which they stimulate just like TSH itself, causing thyroid hypersecretion. These IgG antibodies are also termed long-acting thyroid stimulator.
2. *Specific autoantibodies* causing retro-orbital inflammation and thyroid eye disease.
3. *Thyroglobin antibodies* against the storage molecule, thyroglobin.
4. *Microsomal antibodies* against thyroid peroxidase.

These autoantibodies can be found in the plasma of most cases of Grave's disease.

Fig. 28-3:The pathogenesis of Graves disease, and the clinical manifestations of Graves's disease.

The increased *metabolic rate* and *sympatho-adrenergic activity* dominate the patient. The patient is anxious with warm and sweaty skin, tachycardia, palpitations, fine finger tremor, and pretibial myxoedema (ie, accumulation of mucopolysaccharides - see Fig. 28-3). Typically is a symmetrical, warm pulsating goitre. Lean hyperthyroid females - like female distance runners - have small fat stores and greatly reduced menstrual bleedings (*oligomenorr amenorrhoea*. The high T_3 level increases the density of b-adrenergic receptors on the myocardial cells. The cardiac output is high even at rest and arrhythmias are frequent (eg, atrial fibrillation).

Elderly patients may present with an *apathetic hyperthyroidism,* where they complain of tiredness and somnolence. Measurement of serum TSH with T_3/T_4 reveals that the diagnosis is not hypo- but *hyperthyroidism*. Erroneous treatment with thyroid hormone can kill the patient by causing vasodilatation and *cardiac output failure*.

A suppressed serum TSH confirms the diagnosis of hyperthyroidism, and the serum T3 or T4 is raised.

Several drugs are used in the treatment of hyperthyroidism.

Carbimazole and *methimazole* inhibit the production of thyroid hormone and have immuno-suppressive actions.

Monovalent anions and ouabain *inhibit* the iodide trap.

Thiocarbamide *inhibits* the iodination of tyrosyl residues.

Sulphonamides *inhibit* thyroid peroxidase, which oxidises iodide to iodine.

Large doses of iodide *inhibit* the TSH-receptors on the thyroid gland.

The high activity of the sympatho-adrenergic system is inhibited by *β-blockers*, preferably with central sedative effects.

Subtotal thyroidectomy is used to treat patients with a large goitre, or patients with severe side effects to drug therapy.

Radioactive iodine is stored in the gland and destroys the follicle cells. This therapy is complicated, and some patients develop hypothyroidism.

Toxic goitre and *toxic solitary adenoma* (Plummers disease) are cases of *secondary hyperthyroidism* just as inflammation in *acute thyroiditis* and *chronic thyroiditis.* The cells secrete thyroid hormone without inhibition from the hypothalamo-pituitary axis.

2. Hypothyroidism

Primary hypothyroidism is an abnormally low activity of the thyroid gland with low circulating thyroid hormone levels caused by thyroid disease. *Secondary hypothyroidism* results from hypothalamic-pituitary disease.

Primary hypothyroidism is caused by *microsomal autoantibodies* precipitated in the glandular tissue. Lymphoid infiltration of the thyroid may eventually lead to *atrophy* with abnormally low production of T_4. Another clinical form starts out as *Hashimotos thyroiditis,* often with hyperthyroidism and goitre. Following atrophy caused by *microsomal autoantibodies,* the condition ends as hypothyroidism, or the patient is euthyroid.

When hypothyroidism is congenital both physical and mental development is impaired and *cretinism* is the result. Also iodide deficiency in childhood may also result in a *cretin* or a mentally retarded hypothyroid dwarf.

Myxoedema in the adult is severe thyroid gland hypothyroidism with a puffy swollen face due to a hard, non-pitting oedema (called *myxoedema* or *tortoise skin*). The skin is dry and cold; there is bradycardia, often cardiomegaly (ie, myxoedema heart), hair loss, constipation, muscle weakness and anovulatory cycles in females. A *high* TSH level and a *low* total or free T_4 in plasma confirms the diagnosis primary hypothyroidism. Thyroid autoantibodies are usually demonstrable in the plasma. Hypercholesterolaemia and increased concentrations of liver and muscle enzymes (aspartate transferase, creatine kinase) in the plasma is typical.

As stated thyroid gland high TSH characterises hypothyroidism. A test dose of TSH to a patient with thyroid hypothyroidism will not stimulate the thyroid gland.

A test dose of TRH will result in an increased TSH response in *thyroid gland hypothyroidism* and decrease in *hyperthyroidism*. This is due to the negative feedback of thyroid hormones on the hypophysis. Hypothyroid females often have ex
(menorrhagia and polymenorrhoea). Hypothyroid patients exhibit slow cardiac activity.

Secondary hypothyroidism is caused by reduced TSH drive due to pituitary or hypothalamic insufficiency. A test dose of TRH to a myxoedema patient with hypothalamic or pituitary insufficiency will result in a normal TSH response.

Replacement is given to the hypothyroid patient with approximately 100 mg T_4 daily for the rest of the patients life.

3. Struma

Struma is a *visible* or *palpable enlargement* of the thyroid. Struma is due to iodine *deficiency*,

increased iodine *demand* or *strumagens*. Any prolonged TSH stimulation results in an enlarged thyroid.

Diseases in the thyroid gland including struma are caused by malfunction in the gland itself or by hypothalamic-pituitary *defects*.

4. Thyroid medullary carcinoma

Mutations of a gene located on chromosome 10 can produce an error in receptor tyrosine kinase proto-oncogene associated with *thyroid medullary carcinoma*.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. Mutations of a gene located on chromosome 10 can produce a change in receptor tyrosine kinase proto-oncogene associated with thyroid medullary carcinoma.

B. Thyroid hormones are water-soluble, which is why they pass through cell membranes quite easily.

C. Struma is due to iodine deficiency, increased iodine demand or strumagens.

D. Calcitonin is secreted from the thyroid gland in response to hypercalcaemia and it acts to lower plasma [Ca^{2+}] as opposed to the effect of PTH.

E. Tri-iodothyronine has a strong stimulatory effect on TRH secretion.

Case History A

A female, 62 years of age, suffers from pernicious anaemia for which she has received 1 (one) mg cyanocobalamine intramuscularly every 3.month for the last 10 years. At a routine visit the patient is found with a puffy swollen face due to a non-pitting oedema. Her skin is dry and cold, the heart rate is 55 beats per min, her hair is sparse, and she complains of constipation and fatigue. A series of blood tests reveals the following: High levels of microsomal autoantibodies against the thyroid gland and autoantibodies against her parietal cells. The TSH concentration in the plasma is high, whereas the T_4 is low. The haematological variables are satisfying.

1. What is the probable diagnosis?

2. What are the treatment?

3. Is there any connection between pernicious anaemia and the other condition?

Case History B

A 49-year-old female (weight 52 kg; height 1.69 m) is in hospital and is being examined for thyroid disease. Her distribution volume for iodine is 12 l and her renal clearance is 36 ml plasma per min. In a period where her iodine intake equals her output, she is subjected to the following test. In the morning she receives a small dose of the radioactive isotope ^{131}I, and three (3) hours later she urinates. From that moment she collects her urine for the following two (2) hours. The urine collection has a volume of 0.2 l and an iodine concentration of 65 μg per l. The total urine radioactivity is 1.6×10^7 disintegration per s (Becquerel or Bq). During the two hour test period, her plasma concentration of ^{131}I falls from 38 300 to 26 100 Bq per l. The radioactivity in the thyroid gland (measured with a scintillation counter) increases during the test by 77 500 Bq.

1. *Calculate the concentration of iodide in her plasma at the start of the test and at the end of the test.*

2. *Calculate the uptake of iodide in the thyroid during the 2-hour test and compare the result with a mean value of 2.4 μg per hour for healthy persons.*

3. *Calculate the thyroid plasma clearance for iodide and compare the result to the expected value of 10 ml per min.*

4. *Calculate the elimination rate constant for iodide.*

5. *Calculate the biological half-life for iodide in its distribution volume and compare the result to the physical half-life of ^{131}I (8 days).*

See answers

Highlights

- *T_4 is likely to be a prohormone, which is deiodinised by monodeiodinase to the more potent T_3 just before it is used in the cells. Thus T_3 is probably the final hormone, although it is present only in a very low concentration (10^{-9} mol per l).*

- *Thyroid hormones are synthesised in adult persons as long as the dietary iodine (I_2) supersedes 75 μg daily. This is an adequate supply to prevent goitre formation.*

- *The endoplasmic reticulum synthesises a large storage molecule called thyroglobulin. This compound is build up by a long peptide chain with tyrosine units and a carbohydrate unit completed by the Golgi apparatus. Iodine-free thyroglobulin is transported in vesicles to the apical membrane, where they fuse with the membrane and finally release thyroglobulin at the apical membrane.*

- *Thyroid hormones stimulate oxygen consumption in almost all cells. They stimulate the rate of 1) hepatic glucose output and peripheral glucose utilisation, 2) hepatic metabolism of fatty acids, cholesterol and triglycerides, and 3) the synthesis of important proteins. The many rate-stimulating effects are summarized in an overall increase in oxygen consumption. This slow - but long lasting - calorigenic and thermogenic effect is confined to the mitochondria.*

- *The thyroid hormones and the catecholamines work together in metabolic acceleration. Thyroid hormones increase the number of β-adrenergic receptors. Thyroid hormones modulate the secretion of sex hormones (sex development), growth hormone (growth), and nerve growth factors (CNS development).*

- *The high basal metabolic rate raises the core and shell temperature, so that the peripheral vessels dilatate. This vasodilatation forces the cardiac output to increase. A circulatory shock develops, if the rise in cardiac output is insufficient - so-called high output failure.*

- *Calcitonin is produced by the parafollicular C-cells of the thyroid. Calcitonin inhibits bone resorption by blocking the PTH receptors on the osteoclasts. The result is an extremely effective lowering of plasma [Ca^{2+}] and [phosphate]. Calcitonin is important in bone remodelling and in treatment of osteoporosis.*

- *The classical hyperthyroidism or thyrotoxicosis (Graves thyroiditis, Basedows disease) is a condition characterized by an abnormal rise in basal metabolic rate, struma and eye signs (thyroid eye disease). The eyes of the patient typically bulge (ie, exophtalmus). Patients with thyrotoxicosis have overwhelmingly high metabolic rates.*

- *Primary hypothyroidism is abnormally low activity of the thyroid gland with low circulating thyroid hormone levels caused by thyroid disease. Secondary hypothyroidism results from hypothalamic-pituitary disease.*

- *Primary hypothyroidism is caused by microsomal autoantibodies precipitated in the glandular tissue. Lymphoid infiltration of the thyroid may eventually lead to atrophy with abnormally low production of T_4. Another clinical form starts out as Hashimotos thyroiditis, often with hyperthyroidism and goitre. Following atrophy caused by microsomal autoantibodies, the condition ends as hypothyroidism, or the patient is euthyroid. When hypothyroidism is congenital both physical and mental development is impaired and cretinism is the result. Also iodide deficiency in childhood may result in a hypothyroid dwarf or cretin.*

- *Myxoedema in the adult is severe thyroid gland hypothyroidism with a puffy swollen face due to a hard, non-pitting oedema (tortoise skin called myxoedema). The skin is dry and cold; there is bradycardia, often cardiomegaly (ie, myxoedema heart), hair loss, constipation, muscle weakness and anovulatory cycles in females.*

- *Struma is a visible or palpable enlargement of the thyroid. Struma is due to iodine deficiency, increased iodine demand or strumagens. Any prolonged TSH stimulation results in an enlarged thyroid.*

Further Reading

Griffin, J.E. and S.R. Ojeda. "Textbook of Endocrine Physiology." *Oxford University Press*, N.Y./London, 1992.

Hofstra, R.M.W. et al. "A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid carcinoma." *Nature* 367: 375-377, 1994.

Return to top

Return to Content

Chapter 29.

Sexual Satisfaction, Reproduction And Disorders

Study Objectives

- To *define* amenorrhoea-oligomenorrhoea, dyspareunia, gynaecomastia, hypogonadism, impotence, infertility, menarche, menopause, menstruation and phases of the menstrual cycle, oligospermia-azoospermia, sterility, and virilization.
- To *describe* anticonception, anovulatory cycles, bleeding disturbances, castration, cryptorchism, postmenopausal hormonal alterations, puberty, anabolic steroids and doping, genetic and psychosocial sexual disorders.
- To *explain* the effect of anabolic steroids, the normal menstrual cycle, conception, implantation, pregnancy, pregnancy tests, birth and suckling. To explain the normal ovarian and testicular function, gametogenesis, erection, ejaculation and sexual satisfaction (orgasm). To explain the effect of androgen-binding protein, inhibin, aromatase, and the biosynthesis of steroids.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The gonads are concerned with the well being and preservation of the human race.*
- *The sperm decides the genetic sex (genotype). The Y chromosome is a constant determinant of maleness.*
- *Foetal differentiation of the genital ducts and of the external genitalia requires foetal gonadal hormones. The foetal genital tract will always develop into female genitals, if unexposed to embryonic testicular secretion.*

Definitions

- **Amenorrhoea-oligomenorrhoea** are terms used for absence - irregular, infrequent menstrual periods. These signs suggest *female hypogonadism*, when pregnancy is excluded.
- **Azoospermia** describes absence lack of sperm in the ejaculate
- **Dyspareunia** refers to female pain or discomfort during intercourse.
- **Gametogenesis** is the formation of ova and sperm. The primitive germ cells are divided by *meiosis*, so the number of chromosomes is halved (22 autosomes and one sex chromosome).
- **Genetic sex** is determined by the presence or absence of the Y chromosome. The Y chromosome determines the development of testes and maleness. The Y chromosome contains a *sex determining region* (the SRY gene), which encode the *testis determining factor* (TDF).

- **Genital sex** is the *phenotypic* sex (apparent female or apparent male).
- **Gonadal sex** is determined by the presence of normal *ovaries* or *testes*.
- **Gynaecomastia** refers to the occurrence of female breasts in males. The causes are HCG-producing tumours, oestrogens or oestrogenic drugs.
- **Hypogonadism** (male) refers to a condition with small, soft testes producing little sperm and testosterone. The condition is usually found with subfertility.
- **Impotence** is inability of the male to produce an adequate erection for satisfactory sexual intercourse.
- **Infertility** (subfertility) is a diagnosis used on a couple, which has been unable to conceive during one year of unprotected intercourse. The causes are oligospermia, tubule blockage, ovulatory disorders, or combined problems with both persons in the couple.
- **Menarche** refers to the age at the first menstrual period.
- **Menopause** refers to cessation of periods, which usually occurs around the age of 50 years.
- **Menstruation** is the onset of spontaneous regular uterine bleeding.
- **Oligospermia** refers to reduced numbers of sperm in the ejaculate. The causes are primary testicular disease or blockage of the vas deferens.
- **Puberty** is the transition period from a non-reproductive to a reproductive state.
- **Sterility** refers to individual infertility. Chemotherapy and other drugs may cause sterility. Surgical blockage of the tuba or the vas deferens results in sterility.
- **Virilization** is the occurrence of male secondary sex characteristics in the female.
- *Definitions of other genetic concepts are given in* Chapter 31.

Essentials

This paragraph deals with *1. The sexual drive* , *2. Sex before birth*, *3. The menstrual cycle*, *4. Ovulation/Female orgasm*, *5. Conception*, *6. Breast development*, *7. Labour*, *8. Efferent activity during coitus*, *9. Sex hormones*, *and* *10. Male puberty.*

1. The sexual drive

We feel the *sexual drive* or desire for sex *(libido)*, when *sex-related areas* in the higher brain centres are stimulated. These centres include the limbic system, Stria terminalis and the preoptic region of the hypothalamus. The desire for sex is increased by androgens in both sexes.

The sex desire of females is variable - for some it increases near the time of ovulation, when *oestradiol* secretion is increasing, while others experience a peak drive near menstruation. The CNS cells involved (see above) must contain *sex hormone receptors*. Sex hormones are steroids. They are lipid soluble and pass the cell membrane easily. After binding to *cytoplasmic receptors* (the steroid-thyroid family), the receptor-hormone complex translocates to the cell nucleus. Here the information is transcribed and translated. The result is release of

new proteins with the same information into the cytosol, where the physiologic response is triggered. *Castration* is assumed to reduce female libido minimally, but male libido is most often lost. Removal of one testis need not change the male libido. These clinical observations reflect psychosocial differences, and not necessarily a different libido mechanism in the two sexes. Hypothyroid persons lose their sex drive. The sex desire (libido) is stimulated by a multitude of *sense impres*
to the ability to engage in intercourse.

The *brain* is an important sex organ. Obviously, any natural *body contact* can be considered part of a healthy sex life - including the penetration of the penis in the vagina.

Sexual satisfaction is synonymous with *orgasm* in Western cultures. *Orgasm* is the psychological climax or the culmination of *total commitment* in a sexual act that is accompanied by a series of physiolo
similar to those involved in male ejaculation (see later). One very important reaction is ovulation, which is an automatic consequence of copulation among many animal species and periodically in humans.

Sexual enjoyment covers several phenomena. For example the *fetishist satisfaction* of wearing the clothes of the opposite sex. This is the important part of a transves
transvestites become asexual in the general sense of the term, since they do not need partners. Some individuals prefer *masturbation (onany)* as a substitute for partnership. Many individuals prefer heterosexual contacts; others prefer homosexual activities, while bisexuals may prefer either sex - depending on the circumstances. Sexual activities can vary. Besides, homosexual activity, oral sex, anal sex and many other variants are not uncommon.

2. Sex before birth

Normal sexual development in the embryo involves several processes. The sperm, which can be an X or an Y chromosome sperm, decides the *genetic sex* or sex genotype (Fig. 29-1). The genetic sex is independent of the ovum.

If the ovum is fertilised by an X spermatozoa (22 + X-chromosomes) the offspring is XX, a female. If the ovum fertilises by an Y spermatozoa (22 + Y-chromosomes) the offspring is XY, a male (Fig. 29-1).

Fig. 29-1: The sperm decides the genetic sex. The presence of the Y chromosome is the determinant of maleness.

Sex differentiation in the embryo usually harmonises with the *sex genotype*, but hormonal disturbances can lead to abnormalities. Proliferation of non-germinal and germinal cells in the *genital ridge* creates the gonadal *primordia*, which develops into a cortex surrounding the medulla. Until the 7th week of gestation, each sex has a bipotential system (the sexual indifferent stage) with both Wolffian and Müllerian ducts. The *urogenital sinus* develops into the external genitals in both females and males.

Around the 7th week, the medulla of the primitive gonad begins to differentiate into a *testis*, if an *Y chromosome* is present. This is because the Y chromosome contains the so-called *SRY gene* (the sex determining region of Y), which encodes the *testis-determining factor*.

As the testes grow and their Leydig cells start to produce testosterone, the *Wolffian ducts* develop into the male reproductive tract (epididymis, vas deferens, seminal vesicles and the ejaculatory ducts), whereas the *Müllerian ducts* regress. Testosterone stimulates the growth and differentiation of the Wolffian ducts in the male. The regression of the Müllerian ducts is caused by the *antimüllerian hormone* from the Sertoli cells.

Conversely, in the female, the cortex of the indifferent gonads differentiate into *ovaries*, if only two X chromosomes are present and no Y. In the female foetus, where there is a developing ovary and no antimüllerian hormone, the Müllerian ducts develop into the female

reproductive tract (the uterine tubes, uterus and the upper vagina), and the Wolffian ducts degenerate because the ovary does not secrete testosterone.- When a normal female foetus is exposed to androgens during the period of differentiation of the external genitalia, an *apparent male* can result.

Visible differentiation of the gross anatomy does not appear until late in the second month of embryonic life. Testosterone causes the *differentiation* of the foetus to a male. The foetal genital tract will always develop into female genitals, if unexposed to embryonic testicular secretion. The *genital sex* is a phenotypic female. If testosterone is present, male external sex organs develop and the *genital tubercle* elongates to form the male phallos. If testosterone is absent, female organs develop instead. It is the action of *testosterone* and *5-α-dihydrotestosterone* on the urogenital sinus that is behind the normal development of the male external genitalia. In the last months of gestation the growth of the external genitalia depends upon foetal pituitary LH.

One population of cells in the indifferent gonade develops into the granulosa cells of the ovarian follicle and the Sertoli cells of the testicular seminiferous tubules. These cells support and mature the germ cells. – Another population of so-called *interstitial* cells develop into the theca cells of the ovary and the Leydig interstitial cells in the testis. The Leydig interstitial cells secrete testosterone, in response to *human chorionic gonadotropin* (hCG) from the placenta.

The presence of normal ovaries or testes determines the *gonadal sex*. Without normal ovaries or testes any *genetic sex* will develop into an *apparent female*.

Foetal plasma growth hormone (GH) concentrations are high, but GH-receptors are deficient and foetal GH is not essential for linear growth. Prolactin and placental GH act as growth factors and induce the presence of IGF-1 and IGF-2. A small transfer of maternal thyroid hormone is important for early foetal development. At birth, the babys own thyroid hormone is important for CNS development and somatic growth. Foetal PTH stimulates the Ca^{2+}-transfer across the placenta and controls plasma-Ca^{2+}. Foetal ACTH is important late during gestation in particular at birth, and the cortisol concentration is high in umbilical cord plasma. Foetal pancreatic a-and b-cells are functional by 14 weeks of gestation, but their release of glucagon and insulin is low.

In 1949 Barr et al. found a densely coloured body in the periphery of the nucleus (the *Barr body* or *sex chromatin*) of the buccal mucosa of females. The Barr body is also present in other individuals with two or more X-chromosomes in each cell. Individuals with one sex chromatin (Barr body) also have a *drumstick* attached to a small fraction of their leukocytes (Fig. 29-6). We find sex chromatin and drum sticks in cells, whether they divide or not. Chromosomes are only visible in dividing cells. The maximum number of sex chromatin and drumsticks is always one less than the number of X-chromosomes (Fig. 29-6).

3. The menstrual cycle

The *menarche* is the age at the *first menstrual bleeding.* It often occurs between the 12th and the 14th year.

LH and FSH are coordinators of gonadal function. The secretion of these pituitary gonadotropins is regulated through negative feedback by the plasma concentration of gonadal steroids. LH stimulates the interstitial cells of the ovaries (and testes), but LH also acts on female granulosa cells. LH binds to a LH-receptor, which spans the cell membrane several times. The LH receptor acts via adenylcyclase and with cAMP as a second messenger. Prostaglandins may increase the cAMP effects. Maintained stimulation by LH down-regulates the number of LH-receptors on the surface of gonadal cells.

FSH acts on ovarian granulosa cells (and testicular Sertoli cells) by binding to FSH-receptors,

partially homologue with the LH-receptors. The increase in cAMP following FSH-receptor binding *transcribes* the *aromatase gene* and stimulates oestrogen synthesis. FSH stimulates synthesis of inhibin and peptide/protein products from granulosa and Sertoli cells. FSH amplifies the sensitivity to LH by increasing the number of LH-receptors on granulosa cells.

LH and FSH increase glucose oxidation, lactic acid production and protein synthesis.

The menstrual cycle starts at the *first day of bleeding* (menstruation). The bleeding is due to decrease of oestrogen and progesterone secretion. The FSH and LH secretion start to rise and stimulate the growth of several follicles - in particular following the bleeding. One of these – the dominant follicle – select itself by outstripping the others and grow so fast that the follicle can protrude more than 10 mm from the surface of the ovary. The dominant follicle has an increased oestrogen synthesis due to increased aromatase activity. Oestrogen from the granulosa cells of the dominant follicle binds to specific, *cytoplasmic receptors* (of the steroid-thyroid-family) in the endometrial and other uterine cells. Oestradiol activates and stimulates formation of *oestrogen* and *progesterone receptors*.

Fig. 29-2: The menstrual cycle in a female.

Oestrogen increases the *thickness of the endometrium*, the *size of the myometrial cells* and the *number of gap junctions* thus allowing the myometrium to work as a unit. The oestrogen phase is also called the *proliferative phase*. The concentration of sex hormones in plasma is shown in Fig. 29-2. Oestrogens work synergistically with progesterone to release gonadotropins by *positive feedback* just before ovulation.

Following the rupture of the follicle (*ovulation*), the *corpus luteum* produces increasing amounts of progesterone in addition to oestradiol also from a new developing follicle (Fig. 29-2).

Due to the *priming effect* of oestrogen on progesterone receptors, both hormones stimulate the growth of the endometrial glands, so that they curl like a helix. The progesterone effect in particular provides the endometrial/myometrial tissues with their *high secretion* and *bloodflow*, so the uterus is prepared to receive the fertilised ovum. During sexual stimulation the vaginal fluid secretion increases, as does the bloodflow of the organs involved.

If fertilisation does not occur, the level of oestradiol and progesterone switches off both gonadotropins. The corpus luteum fades out and degenerates with no LH to support it (Fig. 29-2).

The ovarian hormones almost cease to flow, and the uterus is deprived of their stimulating action. Therefore the *uterus shrinks* and sheds its swollen lining.

On the first day of the menstrual bleeding, the low progesterone and *high prostaglandin* level probably releases enough Ca^{2+} to start *spontaneous contractions* of the myometrial cells. Ca^{2+} -ions enter myometrial cells and stimulate their activity in the secretory (progesterone) phase.

The gap junctions *synchronise* these contractions, so that they include the whole myometrium. This can make excretion of blood and necrotic cells (containing prostaglandins) extremely painful. Prostaglandins dominate in menstrual fluid and stimulate the spontaneous activity of the human myometrial cells. A *normal bleeding* corresponds to a *loss of up to 50 ml of whole blood*. The mixture of vaginal fluid and menstrual blood produces a pH close to that of normal blood. The average cycle length is 28 days.

ADH (vasopressin) secretion from the neurohypophysis can cause *pre-menstrual tension* and an unpleasant increase in body fluid volume.

4. Ovulation/Female orgasm

Ovulation

A sudden increase in the plasma level of oestradiol maintained for more than 24 hours can increase FSH output by *positive feedback*. This is called the *positive feedback release ovulation*. The pulsatile release of GnRH from the hypothalamus is possibly stimulated by the high oestradiol concentrations in mid-cycle and oestradiol increase the number of GnRH receptors on the gonadotropic cells of the anterior pituitary. A neural hypothalamic pulse generator has been proposed to be involved in ovulation, and in some cases female orgasm triggers ovulation.

At lower plasma levels oestradiol is a potent inhibitor of GnRH secretion and thus of FSH and LH output (*negative feedback*). The negative feedback forms the basis for the ovulation-inhibition by contraceptives.

LH binds to a membrane LH-receptor and acts via a G-protein, adenylyl cyclase and cAMP. LH mobilises cholesterol and its conversion to progesterone.

FSH acts on ovarian granulosa cells and testicular Sertoli cells by binding to a membrane receptor homologous with the LH-receptor. The binding increases the transcription of the aromatase gene, the oestrogen and the inhibin synthesis.

The primary inhibitor of FSH secretion is the peptide, *inhibin,* that is secreted by the ovary (and testis), and blocks the effect of GnRH.

The oestradiol release from the *dominant follicle* increases sharply in the last part of the follicular phase. This triggers the *preovulatory surge* of gonadotropins (LH and FSH).

The LH surge induces an enzyme that increases the synthesis of leukotrienes, prostaglandins and thromboxanes. These molecules create an inflammation that causes rupture of the follicle. LH continues to act on the follicular granulosa cells, turning them into a yellow endocrine organ, the corpus luteum.

Orgasm

The *time for preplay* including *clitoral and multifocal stimulation* is important for most females. A clitoral orgasm in the preplay often triggers more female orgasms later during the intercourse. Female orgasm is released from the spinal cord reflexes via sympathetic signals in the pudendal nerves.

Two persons with a simultaneous sexual drive must have the necessary time for the sexual act. If they are also in love, it is natural to explore and use all means to satisfy each other.

Years ago, when the Kinsey report was made, the average duration of sexual intercourse was measured in seconds in the US. American males able to ejaculate even faster were assumed to be particularly virile. Today, such a short performance is considered a male disease called *premature ejaculation*.

5. Conception and pregnancy

Conception

Approximately 100-200 million sperms are produced each day of the fertile lifespan. The female foetus may contain 6 million oocytes, but the number decreases throughout her life (less than half a million at puberty and she may have 500 ovulations before the menopause).

The *autonomic moving* spermatozoa passes through the uterus while prostaglandins inhibit their spontaneous activity. The spermatozoa can keep their *vitality* for more than *2 days*, if they reach the fallopian tube. They lose their protective cover in the fallopian tube. The head of the spermatozoa *swell* and liberates *proteolytic enzymes*. These enzymes can dissolve the zona pellucida around the egg (oocyte). All these events in the spermatozoa takes days before it meets with the oocyte. The *oocyte* can only live *12-24 hours* without conception.

Pregnancy

Many sperms bind to the zona pellucida, but only one penetrates the wall – and blocks the entry of other sperms. Fusion of the two sex cell membranes forms the zygote, and the mitosis is complete within 24 hours.

The zygote passes into the uterine tube within a few days, protected against other spermatozoa by an increased permeability for K^+, so that the zygote membrane hyperpolarises. *Peristaltic movements* of the tube and ciliary motion conduct the zygote to the uterine cavity while undergoing *cleavage division*. Each cell is capable of developing into a complete human being up to the *eight-cell stage.*

At the *morula stage,* the cells start to develop into the *inner* cell mass or *blastocyst,* and the trophoectoderm or *trophoblast*. Seven days after conception, the blastocyst loses the zona pellucida and *implants* in the wall of the uterus (*nidation*). Nidation depends on prior conditioning of the endometrial stromal cells by progesterone bringing it into the proliferative phase. The stromal cells accumulate nutrients and swell or *decidualize* around the blastocyst. Endometrial *laminin* and *fibronectin* facilitate adhesion. *Histamine* and *prostaglandins* increase the permeability of the vessels around the nidation site. More than 2/3 of all conceptions result in miscarriage, because of insufficient attachment or other anomalies.

The *foetal trophoblast,* which give rise to the *extra-embryonic tissues* differentiates into two cell types. An inner layer of *cytotrophoblasts*, and an outer layer of *syncytiotrophoblasts*. The cytotrophoblasts synthesise stimulatory hormones such as CRH, GnRH, TRH and steroids.

The syncytiotrophoblasts synthesise first of all *human chorionic gonadotropin* (hCG). The *β-group* of hCG is specific and detected in maternal plasma 6 days following conception by specific antibody methods. The hCG is detectable in the urine within 9 days after conception.

The placenta is a fantastic hormone factory, which produces large amounts of hCG, relaxin. oestradiol, progesterone and human chorionic somatomammotropin (hCS or human placental lactogen, hPL). The hPL is synthesised from the 4. week of gestation. The hPL stimulates maternal lipolysis and inhibits insulin effects, causing hyperglycaemia.

The hCG is chemically related to TSH, FSH and LH. The hCG acts like LH and binds to the LH-receptors. The secretion of hCG is stimulated by GnRH produced by cytotrophoblasts. This is what keeps *corpus luteum* in being, and the pregnancy continues.

During pregnancy, *hCG* thus conserves the *corpus luteum*, taking over the role of LH.

The secretion of hCG stimulates ovarian release of progesterone and oestrogens just like LH. The hCG stimulates production of relaxin, inhibits the maternal secretion of LH and stimulates the maternal thyroid gland causing struma or hyperthyroidism in some pregnant females. Inhibin A from foetal trophoblasts peaks within the first week and suppresses maternal FSH secretion. Inhibin B concentrations remain low throughout gestation. LH and FSH concentrations in foetal plasma peak in mid gestation.

Fig. 29-3: Variations in plasma hormone concentrations during a normal pregnancy (42 weeks).

The plasma [hCG] reaches a peak value after *10 weeks* of pregnancy, when the syncytiotrophoblast count is maximum (Fig. 29-3). - Shortly after delivery hCG disappears.

The first peak on the plasma progesterone curve is progesterone produced by corpus luteum. The placenta takes over the progesterone production during the remaining pregnancy period ending with a peak concentration before birth. Progesterone protects the foetus in the uterine cavity by stimulation of endometrial glands that nourish the zygote and by maintenance of the decidual cells. Progesterone inhibits uterine contractions (inhibits prostaglandin synthesis and oxytocin sensitivity).

The *foetus and the placenta* form a foetoplacental unit. It produces all the hormones necessary

for a successful pregnancy. Steroid precursors are delivered from both the foetus and the mother. Oestrogen (oestradiol, oestrone, and oestriol) concentrations rise steadily throughout pregnancy (Fig. 29-3). Oestrogens stimulate the growth of the myometrium and of the ductal system of the breast. Oestriol production independs on the foetal adrenal cortex, so maternal plasma oestriol provides an estimate of the foetal condition.

The placental progesterone blocks the menstrual cycle of the mother. Pregnant females therefore develop *amenorrhoea*.

6. Breast development

During puberty FSH, LH, growth hormone, and insulin are important for the breast development. The thyroid hormones (T_3/T_4) are *permissive*. Before puberty, plasma LH and FSH concentrations are low. There is no reaction to the low concentrations of gonadal steroids and inhibin.

Oestrogens are growth factors for the myometrium and for the ductal system og the breast during pregnancy. At the end of pregnancy there are other hormonal events. Progesterone secretion reaches a peak and then falls. This fall in progesterone allows the pituitary to release *prolactin* (LTH).

Prolactin from the maternal pituitary rises throughout pregnancy. Prolactin acts on the enlarged mammary glands turning them into *milk producers.* Prolactin develops the milk producing acini in the breasts during pregnancy. *Prolactin Inhibiting Factor* (PIF or *dopamine)* from the brain inhibits the prolactin secretion.

The baby's suckling stimulate the secretion of prolactin and oxytocin, but oestradiol and sexual stimulation is also involved. The mechanical stimulation of the breast increases the secretion of prolactin from the pituitary, but the response is strikingly reduced by alcohol.

Prolactin is important for the development of the *mammary gland tissue*, oxytocin, however, governs the *ejection of milk* during lactation. *Oxytocin* causes contraction of the myoepithelial cells in the milk ducts (just as it does in the myometrial cells).

Mother-milk contains long chain fatty acids that are essential for brain development. Suckling babies are protected against juvenile diabetes in comparison to non-suckling babies. Cow's milk contains much more protein and less lactose than human milk.

7. Labour

When the foetus has reached a *critical size,* the myometrial fibres are stretched, which increase their contractility.

At the end of pregnancy the *uterus* is sensitised by *oestrogen*. After a high peak in progesterone secretion the progesterone output *falls*. This fall in progesterone allows the uterus to respond to *oxytocin,* whose release is the *final trigger for parturition*.

The foetal pituitary-adrenal axis signals to the placenta a decrease in the *progesterone-oestrogen ratio* acting on the myometrium. This increases myometrial contractions that are mediated by prostaglandins (PGE_2 and PGE_2-a). A local increase in prostaglandin concentration increases myometrial cell Ca^{2+} and triggers uterine contractions. The density of oxytocin receptors in the myometrium increases throughout gestation and particular at term.

The role of the stable plasma concentration of *maternal oxytocin* at parturition is an enigma. Oxytocin is released according to a pulsatile pattern. The frequency of oxytocin pulsations increases at labour. This fact does not exclude an important role of oxytocin in normal human parturition.

Therapeutic doses of oxytocin initiate labour in most cases at the end of gestation,

The foetal cortisol production prepares the foetus to adapt to extrauterine life by stimulating lung maturation, by increasing the hepatic glycogen stores, and by promoting closure of the *ductus arteriosus* (Fig. 12-7).

Relaxin is an insulin-like polypeptide produced by the corpus luteum and placenta. The hormone relaxes pelvic articulations, suppresses myometrial contractions and softens the uterine cervix in order to facilitate passage of the foetus.

Several other factors are involved in human labour, but the exact trigger mechanism remains unclear.

8. Efferent activity during coitus

The activity in males is described as an example. The typical sequence of efferent events in the male includes erection, emission of semen and ejaculation.

Erection means penile rigidity and elongation due to parasympathetic vasodilatation. Psychological factors trigger penile rigidity, and sexual thoughts can cause erection, emission and ejaculation. The penis contains erectile tissue located in two dorsal *corpora cavernosa* and in a single ventral *corpus spongiosum*. All the cavernous spaces of the three penile corpora receive blood from thick-walled arteries ending centrally in each corpus. The blood leaves the cavernous spaces through thin-walled veins starting peripherally. Tactile stimuli, especially from the very sensitive *glans penis* activate sensory, somatic fibres in the pudendal nerve, whereby impulses reach the *sacral plexus*. Parasympathetic impulses (S_2-S_4) from the sacral plexus elicit dilatation of the arteries and constriction of the veins in penis. The cavernous spaces are hereby filled with blood under high (arterial) pressure within seconds, causing the penis to become hard and elongated for penetration. - Erection occurs quite normally during the REM phases of sleep.

Emission is caused by sympathetic contraction of smooth muscles (in epididymis, ducts and glands), which drive the fluids into the posterior urethra. Oxytocin ejects sperm into semen. Two exocrine glands near the neck of the bladder (the seminal vesicles and the prostate) secrete fluids that nourish the sperm and transport it through the urethra during the sexual act. The prostate gland supply an alkaline secretion containing Ca^{2+}, Zn, and phosphatase to the ejaculate. The seminal vesicles supply fructose and prostaglandins. These two secretions neutralise the acid semen and help propel the spermatozoa towards the ovum. Seminal fluid also contains gonadotropins, sex hormones, inhibins, endorphins, relaxin, proteases and plasminogen activator. Epididymis supply sperm-coating proteins.

Ejaculation: Ejaculation is a sympathetic response. Contractions of skeletal muscles expel the semen from the urethra in a rhythmic pattern. Signals from glans penis reach the lumbar region of the spinal cord through afferent fibres in the internal pudendal nerves. Filling the posterior urethra with semen triggers sensory impulses that travel through the pudendal nerves to the spinal cord. The spinal cord transmits rhythmic signals to the *skeletal ejaculation muscles* (the ischio-and bulbo-cavernous muscles and those of the pelvis). These rhythmic signals stimulate rhythmic contractions that expel the semen from the urethral meatus into the female genitals. – A typical ejaculate contains 300 million spermatozoa in 3 ml.

Inside the female genitalia the sperm is subject to the process of *capacitation*, which takes place within 6 hours. The sperm head is coated with substances from the ejaculate, Ca^{2+} enters the sperm, sperm motility increases, and the ability to penetrate the ovum is enhanced. The acrosomal membrane fuse with the outer sperm membrane, so that pores are formed and proteolytic enzymes can reach the surface of the sperm head.

9. Sex hormones

Sex hormones are oestrogens, progesterone, androgens and eichosanoids. Steroid synthesis in

the gonads begins with cholesterol from acetyl Coenzyme A, and is almost identical to that of the adrenal cortex.

Oestrogens stimulate the female genitals and act to produce female secondary sex characteristics when a female enters puberty. Oestrogens and progesterone all enters the cell cytosol easily and bind to cytoplasmic receptors of the steroid-thyroid family. Oestradiol increases the synthesis of *oestrogen*- and *progesterone*-receptors.

These sex characteristics include the progressive growth of fallopian tubes, uterus, vagina, and external genitalia; also the fat deposition in breasts, buttocks, and thighs (Fig. 29-4). The ductal and stromal growth of the breasts is initiated just as the general growth at puberty with increased RNA and protein synthesis in the body cells. Oestrogens stimulate secretion of prolactin from the pituitary lactotrophic cells, increase the thickness of the endometrium and the size of the myometrial cell and their number of *gap junctions*. Oestrogens stimulate the hepatic production of essential proteins (eg, TBG, blood clotting factors, plasminogen, and HDL), but they inhibit formation of antithrombin III and LDL. Retention of salt and water can cause oedema.

Oestrogens consist of *oestradiol*, the principal ovarian oestrogen, *oestriol*, the major placental oestrogen, and *oestrone*, an important ovarian and postmenopausal hormone.

At a certain level oestradiol increases GnRH secretion and FSH output by positive feedback. There is also an increased LH sensitivity to GnRH. This feedback is already called *positive feedback release ovulation*, where the leading follicle ruptures. At lower oestrogen levels in the blood, it is a potent inhibitor of gonadotropin releasing hormone (GnRH) secretion and thus of FSH output. This is the reason for the *ovulation-inhibition* by many oral contraceptives. In the blood oestradiol is bound to *sex steroid-binding globulin.*

Fig. 29-4: Feedback loops and targets organs in the hypothalamic-pituitary-ovarian axis.

The hypothalamic GnRH secretion shows a cyclic variation in adult females of approximately 28 days, probably a genetic code imposed by the CNS.

1. Peaks of *GnRH release* reach the adenohypophysis through the portal system, and release both FSH and LH to reach the ovary via the systemic circulation (Fig. 29-4).

2. FSH stimulates follicular growth, *inhibin-release* from stromal cells, and aromatase activity in the ovary. *Aromatase* converts ovarian and other androgens to oestrogens.

3. LH stimulates the ovarian *androgen* production.

4. Inhibin is the primary inhibitor of FSH release by blocking the effects of GnRH on the adenohypophysis.

Oestrogens are responsible for the female secondary sex characteristics, the maintenance of libido, anabolic effects, and the negative feedback on the GnRH secretion and the Gonadotropin secretion of the adenohypophysis (Fig. 29-4).

Progesterone secretion rises sharply in the luteal phase of the menstrual cycle, and modulates the effect on oestrogens on the endometrium and the myometrial cells. Since the oestrogens have primed all the progesterone receptors, both hormones stimulate the growth of endometrial glands so they curl. Progesterone stimulates the secretion and high bloodflow of the uterus, so it is prepared to receive the fertilised ovum. Progesterone increases the basal core temperature by 0.5 ^{o}C, which is used as an indicator of ovulation.

In the absence of pregnancy, progesterone secretion falls and switches off the release of GnRH and both Gonadotropins (Fig. 29-4). The corpus luteum degenerates, resulting is sloughing of the endometrium (ie, menstruation).

Pregnancy is maintained by progesterone, and uterine contractions are inhibited. Progesterone has a certain aldosterone effect by competition for the same receptors. Progesterone has a negative effect on the lipid profile by increasing the LDL and reducing the HDL fractions in the blood plasma.

Androgens, such as testosterone, are anabolic, maintains spermatogenesis and libido, and act to produce male secondary sex characteristics. These characteristics are the deepening of the voice at puberty, beard, body hair, sebaceous glands in the skin, as well as the growth of the skeleton, the striated muscle system, the external genitalia and male behaviour-attitude. The primary sex structures are the testes with seminiferous tubules, epididymis, prostate, and seminal vesicles. Testosterone is responsible for the growth, maturation, and maintenance of the primary sex structures. Androgens stimulate the growth and polyamine synthesis in the prostate and the seminal vesicles. Hereby RNA synthesis is stimulated and the result is often hypertrophia and hyperplasia. Testosterone increases LDL and decreases HDL concentrations in plasma.

Fig. 29-5: The feedback system of the hypothalamic-pituitary- testicular axis.

Testosterone is reduced to two other potent androgens (dihydrotestosterone, 5a-androstendiol) in many tissues. Testosterone is thus a prohormone for these potent androgens. Most of the testosterone in plasma binds to *sex steroid-binding globulin*, a small fraction binds to albumin and only 1% is free testosterone. Thyroid hormone and oestrogens increase the concentration of sex steroid-binding globulin and thus reduces the free fraction. Androgens have the opposite effect. Testosterone diffuses easily into the cell cytoplasm and binds to a cytoplasmic receptor belonging to the steroid-thyroid receptor superfamily.

The male sexual system is controlled in the following way:

1. GnRH (= LHRH) is released from hypothalamic cells in a pulsate pattern, and stimulates release of LH and FSH from gonadotropic cells of the adenohypophysis (Fig. 29-5).

2. LH stimulates the Leydig cells of the testes to produce testosterone. These cell also produce small amounts of oestrogens, oxytocin and subunits of pro-opio-melanocortin.

3. FSH and testosterone stimulate the Sertoli cells of the testicular seminiferous tubules to produce spermatocytes and inhibin.

4. *Inhibin* is a glycoprotein that reduces the pituitary FSH secretion (blocks the effects of GnRH) by negative feedback (Fig. 29-5). *Activins* are synthesized by subunits of inhibin. The *Sertoli cells* produce *inhibin*, as do the granulosa cells in females. Inhibin inhibits FSH but not LH secretion by the pituitary gland. Activin stimulates FSH secretion just as GnRH. Follistatin binds and neutralises activin, so follistatin inhibits FSH-secretion.

5. Testosterone is responsible for the male secondary sex characteristics, the maintenance of libido, anabolic effects, and the negative feedback on the GnRH secretion and the Gonadotropin secretion of the adenohypophysis.

Acne during puberty is due to *testosterone*, but in the female *adrenocortical androgens* are involved. Testosterone promotes protein synthesis (anabolic effect). Anabolic steroids have been synthesized, which have a powerful anabolic action but only a modest androgenic action. These artificial hormones are still used to produce short-term *super-athletes*. Such a misuse of medicine for doping purposes often results in addiction, which has serious psychological, social and physical effects.

The *human hypophysis* produces four sex-related hormones FSH, LH, prolactin and oxytocin. LH is also called *Interstitial Cell Stimulating Hormone* (ICSH) in the male, because it

stimulates the Leydig interstitial cells that produce testosterone, which in turn specifically inhibits LH secretion. Removal of the male pituitary causes complete loss of all testicular functions; administration of FSH and ICSH then restores these functions completely.

Eicosanoids are oxygenated, unsaturated 20-carbon fatty acids that originate primarily from arachidonic acid by activation of *phospholipase* A_2. Eicosanoids exert important effects on most human cells. Arachidonic acid is a major component of the phospholipids of membranes. *Arachidonic acid* is converted to prostaglandins and thromboxanes by *cycloocygenases*, to leucotrienes by three types of *lipogenases*, and to epoxides by cytochrome P-450-dependent *mono-oxygenases*.

Prostaglandins and thromboxane (TxA_2) are synthesized in response to stimuli, and they mainly act *locally* as autocrine or paracrine hormones. Prostaglandins (PG) are abbreviated PGD, PGE, PGF, PGG, PGH, and PGI_2 (prostacyclin). Leucotrienes (LT) are abbreviated LTA, LTB, LTC, LTD, LTE, and LTF. The leucotrienes LTC_4 and LTD_4 are vasoconstrictors.

TxA_2 is not only an activator of platelet aggregation, but also an effective bronchoconstrictor, and TxA_2 constricts both the cerebral and the coronary arteries.

Prostaglandin E_2 (PGE_2) can be used to induce labour just as oxytocin. PGE_2 and PGF_{2a} increase uterine contractility by Ca^{2+}-influx and moderation of cAMP. Prostaglandins are especially useful in second-trimester abortion.

PGE is a potent vasodilatator, which can be used for intracavernous injection for impotence.

10. Male puberty

Young children have low plasma [gonadotropin] from birth. The gonadotropin releasing hormone is formed in the *hypothalamus* (a decapeptide *GnRH* = LHRH), and stimulates *pituitary gonadotropins* to pulsate secretion of FSH and LH (= ICSH). Through childhood they develop pulsate secretion of pituitary gonadotropins, with a LH peak at night in puberty. The nocturnal LH peak disappears when adult status is reached.

At the onset of puberty a timing device in the brain triggers the gonadotropin producing machinery in the *hypothalamic-pituitary-testicular axis*. Puberty is probably triggered by GnRH in a sufficiently mature CNS. The hypothalamic neurons mature in accordance with a genetic (familial) pattern.

Puberty is a maturation process descending from the programmed brain (hypothalamus) to the pituitary gland, the gonads and eventually to the entire body. Hormones are produced at high rates, and the secondary sex characteristics then develop.

Negative feedback control operates both before and after puberty, but the output of FSH and ICSH from the adenohypophysis is more than 100 times greater in young adults than in boys.

Circulating inhibin is the primary inhibitor of FSH secretion by negative feedback on the pituitary gonadotropins. FSH stimulates Sertoli cells to produce more inhibin at puberty (Fig. 29-5).

Circulating testosterone regulates ICSH secretion by negative feedback primarily on the eminentia mediana hypothalami. The plasma [testosterone] is highest during the night and in the morning (circadian rhythm) but there is virtually no seasonal rhythm with testosterone secretion in humans.

Enlargement of the testes is the first clinical sign of male puberty. The testis consists of Leydig cells that produce testosterone. Gap junctions connect adjacent Leydig cells, and their testosterone has local nourishing effects on germ cells. The seminiferous tubules contain the

germ cells (spermatogonia) and Sertoli cells. Each spermatogonium can divide into 64 spermatozoa within 65 days. The Sertoli cells secrete a wide variety of growth factors, activin, inhibin, oestrogens and an androgen-binding protein, all of which nourish the germ cells. The seminiferous tubules drain into rete testis Halleri, which communicates with the epididymis via ductuli efferentes. The epididymis is a maturation chamber for spermatozoa, where they lose their cytosol and become increasingly mobile within a few weeks. The store of mature spermatozoa is emitted into the female genitalia during copulation. The human testes of an adult male are positioned in the scrotum at a temperature around 35^{o}C.

In disease or old age the seminiferous tubules may cease functioning, but the *sexual capacities* (other than fertility) are well maintained as long as testosterone is produced.

Pathophysiology

Aberration of sex development can arise from two different causes. *1.* The *sex chromosomes* can create *genetic sex disturbances*, and *2. hormones* can disturb our *sex differentiation*. This paragraph also deals with *3. Psychosocial sex-deviations*, *4. Cryptorchism*, *5. Castration*, *6. Oral contraception*, *7. Impotence/Prostate disorders*, *8. Menstrual disorders/Occlusion of the fallopian tube*, *9. Menopause*, *10. Osteoporosis*. *11. Breast cancer*, *and 12. Abortion*.

Sexually related *infections* are gathered at the end of Chapter 33.

1. Genetic sex-disturbances

In 1938 Turner described a syndrome in small persons, retarded in growth and in sexual development. They are apparent females with small or no ovaries and a *XO chromosomal karyotype*. Since they have only one sex chromosome (X), their total chromosome number is 45. The *Turner patient* lacks the inputs from two active X chromosomes and from an Y chromosome. The lack of antimullerian hormone and testosterone leads to Mullerian duct development and female genitals, but the ovary is just a fibrous streak devoid of germ cells. The Turner patients have no sex chromatin and no drum stick (Fig. 29-6).

In 1942 *Klinefelter* described a syndrome in persons appearing as men. These males are tall, have small dysgenetic testes, some have female breasts (*gynaecomastia*), and they are sterile. Their cells contain XXY chromosomes (47 instead of the normal 46). Thus Klinefelter patients must have one sex chromatin and one drumstick just like normal females (Fig. 29-6). These *phenotypic XXY-males* have significantly higher LH & FSH, and lower blood [testosterone] than matched XY controls. The seminiferous tubule development and spermatogenesis are deficient in Klinefelter males. The XXY-males did not show more feminine behaviour than matched controls. A similar group of tall males with *XYY chromosomes* were not extraordinarily masculine. Some XYY-males have significantly higher [testosterone] in their blood than matched XY controls.

Some small *super women* have an extra X chromosome: XXX, making a total of 47 chromosomes. We expected them to have two sex chromatin and two drumsticks, and this has been confirmed. The XXX females have deficient germ cell development and often a short reproductive life.

Fig. 29-6: Intersex syndromes.

Apparent men with XXXY (48) chromosomes have Klinefelter *characteristics* with testes, and also two sex chromatin and two drumsticks (Fig. 29-6).

Individuals with *four* X-chromosomes are extremely rare. They are *apparent females* with XXXX (48), and *apparent males* with XXXXY (49). Cells with 4 X-chromosomes contain a maximum of 3 sex chromatin (Barr bodies) and 3 drumsticks, regardless of whether the cells come from apparent females or males (Fig. 29-6).

A very small number of individuals end up being of *indeterminate gonadal sex* (ie, has both

ovarian and testicular tissues present). Some persons have an ovary on one side and a testis on the other - a *true hermaphrodite*. In the Greek mythology *Hermaphrodites* was the child of Hermes and the beautiful Aphrodite. *Pseudo-hermaphrodites* have external genitals from both sexes, but only one gonadal sex. Males have normal XY chromosomes, but small testes with poor sperms (poor spermatogenesis). Some of these genetic (XY) boys are born as apparent girls, but they may change from female to male at puberty if the penis grows. An enzyme defect that blocks the conversion of testosterone to *5-α-dihydrotestosterone* disturbs the development of the external genitals. Female hermaphrodites have ovaries, female ducts, XX chromosomes, and varying degrees of masculine differentiation of the external genitals. Any XY individual with a genetic defect in testosterone synthesis develops testes due to the presence of the Y chromosome, and Mullerian duct regression due to the presence of antimullerian hormone. The Wolffian duct does not develop normally, because of the testosterone deficiency.

Other XY individuals lack the androgen receptor. They develop testes (Y chromosome presence) and the so-called *X-linked testicular feminisation syndrome*. These XY persons show Mullerian duct regression because the antimullerian hormone is present. The lack of androgen receptors and the effects of androgens on the Wolffian ducts prevents masculinization and the external genitals are feminine.

2. Hormonal differentiation disturbances

The virilising effect of testosterone on the *urogenital sinus* in early life causes the *adrenogenital syndrome* in XX individuals. They have ovaries (XX chromosome presence) and the Mullerian duct develops normally, because of the absence of antimullerian hormone. The androgen hypersecretion results in variable development of male external genitalia. The *adrenal hyperplasia* is caused by enzyme defects.

XY individuals with deficient testosterone synthesis ability to convert testosterone to dihydrotestosterone develop testes, but the Wolffian duct structure are underdeveloped to a varying degree ranging from a partial to a complete female pattern.

XY individuals who lack oestrogen receptors or have a mutant gene for aromatase, lack oestrogen effects. The functional lack of oestrogen results in unfused epiphyseal zones, so these males are tall, and they have high plasma concentrations of LH although testosterone is normal.

3. Psycho-social sex-deviations

Sex identity is the individual *perception* of herself or himself as a female or a male. Sex identity is established early, and is not lost by castration. Both psychological and social factors can interfere with normal sexual development on the psychological plane. An imminent urge to change sex (operative sex shifts etc.) characterises *trans-sexual persons*.

The *sex role* is the social behaviour or cultural role played by or forced upon each individual. Some male homosexuals wish to express their femininity while other males clearly signal that they are men. *Transvestites* love to dress like the opposite sex. Transvestites are heterosexual, homosexual or asexual just as others.

4. Cryptorchism

Cryptorchism means *hidden orchids* (testes). The flower orchid (French orchidé) has a root, which is actually shaped like a testis.

If the testes do not descend from the abdominal cavity to the scrotum, heat destroys the sperm-producing seminiferous tubule cells. Heat does not harm the Leydig (testosterone-producing) interstitial cells.

5. Castration

Certain cultures castrate boys to preserve their tenor voices. Puberty and natural sex development does not take place.

Adult males retain their *secondary* sex characteristics and *erection* but they often lose libido. Eunuchs are more or less trustworthy in Harems.

The effects of castration of adult females are surprisingly trivial, as long as the pituitary is working well. Castration, of course, stops their menstrual periodicity (*artificial menopause*), and they are *sterile*.

6. Contraception/ Infertility

Modern contraception is obtained with tablets (pills) containing 20-30 mg ethinyl-oestradiol and variable progesterone. The oestrogen content suppresses the hypophyseal release of gonadotropins, which prevents the maturation of the follicle, the ovulation and the luteinisation. The progesterone content favour the secretion of sperm-hostile mucus in the uterus, inhibits tuba motility and endometrial nidation.

Side effects of the combined tablet are more frequent in smoking female over 35 years, and in all females with cardiovascular risk factors. Side-effects are weight gain, accentuation of cervical and breast cancer, hypertension, acute myocardial infarction, stroke, increased clotting capacity, phlebothrombosis, gallstones, hepatomes, migraine, depression, impaired glucose tolerance (Fig. 27-6), diabetes mellitus, hypercholesterolaemia, and infertility. These serious side-effects are rare but still present. Prescription of even the modern low risk pills necessitates careful control of all risk factors.

Infertility is a diagnosis used on a couple, which have been unable to conceive during one year of unprotected intercourse. The causes are oligospermia, tuba blockage, ovulatory disorders, or combined problems with both persons in the couple.

In some cases ovulation can be elicited by a synthetic *oestrogen receptor antagonist* (clomiphene), which has a high affinity towards hypothalamic oestrogen receptors. Clomifene administration simulates oestrogen deficiency in the infertile patient with a hypothalamic defect, and by negative feedback clomifene increases GnRH and FSH/LH secretion and promote fertility.

7. Impotence/Prostate disorders

Impotence is frequently seen without organic cause such as hypogonadism. Diabetes mellitus, essential hypertension, and neuropathy of the autonomic system cause impotence just as antihypertensive drugs (diuretics, methylDOPA and b-blockers). *Intracavernosal injections* of prostaglandin E, papaverine, phentolamine, and other vasodilatating substances can provide erection for a few hours. The patient can use such a *cocktail* when needed.

Prostate disorders, such as benigh prostatic hypertrophy and prostate cancer, increase in frequency with age above 60. Both disorders interfere with micturition and can obstruct renal function leading to renal insufficiency (Chapter 25). The enzyme *5α-reductase* normally produces dihydrotestosterone from testosterone. Inhibition of this enzyme minimises the hormone conversion and causes the prostate to shrink.

Prostate cancer is frequently present in males with elevated plasma concentrations of *prostate-specific antigen* (PSA). Manifest prostate cancer is removed immediately. In a case where surgery is containdicated, long acting GnRH agonists reduce testosterone secretion and the growth of prostate cancer.

8. Menstrual disorders/Occlusion of the fallopian tube

Amenorrhoea or *oligomenorrhoea* are terms used for absence or irregularity of menstrual periods. Deficient GnRH release prevent FSH secretion from recruiting a dominant follicle, and complete loss of menses (amenorrhoea) may result. In oligomenorrhoea the oestrogen secretion is sufficient for uterine bleeding to occur in an irregular pattern, but often insufficient to induce a midcycle peak of LH and ovulation.

Causes are ovarian disease or absence (*Turners syndrome*, XO), hypothalamic deficiencies, *congenital adrenal hyperplasia* (adrenogenital syndrome), and *starvation amenorrhoea* (anorexia nervosa and excessive exercise), hypothyroid amenorrhoea with increased TRH and prolactin, and withdrawal amenorrhoea (following oral contraception). Starvation amenorrhoea and anovulatory bleeding cycles often occur in female long distance runners and ballet dancers, as well as in *anorexia nervosa* patients (Chapter 7). These females have lost substantial amounts of fat and suffer from a serious oestrogen deficiency, which even may lead to osteoporosis (Chapter 30).

Occlusion of the fallopian tube

From the start of the menstrual cycle the woman is given FSH to stimulate her ovaries before ovulation. On the 12th day she is given hCG. When ovulation occurs (after 30 to 35 hours), egg cells are *sucked out*, placed in a tissue culture and *exposed* to spermatozoa. After 48 hours some eggs fertilise into the 4-8-cell stage. A few of these *fertilised* eggs are placed in the uterus. One in four of these eggs will nidate.

Therapy is directed towards the cause of the disorder.

9. Menopause

The *menopause* is the event in the life of a female, where the menses stop. The last ovulations are anovulatory and conception is no longer possible. The ovaries become atrophic, the concentrations of pituitary gonadotropins (FSH more than LH) in blood plasma and in urine are the highest in the life of the female, because the follicles become more and more insensitive to gonadotropin stimulation and the oestrogen and inhibin production diminishes. Functional changes in other organs are less definitive, but vascular flushing of the head and neck are typical, probably due to the release of large amounts of hypothalamic gonadotropin releasing hormone (GnRH). Attacks of sweating during the night are classical complains.

Adrenal and ovarian stromal cells secrete androgen precursors that are converted to oestrogens by *aromatase* in adipose tissues. This is why menopausal females with sufficient adipose tissue suffer less from oestradiol deprivation than lean females.

Females with severe complains are treated with oestrogen, which ameliorates the disorders and reduces the rate of *heart disease* and of *postmenopausal osteoporosis*.

10. Osteoporosis

Osteoporosis or *thin bones* is a term used for a marked reduction in all elements of bone mass. Postmenopausal females reduce their bone mass progressively with age up till the age of 70-75 years. This bone reduction also occurs in elderly males, but at a much slower rate. Elderly patients living indoors all year round are less exposed to sunshine and do not synthesise vitamin D in the skin. If their diet simultaneously is poor in vitamin D and Ca^{2+}, it is not surprising that their bones get thin.

Oestrogen therapy is beneficial as a preventive strategy in *postmenopausal* osteoporosis. So is increased dietary Ca^{2+} with *vitamin D*. Walking, jogging, golf are exercises retarding bone mass loss. *Calcitonin* has proven of benefit in some studies. A promising approach is the use of oestrogen-receptor modulators to prevent osteoporosis and thrombo-embolic events, without increasing the risk of breast cancer.

11. Breast cancer

Breast cancer tumours can be treated with synthetic blockers of the oestrogen receptor. The blockers suppress the growth of oestrogen-sensitive breast cancer. – Another therapy principle is to diminish oestrogen production. This is done with the drug, aminoglutethimide, which inhibits the desmolase reaction and thereby reduces adrenal steroid synthesis as a whole.

12. Abortion

Synthetic blockers of the progesterone receptor (mifepristone) induces *early abortion* by removing the positive progesterone effects on the conceptus.

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A. The menarche is the last menstrual bleeding.

B. Pseudo-hermaphrodites have external genitals from both sexes, but only one gonadal sex.

C. HIV means Human Immunodeficiency Virus. HIV is the cause of Acquired Immune Deficiency Syndrome. HIV triggers a progressive and irreversible depletion of T-helper lymphocytes.

D. Transvestites love to dress like the opposite sex. Transvestites are heterosexual, homosexual or asexual.

E. At the onset of puberty a timing device in the brain triggers the Gonadotropin producing machinery in the hypothalamic-pituitary axis.

Case History A

A 24-year-old female is going through her last menstrual cycle before pregnancy.

1. *Summarise schematically the most important hormonal events in her menstrual cycle.*
2. *Summarise schematically the most important hormonal events during continued pregnancy and delivery.*

Case History B

A pregnant woman delivers oxygen to her foetus. Her A-haemoglobin (A = adult) is functionally different from that of her foetus (F-haemoglobin).

1. *Why is this difference important? How are the two dissociation curves related?*
2. *FSH and LH are important for this woman. Describe why. Describe the function of the two hormones in her husband.*
3. *Following birth the mother breastfeed her baby and experience a feeling of sexual pleasure including uterine contractions. Describe the mechanism.*

See answers

Highlights

- *The presence of normal ovaries or testes determines the gonadal sex. Without normal*

ovaries or testes any genetic sex will develop into an apparent female.

- *The brain is an important sex organ. The sex desire (libido) is stimulated by a multitude of sense impres*
  ability to engage in intercourse.

- *On the first day of the menstrual bleeding, the low progesterone and high prostaglandin level probably releases enough Ca^{2+} to start spontaneous contractions of the myometrial cells. Ca^{2+} -ions enters myometrial cells and stimulates their activity in the secretory (progesterone) phase.*

- *At certain high plasma level of oestradiol can increase FSH output. This is called the positive feedback release ovulation. At lower levels oestradiol is a potent inhibitor of Gonadotropin-RH secretion and thus of FSH output (negative feedback). The negative feedback forms the basis for the ovulation-inhibition by contraceptives.*

- *The primary inhibitor of FSH secretion is the peptide, inhibin that is secreted by the ovary and testis, and blocks the effect of Gonadotropin-RH.*

- *The plasma [oestradiol] increases sharply in the last part of the follicular phase, while the [LH] also increases. The sharp rise in LH and a modest rise in FSH coincide with ovulation. The LH not only causes rupture of the follicle; it continues to act on the follicular cells, turning them into a yellow endocrine organ, the corpus luteum.*

- *The spermatozoa can keep their vitality for more than 4 days if they reach the tube. They lose their protection cover in the uterine tube. The head of the spermatozoa swell and liberates proteolytic enzymes. These enzymes dissolve the corona radiata around the egg (oocyte). The oocyte can only live 14 hours without conception.*

- *Due to the priming effect of oestrogen on progesterone receptors, both hormones stimulate the growth of the endometrial glands, so that they curl like a helix. The progesterone effect in particular provides the endometrial/myometrial tissues with their high secretion and blood perfusion, so the uterus is prepared to receive the fertilised ovum.*

- *The β-group of hCG is specific and found in the blood by specific antibody methods even before the first menstrual bleeding fails to appear. The hCG is detectable in the urine 8-12 days after the first missing vaginal bleeding.*

- *During puberty FSH, LH, growth hormone, and insulin are important for the breast development. The thyroid hormones (T_3/T_4) are permissive. At the end of pregnancy there are other hormonal events. Progesterone secretion reaches a peak and then falls. This fall in progesterone allows the pituitary to release prolactin (LTH).*

- *Relaxin is a pro-insulin-like polypeptide produced by the corpus luteum. The hormone relaxes pelvis articulations and softens the uterine cervix in order to facilitate passage of the foetus. These and several other factors are involved in human labour, but the exact trigger mechanism remains unclear.*

- *Turner described a syndrome in small apparent females, retarded in growth and in sexual development, and with small or no ovaries. Since they have only one sex chromosome (X), their total chromosome number is 45. They have no sex chromatin and no drumstick.*

- *Klinefelter described a syndrome in persons appearing as males. They are tall, have small testes, some have female breasts (gynaecomastia), and they are sterile. Their cells contain XXY chromosomes (47 instead of the normal 46).*

- *Amenorrhoea or oligomenorrhoea are terms used for absence or irregularity of menstrual periods. Causes are ovarian disease or absence (Turners syndrome, XO), hypothalamic deficiencies, congenital adrenal hyperplasia (adrenogenital syndrome), starvation amenorrhoea such as in anorexia nervosa and excessive exercise, hypothyroid amenorrhoea with increased TRH, and withdrawal amenorrhoea (following oral contraception).*

Further Reading

Johnson MH and BJ Everett. *Essential reproduction*. Blackwell Science, Oxford, 1995.

Return to top

Return to Content

Chapter 30.

Other Hormones And Disorders

Study Objectives

- To *define* or *describe* concepts such as catecholamines, nanism, pituitary dwarfism, pseudo-hypo-para-thyroidism, somatomedins, somatostatin, and somatotropin.
- To *describe* the regulation of extracellular and intracellular Ca^{2+}, the bone structure, remodelling and the function of osteoblasts, osteoclasts and osteocytes. To describe the biosynthesis of calcitonin, parathyroid hormone, vitamin D, mineralocorticoids, glucocorticoids, adrenaline and noradrenaline. To describe osteoporosis, osteomalacia, and rachitis. To describe the Ca^{2+} and phosphate balance in healthy persons.
- To *explain* the effects of calcitonin, parathyroid hormone, vitamin D, mineralocorticoids, glucocorticoids, androgens and oestrogens from the adrenal cortex, and adrenaline/noradrenaline. To explain phenomena such as phaeochromocytoma, shock, adrenogenital syndrome, virilization and pubertas praecox.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The sympatho-adrenergic system give rise to the fright, flight or fight reactions that are all acute stress situations.*
- *The adrenal cortex is concerned with the carbohydrate metabolism and with the electrolyte balance.*
- *The parathyroid hormone and vitamin D maintain normal levels of calcium and phosphate in the body.*

Definitions

- **Calcium concentration** (total) in plasma: 2.2-2.7 mM. Half of the total calcium is ionized.
- **Catecholamines** are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The important catecholamines in humans are adrenaline, noradrenaline and dopamine.
- **Cushing's disease** is hypercorticism (increased glucocorticoid production) caused by a pituitary basophilic adenoma.
- **Cushing's syndrome** refers to the consequences of increased plasma glucocorticoid concentration from any source.
- **Growth hormone releasing hormone** (GHRH) is a peptide produced by GHRH cells in the hypothalamus. GHRH stimulates the release of growth hormone (Somatotropin) from the adenohypophysis.

- **Hypercalcaemia** is an increase in the plasma concentration of ionized calcium. The condition is characterised by neurological signs (hyporeflexia, lethargy or coma). When Ca^{2+} is deposited in the inner medulla of the kidney, ADH resistance develops with polyuria and polydipsia.

- **Hypocalcaemia** is a decrease in the plasma concentration of ionized Ca^{2+} (cramps, tetany, twitching and tingling of the fingers). Hypocalcaemia is caused by renal failure or by thyroidectomy and parathyroidectomy.

- **Hyperparathyroidism.** Primary hyperparathyroidism is caused by parathyroid adenomas or by hyperplasia. Secondary hyperparathyroidism is compensatory hypertrophia due to hypocalcaemia.

- **Hypoparathyroidism.** The idiopathic form is a rare autoimmune destruction of the parathyroid (with hypocalcaemia).

- **Infantile nanism** is a term used to indicate that the somatic age of a child is below its chronological age.

- **Nanism** or *dwarf growth* expresses that the height of an adult is below a certain limit. In the Anglo-Saxon countries the limit is 1.40 m for females and 1.50 m for males.

- **Pituitary dwarfism** refers to low pituitary secretion of GH or insensitive target organ receptors (African pygmies).

- **Pseudo-hypo-para-thyroidism** is used for conditions, where target-organs (bone, kidneys, and gut) are *resistant to PTH*. The plasma- [Ca^{2+}] is low, plasma-[phosphate] is high, and basic phosphatase activity is high.

- **Somatomedins** are small peptides (5 kDa) produced in the liver, promote bone growth and protein synthesis. Somatomedin C is also called insulin-like growth-factor -1 (IGF-1). The hepatic production of somatomedins is stimulated by growth hormone. Somatomedins stimulate the hypothalamic secretion of somatostatin.

- **Somatotropin or growth hormone** (GH) from the anterior pituitary stimulates body growth in humans. GHRH stimulates and somatostatin inhibits the release of GH. Growth hormone stimulates the hepatic production of Somatomedin C (IGF-1).

- **Somatostatin** or growth hormone inhibiting hormone (GHIH) is a cyclic peptide, which is produced in the pancreatic islets and in the CNS, where it is co-transmitter with noradrenaline. GHIH inhibits the release of growth hormone from the adenohypophysis and is therefore also called somatotropin releasing inhibiting factor, SRIF.

Essentials

This paragraph deals with 1. Growth hormone, 2. Parathyroid hormone. Calcium and phosphate homeostasis, 3. Adrenal corticoids, 4. Adrenal catecholamines, and 5. Vitamin D and bone minerals.

1. Growth hormone (GH)

Growth hormone is the main stimulator of body growth in humans. Growth hormone is produced in the pituitary, stimulated by hypothalamic growth hormone releasing hormone (GHRH) and inhibited by hypothalamic *somatostatin* (Fig. 30-1). Pituitary growth hormone

stimulates the hepatic production of Somatomedin C (IGF-1) and its specific *IGF-binding protein 1* (IGF-BP1). Actually, it is somatomedin C that stimulates body growth - both bone and muscle growth.

Fig. 30-1: Pituitary GH and hepatic IGF-1 secretion in the hypothalamic-pituitary feedback system.

Growth hormone releasing hormone (GHRH) is produced by GHRH cells in the hypothalamus and reaches the adenohypophysis via the portal system. GHRH stimulates the release of growth hormone (GH) from the adenohypophysis (Fig. 30-1).

Growth hormone and insulin are important anabolic hormones in humans, although growth hormone has many insulin-antagonistic effects. Growth hormone is released as *hunger spikes* during the day and during sleep.

Growth hormone (GH) produced in the placenta differs from the *pituitary GH* by a few Amino acid residues. *Placental GH* suppresses release of maternal, pituitary GH during Pregnancy. Placental GH stimulates maternal metabolism and foetal cell proliferation and Hypertrophia.

Foetal thyroid hormones stimulate brain development, and *foetal insulin* stimulates foetal Growth, cellular glucose uptake and glucose utilisation. Paracrine and autocrine growth Factors are also important for foetal growth: Insulin-like growth factor-II (IGF-II), nerve Growth factors (NGF), epidermal growth factor (EGF), and platelet derived growth factor (PDGF).

The *pituitary GH* has an endocrine effect on the production of the growth factor *Somatomedin C* (IGF-1) and its specific binding protein in the liver. Hepatic IGF-1 circulates in plasma bound to *IGF-binding protein 1* (IGF-BP1). The binding proteins are important, not only as a vehicle in plasma, but also for the final binding of the hormone to its cell membrane receptor. *Hepatic IGF-I* stimulates bone formation, and GH stimulates the precondrocytes directly. *Locally released IGF-I* stimulates the condrocytes in healthy states, but it also has a regenerative function in damaged tissues.

Pituitary GH also stimulates the production of receptors for other growth factors. GH serves to commit a precursor cell to a specific pathway, and IGF-I enhances its growth and replication. Somato
adenohypophysis. Somatomedins also stimulate the hypothalamic secretion of *somatostatin* (GHIH or somatotropin releasing inhibiting factor, SRIF).

Tissue specific growth factors - just like those important for foetal growth - are produced in damaged tissues, where they are important for regeneration. Nervous tissues produce NGF, epidermal tissues produce EGF, thrombocytes produce PDGF, and fibroblasts produce both *fibroblast growth factor* (FGF) and *transforming growth factors* (TGF-a, and TGF-b). Hepatocyte proliferation after liver damage is induced by both *hepatocyte growth factor* (HGF) and by *hepatocyte stimulating substance* (HSS). Growth factors are used in regenerative therapy.

Fig. 30-2: 68-1: Growth curves in boys - a similar pattern is found in girls.

The relative large growth of the testes at puberty shows the importance of *sex hormones* for early puberty growth (Fig. 30-2). In general, the growth spurt in early puberty is caused by increased production of *sex steroids* (stimulated by luteinizing hormone and by follicle stimulating hormone, FSH), because the sex steroids stimulate hypothalamus to release *more growth hormone.*

Growth hormone and insulin are probably the most important *anabolic hormones* in the body - more effective than oestradiol and testosterone; however, GH has many insulin-antagonistic effects. Human GH stimulates protein synthesis, mitosis in cells, chondrogenesis, ossifica and phosphate balance, while increasing glycolysis (ie, anaerobic breakdown of glycogen). GH

stimulates hepatic glucose production (glycogenolysis) but not its gluconeogenesis. GH enhances RNA synthesis, accelerates glucose uptake and antagonise the lipolytic effect of adrenaline. - GH is produced in spikes during the day (hunger spikes) and during deep sleep (EEG stage III and IV), and the half-life of GH is 20 min.

Children with *pan-hypopituitarism* become *infantile dwarfs* and later turn into *juvenile dwarfs,* as they do not develop sexually.

Children with *hyperpituitarism* become giantesses or giants. The clinical diagnosis is *gigantismus.*

Hypothyroid children (cretins) also become dwarfs. The *thyroid hormones* are necessary or *permissive* for normal growth and development. *Cretins* are *mentally retarded, hypothyroid dwarfs.*

Sex steroids in high concentrations close the epiphyseal lines. If this occurs early in life, the child also becomes an infantile dwarf, but such an individual is often sexually active. The clinical diagnosis is *precocious puberty* or *pubertas praecox.*

Without proper *insulin treatment*, children with diabetes (mellitus) become dwarfs (diabetic, infantile nanismus), because of the intracellular hypoglycaemia. Poorly regulated diabetics also develop *osteopenia* (bone decalcification).

Genetic factors are essential for optimal growth and development, as shown by the strong correlation between the height of the parents and the final height of the child. Tall people are high before the puberty growth spurt, which is more or less the same for tall and shorter people. - Persons with only one sex chromosome (X, 0) show retarded growth from birth, whereas persons with an extra sex chromosome (XXY; XYY) become tall.

Optimal growth depends upon *optimal nutrition* (essential amino acids, vitamins, minerals, and fatty acids) and optimal Neuroendocrine, metabolic control.

Optimal growth also depends upon an *optimal health.* Most disease states and all immuno-defence threatening treatments retard growth or imply weight loss. Such disorders cause catabolic hormones to dominate and induce anorexia.

2. Parathyroid hormone (PTH). Calcium and phosphate homeostasis.

We overlooked the existence of the parathyroid glands until the consequence of their surgical removal was realised. Without the parathyroid, a person develops *tetanic cramps* due to a fall in the concentration of ionised calcium ($[Ca^{2+}]$) in the blood plasma (explained in Chapter 17). - Already in 1909, MacCallum treated this condition successfully with calcium salts.

The *chief cells* (C cells) of the parathyroid glands produce *pre-proparathyroid hormone,* which is cleaved in their endoplasmic reticulum to *proparathyroid hormone*. This in turn is cleaved in the Golgi apparatus to *parathyroid hormone* (PTH). PTH is a single chain peptide with a molecular weight of 9500 Da.

PTH is water-soluble and binds to membrane receptors on the surface of the target cells. Thus the PTH action is dependent on second messengers. Both intracellular Ca^{2+} and cyclic adenosine monophosphate (cAMP) are used.

Humans carry a single PTH gene on *chromosome 11*. The chief cells also produce a *parathyroid secretory protein* of unknown function. In man, the four parathyroid glands are located just behind the thyroid gland. Ectopic tissue sometimes develops in the mediastinum or in the neck.

PTH binds to *membrane receptors* on all target cells. The major effects of PTH are on *three target organs:*

1. Bone: PTH accelerates the removal of Ca^{2+} and phosphate from bones (osteolysis by

surface osteocytes, resorption of bone by osteoclasts). PTH stimulates the osteolysis by surface osteocytes causing the release of Ca^{2+} for rapid equilibrium with the ECV. After 12 hours, the delayed effect of PTH, which stimulates the osteoclasts to reabsorb mineralised bone, sets in.

2.Kidney: PTH reduces the reabsorption of Ca^{2+} and phosphate from the proximal tubules, and increases the reabsorption of Ca^{2+} from the distal tubules, frequently resulting in an increased net loss of Ca^{2+} in the urine. PTH binds to the basolateral membrane of the tubule cell, and stimulates cAMP, which in turn diffuses through the cell to the luminal membrane. Here cAMP activates a Ca^{2+} -reabsorption port.

The glomerular filtration of Ca^{2+} is easy to calculate, since approximately half the total plasma concentration is free and filterable (2.5/2 mM). Since 0.94 parts of plasma is water, the $[Ca^{2+}]$ in the ultrafiltrate is (1.25/0.94) 1.33 mM. A person with an average *glomerular filtration rate* of 0.125 l per min will produce a 24-hour ultrafiltrate of 180 litre.

Thus, a total Ca^{2+} flux of ($1.33 \times 180 =$) 239 mmol or almost *10 000 mg daily*, will pass the glomerular *ultrafilter*.

Fortunately, almost all Ca^{2+} is reabsorbed in the kidney tubules (about 67% is reabsorbed in the proximal, and the reabsorption of the balance in the distal tubules is regulated by PTH). We only excrete 100 to 200 mg daily (or 2.5 to 5 mmol daily) in the urine and 50 mg or 1.1 mmol through the skin. This is a daily maximum of *250 mg* (or 6 mmol) *Ca^{2+} excretion from the body.*

Another drastic action of PTH is on the proximal tubules, where PTH *inhibits* phosphate reabsorption so efficiently that its excretion in the urine increases within 5 minutes.

3.Gut: The PTH action on the gut is indirect. PTH stimulates the renal production of biologically active vitamin-D (1,25-dihydroxy-vitamin D), which stimulates the active absorption of Ca^{2+} and phosphate across the gut mucosa (Fig. 30-3), and potentiates the action of PTH on bone resorption.

Fig. 30-3: The normal calcium transfer (all numbers are in mg per day or in [mmol per day]). The balance is 250 mg per day.

The end result of these three actions is an increase in plasma $[Ca^{2+}]$ and a decrease in plasma [phosphate]. Simultaneously, the bone resorption activity is illustrated by a high basic phosphatase concentration. *Cystic areas*, corresponding to periostal reabsorption, are visible on bone radiographs (osteitis fibrosa cystica). Hypercalcaemia predisposes to kidney stones and to *metastatic calcification* in synovial membranes and meninges, in the lungs, kidneys, pancreas and elsewhere. Certain tumours produce *PTH-related protein* with sequence homology to PTH. Perhaps this substance is related to the *hypercalcaemia of malignant processes* (see ectopic tumours).

The daily need of phosphate is 18 mmol just as the daily need of Ca^{2+} .The human body contains 25 mol of phosphorous (31 Da) as phosphate. Out of the 25 mol, 20 mol is located in bones and about 4 mol in muscle and other soft tissues (Fig. 30-4). A daily food intake of 46 mmol of phosphate (equal to the Ca^{2+} intake) with a secretion of 3 mmol and an intestinal absorption of 39 mmol, leaves (10 + 36) mmol daily for faecal and renal excretion in order to be in balance. The *normal* plasma concentration of phosphate is 1 mM. This implies a daily phosphate filtration rate (net-flux) of 180 mmol (1 mM*180 l of plasma) - at a glomerular filtration rate of 0.125 l per min. Phosphate is a threshold substance with a tubular reabsorption capacity (T_{max}) of 0.1 mmol per min; this means that 80% of the filtered load is reabsorbed. The *maximal* daily filtration- reabsorption- and excretion flux is calculated in Fig.

30-4.

Most of the reabsorption of phosphate takes place in the proximal tubules, where PTH inhibits phosphate reabsorption. Renal control maintains the phosphate concentration in blood plasma. The T_{max} for phosphate is up regulated by high phosphate intake and down regulated by low phosphate intake in the food. The daily exchange of phosphate from bone is 8 mmol and from soft tissue 16 mmol (Fig. 30-4). – Phosphate depletion may leads to muscle weakness (cardiac and skeletal muscles) and pathological bone formation.

Fig. 30-4: The daily phosphate balance.

The binding of PTH activates adenylcyclase and this raises the [cAMP], which interacts with protein kinase A. The enzyme then catalyses the *phosphorylation of effector proteins.* PTH is secreted in response to hypocalcaemia, in particular, a low [Ca^{2+}] (or low [Mg^{2+}]). The major effects of PTH are to increase plasma [Ca^{2+}] and decrease plasma [phos
effect on three target organs: bone, kidney and gut. – Magnesium Mg^{2+} is a cofactor in enzyme reactions and important for the neuromuscular transmission.

3. Adrenal corticoids

The adrenal cortex of the adult human has three layers: The *outer* zona glomeru
small cells and is narrow, the *middle* zona fasciculata is wide, and the *inner* zona reticu
contains a reticulum of interconnected cells. The adrenal cortex is of mesodermal origin.

Steroids

The human adrenal cortex produces three types of steroids: Glucocorticoids, mineralocorticoids, and a minimal amount of sex steroids (androgens and oestrogens). All steroids are lipid-soluble and easily cross the lipid membrane. All steroids represent chemical modifications of four-ring structures (A-D) forming the *cyclopentano-perhydro-phenantren.*

Steroids bind to specific *cytosol-receptor proteins* that are then translocated to the cell nucleus. Here they reversibly bind to DNA.

The synthesis of *glucocorticoids* (cortisol and corticosterone) occurs in the *zona fasciculata* with a small contribution from zona reticularis. The *mineralocorticoid*, aldosterone, is produced in no other region of the cortex than *zona glomerulosa.* The synthesis of *sex hormones* (androgens and oestrogens) occurs mainly in *zona reticularis.* The precursor for these hormones is *cholesterol* absorbed from the blood HDL and LDL fractions by the cortex cells. Most of the synthetic reactions involve *mixed oxygenases* (belonging to the cytochrome P-450 enzymes) localized in the endoplasmatic reticulum and in the mitochondria.

During ACTH stimulation the size and number of cells in the zona fasciculata and zona reticularis increase, mainly because the cortisol and the sex hormone production increase. The mitochondria, central ribosomes, vesicular cristae and endoplasmic reticulum grow in these cells. ACTH activates all steps in corticosteoid hormone synthesis.

A *microsomal desmolase* removes C_{20-21} from the precursors, pregnenolone and progesterone (C_{21} steroids). The residues are dehydroepi-androsterone and androstenedione (C_{19}). These androgens are weak and are converted to a more potent form, testosterone, in peripheral tissues.

In the zona reticularis, testosterone is converted to oestradiol (C_{18}), due to removal of a CH_3 group by *aromatase.* The gene for the human androgen receptors is found on the *X chromosome*. The receptor protein has a molecular weight of 98 kDa and is found both in the cytosol and the nucleus.

CRH & ACTH

The hypothalamic *Corticotropin Releasing Hormone (CRH)* stimulates the secretion of ACTH.

Stress stimulates not only the sympatho-adrenergic system with catecholamine release, but also the CRH/ACTH release with increased secretion of neuroregulatory peptides and cortisol. This is important, since small amounts of *cortisol* have *permissive effects* on catecholamines, while inhibiting TSH. Stress also releases *growth hormone* (GH that stimulates glycogenolysis/glycolysis) and *prolactin*, both from the acidophilic cells in the adenohypophysis. Prolactin released by stress is possibly mediated by *hypothalamic histaminergic neurons.*

ACTH binds to the cells of zona fasciculata and activates adenylcyclase, which results in a rise in the cAMP level. The major effect of ACTH - through increased cAMP level - is to stimulate the conversion of cholesterol to *pregnenolone* by desmolase. This is the rate-limiting step in the production of cortical steroids! Plasma levels of *cortisol* (hydrocortisone), *adrenal androgens* and their precursors rise within three minutes of intravenous ACTH injection. Cortisol inhibits the release of CRH and ACTH by *negative feedback.*

Glucocorticoids

Kendall isolated cortisone, and Hench & Reichstein directed the first administration of cortisone and ACTH to patients with rheumatoid arthritis. The result was a dramatic improvement. They shared the Nobel Prize in 1950. The serious side effects were recognized later.

The gene for the *human glucocorticoid receptor* is present in *chromosome 5*. The receptor protein (94 kDa) is found in both the cytosol and the nucleus.The effects of a glucocorticoid such as cortisol are physiologic or pharmacological dependent upon the dose.

Cortisol is the main endogenous *glucocorticoid* synthesized in the adrenal cortex.

Cortisol is essential for life and acts permissively to facilitate the mobilisation of fuels.

Cortisol defends the body against hypoglycaemia evoked by insulin.

Cortisol stimulates hepatic glucose production (gluconeogenesis).

Cortisol augments the glucagon stimulation of glycogenolysis.

Cortisol inhibits the glucose uptake in target cells (GLUT 4 in muscle cells, heart cells and adipocytes).

Cortisol is diabetogenic.

Cortisol is lipolytic and acts permissively on adrenaline, GH and other lipolytic substances to mobilize triglycerides. Cortisol also suppresses CRH release and an excess of cortisol may lead to truncal obesity.

Cortisol induces leptin synthesis in fat cells. This synthesis limits the appetite by negative feedback.

Cortisol maintains the contractility of striated and cardiac muscle (the Na^+- K^+-pump, b-adrenergic receptors etc). Excess cortisol increases muscle metabolism and reduces muscle mass and strength. Cortisol decreases the ratio of insulin-sensitive slow oxidative type I muscle fibres to the fast glycolytis type II-B muscle fibres.

Cortisol reduces the differentiation to active osteoblasts and the synthesis of collagen in bone matrix and connective tissue. Cortisol antagonises the action of 1,25-dihydroxy-cholecalciferol and thus the absorption of Ca^{2+} from the gut – cortisol excess leads to osteoporosis.

Cortisol is permissive to the vasoconstriction of catecholamines and angiotensin II.

Cortisol increases renal bloodflow and GFR.

Cortisol is important for the development of CNS, skin, gut, bone marrow and lungs.

Glucocorticoids stimulate erythropoiesis.

Cortisol inhibits processes involved in inflammation, infections, tissue damage, and immune system reactions. Glucocorticoids inhibit most responses mediated by leucotrienes, NO, PAF, prostaglandins, and thromboxanes. Therapeutic doses of glucocorticoids are used to treat more diseases than any other group of drugs, because the hormones act anti-inflammatory and anti-allergic. The negative effects are delayed healing of wounds and increased gluconeogenesis with destruction of tissue proteins.

In plasma, most of the cortisol binds to *transcortin,* to corticosteroid binding globulin (CBG), or to albumin and only 5% is free. The protein binding means that cortisol has a long plasma-$T_{½}$ of more than 70 min. Oestrogens stimulate the production of cortisol binding proteins. Patients with liver and kidney diseases do not produce enough. – Transcortin-cortisol complexes activate adenylyl cyclase, and the complex liberates free colesterol at sites of inflammation.

A low concentration of *hormone specific globulin* implies a low concentration of the *bound* hormone fraction, but the *free* hormone concentration can still be the same.

Cortisol is converted into tetrahydrocortisol, cortisone or to 17-ketosteroids. Cortisone is a biologic analog to cortisol. Exogenous cortisone is an important source of cortisol in many tissues due to the presence of the enzyme, 11beta-hydroxy-dehydrogenase.

Conjugation with glucuronic acid or sulphate forms water-soluble products. Only minimal amounts of the daily cortisol secretion is excreted in the urine. The fraction excreted in the bile is released to the enterohepatic circuit.

Half of the daily cortisol secretion is excreted in the urine as 17-hydroxycorticoids.

The plasma cortisol concentration is increased by anxiety, burns, fever, exercise, hypoglycaemia, pain, severe physical or psychiatric disease, stress and surgery.

The plasma cortisol concentration decreases promptly following administration of synthetic glucocorticoids (dexamethasone), which suppress ACTH secretion by negative feedback.

Mineralocorticoids

Aldosterone is the major *mineralocorticoid* with corticosterone contributing only little. Aldosterone maintains ECV by conserving body Na^+. When body Na^+ is depleted, the fall in ECV and plasma volume reduces renal bloodflow and pressure. A reduction in circulating fluid volume recorded in the kidneys releases aldosterone.

Aldosterone promotes the reabsorption of Na^+ and increases the secretion of K^+ and H^+ in the distal tubules and the cortical collecting ducts. A rise in plasma-[K^+] from normal or a fall in plasma-[Na^+] also releases aldosterone by direct action on zona glomerulosa.

The *renin-angiotensin-aldosterone cascade* (Chapter 24) controls the adrenal aldosterone secretion, not the ACTH.

Only a small fraction of the aldosterone binds to proteins. This makes its half-life *short* (20 min). Aldosterone is reduced to *tetra-hydro-aldosterone* in the liver and is conjugated with glucuronic acid.

Aldosterone like all other steroids is lipophilic, so that it passes easily through the cell membrane. The human mineralocorticoid receptor is found primarily in the cytoplasm, and its gene is present on *chromosome 4.* This receptor has been cloned revealing a molecular weight of 107 kDa. When aldosterone is bound to the receptor, the receptor reveals a *DNA-binding site.*

The receptor-aldosterone complex translocates from the cytosol to the nucleus, where it binds to chromatin by means of the DNA-binding site. This process activates the transcription of the specific gene producing mRNA, which is then translated into specific proteins. Following complete transmission, the *receptor-aldosterone complex* dissociates from chromatin and each other. The receptor returns to the cytosol with a *hidden DNA-binding site.*

By this process aldosterone promotes the synthesis of new proteins, that may stimulate the Na^+-K^+-pump or facilitate Na^+ entry into the tubular cell through integral Na^+ -channel proteins in the luminal membrane.

Increased Na^+-reabsorption from the tubules due to aldosterone causes a simultaneous reabsorption of water and thus an increased ECV, with increased arterial blood pressure. The pressure rise leads to increased GFR. The rapid filtration flow overrides the high reabsorptive effect of aldosterone, which down-regulates the size of ECV. – Beta-adrenergic blockers lower the arterial pressure in part by reducing the Na^+ -retention caused by the renin cascade. Inhibitors of angiotensin converting enzyme or angiotensin II receptor blockers also lower the arterial blood pressure. Antagonists of aldosterone at the renal distal tubule cells directly prevent Na^+ -reabsorption and reduce hypertension.

The ACTH effect is only a moderate stimulation in acute situations. The action of *angiotensin II* on aldosterone secretion is much more important.

The adenohypophysis probably produces an *aldosterone-stimulating factor*. Dopamine (Prolactin inhibiting hormone, PIH) is an *aldosterone inhibitor.*

Adrenal sex corticoids are mainly weak adrenal androgens produced largely in zona reticularis. By conversion to testosterone and dihydrotestosterone in peripheral tissues, the adrenal androgens play a physiological role. Also a small oestrogen production is present in healthy persons. In females, adrenal androgens produce testosterone for normal pubic and axillary hair development and erythrogenesis.

4. Adrenal catecholamines

The *medulla* is a *modified sympathetic ganglion* derived from neuroectodermal cells. The development of the sympathoadrenergic nervous system is stimulated by neural growth factors.

The adrenal medulla is the source of the circulating catecholamine, adrenaline. The medulla also secretes small amounts of nor-adrenaline, normally a neurotransmitter.

The *sympathetic system* consists of short preganglionic fibres, which have tropic centers in the lateral horn of the spinal cord from the first thoracic segment to the third lumbar segment.

The myelinated nerve fibres form ramus communicans albus and pass to the paravertebral or *lateral* sympathetic ganglion chain. Here most of the fibres end on the postganglionic cell bodies. Some preganglionic fibres pass the sympathetic chain and reach *collateral ganglia* such as the *solar plexus*, and the *ganglion mesentericus superior* and *inferior.*

In these lateral and collateral relay stations each preganglionic fibre is in contact with many cell bodies. These cells have long, postganglionic, unmyelinated fibres serving the different organs.

The preganglionic fibres to the adrenal medulla pass all the way to the chromaffin cells in the adrenal medulla. These *chromaffin cells* (high affinity for chromium cell stains) are embryological analogues to postganglionic fibres and the synapse is cholinergic.

Synthesis of catecholamines

Catecholamines are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The catecholamines are synthesises within the chromaffin cell.

The dietary amino acid *tyrosine* is absorbed as tyrosine from the gut (Fig. 30-5). Tyrosine is

also produced in the liver from dietary phenylalanine.

The L-tyrosine is hydroxylated by *tyrosine hydroxylase* in the cytosol of the chromaffin cell, and the product is L-dihydroxy-phenyl-alanine (Dopa, Fig. 30-5). Tyrosine hydroxy rate-limiting step in the catecholamine production. Dopa is decarboxylated to dopamine, catalysed by Dopa-decarboxylase. Dopamine is produced in the cytosol, and then taken up by the chromaffin granules, which contain the enzyme, dopamine b-hydroxylase. The enzyme catalyses the formation of nor-adrenaline from dopamine. In most granules noradrenaline diffuses back into the cytosol, where it is N-methylated by the enzyme, phenyl-ethanolamine N-methyl-transferase to form adrenaline. Adrenaline is then stored in the granules as the most important adrenomedullary hormone. The storage process requires ATP, and catecholamines are stored with ATP in the granules. - In the newborn the primary end product in the medulla is nor-adrenaline, but with advancing age there is a dramatic rise in the adrenaline content. This conversion depends upon cortisol.

The three important catecholamines in humans are *dopamine, epinephrine* or *adrenaline* (Ad), and *norepinephrine* or *noradrenaline* (NA in Fig. 30-5).

Dopa is converted to *melanin* in the melanocytes of the skin, stimulated by *melanocytic stimulating hormone* (MSH in Fig. 30-5).

Fig. 30-5: Synthesis of catecholamines within the chromaffin cell. Tyrosine hydroxylase and dopamine β-hydroxylase is activated by sympathetic stimulation.

Catecholamines are formed in *adrenergic, noradrenergic and dopaminergic neurons.*

The neurons of substantia nigra produce dopamine, which by axonal transport reaches the striatum (corpus striatum in Fig. 30-5). Here, dopamine is stored in the granules of the terminals - ready for liberation in the synapses.

In the sympathetic ganglion cells, dopamine is just an intermediary step, which is oxidised (dopamine b-hydroxylase) to form noradrenaline.

Euler identified noradrenaline as the chemical transmitter from adrenergic nerves, and Axelrod demonstrated the reuptake of the transmitter after its release from nerve terminals. In 1970 they shared the Nobel Prize with Katz, who explained the role of calcium in synaptic transmission.

Actions of catecholamines

Catecholamines act on adrenergic membrane receptors designated a_1-, a_2-, b_1- b_2- and b_3-receptors. The a-receptors are located on effector cells that are highly sensitive to noradrenaline, less so to adrenaline, but much less so to isoprenaline (a synthetic catecholamine). The b-receptors are located on effector cells that are most sensitive to isoprenaline, but less so to adrenaline (and noradrenaline). The a_1-receptor acts via calcium and protein kinase C. The a_2-receptor acts through G-protein inhibition of membrane adenylyl cyclase. The b-receptors are transmembrane glycoproteins, and they use cAMP as second messenger. Catecholamines prevent hypoglycaemia and restore glucose delivery to the CNS. Catecholamines increase glucose production by gluconeogenesis in the liver, and stimulate glycogenolysis in liver and muscle cells. In the absence of glucagon, adrenaline is important for recovery from hypoglycaemia. Adrenaline inhibits the insulin-mediated glucose uptake by muscle and fat cells. Catecholamines stimulate glucagon secretion and inhibit insulin secretion.

Adrenaline activates lipase in fat cells and thus increases FFA release, b-oxidation of FFA in muscle and liver and ketogenesis.

Adrenaline increases basal metabolic rate, nonshivering thermogenesis, and diet-induced thermogenesis during cold exposure. Catecholamines stimulate proton transfer into the

mitochondria and uncouple ATP synthesis from oxygen utilisation.

Catecholamines increase the heart rate, contractile force and the cardiac output by stimulation of the *adrenergic β_1-receptors* in the myocardium. Noradrenergic nerve fibres innervate vessels all over the body, and this system usually has some tonic, vasoconstrictor activity. The a_1-receptors are located on the surface of vascular smooth muscles. Adrenaline constricts the arterioles in the cutaneous, renal and splancnic beds.

Catecholamines dilatate the bronchial airways by stimulation of their adrenergic b_2-receptors. They increase both tidal volume and respiratory frequency. The result is an increased ventilation with an increased cardiac output.

Catecholamines relax the *smooth muscles of the digestive tract* (b_2-receptors), but *contract the sphincters* just like the sympathetic nerve system.

Finally, adrenaline stimulates the *ascending reticular system* (ie, the reticular activating system or RAS) in the brain stem, keeping us alert and causing *arousal reactions* with desynchronisation of the EEG.

Stress responses

Stress comprises severe emotional and physical burdens (fear, pain, hypoxia, hypothermia, hypoglycaemia, hypotension etc).

The adrenal medulla functions in concert with the adrenal cortex during stress and during immune system disorders. The two systems provide useful local cytokines and avoid systemic damages from accumulated cytokines.

Stress activates the CRH-ACTH-corticoid axis and adrenergic neurons in the hypothalamus with connections to the sympathetic nervous system. ACTH is released whereby plasma cortisol is elevated; the adrenergic stimulus elevates plasma adrenaline. These hormones increase glucose production.

Noradrenaline and CRH induce a state of arousal and aggressiveness. CRH inhibits sexual activity and feeding behaviour.

The sympatho-adrenergic system gives rise to the *fright, flight or fight-reactions*, which are all *acute* stress situations. During stress adrenaline induces hyperglycaemia and ketoacidosis (a diabetogenic hormone). Prolonged *stress* liberates ACTH via hypothalamic signals. ACTH stimulates the glucocorticoid secretion through cAMP. Small amounts of glucocorticoids are *permissive* for the actions of catecholamines.

Acute stress activates the splancnic nerves and liberates large amounts of adrenaline from the medulla. Diabetics, who are developing acute hypoglycaemia, secrete large amounts of catecholamines.

In the endangered individual catecholamines dilatate bronchioles, inhibit gastrointestinal motility, and dilatate pupils to improve distant vision.

Exercise

Catecholamines support the sympathetic system in modifying the circulation during exercise. During *exercise* catecholamines promote the use of muscle glycogen, the gluconeogenesis from lactate, and the mobilisation of FFA as alternative fuel.

During exercise the blood is directed to the working muscles from the other parts. The resistance vessels of the striated muscles in hunting predators (and perhaps in humans) are also innervated by another system. This is the *cholinergic, sympathetic vasodilator system.* It is capable of a rapid and appropriate bloodflow response during hunting. The fall in the *α-adrenergic tone* in the muscular arterioles is probably the most important exercise response in

humans.

Metabolism of catecholamines

Plasma catecholamines are rapidly removed from the blood and have half-lives in plasma of 20-60 s. This is the combined result of *rapid uptake by tissues* and *inactivation* in the liver and vessel walls.

Enzymes inactivate catecholamines by methylation or by oxidation. *Catechol-O-methyl-transferase* (COMT) on the surface of the target cells catalyses methylation to metanephrine, normetanephrine.

Monoamine oxidase (MAO) in the mitochondria catalyses oxidative removal of the amino group with formation of vanillylmandelic acid (VMA) and methoxy-hydroxy-phenyl-glycol (MOPG). The final products in the urine areVMA and MOPG. The normal excretion of VMA and MOPG in the daily urine averages 4 and 2 mg.

5. Vitamin D and bone minerals

Bones are compared to iron mixed with concrete. The *elasticity* is due to the collagen (iron) network, and the *stiffness* (concrete) is due to calcium salts (Ca - hydroxyapatite complex). - Bone tissue consists of two compartments, bone cells (metabolically active) and the bone matrix (metabolically inert extracellular compartment). The organic part of the bone matrix consists of *collagen fibres* and *ground substance,* and the inorganic part consists of *calcium-phosphate-hydroxyapatite.*

The atomic weight of calcium is 40, and we consume 1000 mg (25 mmol) per day. However, we absorb only 400 mg in total. The *net-absorption* is only *250 mg,* since we secrete 150 mg daily to the intestines. Thus, 750 mg (19 mmol) must be excreted in the faeces every day (Fig 30-3). The net-absorption is saturable (an active process), since it depends on available Ca^{2+} -*binding protein* in the brush border and in the cytosol of the enterocyte. The synthesis of this protein in the mucosal cells, and thus intestinal Ca^{2+} -absorption, is induced by active vitamin D and by the parathyroid hormone. Steroid hormones like vitamin D, exerts their major effects after binding to specific nuclear receptors and then stimulating the synthesis of mRNA that codes for *cytosolic* Ca^{2+} *-binding protein.* The basolateral membrane contains two transporters of Ca^{2+}: A Na^{+}/Ca^{2+} *exchanger,* which is more effective at high intracellular [Ca^{2+}], and a Ca^{2+}*-ATPase,* which is the major mechanism at low levels of intracellular [Ca^{2+}]. The duodenum-jejunum can concentrate Ca^{2+} against a tenfold concentration gradient.

The amount of phosphate absorbed and its concentration in plasma is determined by the amount available in the diet, but the active transport is somewhat dependent on vitamin D.

High plasma [Ca^{2+}] and [phosphate] promote bone formation in children. The children increase the precipitation of Ca-hydroxy-apatite in their bone matrix.

Intracellular phosphate is an essential component of nucleic acids, high-energy molecules, cofactors, regulatory phospho-proteins, and glycolytic intermediates.

The [total calcium] in the blood plasma of healthy persons is normally *2.5 mM* or 100 mg per l. Half the calcium binds to plasma proteins and 5% binds to a citrate-phospha complex. A significant portion (45-50%) is free ionised calcium ([Ca^{2+}]); this part has critical roles in neuromuscular function and coagulation. If the plasma [Ca^{2+}] falls to half its normal level, the body develops increased neural excitability with *tetanic cramps* (see below). On the other hand, an increase in [Ca^{2+}] leads to *calcification* of soft tissue with heart, kidney and intestinal diseases. The [Ca^{2+}] is the one variable regulated by parathyroid hormone (PTH) - see Fig. 30-3.

The normal adult contains 27 500 mmol or 1100 g of calcium and 600 g of phosphate. The

major part exists in the bones (1080 g fixed as hydroxyapatite and only 4 g or 100 mmol as *exchangeable bone calcium*). We exchange this 4 g of calcium five (5) times a day. The flux between the bones and the ECV (16 l) is approximately 500 mg or 13 mmol daily. This fast exchange regulates [Ca^{2+}] in plasma. But we also have a small amount of daily bone formation and resorption (*bone remodulation*).

In our soft tissue cells 99% of all calcium is complexed with phosphate in the *mitochondria* (10^{-5} mM), or bound to membranes and the endoplasmic reticulum. The cytosol contains an extraordinarily low [Ca^{2+}]: 10^{-7} mol per l. The intracellular [Ca^{2+}] is of extreme physiologi importance as *regulator* of enzymatic reactions and secretion, as well as a *secondary messenger* for peptide hormones. In general Ca^{2+} is important to almost all biological systems: action and membrane potentials, blood clotting, cell division, cellular secretion, contraction, cytoskeletal rearrangements, and motility.

Both calcium and phosphate ions are in equilibrium with calcium-proteinate:

$$\text{Protein} + Ca^{2+} \Rightarrow 2\ H^{+} + \text{Calcium-proteinate}$$

A hyperventilating patient eliminates carbonic acid, liberates H^+ and reduces [Ca^{2+}] in plasma, because more Ca^{2+} binds to protein. Low [Ca^{2+}] in plasma leads to increased excitability of neuromuscular tissue and to tetany (tetanic muscle cramps). The treatment is simply to reinhale the exhaled air (rebreathing) resulting in carbonic acid accumulation (Chapter 17). – As seen from the reaction above, alkalosis reduces and acidosis increases [Ca^{2+}] in plasma.

Fig. 30-6: Vitamin D and the Ca^{2+} balance.

Vitamin D_3 (*cholecalciferol*) from the skin, and vitamin D_2 (*ergocalciferol*) from plant diet are concentrated in the liver. Both are produced in the skin by ultraviolet irradiation (Fig. 30-6). In the liver microsomal oxidase simply hydroxylates the molecules to the weakly active *25-hydroxy-cholecalciferol* (25-OH-D). Lack of vitamin D leads to insufficient bone formation, because the osteoid matrix does not calcify. This disease is called rickets (rachitis) in children and osteomalacia in adults.

This substance is transferred with the blood to the kidneys, where it is further hydroxylated at the C-1 position to the most potent form *1,25-dihydroxy- cholecalciferol* (Fig. 30-6). This is what makes vitamin D a potent *steroid hormone*. Vitamin D is stored in muscle and adipose tissue and circulates in plasma bound to *vitamin D-binding protein*.

Vitamin D_2 is derived from vegetable ergosterol and its chemical name is *ergocalciferol.* Vitamin D_3 or *cholecalciferol* is produced in the skin by ultraviolet irradiation. The two molecules have equal biological capacity in humans. Vitamin D must be hydroxylated to the weakly active *25-hydroxy-cholecalciferol* in the liver. This substance is transferred with the blood to the liver, where it is further hydroxylated at the C-1 position to the most potent form *1, 25-dihydroxy-cholecalciferol.* This is exactly what is necessary to make vitamin D a potent *steroid hormone.* Vitamin D circulates in plasma bound to *vitamin D-binding protein* and is stored in liver, muscle and adipose tissue.

Vitamin D deficiency causes rickets in children and *osteomalacia* in adults. The deficiency is caused by insufficient diet, inadequate absorption as in fat malabsorption, insufficient synthesis in the skin, or abnormal conversion to its potent form in the liver and kidneys.

Rickets (rachitis) is extremely rare in the rich part of the world. The epiphyseal plate of the growing skeleton is inefficiently mineralised and considerably thickened. Rachitic children have a *flat skull* with prominent frontal bones. The sternum protrudes as *pigeon breast* or *pectus carinatum*. The costal cartilage is enlarged forming the *rachitic rosary*. The limbs are shortened and deformed.

Osteomalacia or *soft bones* are inadequate mineralisation of the organic bone matrix. Patients are asymptomatic or they suffer from diffuse pains. Bone X-rays are *osteopenic.*

High plasma concentrations of parathyroid hormone, phosphate, and oestrogens stimulate the renal hydroxylation to active vitamin D. The active vitamin D-metabolite, the 1,25-dihydroxy-cholecalciferol, stimulates the transport of Ca^{2+} across the cell and mitochondrial membranes, and has the following two effects: 1. Enhanced effect on gut absorption of Ca^{2+}, so that plasma-[Ca^{2+}] increases, 2. Enhanced effect of PTH on bones. In vitamin D deficiency the low plasma-[Ca^{2+}] stimulates the parathyroid glands, which leads to secondary hyper-parathyroidism.

Ca^{2+} loss continues for months following space flight. Stimulation of bone deposition requires the stress of muscular activity as when a person is working against a gravity field. Appropriate exercise during space missions reduces the total Ca^{2+} loss.

Puberty growth is essentially stimulated by the increased production of sex hormones, because the sex steroids stimulate hypothalamus to secrete more GH. Testosterone and oestradiol are generally important anabolic hormones.

GH and insulin are also important anabolic hormones in humans, although GH has many insulin-antagonistic effects. GH is released as *hunger spikes* during the day and during sleep.

Primary hormones for bone remodelling are *GH, PTH, calcitonin and 1,25-D_3.*

Secondary hormones for bone remodelling are glucocorticoids, sex hormones, thyroid hormones, prostaglandins (PGE_2), insulin and IGF-I to III (somatomedins).

Bone remodelling is a balance between bone formation by osteoblasts and bone resorption by osteoclasts and mononuclear cells. This balance involves the following hormones besides GH:

1. 1,25-Dihydro-cholecalciferol (1,25-D_3) is a D-vitamin, but also a steroid hormo Kidneys produce 1,25-D_3 when stimulated by PTH. The steroid hormone increases the calcium absorption from the gut and mobilises Ca^{2+} from the bones. This increases the [Ca^{2+}] and [phosphate] in plasma. The 1,25-D_3 also increases the renal reabsorption of Ca^{2+} indirectly.
2. PTH also increases the plasma [Ca^{2+}]. PTH mobilises Ca^{2+} from the bones and increases the reabsorption of Ca^{2+} in the distal, renal tubule cells while inhibiting phosphate reabsorption in the proximal tubules of the kidneys.
3. Calcitonin from the thyroid inhibits bone resorption by blocking the PTH receptors on the osteoclasts. Thus plasma [Ca^{2+}] and [phosphate] fall.

Cyclic treatment with calcitonin or combined treatment with 1,25-D_3 and PTH *improve bone formation* in *osteoporosis* (ie, bone atrophia involving both minerals and matrix). Postmenopausal osteoporosis is treated successfully with calcitonin. Calcitonin is important in *bone remodelling* and in *treatment of osteoporosis*.

A parallel to the bone formation by osteoblasts is the dentin formation by odontoblasts in our teeth. Dentists stimulate dentin formation with $Ca(OH)_2$, when treating caries.

Pathophysiology

This paragraph deals with: *1. Acromegaly and gigantism, 2. Nanism, 3. Hyperparathyroidism, 4. Hypoparathyroidism, 5. Hyperaldosteronism (Conn´s disease), 6. Cushings disease and syndrome, 7. Congenital adrenal hyperplasia, 8. Primary hypoadrenalism, 9. Secondary*

hypoadrenalism, *10. Phaeochromocytoma*.

1. Acromegaly and gigantism

Pituitary acidophilic adenomas that secrete excess growth hormone (GH) causes gigantism in children and acromegaly in adults. Rare cases are caused by GHRH excess release from the hypothalamus. Since pituitary acidophilic adenomas often contain somatotropic as well as mammotropic cells, the combined adenomas secrete an excess of both GH and prolactin (causing galactorrhoea in males).

GH excess before the long bone epiphyses have closed, results in *gigantism*. The arms and legs are long, and the patients are tall. Giants may suffer from deficient sexual development, if the gonadotropic pituitary function also suffers.

The hypophyseal tumour is demonstrated by CT scans or by magnetic resonance imaging.

GH stimulates both soft-tissue and skeletal growth. In adults only bones in the cranium, the face, and the hands and feet respond the GH excess. Acromegaly often develops around the age of 30 years in both females and males.

Fig. 30-7: Clinical features of acromegaly.

The calvarium, the maxilla and the mandible grow (Fig. 30-7). The teeth are separated by spaces and may fall out. The hands and feet are large and *spade-shaped.* This is why the adult disorder is called *acromegaly* (from Greek: extremity-great). Visual field defects are present when the enlarged hypophysis damages the crossed fibres in chiasma opticum, resulting in *bilateral, heteronymous hemianopsia* (Fig. 5-6). The GH concentration in plasma is only increased during stress and as occasional peak values, but IGF-1 is increased, and *impaired glucose tolerance* is diagnostic for a secondary, diabetic condition. There is often polyuria and glucosuria (Fig. 30-7).

Therapy in cases with growing and pressure-damaging tumours is trans-sphenoidal or trans-frontal surgery with post-operative radiotherapy. Medical treatment with *octreotide,* an analogue to somatostatin (GHIH), is beneficial in some cases.

2. Nanism

Nanism or dwarf growth expresses that the height of an adult is below a certain limit. In the Anglo-Saxon countries the limit is 1.40 m for females and 1.50 m for males. Infantile nanism is a term used to indicate that the somatic age of a child is below its chronological age.

Low pituitary secretion of GH or insensitive target organ receptor (African pygmies') result in *pituitary dwarfism.* Many cases of dwarfism are idiopathic (of unknown origin). An *achondroplastic dwarf* has short limbs and a large head with a flat nose, because of abnormal growth of cartilage and bone. The abnormal gene is autosomal and dominant.

3. Hyperparathyroidism

Primary parathyroid hyperfunction is almost inevitably due to *parathyroid adenomas* or *hyperplasia* that secrete an excess of *parathyroid hormone*, PTH. Ectopic tumours have been found in the mediastinum and elsewhere. Excessive secretion of PTH leads to: bone resorption, high [Ca^{2+}] in plasma, high Ca^{2+} -excretion in the kidneys with renal stone formation, bone lesions, and metastatic calcification (Fig. 30-8).

The action of PTH on bone results in osteitis fibrosa cystica, there is *hypercalcaemia* with low plasma phosphate, metabolic acidosis and raised PTH levels in plasma.

Secondary hyperparathyroidism is an adequate physiological response to hypocalcaemia by any cause such as renal failure or malabsorption.

Hypercalcaemia has many other causes: hypervitaminosis D, excessive Ca^{2+} -intake, leukaemia, myelomata, sarcoidosis and cancer with bone metastases or malignant tumours

producing PTH.

Fig. 30-8: Clinical manifestations of hyperparathyroidism.

4. Hypoparathyroidism

Hypoparathyroidism is characterised by *hypocalcaemia*, but most cases of hypocalcaemia are caused by renal failure with phosphate retention, vitamin D deficiency or resistance to vitamin D, resistance to PTH, pregnancy and lactation, or calcitonin administration.

Most cases of hypoparathyroidism are *iatrogenous* (ie, caused by doctors during thyroidectomy or parathyroidectomy). After seemingly successful surgery there is a dramatic fall in plasma [Ca^{2+}] and a rise in [phosphate], leading to *hypocalcaemia cramps.*

The patient is in tetany with paresthesia, convulsions, and laryngeal stridor. Without therapy death ensues.

Primary, idiopathic hypoparathyroidism is an extremely rare *autoimmune disease* often found in combination with other autoimmune disorders.

Pseudo-hypo-para-thyroidism is used for conditions, where target-organs (bone, kidneys, and gut) are *resistant to PTH.* The plasma- [Ca^{2+}] is low, plasma-[phosphate] is high, and basic phosphatase activity is high. A *PTH injection* does not increase renal phosphate excretion, because the kidneys are resistant.

Hypoparathyroidism is treated with Ca^{2+} -infusion and the continuation is a lifelong treatment with PTH and vitamin D.

5. Hyperaldosteronism (Conn´s disease)

Patients with adenoma or hyperplasia of the zona glomerulosa secrete large amounts of aldosterone - they suffer from *primary hyperaldosteronism*. Here, aldosterone works directly on the highly Na^+ -loaded distal tubules, so the patient becomes severely K^+-depleted. The renal retention of salt and water leads to *hypertension*. Loss of K^+ and H^+ leads to *hypokalaemia* with muscle weakness and cardiac arrhythmias and to metabolic alkalosis with hypocalcaemia and tetany.

Plasma aldosterone is elevated although the patient has a high intake of NaCl. Removal of the neoplasm cures Conn´s disease.

Secondary hyperaldosteronism. A patient with serious congestive heart failure may develop *secondary hyperaldosteronism* and become *K^+-depleted*. At rest the GFR, the renal bloodflow, and the cardiac output are close to normal. The cardiac patient has an abnormally high Na^+ retention and the accompanying high water retention that increases the ECV. This leads to oedema and increased venous return to the heart with a small rise in cardiac output. During *exercise* the rise in cardiac output is insufficient. The insufficient rise in bloodflow and blood pressure elicits (via the baroreceptors) a marked vasoconstriction of the renal circulation. Thus the RBF must decline. Other vascular beds also constrict markedly during exercise. The key factor to the *abnormal Na^+ retention* is the *reduction in RBF*. The major Na^+ reabsorption normally takes place in the proximal tubules. Therefore the distal tubular fluid contains a small load of Na^+, allowing only a small K^+ secretion here. This is in contrast to the patient with Conn's disease.

6. Cushings disease and syndrome

Cushing's disease is a pituitary disorder with increased secretion of pituitary ACTH.

Cushing's syndrome is used as a common term for all clinical cases of abnormally high glucocorticoid concentration, [cortisol], in the blood plasma.

All clinical cases are divided into two groups: ACTH-dependent Cushing and non-ACTH-dependent Cushing.

1. *ACTH-dependent Cushing* is caused either by pituitary disorder (Cushing's disease, see above) or by an ectopic ACTH-producing tumour. This group of cases is characterized by a high ACTH concentration in the plasma.
2. *Non-ACTH-dependent Cushing*. Most clinical cases are caused by glucocorticoid administration for long periods, but also adrenal tumours (adenomas and carcinomas) produce excess glucocorticoid. The ACTH secretion is suppressed.

Clinical manifestations

The high [cortisol] in the blood has two major effects on the body:

1. *Salt and water-retention* with renal loss of K^+ results in a plethoric *tomato face* or *moon face* (Fig. 30-9). The fluid-retention eventually leads to *cardiac hypertrophy* due to prolonged *hypertension.* There is often *peripheral oedema,* if not eliminated by polyuria due to the *glucocorticoid-induced diabetes*. This type of diabetes is typically resistant even to large doses of insulin.
2. *Catabolism* causes muscle wasting, fat accumulation, osteoporosis with kyphosis, buffalo hump, and fractures. The skin is thin with ulcers and red stirrer, and there is poor wound healing. There is impaired fibrocyte formation and capillary resistance.

Fig. 30-9: Manifestations of Cushing's syndrome.

In *Cushings disease* there is a *pituitary* tumour with enlarged sella turcica or pituitary hyperplasia. The pituitary corticotrophins produce excessive amounts of ACTH, which leads to *bilateral* adrenal cortical *hyperplasia.*

In other cases of *Cushing's syndrome* the primary cause is an *adrenal* tumour or *adrenal* hyperplasia. There is a *hypersecretion* in cortical zona reticularis and fasciculata.

The patient may suffer from *muscular and bone atrophy* due to the *catabolic* effect of the cortisol surplus. The bone atrophy leads to osteoporosis, vertebral fractures and necrosis of the hips. There is a *delayed healing* of lesions and wounds due to the slow fibrocyte formation. The skin is thin with purple striae and fragile capillaries. Sodium and water accumulate in the body, whereas potassium is lost. The nitrogen balance is negative. The Cushing patient is often depressed and suffers from insomnia.

The arms and legs are thin and weak due to the protein catabolic effect, but fat collects as truncal obesity and in the back and neck regions (*bulls neck* and *buffalo hump*). The fluid accumulation leads to a special look called *moon face* or *tomato face.*

The high blood [cortisol] also increases the blood [glucose]. A *diabetic condition,* which is resistant to even large doses of insulin, develops. The patient suffers from *hypertension* and rapidly develops *arteriosclerosis*. Female patients complain of amenorrhoea and poor libido.

The primary therapy is to reduce plasma [cortisol] to approximately 350 nM also as a preparation for surgery. In cases of *Cushing's disease* the tumour is typically removed trans-sphenoidally, and adrenal adenomas are resected.

The best assessment of increased ACTH production is by measuring the 24-hour urinary cortisol excretion. *Suppression* of ACTH and cortisol is performed with a standard dose of synthetic Glucocorticoid (*dexamethasone*), which is potent enough to suppress the ACTH production, without influencing the plasma [cortisol].

The principle of the *suppression test* is to *challenge the hypothalamic-pituitary feedback system.* If the system is intact a fall in blood [cortisol] is expected, and if not (hypothalamic-pituitary Cushing) the high [cortisol] is maintained.

In healthy persons such a standard dose will *suppress ACTH secretion,* and lead to a reduced blood [cortisol]. Such a normal observation is also typical for patients with a *primary adrenocortical hypertrophy*. Dexamethazone does not suppress the high cortisol level of patients with a *damaged feedback system* producing excess ACTH from a pituitary tumour.

Cortisol excess has been treated with ketoconazole, which inhibits several steps in the corticosteroid synthesis.

Differential diagnosis:

Alcoholic patients with so-called *Pseudo-Cushing* look exactly like *Cushing's syndrome,* but the glucocorticoid concentration in the plasma is within normal limits. The cause is unknown. Cases with plasma [cortisol] in the upper end of the normal scale can be explained in the following way. Their progressive hepatic failure probably impairs the normal hepatic destruction of glucocorticoids, whereby the plasma [cortisol] is rising. Such patients should be controlled, because the liver destruction is likely to proceed, and they become more and more Cushingoid with continued alcohol abuse.

7. Congenital adrenal hyperplasia

This is an autosomal, recessive enzyme deficiency blocking steroid synthetic pathway of the adrenal cortex. The most frequent of these rare genetic disorders is the *21-hydroxylase deficiency* on *chromosome 6*. Lack of hydroxylase blocks the conversion of 17-hydroxyprogesterone into *11- deoxycortisol* and further to *cortisol*. Instead androstenedione and testosterone is produced and virilization is an inevitable result. The impaired cortisol release from the adrenal increases the pituitary ACTH secretion by negative feedback and more precursor molecules are accumulated (ie, hydroxyprogesterone, androstenedione and testosterone).

The clinical features are those of an *adrenogenital syndrome.* Newborn girls may show clitoral hyperplasia and labioscrotal fusion, and later some boys develop *precocious puberty*. Varying degrees of virilism in females, *birdied ladies* in circus, or perhaps only hirsutism before puberty occur. The lack of aldosterone causes salt-and water-loss, which may be life threatening.

A defect in the *11-hydroxylase enzyme gene* also leads to overproduction of androgens and 11-deoxycorticosterone from accumulated precursors. The later has a massive mineralocorticoid effect.

The therapy is exact replacement of glucocorticoid activity and any other inefficiency.

8. Primary hypoadrenalism

Primary hypoadrenalism is a complete destruction of the adrenal cortex, and thus it impairs all three lines of steroid production (ie, glucocorticoids, mineralocorticoids and sex hormones). Primary hypoadrenalism occurs in an acute and a chronic form.

Acute hypoadrenalism (ie, *adrenal crisis* or Waterhouse-Friderichsens syndrome) is a fulminant infection with shock and massive bleeding all over. The large internal bleedings are visible at the skin as areas of purpura. The patient dies within a few hours, if not treated urgently. *Hydrocortisone* (100 mg) is given intramuscularly. Antibiotics and cortisol is administered.

Chronic hypoadrenalism (ie, Addison's disease) is hypocorticism due to destruction with atrophy of the entire adrenal cortex by autoimmune processes, malignancy, infarction or infection. Most cases develop *organ-specific autoantibodies;* these cases are associated with many other autoimmune disorders (eg, diabetes mellitus, hypoparathyroidism, pernicious anaemia, vitiligo and thyroiditis).

+

Addison's disease is a life-threatening condition with loss of Na and thus also of ECV. Symptoms and signs include: reduced blood pressure, reduced blood [glucose], tiredness and skin pigmentation caused by MSH, whose level is increased concurrent with the *overproduction of ACTH*, due to the decreased negative feedback. Severe hypotension may develop into cardiovascular collapse, which is called an *Addisonian crisis*.

The clinical manifestations are consequences of the impaired secretion of the three hormone groups (Fig. 30-10).

Fig. 30-10: Clinical manifestations of primary, chronic adrenocortical insufficiency (Addison).

The low cortisol increases the hypothalamic CRH secretion through feedback and thereby increases ACTH secretion. Fragments of the ACTH molecule form MSH, which is responsible for the skin hyperpigmentation.

When Addison's disease is suspected a *single* plasma ACTH level is often diagnostic, as it will be elevated if hypocorticism originates in the *adrenal gland* (primary) and normal or low in *hypothalamic-pituitary hypocorticism* (secondary). - Differentiation between the hypothalamic and the pituitary hypocorticism can be accomplished by *injection of CRH*, which stimulates the pituitary directly. - *Insulin-induced hypoglycaemia* normally recruits the combined hypothalamic-pituitary axis. Hypoglycaemia is a prolonged stress, which will normally increase ACTH and glucocorticoid secretion. If there is adrenal gland insufficiency then this increase in glucocorticoid secretion is not observed. *Hypothalamic and pituitary failure* is detected when a dose of CRH restores glucocorticoid release. - *A single ACTH-injection* along with measurement of plasma [cortisol] is a simple screening procedure for adequate endogenous ACTH secretion. Synthetic ACTH is injected intravenously. If plasma [cortisol] is normal 30 minutes later, the disease is not a *primary* adrenocortical atrophia. - *The metyrapone test* is used when patients are suspected to have Addison's hypocorticism, but have reacted normally to the ACTH-test. Metyrapone inhibits the last step in cortisol synthesis, so 11-dehydrocortisol is accumulated and an acute cortisol deficiency is created. Plasma-[Cortisol] and -[11-deoxycortisol] is measured two mornings in succession, and the *adrenal 11-hydroxy*
plasma- [cortisol] must fall while -[11-deoxycortisol] accumulates. The fall in cortisol biosynthesis increases the CRH and ACTH secretion by negative feedback, provided the hypothalamic-pituitary system is *intact*. Failure to respond to metyrapone suggests a hypothalamic-pituitary defect that has abolished this feedback.

Substitution therapy with gluco- and mineralo-corticoids is necessary for the rest of the patient's life.

9. Secondary hypoadrenalism

This is usually caused by prolonged corticoid therapy of chronic disorders such as rheumatoid arthritis or COLD. A rare form of secondary hypoadrenalism is *panhypopituitarism*.

10. Phaeochromocytoma

Rarely some patients have attacks of *severe hypertension*., which are caused by a *medullary tumour* (a *phaeochromocytoma* of chromaffine cells), which liberates large amounts of catecholamines.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

A. Atropine blocks muscarinic receptors, and d-tubocurarine blocks nicotinic receptors.

B. b-Receptors are blocked by propranolol.

C. Catecholamines have a half-life in plasma of more than 20 min.

D. The three important catecholamines in humans are adrenaline, noradrenaline, and DOPA.

E. b_1-receptors are stimulated by adrenaline and located in the myocardium.

II. Each of the following five answers have False/True options:

ACTH stimulates the liberation of**:**

A**.** Hydrocortisone or cortisol

B. Adrenaline

C. Aldosterone

D. Adrenal androgens

E. Dopamine.

Case History A

A man, 49 years of age (height 1.86 m; weight 62 kg) is in hospital due to the following symptoms and signs. He is nervous and has a diffuse struma. A characteristic, blowing sound is heard from the thyroid gland with a stethoscope. The blood pressure is 145/70 mm Hg. An attack of cardiac arrhythmia is recorded with an ECG. A P-wave frequency above 400 per min is present during the attack. The concentration of thyroxine in blood serum is 180 nM. The distribution volume for radioactive thyroxine is 8 l. The elimination rate of this thyroxine is 14% of the total content per 24 hours.

1. *Calculate the serum [thyroxine] in μg per l.*

2. *How much thyroxine is eliminated per 24 hours expressed in μg daily?*

3. *Calculate the thyroid plasma clearance of iodide, when the concentration of free iodide in plasma is 4 μg per l.*

4. *Explain the condition of this patient***.**

Case History B

A 22-year-old medical student is treated with an intravenous dose of PTH. This changes his renal excretion flux of two substances and their plasma concentrations.

1. *Describe the alterations and explain the mechanisms.*

2. *What is the diagnosis?*

3. *Describe the most likely symptoms and signs of this patient before treatment.*

Case History C

A female (52 years of age; height 1.68 m; weight 62 kg) is in hospital due to her third attack of kidney stone pains. The first routine examination with arterial blood analysis reveals the

following. Her blood pH is 7.21 and her plasma [Ca^{2+}] is 2 mmol per l (mM) in ionised form. Her ionised [Ca^{2+}] constitute 62% of the total calcium concentration in plasma. The inorganic phosphate concentration (total) is 0.84 mmol per l of plasma. The patient excretes 2-3 l of urine per day. pK_2 = 6.8 for H_3PO_4.

1. *Calculate the fractions of primary ($H_2PO_4^-$) and secondary (HPO_4^{--}) phosphate in her plasma.*

   Compare the results to the normal mean value of 1 mM of inorganic phosphate with 20% primary and 80% secondary phosphate at pH = 7.40.

2. *Calculate the total plasma [calcium] of this patient and compare the result to the normal mean value of 2.5 mM.*

3. *Why is the ionised calcium fraction much higher than normal (0.45)?*

4. *Could this condition be the result of a classical endocrine disease?*

5. Why did this patient develop kidney stones? Was her diuresis normal? If not explain why.

Try to solve the problems before looking up the answers.

Highlights

- *Growth hormone (GH) produced in the placenta differs from the pituitary GH by a few amino acid residues. Placental GH suppresses release of maternal, pituitary GH during pregnancy. Placental GH stimulates maternal metabolism and foetal cell proliferation and hypertrophia.*
- *Foetal thyroid hormones stimulate brain development, and foetal insulin stimulates foetal growth, cellular glucose uptake and glucose utilisation. Paracrine and autocrine growth factors are also important for foetal growth: Insulin-like growth factor-II (IGF-II), nerve growth factors (NGF), epidermal growth factor (EGF), and platelet derived growth factor (PDGF).*
- *Growth hormone and insulin are the most important anabolic hormones in the human body.*
- *PTH binds to membrane receptors on target cells in bone, kidney and gut. The PTH actions result in hypercalcaemia and hypophosphataemia. The bone resorption is illustrated by a high basic phosphatase concentration in blood.*
- *Cortisol stimulates hepatic glucose production, both glycogenolysis and gluconeogenesis, lipolysis, formation of FFA and of ketone bodies. Cortisol inhibits the glucose uptake in target cells (GLUT 4 in muscle cells, heart cells and adipocytes).*
- *Therapeutic doses of glucocorticoids are used for a multitude of diseases such as inflammations, allergy, malignancy and aplastic anaemia. The negative effects are delayed healing of wounds and increased gluconeogenesis with destruction of tissue proteins.*
- *Aldosterone is the major mineralocorticoid with corticosterone contributing only little. Aldosterone promotes the reabsorption of Na^+ and increases the secretion of K^+ and H^+ in the distal tubular system (ie the cortical collecting ducts and the connecting segment). A rise*

in serum - $[K^+]$ from normal or a fall in serum- $[Na^+]$ releases aldosterone. The renin-angiotensin-aldosterone cascade controls the adrenal aldosterone secretion, not ACTH.

- *Sex steroids are mainly weak adrenal androgens, which are metabolised to testosterone and dihydrotestosterone. Also a small oestrogen production is present in healthy persons.*
- *Catecholamines increase the heart rate and the cardiac output by stimulation of the adrenergic β_1-receptors in the myocardium.*
- *Noradrenergic nerve fibres innervate vessels all over the body, and this system usually has some tonic, vasoconstrictor activity. The α_1-receptors are located on the surface of vascular smooth muscles.*
- *Catecholamines dilatate the bronchial airways by stimulation of their adrenergic β_2-receptors. Catecholamines increase both tidal volume and respiratory frequency. The result is an increase in ventilation together with an increase in cardiac output.*
- *Catecholamines relax the smooth muscles of the digestive tract (β_2-receptors), but contract the sphincters just like the sympathetic nerve system.*
- *Catecholamines stimulate metabolism (T_3). Adrenaline stimulates hepatic glycogenolysis and lipolysis in adipose tissue. Adrenaline increases the plasma concentrations of glucose, FFA and ketoacids.*
- *Adrenaline stimulates the ascending reticular system (ie, the reticular activating system or RAS) in the brain stem, keeping us alert and causing arousal reactions with desynchronisation of the EEG.*
- *Pituitary acidophilic adenomas that secrete excess growth hormone (GH) causes gigantism in children and acromegaly in adults. Rare cases are caused by GHRH excess release from the hypothalamus. Since pituitary acidophilic adenomas often contain both somatotropic and mammotropic cells, the combined adenomas secrete an excess of both GH and prolactin (causing galactorrhoea in males).*
- *Primary parathyroid hyperfunction is almost inevitably due to parathyroid adenomas or hyperplasia that secrete an excess of parathyroid hormone, PTH. Ectopic tumours have been found in the mediastinum and elsewhere. Excessive secretion of PTH leads to: bone resorption, high $[Ca^{2+}]$ in plasma, high Ca^{2+} -excretion in the kidneys with renal stone formation, bone lesions, and metastatic calcification.*
- *Chronic hypoadrenalism (ie, Addison's disease) is hypocorticism due to destruction with atrophy of the entire adrenal cortex by autoimmune processes, malignancy, infarction or infection. Most cases develop organ-specific autoantibodies; these cases are associated with many other autoimmune disorders (eg diabetes mellitus, hypoparathyroidism, pernicious anaemia, vitiligo and thyroiditis).*

Further Reading

Costanzo, L S. Regulation of calcium and phosphate homeostasis. *Am. J. Physiol.* 275 (*Adv. Physiol. Educ.*): S206-S216, 1998.

Änggård, E. "Nitric oxide: mediator, murderer, and medicine." *Lancet* 343: 1199-1206, 1994.

Tonegawa, S. "The molecules of the immune system," *Sci. Am.* 253: 122-131, 1985.

Return to top

Return to Content

Chapter 31

The Human Genome And Genetic Disorders

This Chapter was written following fruitful discussions with my colleague, Erik Niebuhr, MD, DSc, University of Copenhagen.

Study Objectives

- To *define* allele, anaphase, anticodon, autosomes, chromatid, chromatin, chromosome, clone, codon, diploid, exons, gene, gene frequency, haploid, haemophilia, introns, nucleotides, post-translational modification, probability, promoter sequence, ribosomal translation, sex-linked genes, splicing, transcription, translocation, and transversion.
- To *describe* the human genome, gene therapy, and the roles of DNA, RNA, messenger RNA, transfer RNA and recombinant DNA
- To *calculate* the frequency of abnormal genes in the total gene pool of a population.
- To *explain* protein synthesis, DNA transcription, mutations, and genetic disorders.
- To *use* the above concepts in problem solving and case histories.

Principles

- *Many hormones regulate genes expressed by some cells. This mechanism controls the synthesis of enzymes, receptor proteins, structural proteins, and transcription proteins. This is why steroid and thyroid hormones require hours for their biological effect.*
- *The partial or total lack of specificity in the third base of some triplet codons may allow 2-4 codons, differing only in the third base, to code for the same amino acid.*

Definitions

- **Allele** is an alternative form of a gene occupying the same locus on a particular chromosome.
- **Anaphase** refers to the stage of nuclear division, which is characterized by movement of chromosomes from spindle equator to spindle pools.
- **Anticodon** is the group of three nucleotides in transfer RNA that pairs complementarily with three nucleotides of messenger RNA during protein synthesis.
- **Autosomes** refers to any chromosome which is not a sex chromosome or mitochondrial chromosome. Humans possess 22 pairs of autosomes.
- **Bacteriophag** is a bacterial virus. Bacterial viruses are modified and used as vectors for DNA cloning.
- **Chromatid** results from the replication of chromosomes during interphase. A chromatid is

one of the two identical halves of a chromosome, which shares a common centromere with a sister chromatid.

- **Chromatin** refers to the nuclear material, which comprises the chromosomes: The DNA complex.
- **Chromosomes** are nucleoprotein structures, which are the sites of nuclear genes arranged in linear order.
- **Clone**. A group of cells or organisms derived from a single ancestral cell or individual by asexual multiplication (repeated mitoses). All members of a clone are genetically identical.
- **Codon** or *triplet code* is determined by a base triple (three adjacent nucleotides that code for one amino acid). A codon encodes a specific amino acid residue to be added into the peptide chain or specifies termination of translation. Three codons in mRNA (UAA, UAG and UGA) are *stop codons*. The signals encoded by the codons form the *genetic code*.
- **Dalton** refers to a weight unit equal to the mass of the hydrogen atom.
- **Deoxyribonucleic acid** (DNA) is a helical coiled nucleic acid molecule composed of deoxyribose-phosphate units connected by paired bases attached to the deoxyribose sugar. DNA is the genetic material of all living organisms and vira.
- **Diploid** refers to the number of chromosomes found in somatic cells (ie, two sets).
- **DNA polymerase** is the enzyme that replicates DNA.
- **Dominant** describes a trait expressed in individuals who are heterozygous for a particular gene
- **Exon** is a segment of a gene that is represented in the final spliced mRNA product.
- **Expressivity** is the degree to which the effect of a gene is expressed.
- **Frequency** refers to the relative number of actual cases per 100 000 persons in a population.
- **Gene** is a part of a DNA molecule that codes for the synthesis of a specific polypeptide chain through its sequence of nucleotides. Each human gene extends over 40 kb in general, but we also posses longer genes. The gene, located in the chromosome, is the particulate determiner of hereditary trait.
- **Gene frequency** refers to the number of loci at which a gene occurs, divided by the number of loci at which it could occur. This is the proportion of one allele of a pair of genes present in the population.
- **Gene therapy** is the application of gene transfer (DNA delivery to cells) in an attempt to treat genetic or acquired disorders. The transferred DNA must contain promoter zones for transcription and the protein-coding region of the gene.
- **Genome** is the total amount of genetic material in the cell.
- **Genotype** is the genetic constitution of an individual.

- **Haploid** refers to an individual or germ cell having a single complete set of chromosomes (one set).
- **Heterozygote** is an individual possessing two different alleles at the corresponding loci on a pair of homologous chromosomes.
- **Homozygote** is an individual possessing identical alleles at the corresponding loci on a pair of homologous chromosomes.
- **Haemophilia** is an X-linked, recessive genetic disorder characterised by free bleeding from even slight wounds because of lack of formation of clotting substance.
- **Incidence** refers to the new cases of a disorder diagnosed per year in a population group.
- **Intron** is a segment of a gene not represented in the final mRNA product. The segment has been removed through splicing together of exons on either side of it.
- **Karyotype** is the number, size and shape of the chromosomes in a cell.
- **Meiosis** is a nuclear division in which the diploid chromosome number is reduced by half.
- **Mutagens** mean all agents that bring about a mutation.
- **Mutation** is a sudden change in genotype without relation to the ancestry of the individual.
- **Nucleotides** are the basic units of nucleic acids and contain a 2-deoxyribose (a pentose sugar molecule), a phosphate group, and a nitrogenous base. DNA has only four types of nitrogenous base, namely: Adenine (A), Cytosine (C), Guanine (G), and Thymine (T): A C G T.
- **Oncogenes** are genes which are altered in structure or expression, and contribute to the neoplastic transformation of cells (cancer cells).
- **Phenotype** is the appearance of an individual, resulting from the effects of both environment and genes.
- **Post-translational modification:** Following ribosomal translation, proteins are modified by addition of carbohydrates, cleavage of bonds within the new protein, shortening or folding.
- **Polyploid** refers to an individual having more than two complete sets of chromosomes (triploid = 3, tetraploid = 4).
- **Probability** is the likelihood of occurrence of a given event.
- **Promoter** is a region of DNA that plays an essential role in the initiation of transcription of a gene.
- **Recessive** describes a trait expressed in individuals, who are homozygous for a particular gene but not found in the heterozygote.
- **Recombinant DNA** refers to a DNA strand resulting from the physical joining of two or more pre-existing DNA strands.

- **Ribonucleic acid** (RNA) is similar to DNA except that it contains ribose instead of pentose and Uracil (U) instead of T: **A C G U**. RNA is usually single stranded.
- **Ribosomal translation** is programmed protein synthesis. Inside the ribosome the particular sequence of the mRNA is read, and the particular sequence of nucleotides is build into a polypeptide. This is the so-called *translation process*, whereby the amino acids are always linked to the polypeptide chain in the proper sequence.
- **RNA polymerase** is the enzyme that synthesises RNA based on a DNA template.
- **Sex-linked gene** is a gene located on the X- or Y-chromosome in XY species. The genes of well-known sex-linked disorders are located mainly on the X-chromosome.
- **Splicing** takes place inside the cell nucleus. The *coding sequences* or *exons* of the immature RNA molecule are cut and spliced together by enzymes eliminating intervening *introns*. The final product is called mature messenger RNA (mRNA).
- **Transcription** (copying) is the copying of the DNA code into a single stranded immature messenger RNA (mRNA). RNA is ribonucleic acid. In the TATA box, the RNA polymerase transcribes a single stranded copy of the DNA sequence. The copying stops at the end of the gene.
- **Transfer RNA** (tRNA) is a molecule which carries a single amino acid and transfer the amino acid to the ribosome. - There are at least 20 different types of tRNA, each carrying only one amino acid.
- **Translation** is the process by which a specific mRNA nucleotide sequence is responsible for a specific amino acid sequence of a polypeptide. The genetic information is translated into protein synthesis.
- **Translocation** refers to the transfer of a portion of one chromosome to another non-homologous chromosome.
- **Transversion** is the substitution in DNA or RNA of a purine for a pyrimidine or vice versa.
- **Trisomy** is the representation of a chromosome three times rather than twice, yielding a total of 47 chromosomes.
- **Vector** is a DNA molecule used to carry DNA regions.
- **Zygote** is the result of fusion of two gametes in sexual reproduction (a fertilised egg is a zygote).

Essentials

This paragraph deals with *1. The human genome*, *2. Protein synthesis*, *3. DNA transcription*.

1. The human genome

Genetic coding is stored in DNA or *deoxyribonucleic acid* (Fig. 31-1). Nuclear DNA is the dominant controller of protein synthesis in the ribosomes. The codes of special enzymes and other proteins are based on at least 20 amino acids arranged in sequence. The nucleus contains a nucleolus and the chromatin network. Chromatin is DNA molecules rolled up in a pearl

chain structure with proteins called *nucleosomes*.

DNA consists of two nucleotide strands twisted around each other to form the so-called *double helix* (Fig.31-1), which can be several cm in length. The double helix is folded to form *chromosomes* with a length of approximately 10 mm (Fig. 31-1).

Nucleotides contain a *2-deoxyribose* (a pentose sugar molecule), a *phosphate* group, and a *nitrogenous base*. DNA has only four types of nitrogenous base, namely: Adenine (A), Cytosine (C), Guanine (G), and Thymine (T): A C G T (Fig. 31-1)

The *human body cell* contains 46 chromosomes in a normal nucleus. Females have two X-chromosomes and two sets of 22 autosomes, whereas males have one X- and one Y-chromosome plus two sets of 22 autosomes.

The length of a DNA molecule is measured as the number of base pairs (bp), and a fraction of DNA which is 1000 bp long is called *one kb* (1 kilobase pair).

A *gene* is a part of a DNA molecule that codes for the synthesis of a specific polypeptide chain through its sequence of nucleotides. Each human gene extends over 40 kb in general, but we also posses longer genes. A portion of a human gene - showing 1.5 kb - is drawn in Fig. 31-1.

Fig. 31-1: DNA sequence with nitrogenous bases inside the double helix. The complementary base pairs are held together by weak hydrogen bonds. The drawing is 2-dimensional, whereas reality is 3-dimensional. Thus, the distance between the two DNA strands is approximately identical at any point along the DNA spiral. The details are magnified from the promoter zone of the gene.

The entire *human genome* on the 24 different chromosomes is estimated to be $3*10^9$ bp long and to contain 10^5 genes.

2. Protein synthesis

Each gene has its own promoter sequence. A *promoter sequence* is a particular sequence of A, C, G, and T (Fig. 31-1). The nuclear enzyme *RNA polymerase* recognises a particular sequence in the promoter region, and binds to a region of TATAAT, also called the *TATA box* of the promoter region.

2.1. Transcription (copying). This is the copying of the DNA code into a single stranded immature *messenger RNA* (mRNA). RNA is ribonucleic acid. In the TATA box, the RNA polymerase transcribes a single stranded copy of the DNA sequence. The *transcription process* (copying) stops at the end of the gene. RNA is similar to DNA except that it contains ribose instead of pentose and Uracil (U) instead of T: **A C G U**. RNA is usually single stranded. The RNA polymerase activity is shown in two stages, first with 2 base units and later with 4 (Fig. 31-1).

2.2. Splicing. The so-called *splicing* also takes place inside the cell nucleus (Fig. 31-2). The *coding sequences* or *exons* of the *immature* RNA molecule are cut and spliced together by enzymes eliminating intervening *introns* (ie, *non-coding* sequences). The final product is called *mature messenger RNA* (mRNA).

2.3. Diffusion. The mRNA diffuses through the nuclear membrane into the cytoplasm, where it is bound to ribosomes (Fig. 31-2).

2.4. Ribosomal translation is programmed protein synthesis. Inside the ribosome the particular sequence of the mRNA is read, and the particular sequence of nucleotides is build into a polypeptide (Fig. 31-3). This is the so-called *translation process*, whereby the amino acids are always linked to the polypeptide chain in the proper sequence (Fig. 31-2). Following ribosomal translation, proteins are modified by addition of carbohydrates,

cleavage of bonds within the new protein, shortening or folding. These processes are called *post-translational modification* (Fig. 31-3).

Fig. 31-2: Four steps of genetic copying from DNA in the nucleus to new protein in the cytoplasm via mRNA.

The first nucleotides on the mRNA form a *regulatory sequence*, and do not code for amino acids. This sequence is also called the *5'untranslated region* (Fig. 31-2, green 5'). The nucleotide sequence in the middle of the mRNA code for amino acids - the so-called *coding region*. The end of the mRNA does not either code for amino acids, and makes up the *3'untranslated region* (Fig. 31-2, green 3').

A *codon* or *triplet code* is determined by a base triple (three contiguous nucleotides). A codon provides information about a certain amino acid to be added into the peptide chain or for the chain to stop. Three codons in mRNA (UAA, UAG and UGA) are *stop codons* (Fig. 31-2). The signals encoded by the codon form the *genetic code*.

The amino acids are carried on small RNA molecules called *transfer RNA* (tRNA). Each of these tRNA molecules contains an *anticodon* consisting of 3 unpaired nucleotides, which carry complementary bases to the bases on the mRNA. Additionally, each tRNA carries a specific amino acid for the prolonged polypeptide. There are at least 20 different types of tRNA, each carrying only one amino acid.

Ribosomal RNA (rRNA) molecules diffuse from the nucleus to the cytoplasm and into the ribosomes, where it plays an essential role in the protein synthesis (Fig. 31-3). Every *anticodon* or *base triplet* recognizes its complementary codon in the mRNA. The tRNA deposits its amino acid in the peptide chain.

Fig. 31-3: Increased protein copying in the ribosomes - ribosomal translation.

3. DNA replication

In the cells the nucleotide strands of DNA are separated bit by bit and a special enzyme, DNA polymerase, synthesizes new DNA. DNA replication is normally an accurate process at which exact copies of the DNA molecule are made. Still, a 10-30 errors occur in the DNA replication of each body cell every day. Fortunately, efficient gene reparation processes correct the majority of these errors.

Pathophysiology

This paragraph deals with *1. Mutations*, *2. Genetic disorders*, *3. Gene therapy*, *and* *4. Other strategies.*

1. Mutations

Rarely occurring errors in DNA replication are called *mutations*. Several types of mutation occurs:

1.1. Termination mutations. Mutations involving *stop codons* affect the normal polypeptide chain termination. A nucleotide change allows the insertion of an extra amino acid before termination, or delete one amino acid from the normal sequence. A typical example is the rare type of *α-thalassaemia* called *Haemoglobin constant Spring*. The globin part of normal haemoglobin consists of two a- and two b-chains. In contrast, mutations in the stop codon produce *α-chains* with many extra amino acids in thalassaemia. *Thalassa* means *sea* and originally referred to the distribution of this particular anaemia along the coastline of the Mediterranean Sea. Today thalassaemia is found everywhere.

1.2. Splicing mutations. When mutations occur in the DNA sequence, which normally code the splicing enzymes and thus direct the splicing of introns from RNA, the result may be

abnormal splicing with introns in the final mRNA. Such a mRNA is translated into a protein molecule, which is abnormal because there is an erratic amino acid in the polypeptide chain.

1.3. Insertion-deletion mutations. Insertion or deletion of one or more nucleotides in DNA by mutation results in an abnormal sequence. The gene for a-chains in haemoglobin is duplicated on both *chromosomes 16*. Mutations may result in deletion of one or both genes on each chromosome 16. When all four a-chain genes are absent, the a-chain synthesis is impossible and only *gamma 4-chains* are present. These chains cannot carry oxygen, and infants are stillborn or die shortly after birth (the new-born is pale and oedematous: *hydrops foetalis*). When three genes are deleted, there is anaemia called *α-thalassaemia*. When only one or two genes are deleted there is a microcytosis and polycythaemia called *α-thalassaemia traits* (ie, essentially healthy carriers). Possession of an *angiotensin converting enzyme* gene deleter (ACE-D), which delete a 287 bp *repeat sequence* in the enzyme, results in high concentrations of circulating enzyme, and a significantly higher frequency of myocardial infarction in genotype DD. Deletions in the dystrophin gene of the X chromosome remove coding sequences, so the muscles are deprived of the cytoskeletal muscle protein, *dystrophin.* Dystrophin is the normal gene product, which is normally linked to the sarcolemma as a network to the sarcomers. Lack of dystrophin is the cause of *Duchenne Muscular Dystrophy* (DMD). DMD is a *X-linked recessive disease* caused by the defect dystrophin gene. DMD also occurs spontaneously by mutation in the DMD locus on the X chromosome (ie, the Xp21 region). There is proximal limb weakness with pseudohypertrophy of the calves. The suffering boys have to climb up their legs with their hands, and they have difficulties in walking and running. The creatine phosphokinase concentration in the plasma is substantially elevated. Muscle biopsies show phagocytosis, fibre necrosis, and absence of dystrophin, regeneration, and fatty patches. Exercise helps to preserve muscle function by activation of otherwise inactive synergists. The boys usually die before adolescence (20 years of age).

*1.4. Point mutations***.** Substitution of one nucleotide for another may totally change the codon in a coding sequence. *Carcinogens* can cause point mutations in genomic DNA, and if the mutations occur in the coding region they may be pathogenic. In *sickle cell anaemia* a mutation in the *β-chain gene,* changes the codon from GAG to GTG. The codon GAG initiates translation of a polypeptide chain with glutamic acid and the codon GTG incorporate valin in each of the two b-chains. The product is a highly unstable b-chain, which cannot be utilised for oxygen transport. This particular haemoglobin is called *Haemoglobin S* or *Sickle cell Haemoglobin*. Exposed to low oxygen tensions these molecules form elongated crystals inside the red cells. The spiked ends of the crystals rupture the cell membrane leading to haemolysis and sickle cell anaemia. *Sickle cell anaemia* is the homozygous state, where both genes are abnormal (HbSS), whereas *sickle cell trait* is the heterozygous state (HbAS), with only one chromosome carrying the abnormal gene. The disease (HbSS) manifests itself at about 6 months of age, where the concentration of haemoglobin F decreases towards adult levels.

Different degrees of *β-thalassaemia* also occur following point mutations. The b-chains are not produced or produced to a limited extent, whereas there is an excess of a-chains. The precipitation of a-chains causes haemolysis of erythroblasts and erythrocytes and inefficient erythropoiesis. *Homozygous β-thalassaemia* cases have no b-chains (bo) or too few (b+). In *heterozygous β-thalassaemia* there may be a mild anaemia and usually symptomless microcytosis.

2. Genetic diseases

Genetic disorders are classified into *2.1. chromosomal, 2.2. Single gene defects, and 2.3. Multifactorial disorders*

2.1. Chromosomal defects.

When a chromosome fail to separate during meiosis, the ovum or sperm gets an extra chromosome and become trisomic, or no chromosome and become monosomic. *Sex chromosome trisomy (XXY)* is called *Klinefelter's syndrome*. Sex chromosome monosomy (only one X and no Y) is called *Turner's syndrome* (Chapter 29).

2.2. Single gene defects.

Each diploid cell contains two copies of all autosomes. If one of the two copies has a mutation and the normally produced protein cannot compensate, then an autosomal dominant disorder occurs. Examples are *achondroplasia, porphyria, α_1-antitrypsin deficiency, Huntingtons's chorea,* and *von Willebrand's disease.*

When both chromosomes carry the mutated gene, an autosomal recessive disorder appear. This is not the case when only one mutated gene is present. These patients are heterozygous for the mutated gene and thus unaffected healthy carriers. Examples are *mucoviscidosis* or *cystic fibrosis, Wilson's disease*, and many other inborn errors of metabolism.

Pancreatic cystic fibrosis (mucoviscidosis)

This is a recessive genetic defect with dysfunction of exocrine glands.Cystic fibrosis is an autosomally recessive genetic disorder caused by a cystic fibrosis gene. There is a *gene mutation* in chromosome 7, which result in a defect in a regulator protein (*Cystic Fibrosis Transmembrane Conductance Regulator*, CFTR) - a defect b-adrenergic gated *chloride-channel.* The defective chloride-channel fails to open in response to an increase in intracellular cAMP in the pancreatic ducts, the airways and the sweat glands. The patients have a minimal chloride excretion. The decreased excretion of chloride and increased reabsorption of Na^+ and water produce a small viscoid secretion that closes and dilatates the duct systems. Finally, the ducts are destroyed (ie, chronic respiratory disease and pancreatic insufficiency). Fully manifested cases suffer from defect mucous secretion (*mucoviscidosis*) with chronic respiratory disease, cystic pancreatic fibrosis with pancreatic insufficiency, and abnormally high [NaCl] in sweat.

This life-threatening condition is thus a genetic defect in the *β-adrenergic-gated Cl^--channels* of the glands in the airways, the pancreas, and in the sweat glands.

Bronchopulmonary disorders result in chronic hypoxia with *finger clubbing*.

Pancreatic failure with lack of digestive enzymes results in steatorrhoea and cholesterol gallstones.

A *sweat test* resulting in a Na^+-concentration above 60 mM in the sweat is strongly indicative of cystic fibrosis.

Cystic fibrosis is treated with *amiloride*, which block the ductal Na^+- reabsorption, or with ATP, which stimulates Cl^--secretion by a pathway different from the missing cAMP.

In some cases of cystic fibrosis, with severe pulmonary insufficiency, *lung transplantations* have been performed successfully.

Albinism (AMELANOSIS) is inherited as an autosomal recessive disorder of melanin synthesis. The biosynthesis of the enzyme tyrosinase is defective, which results in lack of melanin. Amelanosis is manifest by white hair, pink-white skin, blue eyes and photophobia.

Phenylketonuria (PKU) is also an autosomal recessive disorder. There is a defect conversion

of phenylalanine to tyrosine and thus hypopigmentation. The genetic defect results in lack of the enzyme phenylalanine 4-hydroxylase. PKU must be diagnosed and treated soon after birth in order to avoid severe mental retardation. PKU patients almost never reproduce. PKU occurs once in 25000 live births in the population.

Carbohydrate malabsorption:

The most common chronic disorder in humans is *lactose malabsorption* or *hypolactasia* (lactose-induced diarrhoea or lactose intolerance), which is due to a genetically deficiency of *lactase* in the brush-border of the duodeno-jejunal enterocytes. More than 50% of all adults in the world are lactose intolerant. Infants with the rare *congenital* lactase intolerance are borne without lactase in their brush border. They develop diarrhoea, when they are breast-fed. This can result in a life threatening dehydration. The amount of lactose entering the colon determines the size of the *osmotic diarrhoea*. Milk made from Soya beans and fructose is well tolerated.

Sucrase-isomaltase deficiency is an autosomal recessive genetic condition with sucrose intolerance. This disorder is found in 10% of Eskimos (Inuits), which is not surprising, since they must have lived for thousands of years of a natural diet without sucrose.

Glucose-galactose malabsorption is a rare genetic disorder caused by a defect in the brush border system for glucose and galactose absorption (GLUT-5). Fructose is well tolerated.

2.3. Multifactorial gene defects.

These disorders involve many genes and often also environmental factors. Examples are congenital pyloric stenosis, asthma, hypertension, schizophrenia, congenital heart disease and genetic cancer.

Cancer is often multifactorial. Cancer arises from only a single cell, which suddenly starts to proliferate out of control. *Oncogenes* encode proteins that normally are involved in cellular proliferation and enhance cellular growth.

Tumour suppressor genes suppress undue cellular proliferation. Mutations in the normal *RB gene* result in *retino-blastoma*.

Mutations in the normal *p53 gene* may lead to brain tumours, carcinomas of the breast and lungs, osteosarcomas, colon carcinomas and leukaemia. A nuclear phosphoprotein (53-kDa) encoded by p53, is involved in DNA repair and synthesis. In many cases of different cancer types there is a secondary *mutation* of the *p53 allele*.

3. Gene therapy

Gene therapy is the application of gene transfer (RNA or DNA delivery to cells) in an attempt to treat genetic or acquired disorders. The transferred DNA must contain promoter zones for transcription and the protein-coding region of the gene.

Insertion of DNA into cells is performed with a large number of techniques, among which plasmid-based (plasmid mixed with lipid micelles), and virus-based (retrovirus, adenovirus, herpes simplex virus) vectors are the most effective. The most widely used vehicle is genetically engineered retrovirus from the Moloney murine leukaemia retrovirus.

Retroviruses can cause cancer, when they infect a cell and activate resting or proto-onchogenes. Retrovirus may also be used in future gene therapy, as described in the following:

1. A normal gene, which can activate the immunodefence system, is placed inside a manipulated, pacified retrovirus (Fig. 31-4).

Fig. 31-4: Gene therapy of cancer.

2. Cancer cells, removed from the tumour of a patient, are cultured in vitro and infected with

the retrovirus carrying the normal gene.

3. Selected cells from the culture are implanted in the cancer tumour of the patient, where the retrovirus attacks some of the cancer cells.

4. The immune-activating genes signal to the immune defence system to forward lymphocytes. The lymphocytes enter the cancer cells and destroy them (Fig. 31-4).

4. Other strategies are the following:

Retrovirus is also used for *gene transfer* of the cytokine, interleukin-2 and of *tumour necrosis factor*, in attempts to support the immune response to cancer.

Bone marrow cells from a patient with *β-thalassaemia* are infected with a *pacified retrovirus* carrying a normal human b-globin gene. Hereby, cells capable of normal erythropoiesis can replace a sufficient number of abnormal bone marrow cells.

Low density lipoprotein (LDL) *receptor deficiency* causes hereditary hypercholesterolaemia. Insertion of the normal gene for the *LDL- receptor* into the patients' abnormal hepatocytes may have beneficial effect.

Gene transfer can also result in the production of a cell membrane protein, such as the *chloride-channel* transmembrane *regulator gene* that has mutated in cystic fibrosis. This is transferred to pulmonary cells by aerosol technique or by vectorial insertion into the target cells.

Equation for Gene Frequency Calculation

In the absence of mutation and random genetic drift, genotypes in a population of random mating will be given by the equation:

$$\text{Eq. 31-1: } (p + q)^2 = 1 \text{ or expanded: } (p^2 + q^2 + 2pq) = 1,$$

where **p** is the *frequency of the normal gene* in the population, and **q** is the *frequency of the abnormal gene*. Obviously, $\mathbf{p^2}$ is the frequency of the normal homozygote, $\mathbf{q^2}$ is the frequency of the abnormal homozygote (all affected by the abnormality), and 2pq is the frequency of healthy carriers. Note that (p + q) is 1.

This is the *Hardy-Weinberg theorem*, which is used to calculate the frequency of abnormal genes in the total gene pool of a population.

Self-Assessment

Multiple Choice Questions

Each of the following 5 statements have True/False options:

A. Phenylketonuria (PKU) is an autosomal dominant disorder.

B. Inside the ribosome the particular sequence of the mRNA is read, and the particular sequence of nucleotides is build into a polypeptide.

C. Sex chromosome trisomy (XXX) is called Klinefelter's syndrome.

D. When carcinogens cause point mutations in genomic DNA inside the coding region, they are often pathogenic.

E. The first nucleotides on the mRNA form a regulatory sequence is called the 5'untranslated region, and does not code for amino acids.

Case History A

A female and a male, who plan to have children, want advice concerning a certain genetic disease. The incidence of the recessive disease is 10^{-4} in the population.

What is the frequency of the recessive gene occurring in one human being?

Case History B

A newborn girl suffers from almost continuous coughing. The mother contacts her doctor, who - at the third consultation - suspects cystic fibrosis and arrange a sweat test to be performed. Pilocarpine iontophoresis facilitates the collection of sweat. The NaCl concentration in sweat is found to be 70 mM, which is several fold the normal value. The patient suffers from bronchopulmonary infection with mucoviscidosis of the exocrine gland ducts, malabsorption, fatty stools (steatorrhoea), and deficiency of fat soluble vitamins (A, D, K). - The incidence of cystic fibrosis is approximately 1 out of 1600 live births.

1. What is cystic fibrosis?
2. Calculate the frequency of the abnormal cystic fibrosis gene.
3. Calculate the frequency of the normal homozygote.
4. Calculate the frequency of the heterozygous carrier of cystic fibrosis.

Case History C

A female, whose father is an albino, plan to marry a male albino (genotype aa). They wish to know the probability of having albino children and albino carriers.

1. What is albinism?
2. What are the probability of having an albino child?
3. What are the probability of having albino carriers among their children?

Case History D

The brother of a Phenylketonuria (PKU) patient seeks genetic advice before marriage. The brother is normal and cases of PKU are excluded for generations in the family of the female.

1. What is PKU?
2. What are the probability that the brother is heterozygous or normal?
3. What are the probability of the couple having a PKU child?
4. PKU patients rarely reproduce. How can the gene persist in the population?

Case History E

A 3-year-old boy has difficulty in standing and walking. It is particularly difficult for him to reach the standing position from the floor, and he has to climb up his legs with his hands. The thighs of the boy are abnormally thin, but the calves are muscular and almost hypertrophic. The creatine phosphokinase concentration in the plasma is 100 times the normal level. – The mother of the boy had a brother, who died at the age of 20 by the same disorder.

1. What is the diagnosis?

2. Why are the creatine phosphokinase in plasma elevated?

3. Why are such disorders confined to boys?

4. Explain the hypertrophy of the calves.

5. What are the risk that the sister to a boy with this muscular disease gets a child with manifest muscular disease?

Try to solve the problems before looking up the answers.

Highlights

- *Nuclear DNA is the dominant controller of protein synthesis in the ribosomes. The codes of special enzymes and other proteins are based on at least 20 amino acids arranged in sequence.*
- *The nucleus contains a nucleolus and the chromatin network. Chromatin is DNA molecules rolled up in a pearl chain structure with proteins called nucleosomes.*
- *DNA consists of two nucleotide strands twisted around each other to form the so-called double helix, which can be several cm in length. The double helix is folded to form chromosomes with a length of approximately 10 μm.*
- *The human body cell contains 46 chromosomes in a normal nucleus. Females have two X-chromosomes and two sets of 22 autosomes, whereas males have one X- and one Y-chromosome plus two sets of 22 autosomes.*
- *The entire human genome on the 24 different chromosomes is estimated to be 3 10^9 bp long and to contain 10^5 genes.*
- *Dystrophin is the normal gene product, which is normally linked to the sarcolemma as a network to the sarcomers. Lack of dystrophin is the cause of Duchenne Muscular Dystrophy (DMD). DMD is a X-linked recessive disease in boys caused by the defect dystrophin gene.*
- *Sickle cell anaemia is the homozygous state, where both genes are abnormal (HbSS), whereas sickle cell trait is the heterozygous state (HbAS), with only one chromosome carrying the abnormal gene. The disease (HbSS) manifests itself at about 6 months of age, where the concentration of haemoglobin F decreases towards adult levels.*
- *Homozygous β-thalassaemia have no β-chains (βo) or too few (β+). In heterozygous β-thalassaemia there may be a mild anaemia and usually symptomless microcytosis.*
- *Each diploid cell contains two copies of all autosomes. If one of the two copies has a mutation and the normally produced protein cannot compensate, then an autosomal dominant disorder occurs. Examples are achondroplasia, porphyria, α_1-antitrypsin deficiency, Huntingtons's chorea, and von Willebrand's disease.*
- *When both chromosomes carry the mutated gene, the autosomal recessive disorders appear. This is not the case when only one mutated gene is present. These patients are heterozygous for the mutated gene and thus unaffected healthy carriers. Examples are mucoviscidosis or cystic fibrosis, Wilson's disease, and many other inborn errors of metabolism.*

- *Pancreatic cystic fibrosis is a recessive genetic disease caused by dysfunction of exocrine glands. The defect is in a transmembrane regulator protein called the cystic fibrosis transmembrane conductance regulator (CFTR). The CFTR represents a β-adrenergic gated chloride channel, which is normally opened by elevated intracellular cAMP.*

- *The cystic fibrosis patients have a minimal chloride excretion and thus as minimal excretion of salt and water into the duct systems. This is what makes all exocrine secretions viscid, the duct systems are occluded and dilatated; finally the ducts are destroyed (ie, chronic respiratory disease and pancreatic insufficiency).*

- *Albinism (AMELANOSIS) is inherited as an autosomal recessive disorder of melanin synthesis. The biosynthesis of the enzyme tyrosinase is defective, which results in lack of melanin. Amelanosis is manifest by white hair, pink-white skin, blue eyes and photophobia.*

- *Phenylketonuria (PKU) is also an autosomal recessive disorder. There is a defect conversion of phenylalanine to tyrosine and thus hypopigmentation. The genetic defect results in lack of the enzyme phenylalanine 4-hydroxylase. PKU must be diagnosed and treated soon after birth in order to avoid severe mental retardation. PKU patients almost never reproduce. PKU occurs once in 25 000 live births in the population.*

- *Multifactorial gene defects. These disorders involve many genes and often also environmental factors. Examples are congenital pyloric stenosis, asthma, hypertension, schizophrenia, congenital heart disease and cancer.*

- *Genetic cancer is often multifactorial. Cancer arises from only a single cell, which suddenly starts to proliferate out of control. Oncogenes encode proteins that normally are involved in cellular proliferation and enhance cellular growth.*

- *Tumour suppressor genes suppress undue cellular proliferation. Mutations in the normal RB gene result in retino-blastoma.*

- *Gene transfer can result in the production of a cell membrane protein, such as the chloride-channel transmembrane regulator gene that has mutated in cystic fibrosis. This is transferred to pulmonary cells by aerosol technique or by vectorial insertion into the target cells.*

Further Reading

Nature. Weekly journal published by Macmillan Magazines Ltd, Porters South, 4 Crinan Street, London N1 9XW, UK.

Jorde LB, Carey JC, Bamshead MJ and RL White. *Medical Genetics*. Mosby, St Louis, 1999,

Scientific American. - Example: Verma, IM: Gene therapy. *Sci Am* pp 68-72, 81, 82, 84. Nov 1990.

Return to top

Return to Content

Chapter 32

Immunology And Immune System Disorders

Study Objectives

- To *define* the following concepts and disorders: AIDS, allergy, autoimmune reactions, HIV, interleukins, neutropenia, the reticulo-endothelial system (RES), serum disease, and vaccination.
- To *describe* anaphylactic reactions, the role of natural killer cells, of soluble lysozymes, active and passive immunisation.
- To *explain* the mechanisms involved in allergy, acquired and congenital immunity, haemopoiesis, phagocytotic killing, myasthenia gravis, ulcerative colitis and Crohns disease, rheumatoid arthritis, insulin-dependent diabetes mellitus, pernicious anaemia, and Graves hyperthyroidism.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The immune system defends the body against invading agents, participates in autoimmune and hypersensitivity disorders, and determines transplant tissue reactions.*
- *The ability to recognise foreign antigens allows destruction and removal of invading organisms by various effector mechanisms.*
- *Inappropriate immune reactions against self-antigens or host cells result in autoimmune disorders.*
- *Overt responses to an antigen result in hypersensitivity disorders.*

Definitions

- **Acquired Immuno-Deficiency Syndrome** (AIDS) results from infection with *Human Immuno-Deficiency Virus* (HIV). This is a substantial T-lymphocyte defect. The HIV is bound to receptors on $CD4^+$ T-lymphocytes, but also to monocytes and macrophages.
- **Autoimmune disorders** occur when the immune system kills the body's own cells or *self-antigens*, because the system fails to recognise them. Autoantibodies are produced and immune complexes are deposited in the tissues. Here, complement and neutrophils are activated and accumulated. The neutrophils release proteolytic enzymes and toxic cytokines.
- **Delayed allergic reactions** occur when highly resistant allergens (mycobacterium such as TB, fungi, contact allergens) get into contact with T-lymphocytes, and do not involve production of antibodies. Repeated contact with antigen *activates* helper T- and killer T-cells.
- **Interleukins** are toxic lymphokins and monokines from lymphocytes and monocytes.

- **Natural killer cells** destroy tumour cells and virus-infected cells. They are unspecific, non-phagocytotic lymphocytes that are activated by *interferon* produced by the affected cell. Interferon induces a high degree of resistance in the affected cell.

- **Neutropenia** are too few neutrophils in the peripheral blood.

- **The reticuloendothelial system** (RES) is also called the *mononuclear phagocytotic system* (MPS). Lymphoid organs belonging to the RES are the bone marrow, the liver, the spleen, the lymph nodes, the microglia of the brain, the thymus, tonsils, as well as MALT, BALT, GALT, and SALT (see below).

- **Serum disease** results from passive immunisation with too many antibodies. Antigen-antibody complexes accumulate in the blood. They precipitate in the slow bloodstream of the capillaries.

- **Soluble lysozymes** are enzymes in plasma, lymph, extracellular fluid, saliva, gastric fluid and other secretions, which can destroy the bacterial wall.

- **Vaccination** is iatrogenous immunity. At the first vaccination with a certain antigen, some plasma cells transform to memory B-cells that remain in the RES. At the second vaccination, the memory B-cells evoke an exaggerated antibody production that rapidly deactivates the antigens.

Essentials

This paragraph deals with *1. The immune system*, *2. Congenital immunity*, *3. Acquired immunity*, *and 4.Vaccination.*

1. The immune system

Burnet received the Nobel Prize in 1960 for his *clonal selection theory*. Pluripotent stem cells differentiate into millions of different B-lymphocytes, T-lymphocytes, erythrocytes, polymorphonuclear leucocytes, monocytes, macrophages and mast cells.

Lymphocytes are the most important cells involved in immune responses. Without exposure to all the antigens each of the B-lymphocytes have inherited the ability to divide into a clone of plasma cells. The first contact with the specific antigen starts the clone production. The clone of plasma cells produces the specific immunoglobulins. This understanding of the immuno-reaction made transplantation possible.

Thomas made the first *transplantation* of kidneys in 1956, and received the Nobel Prize in 1990 for his contribution to science and therapy. The second successful transplantation was the transplantation of bone marrow to treat *leukaemia* (ie, uncontrolled proliferation of *impotent* leucocytes).

Overactivity in the immune system causes allergic and autoimmune disorders, whereas underactivity results in immunodeficiency.

The *immune system* is a complex of cells and humoral factors controlled by the hypothalamo-hypophysary axis in concert with the adrenal and probably other endocrine glands (Fig. 32-1). The major organs of the *reticuloendothelial system*, RES (bone marrow, lymph nodes, spleen, and thymus) receive sympathetic efferents - in particular the T-lymphocyte regions (see below). Hereby, the CNS (via the hypothalamus) modulates the intensity of immunoreactions. Endotoxins from the normal bacterial flora of the intestine constantly enter the blood. Ordinarily they are inactivated by phagocytic activity of the RES mainly in the liver.

Macrophages not only inactivate endotoxins; they also release hydroxylase, proteases, certain coagulation factors and arachidonic acid derivatives (ie, prostaglandins, thromboxanes, leucotrienes and monokines). *Monokines* are control proteins that modulate metabolism (Chapter 20), temperature control (Chapter 21), hormone secretion (Chapter 26>30), and the immune defence systems. The important role of RES in haemorrhagic and endotoxic shock is described in Chapter 12.

Fig. 32-1: Control of the immune system by the hypothalamo-pituitary axis during an antigen attack.

The *immune system* protects us against disease. The system confers congenital (inborn) and acquired immunity. Both subsystems activate soluble humoral factors as well as fixed and mobile cells.

Congenital immunity involves T-lymphocytes, which are derived from the thymus, whereas acquired immunity involves B-lymphocytes and the production of antibodies.

The *bone marrow* is the site of haemopoiesis, since all blood cells are derived from the *pluripotent stem cell* or haemocytoblast (Fig. 32-1). This is a primitive cell type, which can divide rapidly and differentiate into *committed stem cells*. The committed stem cells are colony- forming in that they are committed to produce large quantities of *erythrocytes, granulocytes* (neutrophils, eosinophils and basophils), *monocytes-macrophages, megacaryocytes-blood platelets*, and *B- & T-lymphocytes* (Fig. 32-1) depending upon various growth inducers or cytokines.

Interleukines are *toxic lymphokines* and *monokines* from lymphocytes and monocytes. They inhibit the hypothalamic production and release of *corticotropin-releasing- hormone* (CRH) just as *cortisol* - thus reducing the immuno-response. Cortisol also inhibits lymphocyte and monocyte production from stem cells (Fig. 32-1). Normally, CRH stimulates the synthesis and release of *adrenocorticotropic hormone* (ACTH). Stimulation of the sympathetic nervous system by immunological stress releases *adrenaline* from the adrenal medulla. Adrenaline probably stimulates most of the blood cell formation from stem cells (erythro- granulo- lympho- monocyto- thrombo- poiesis) via a_2-adrenergic receptors.

2. Congenital immunity

The inborn immune defence system is *unspecific* and responsible for *immediate responses* to infection (bacteria, fungi, parasites, and viruses) and other pathogens (from tumours or other sources).

The inborn system is immediately activated with all the elements of congenital immunity: *Phagocytes* (neutrophils and macrophages), *cytotoxic eosinophils*, histamine-containing *basophils* and *mast cells*, and *essential* proteins (complement, acute phase proteins, heat shock proteins).

Phagocytes comprise a large number of neutrophils, which are released from the bone marrow during acute infection. *Neutrophilic granulocytes* have an extremely short life cycle, namely 24 hours. They are leucocytes formed in the bone marrow. The production of neutrophils is increased by the action of *granulocyte-colony stimulating factor* (G-CSF) and *granulocyte-macrophage-colony stimulating factor* (GM-CSF). During severe long-lasting infections the bone marrow is exhausted and too few neutrophils are released to the blood (ie, *neutropenia*). Neutrophils are important in the defence against microorganisms.

Fig. 32-2: Congenital immune defence against bacteria.

Granulocytes can leave the blood by moving between endothelial cells to reach the interstitial space of different tissues.

2.1. Steps of microbial destruction

Bacterial invasion. When bacteria invade the body, macrophages release the complement cascade, and B-lymphocytes release immunoglobulins (Fig. 32-2).

Neutrophilic chemotaxis. Complement cascade products and leucotriene B_4 (see later) are released from cells in the infection area. These molecules attract neutrophils from the blood into the infected tissue by so-called *chemotaxis* (ie, attraction of cells by foreign chemical substances). The neutrophils pass the endothelial wall by *diapedesis* (ie, they squeeze through the capillary wall - see Fig. 32-2). Neutrophils surround the microbe with their *pseudopodia* and engorge them. Neutrophils are large enough to phagocytize bacteria and fungi, but they cannot phagocytize larger organisms such as parasites.

Phagolysosomes. The pseudopodia form a membrane bound vesicle around the microbe, and the vesicle is then released as a free-floating *phagosome*. Inside the neutrophil, the phagosome fuses with *neutrophil granules* to form *phagolysosomes*, where the killing occurs. Phagocytes get hungry from *opsonization* of the pathogen surface with complement or with specific immunoglobulins such as IgM and IgG.

Microbial perforation. The complement released from many macrophages also fights its own battle. Besides being bound to immunoglobulins, complement is also bound to the surface of bacteria, whereby they get leaky.

Microbial breakdown. Phagocytotic killing occurs in the *phagolysosomes*. The method of execution is by a *respiratory burst* or by *gas*. Oxygen is reduced to reactive oxygen metabolites by an NADPH oxidase. These reactive metabolites are hydrogen peroxide and oxygen radicals. Many toxic proteins or enzymes (lipases, proteases) take part in the destruction. Immuno-stimulated macrophages produce nitrite and nitrate, and their killer activity is related to the unstable gas, *nitric oxide* (NO). NO is produced in large quantities by the macrophages, kills microbes and cancerous cells. NO is synthesized from one of the guadino nitrogens of L-arginine by the enzyme *nitric oxide synthase*. Several synthases have been purified and cloned. The enzymes represent a new family that contains a haeme moiety.

Soluble lysozymes are enzymes in plasma, lymph, extracellular fluid, saliva, gastric fluid and other secretions, which can destroy the bacterial wall (Fig. 32-2).

Specific antibodies and complement *cascade* substances ease the execution of microbes. Neutrophils carry receptors for immunoglobulins and complement on their surfaces, which increase the binding force between the cell and the microbe, and simultaneously transduct signal molecules to increase the enzymatic killing activity. This is a typical co-operation between congenital and acquired immunity. The capillaries in the area dilate and get leak for proteins. This is why the site of invasion gets hot, red, swells, and pains (Latin: calor, rubor, tumour, and dolor -the classical signs of inflammation).

2.2. Cytotoxic eosinophilia

Eosinophils contain granules with substances, which become cytotoxic, when they are released on the surface of parasites. Thus, eosinophils are mainly involved in reactions against parasitic infections. Eosinophils are not phagocytic, but they intoxicate nematodes and other parasites and bacteria. The cytotoxic substances are *major basic protein*, which kill helminths, *eosinophil cationic protein* (an extremely efficient killer of parasites and potent neurotoxins) and *eosinophil peroxidase* (kills bacteria, helminths and tumour cells). Eosinophils are involved in hypersensitivity reactions.

2.3. Histamine containing cells

Circulating basophils and mast cells residing in the tissues are morphologically similar with granules that contain histamine and other vasoactive amines. These histamine containing cells are involved in hypersensitivity reactions (see IgE-mediated allergy). The binding of IgE to the cells stimulate the release of histamine, but also of prostaglandins, leucotrienes and cytokines. These substances cause immediate (Type I) hypersensitivity. The T- mast cell contains trypsin and cytoplasmic IgE, and the TC- mast cell contains both trypsin and chymotrypsin.

2.4. Natural killer cells

Such cells destroy tumour cells and virus-infected cells. They are unspecific, non-phagocytotic lymphocytes that are activated by *interferon* produced by the affected cell. Interferon induces a high degree of resistance in the affected cell.

2.5. Essential proteins

Complement extirpates microbes and immune complexes. The complement system includes several serum glycoproteins that are activated in a *cascade* similar to the *coagulation cascade* (Chapter 8). Complement activation destruct and removes microbes, immune complexes and tumour cells, recruits cells and proteins to infection sites by chemotaxis, and modulates the *B-cell immune response*.

Acute phase proteins (C-reactive protein, complement complex, fibrinogen, haptoglobulin, caeruloplasmin, a_1-antitrypsin) are produced in response to infection and inflammation (ie trauma, necrosis, tumours etc). The disease activity is measured in blood serum as *C-reactive protein*.

Heat shock proteins preserve the protein structure of cells during infections. They resemble antigens and are involved in immunity and autoimmunity.

3. Acquired immunity

Antigen stimulation of inactivated lymphocytes results in development of *humoral*- or *cell-mediated* immune responses. Humoral responses involve antibodies from B-lymphocytes activated to large antigen-producing plasma cells. Also macrophages and T-helper cells are required. Cell-mediated responses require cells that produce antigens and cytokines to T-helper cells.

Some pathogens can prevent phagocytosis or suppress the formation of lysosomes or kill the neutrophils. When attacked by such pathogens we must rely on acquired or specific immunity. This is produced by rearrangements of germ-line DNA in B- and T-lymphocytes. Hereby, specific *antibodies* and specific antigen-binding T-cell *receptors* are created. Tonegawa has shown how a rearrangement of DNA in only a few genes can produce millions of different antibodies in an individual. This is enough to cover all antigens encountered.

In foetal life, cells from the bone marrow pass through the gastrointestinal lymph nodes. Here the inactive cells become immunologically *active B-lymphocytes*. The cells re-enter the blood and migrate to the foetal spleen, liver and other lymph nodes. When an antigen binds to receptors on these cells, the lymphocytes divide, and from now on the whole clone of plasma cells can produce the specific antibody.

3.1. The reticuloendothelial system (RES)

This system is also called the *mononuclear phagocytotic system* (MPS). Lymphoid organs belonging to the RES are the bone marrow, the liver, the spleen, the lymph nodes, the microglia of the brain, the thymus, tonsils, as well as MALT, BALT, GALT, and SALT (see below). These organs contain macrophages originating from monocytes. Inactivated and

circulating macrophages are called monocytes, but when they nitigate into extravascular tissues they are known as macrophages. Macrophages contain lysosomes filled with various catabolic enzymes. The macrophage membrane contains receptors for binding complement components and immunoglobulins. Macrophages destroy other phagocytized organisms or molecules by production of free radicals and digestive enzymes. Tumour necrosis factor (TNF) is produced by macrophages stimulated by bacterial cell wall components. TNF turns a tumour into haemorrhagic necrosis. Recombinant TNF is available in the form of TNF-a and TNF-b (lymphotoxin).

The cell content of the RES organs covers fixed and locally wandering macrophages as well as B-lymphocytes, which produce the antibodies after antigen exposure, and are now called plasma cells. B-lymphocytes comprise 25% of all lymphocytes. The remaining lymphocytes (75%) are T-lymphocytes, which are undergoing a maturation process in the thymus. T-lymphocytes possess distinct cell surface antigens. RES receive sympathetic efferents. Hereby, the hypothalamus can modulate the intensity of immunoreactions. This is what is termed the *psycho-immune coupling process* (Fig. 32-1).

The *spleen* is the largest lymphoid organ in the body containing both B- and T-lymphocytes. The *lymph nodes* are distributed all over the body. The *thymus* contains cells that originate from the bone marrow. The lymphocytes derived from the thymus are called T-lymphocytes. The immature T-lymphocytes are matured to $CD4^+$ and $CD8^+$ by thymic hormones. The thymus also deletes cells that are reactive to the body's own tissues (*clonal deletion*). MALT, BALT, GALT, and SALT are lymphoid tissues found in the intestinal mucosa (mucosa-associated lymphoid tissue = MALT), in the wall of the main bronchi (bronchus-associated lymphoid tissue = BALT), in the gut (gut-associated lymphoid tissue = GALT), and in the skin (skin-associated lymphoid tissue = SALT). *Tonsils* combat airborne antigens by the help of antigen producing B-lymphocytes.

Fig. 32-3: Formation of sensitised lymphocytes, lymphokines and antibodies. B-lymphocytes are involved in acquired, humoral immunity, and T-lymphocytes in congenital, cellular immunity.

3.2 T-lymphocytes

acquire their immune competence in the thymus (Fig. 32-3). They are divided into *helper T-cells* and *killer T-cells*. Helper T-cells carry CD4 protein on their surface and produce *lymfokines* (interferon and interleukin-2 and -4). Helper T-cells help the killer T-cells to proliferate, to destruct antigen and to reinforce antibody production. Some external antigen molecules are processed in macrophages before they bind to the lymphocytes (Fig. 32-3). Helper T-cells activate *resting B-lymphocytes*, so they differentiate to *plasma cells* or to *active B-lymphocytes*. Some new cells develop to plasma cells and remain in the lymph nodes.

Interleukin-2 is a peptide of 133 amino acid moieties. This substance stimulates the production of lymphokine-activated killer cells that destroy tumour cells without affecting normal cells. Interleukin-3 stimulates the primitive stem cell.

Interferon is called so, because they interfere with viral RNA and protein synthesis; interferon probably induces enzymes that destroy viral RNA and other proteins. Human can produce at least 3 types of interferon (a, b, g); they are glycoproteins with a molecular weight of 20-25 kD.

With viral or RNA stimulation, a-interferon is synthesized in macrophages (Fig. 32-4) and b-interferon in fibroblasts and macrophages. The g-interferon (no sequence homology to the two

other forms) is produced in antigen-stimulated T-lymphocytes. The g-interferon stimulates the antigen production in macrophages and B-lymphocytes.

Recombinant interferon is commercially available. Interferon is used in the treatment of severe attacks of condylomata acuminata, chronic hepatitis B or C, and certain types of sarcomata.

Fig. 32-4: Production of interferon in macrophages following stimulation with RNA or virus.

T-lymphocytes constitute the majority of blood lymphocytes. The lymphocytes proliferate at first contact between antigen and T-lymphocytes. Some new cells bind the antigen in an antigen-antibody reaction and destroy the antigen. *Killer T-cells* is the proper name for these cells, but the destruction of antigen requires the co-operation of *helper T-cells*. Helper T-cells stimulate the proliferation and differentiation of killer T-cells to increase their number. A subgroup of *effector T-cells* can suppress antibody formation by *B-lymphocytes* and inhibit other effector T-cells. the so-called *suppressor T-cells*. Congenital immunity is a delayed form of immunity. The response reaches a peak after 2 days. Delayed immunity reaction encompasses the rejection of transplants, contact allergies and defence reactions against certain viruses and fungi. The T-cell number is deficient in AIDS victims (*Acquired Immune Deficiency Syndrome*).

The T-lymphocytes recognise self-antigens, known by the body's own cells and *non-self antigens* from cancer cells, foreign cells (transplantates) and foreign molecules (external antigens). This recognition ability is acquired early in life, when lymphoid stem cells migrate into the thymus, where a few are modified into *memory T-cells* and released to the blood.

At the first contact between antigen and T-lymphocyte the cell proliferates. Some of the new lymphocytes are killer T-cells. They bind the antigen in an antigen-antibody complex and destroy the antigen. Killer T-cells carry *CD8 protein* on their surface and kill other cells suffering from cancer or virus infection. Some of the killer T- cells are actually *suppressor T-cells*, because they can suppress antibody formation by the B-lymphocytes and inhibit other effector T-cells. Hereby they can down-regulate the immune response when necessary to prevent autoimmune responses.

3.3. B-lymphocytes

produce *specific antibodies* or immunoglobulins to antigens (Fig. 32-3). The immunoglobulins form part of the gamma-globulin fraction in plasma. The B-lymphocytes multiply by *clonal expansion*. A specific antigen is bound and recognized to the B cell through special surface immunoglobulins. B cells also contain CD19 and CD20 proteins. Cytokines activate the B-lymphocyte, so it divides and the resulting cells differentiate to enormous plasma cells with an overwhelming surplus of protein-producing endoplasmic reticulum (ER). This is why plasma cells produce large amounts of antibodies and release them into the blood as Y-shaped molecules (Fig. 32-3). The plasma cells have a short life cycle, and die when they have fulfilled their defence mission. Hereby, the B-lymphocyte population is reduced to its normal size apart from a few cells remaining as *memory cells*. The antibodies are also called *immunoglobulins* (Ig). They are specific serum glycoproteins. Each antibody is Y-shaped and consists of heavy and light polypeptide chains. The heavy chains with complement receptors provide the *constant domain* of the Ig molecule, which is the same in all antibodies (Fig. 32-5). The light chain region constitutes the *variable domain*, which is functionally important. Antibodies deactivate antigens by forming a complex, which causes agglutination and precipitation, by masking the active sites of the antigens, or by activating the complement cascade. A single Ig with its antigen activates a complement cascade with mobilisation of up to 10^9 new complement molecules carrying lots of enzymes that rapidly lyse the antigen-carrying microbe.

The most abundant is *IgG*, which has a high antigen affinity and is the antibody of the secondary response to protein antigens (viruses and tetanus toxin). IgG can cross the placental barrier and protect the newborn for a couple of months.

IgM is confined to the blood, because it is a pentameric molecule (5 IgM molecules joined together). IgM cannot cross the placental barrier, and is responsible for the primary immune response.

IgA_1 predominates in serum, whereas IgA_1 and IgA_2 are present in equal amounts in secretions such as saliva, gastric juice, pancreatic and intestinal juice. IgA protects mucosal surfaces in the gut, respiratory and urinary tracts, by preventing the attachment of poliovirus, enterovirus, bacteria, and enterotoxin.

The concentration of *IgD* in serum is high in disorders with B-lymphocyte activation such as *AIDS* and *Hodgkin's disease*.

IgE is mainly bound to basophils and mast cells, and involved in the pathogenesis of *allergic* and *nematode diseases*.

Fig. 32-5: An immunoglobulin (Ig) or antibody molecule with two antigen molecules attached (left). The immunoglobulins IgA and IgM are build up of two or more immunoglobulins moieties connected with disulphide bonds (right).

4. Vaccination

This is iatrogenous immunity. At the first vaccination with a certain antigen, some plasma cells transform to memory B-cells that remain in the RES. At the second vaccination, the memory B-cells evoke an exaggerated antibody production that rapidly deactivates the antigens.

Vaccine from death microbes is used for bacterial diseases such as diphtheria and typhoid fever. Other vaccines are derived from toxins that are deactivated without losing their antigen specificity (tetanus, botulism). Vaccines against viral disease have passed through a series of other organisms, until a mutant originates without pathogenic activity but with intact antigen specificity (measles, polio, smallpox, and yellow fever).

Tumour-antigen vaccines are under development in order to stimulate an immune reaction against tumour cells.

Pathophysiology

This paragraph deals with *1.Congenital and acquired immuno-deficiencies due to underactivity of the immune system, and 2. Allergic and autoimmune diseases caused by overactivity of the immune system.*

1. Congenital and acquired immuno-deficiencies

The 5 major types of *congenital immuno-deficiency* are:

1. Chronic granulomatous disease, which is a congenital defect of the neutrophil killing mechanism.
2. Complement cascade deficiencies.
3. B-lymphocyte defects with antibody deficiency.
4. Absent thymus with T-lymphocyte deficiency.

5. Combined immunodeficiency with severe defects of both B- and T- lymphocyte function.

Acquired immunodeficiency (AID)

Iatrogenic AID is caused by splenectomy, chemotherapy or iatrogenically induced malnutrition. Glucocorticoids suppress immune responses and the movement of lymphocytes from the blood to the tissues.

Autoimmune suppression by infection is *Acquired Immuno-Deficiency Syndrome*. AIDS results from infection with *Human Immuno-deficiency Virus* (HIV). This is an overwhelming T-lymphocyte defect. The HIV is bound to receptors on $CD4^+$ T-lymphocytes, but also to monocytes and macrophages. HIV and AIDS are described in Chapter 33.

2. Allergy & autoimmunity

Allergic and autoimmune disorders are typically *hypersensitivity reactions*.

Allergic reactions are either *immediate anaphylactic (type I) or delayed (type II). Autoimmune conditions are also called type III hypersensitivity reactions.*

2.1.Immediate anaphylactic reactions (type 1)

occur immediately when the allergen sensitises B-lymphocytes with allergen specific IgE antibodies (ie, within minutes). At the next exposure to the same allergen, plasma cells recognise the allergen and release large amounts of IgE (Fig. 32-6). The allergen-IgE complex is bound to IgE-receptors on the surface of the mast cells and basophils. Hereby, histamine, serotonin (a vasoconstrictor), lymphokines, *platelet activating factor* (PAF) and *toxic eichosanoids* (prostaglandins and leucotrienes) are released (Fig. 32-6). PAF activates both thrombocytes and phagocytes. Histamine is a powerful vasodilatator and bronchoconstrictor. *Leucotrienes* were previously called *slow reacting substances for anaphylaxis* (SRS-A). Leucotrienes are strong bronchoconstrictors causing bronchial asthma attacks. Cells participating in inflammatory and immune responses are suppressed by prostaglandin E_2 (PGE_2). When PGE_2) stimulates adenylcyclase in activated neutrophils, they release less leucotriene (LTB_4) than else; the same mechanism is probably functioning in other cells.

Atopic allergy is genetic and characterized by large amounts of IgE in the blood. An antigen molecule with multiple binding sites can bind to many IgE molecules on the mast cell and release large amounts of anaphylactic substances (Fig. 32-6). Three typical disorders are the following:

a. The so-called *anaphylactic shock* is often fatal shortly after the histamine release. Adrenaline injected intravenously or intracardially may save the victim.

b. *Urticaria* and *hay fever* are anaphylactic reactions in the skin and the nasal mucosa, respectively. The histamine released causes vasodilatation, with increased capillary pressure and ultrafiltration causing oedema and red colour.

c. *Bronchial asthma* in an atopic person is a bronchiolar anaphylactic response. Leucotrienes from the bronchiolar mast cells cause bronchoconstriction, mucosal infiltration with inflammatory cells, mucosal oedema and hypersecretion. Leucotrienes are blocked by a variety of bronchodilatators.

Fig. 32-6: Anaphylactic reaction as it occurs in mast cells and basophils.

2.2. Type II:

Delayed allergic reactions occur when highly resistant allergens (mycobacterium such as TB, fungi, contact allergens) get into contact with T-lymphocytes, and do not involve production of antibodies. Repeated contact with antigen *activates* helper T- and killer T-cells. These cells diffuse into the skin, lungs or other organs and there release a cellular immune response.

2.3 Type III:

Autoimmune disorders occur when the immune system kills the body's own cells or *self-antigens*, because the system fails to recognise them. Autoantibodies are produced and immune complexes are deposited in the tissues. Here complement and neutrophils are activated and accumulated. The neutrophils release proteolytic enzymes and toxic cytokines. They may kill cells in A) a single organ system or B) affect multiple systems.

2.3.A) Single organ disorders:

Insulin-dependent diabetes mellitus (IDDM). Newly presenting patients possess islet-cell antibodies that have destroyed all the insulin producing b-cells of the pancreatic islets (Chapter 27). There is an inherited increased susceptibility to IDDM.

Pernicious anaemia. Typically parietal cell antibodies are found in the blood. They may kill the entire parietal cell population leading to atrophy of the mucosa. The atrophic gastric mucosa fails to produce hydrochloric acid and intrinsic factor for vitamin B_{12} (Chapter 8). Also intrinsic factor antibodies have been found, either able to block the complex binding between intrinsic factor and vitamin B_{12}, or capable of blocking the binding of the intrinsic factor and vitamin B_{12} -complex to its ileal receptors.

Graves's or Basedow's disease is hyperthyroidism combined with eye signs (*exophtalmus*). Normally, the thyroid stimulating hormone (TSH) increases the thyroid hormone production after its binding to the *thyroid TSH receptors*. Bacterial infection in a genetically susceptible person may be the cause of autoimmune production of the TSH-receptor antibodies. These IgG -antibodies behave exactly like TSH itself, and thus stimulate the thyroid hormone production (Chapter 28). Retroorbital swelling and damage of the extraocular muscles cause the thyroid eye disease from specific antibodies.

Atrophic hypothyroidism is caused by microsomal autoantibodies in the thyroid. The hypersensitivity reactions result in thyroid atrophy and hypothyroidism (Chapter 28).

Hashimoto's thyroiditis is caused by other microsomal autoantibodies. The inflammatory reaction produces goitre (struma) with or without hypothyroidism.

Primary hypo-adrenalism (Addison's disease). The entire adrenal cortex is destroyed by organ-specific autoantibodies, and its steroid hormone production (sex hormones, glucocorticoids, and mineralocorticoids) is lost (Chapter 30). Autoimmune disease is the cause of destruction in 4 of 5 patients. Other causes are TB, cancer, infections, and AIDS.

Myasthenia gravis. This is a rare disease caused by autoantibodies against the *acetylcholine receptors* of the neuromuscular endplate (Chapter 2). Complement complexes and IgG molecules are deposited at the endplates. Many patients have thymic hyperplasia.

Crohns disease and ulcerative colitis

These two disorders may be different manifestations of a single disease, *non-specific inflammatory bowel disease.*

Ulcerative colitis is a disorder located to the *colon* only, whereas *Crohns disease* can involve any part of the gastrointestinal tract.

Fig. 32-7: Inflammatory bowel disease (Crohns disease).

Transmissible agents are suspected but not found. Both antigen specific, cellular and autoimmune responses have been postulated. The immune-hypersensitivity is evident from activation of T-lymphocytes, neutrophils, eosinophils, basophils and mast cells. A variety of cytokines, Eicosanoids, oxygen radicals and NO are produced. Dysfunction of the terminal ileum has serious consequences, because bile salts are not reabsorbed and vitamin B_{12} is not absorbed.

The cause of inflammatory bowel disease is unknown, but some features are common to the two conditions:

1. *Autoimmune disease* in the gut lumen and wall.
2. *Cytotoxic T lymphocytes* sensitised by antigens destroy the mucosa or the whole gut wall.
3. *Genetic predisposition* (concordance in monozygotic twins, familial occurrence),
4. Infective agents are *suspected* but not identified.
5. Uveitis, episcleritis, arthritis, anchylosing spondylitis, erythema nodosum occur in both types of bowel disease.
6. The clinical picture includes abdominal pain and diarrhoea often with blood and mucus. *Steatorrhoea* typically points to ileal involvement, whereas frequent *bloody diarrhoeas* indicate colic involvement as in ulcerative colitis. There is an increased risk of colorectal carcinoma in ulcerative colitis.

Crohns disease is a chronic infection or inflammation of the gut with a particular prevalence for the *terminal ileum*, but it can be located all the way along the tract. The patients with terminal ileitis suffer from malabsorption of bile salts and vitamin B12.

Ulcerative colitis is always confined to the colon and it is a mucosal inflammation with haemorrhage and rectal bleeding.

Special hospital units accomplish different diagnosis and management.

Gluten-sensitive enteropathy or *coeliac disease* (sprue) describes a condition where the duodenal and jejunal mucosa is more or less destroyed by hypersensitivity towards gluten. Gluten is found in barley, rye, wheat, and oats. Gluten consists of 4 *gliadin peptides* out of which *α-gliadin* destroys the jejunal mucosa. There is a sequence homology between a-gliadin and *adenovirus 12*.

Nontropical sprue. In allergic persons *gluten* causes an immunological reaction with desquamation of the luminal part of the intestinal mucosa - in particular most of the microvilli. The marked fall in area available for absorption causes malabsorption. In severe cases the malabsorption involves fats causing steatorrhoe, Ca^{2+} causing osteomalacia, vitamin K causing bleeding disturbances, vitamin B_{12} causing pernicious anaemia, and folic acid causing folic acid deficiency.

The inheritance is unknown, but this is a disease occurring in atopic families, and an immunogenetic mechanism is likely - in particular after infection with *adenovirus 12*. Symptoms and signs include stomatitis with ulcers, diarrhoea, steatorrhoea, abdominal pain, osteomalacia and malnutrition with oedema. The desquamation of the mucosa sometimes descends along the villous region in the small intestine and thus includes the terminal ileum.

A gluten-free diet is necessary for successful treatment.

Tropical sprue is found in tropical areas, and probably caused by gastrointestinal infection, although the bacterial diagnosis is seldomly confirmed. Tropical sprue is often curable with antibacterial agents.

2.3.B) Multiple system disorders:

Rheumatoid arthritis (symmetrical inflammatory polyarthritis with progressive joint damage and lung, cardiac, renal and many other organ manifestations) is an autoimmune disease. The synovial fluid contains IgG, lymphokines and immune complexes. The cause is probably a persistent external antigen, which is not removed.

Connective tissue diseases are *Systemic Lupus Erythematosus* (SLE = Arthralgia, rash, cerebral and renal dysfunction), systemic sclerosis (sclerosis of the skin and oesophagus, organ fibrosis, Raynaud's phenomenon), polymyositis (necrosis of muscle fibres with proximal weakness), dermatomyositis (polymyositis with rash) and other rare disorders. The aetiology is unknown, but many of these conditions have circulating autoantibodies and immune complex deposition.

Several autoimmune disorders show a strong association.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. The pluripotent stem cell is not identical with the haemocytoblast.

B. Interferon induces a high degree of resistance in the affected cell.

C. The hypothalamus can modulate the intensity of immunoreactions through sympathetic efferents to the RES.

D. Hyperthyroidism is always combined with exophtalmus.

E. The terminal ileum does not have essential functions.

Case History

A female, 39 years of age, works as a secretary. For 6 months she has been working with ***a*** *new copy machine. Over the last three months she has developed a blistering rash on both hands and the rash is now turned into deep scaling of the skin. There is a mild scaling at the site of both her earrings. The patient has used a bland cream without effect. The dermatologist examines her with patch tests, where solutions of common allergens, chemicals, metals including nickel are placed on her back and occluded with dressings. The next day (24 hours later) the dressings are removed and the exposed areas examined. The examination is repeated 48 hours after the beginning of the exposure. There is a severe rash at the site of the colour powder used in the copy machine, and a mild rash at the nickel spot.*

1. What is the diagnosis?

2. Which therapeutic strategy is recommendable?

Try to solve the problems before looking up the answers.

Highlights

- *The immune system is a complex of cells and humoral factors controlled by the hypothalamo-hypophysary axis in concert with the adrenal and probably other endocrine*

glands.

- *The inborn immune defence system is unspecific and responsible for immediate responses to infection (bacteria, fungi, parasites, and viruses) and other pathogens (from tumours or other sources).*
- *Congenital immunity involves T-lymphocytes, which are derived from the thymus, whereas acquired immunity involves B-lymphocytes and the production of antibodies.*
- *The bone marrow is the site of haemopoiesis, since all blood cells are derived from the pluripotent stem cell or haemocytoblast, which can divide rapidly and differentiate into committed stem cells.*
- *The committed stem cells are colony- forming in that they are committed to produce large quantities of erythrocytes, granulocytes (neutrophils, eosinophils and basophils), monocytes-macrophages, megacaryocytes-blood platelets, and B- & T-lymphocytes depending upon various growth factors or cytokines.*
- *Anaphylactic shock is often fatal shortly after the histamine release. Adrenaline injected intravenously or intracardially may save the victim.*
- *Urticaria and hay fever are anaphylactic reactions in the skin and the nasal mucosa, respectively. The histamine released causes vasodilatation, with increased capillary pressure and ultrafiltration causing oedema and red colour.*
- *Bronchial asthma in an atopic person is a bronchiolar anaphylactic response. Leucotrienes from the bronchiolar mast cells cause bronchoconstriction, mucosal infiltration with inflammatory cells, mucosal oedema and hypersecretion. Leucotrienes are blocked by a variety of bronchodilatators.*
- *Iatrogenic AID is caused by splenectomy, chemotherapy or iatrogenically induced malnutrition.*
- *Autoimmune suppression by infection is the Acquired Immuno-Deficiency Syndrome (AIDS) resulting from infection with Human Immuno-Deficiency Virus (HIV). This is a substantial T-lymphocyte defect.*
- *Immuno-stimulated macrophages produce nitrite and nitrate, and their killer activity is related to the unstable gas, nitric oxide (NO). NO, produced in large quantities by the macrophages, kills microbes and cancerous cells.*
- *Phagocytotic killing occurs in the phagolysosomes. The method of execution is by a respiratory burst or by gas.*
- *Myasthenia gravis. This is a rare disease caused by autoantibodies against the acetylcholine receptors of the neuromuscular endplate. Complement complexes and IgG molecules are deposited at the endplates. Many patients have thymic hyperplasia.*
- *lcerative colitis and Crohns disease (transmural enteric ulceration with a particular affinity for the terminal ileal cells) are of unknown aetiology, but they may represent two aspects of the same disease. Both antigen specific, cellular and autoimmune responses have been postulated.*

- *Rheumatoid arthritis (symmetrical inflammatory polyarthritis with progressive joint damage and lung, cardiac, renal and many other organ manifestations) is an autoimmune disease. The synovial fluid contains IgG, lymphokines and immune complexes. The cause is probably a persistent external antigen, which is not removed.*

- *Insulin-dependent diabetes mellitus (IDDM). Newly presenting patients possess islet-cell antibodies that have destroyed all the insulin producing β-cells of the pancreatic islets. There is an inherited increased susceptibility to IDDM.*

- *Pernicious anaemia. Typically parietal cell antibodies are found in the blood. They may kill the entire parietal cell population leading to atrophy of the mucosa. The atrophic gastric mucosa fails to produce hydrochloric acid and intrinsic factor for vitamin B_{12}.*

- *Graves's or Basedow's disease is hyperthyroidism combined with eye signs (Exophtalmus). Normally, the thyroid stimulating hormone (TSH) increases the thyroid hormone production after its binding to the thyroid TSH receptors. Bacterial infection in a genetically susceptible person may be the cause of autoimmune production of the TSH-receptor antibodies. These IgG -antibodies behave exactly like TSH itself, and thus stimulate the thyroid hormone production.*

Further Reading

*Scandinavian Journal of Immunology***.** Monthly journal published by Blackwell Science Ltd., Osney Mead Oxford OX2 OEL, UK.

*Scientific American***.** Monthly journal published by Scientific American Inc., 415 Madison Avenue, N.Y., USA.

Mims C, Playfair J, Wakelin D, and R Williams. *Medical Microbiology*. Mosby, London, 1998.

Return to top

Return to content

Chapter 33

Infectious Diseases

Study Objectives

- To *define* bacteraemia, bactericidal antibiotics, bacteriostatic drugs, inflammation, infection, active and passive immunity, opportunistic infections, parasitism, pyaemia, septicaemia, symbiosis, and viruses.
- To *describe* the effects of aminoglycosides, b-lactam drugs chloramphenicol, erythromycin, nitroimidazol, quinolones, sulphonamides, and tetracycline.
- To *explain* well-known infectious disorders caused by bacteria, virus, fungi, protozoan, and helminths.
- To *use* the concepts in problem solving and case histories.

Principles

- *Both unicellular microorganisms and multicellular organisms can live within the human body with beneficial (symbiosis) or detrimental (parasitism) consequences.*
- *Febrile infectious diseases increase sympathetic nervous activity and raise the heart rate by release of nor-adrenaline.*

Definitions

- **Aminoglycosides** are compounds of amino sugars derived from Streptomyces. Streptomycin is seldomly used due to oto-and nephrotoxicity. Neumycin is used for local infections. Gentamicin and tobramycin are bactericidal for many Gram-negative microbes. Netilmicin and amikacin are resistant to aminoglycoside-inactivating enzymes.
- β--**lactam drugs:** The general structures of penicillin and cephalosporins include a *β-lactam square*. Lactam is a cyclic amide, and here the amine group is located on the second carbon from the carbonyl.
- **Bacteraemia** refers to the presence of bacteria in the blood.
- **Bactericidal drugs** possess the capacity to kill microbes.
- **Bacteriostatic drugs** prevent the growth of microbes.
- **Chloramphenicol** is a nitrobenzene-containing antibiotic used for the treatment of *plague* (Yersinia pestis) and *enteric fevers* (Salmonella typhi murium, Salmonella paratyphi). The most serious toxic effect of chloramphenicol is irreversible bone marrow depression.
- **Erythromycin** and other macrolides (azithromycin, clarithromycin) are used against infections with Bordetella pertussis (whooping cough), Campylobacter, Clamydia, Coxiella, Legionella, Listeria, Mycoplasma and Toxoplasma gondii. Macrolides bind to and prevent

translocation on bacterial ribosomes but not on mammalian ribosomes.

- **HIV** means *Human Immuno-deficiency Virus*

- **Immunity.** Active immunity is achieved by disease or by vaccination with live attenuated microbes, dead organisms or fractions of organisms, microbial toxins or antigen preparations produced by recombinant DNA technology. *Childhood immunisation* is maintained with booster doses throughout life.

- **Infections** are diseases caused by microorganisms, viruses, protozoan and helminths.

- **Inflammation** is the response of the body to damage or irritation. Inflammation is a vascular reaction, where leucocytes and fluid is transported from the blood into extravascular tissues.

- **Nitroimidazoles,** such as metronidazole, inactivate anaerobic bacteria and certain protozoa by breaking up the DNA of the microbes. Indications are amoebiasis, giardiasis, trichomoniasis and anaerobic bacterial infections.

- **Macrolides** are bactericidal antibiotics that bind to and prevent translocation on bacterial ribosomes but not on mammalian ribosomes.

- **Opportunistic infection** refers to side-infection in immuno-deficient patients with TB, AIDS or other serious disorders. The most common opportunistic infection in AIDS is the pneumocystis carinii infection, which attacks the lungs.

- **Parasitism** refers to a condition, where two dissimilar organisms live together to the benefit of the parasite and to the detriment of the host (protozoan and helminthic infections).

- **Prions** (small proteins) and *viroids* (small nucleic acids) can live in humans and may contribute to human disease (Creutzfeld-Jacobs disease, dementia etc).

- **Pyaemia** is septicaemia with emboli causing abscesses in the brain, lungs, liver or other sites in the body.

- **Quinolones** inhibit bacterial *DNA gyrase*, which normally maintains the helical structure of DNA. These antibiotics are used for infections with resistant microbes.

- **Rickettsias** are extremely small bacteria that are transferred to humans by rat fleas, human lice, mites and tics - often by scratching. They cause different types of Typhus fever.

- **Septicaemia** refers to a condition, where microorganisms multiply in the blood, cause high fever, fall in blood pressure and shock.

- **Sulphonamides** contain an essential para-amino group and inhibit folic acid synthesis in the microbes. These drugs are used for urinary tract infections and for inflammatory bowel disease

- **Symbiosis** refers to a condition, where two dissimilar organisms co-operate in order to benefit or survive.

- **Tetracycline** contains a four-ring structure and act by preventing attachment of tRNA to

the mRMA-ribosomal complex. Tetracycline inhibits a wide spectrum of aerobic and anaerobic Gram-positive and Gram- negative bacteria, Clamydia, Mycoplasma and rickettsia.

- **Viruses** are small infectious agents, which contain either DNA or RNA, and only survive intracellularly. The general structure is a nucleic acid core and an antigen protein shell that is specific for each virus. *Viruses* are the only a-cellular, proteinaceous agents known to infect humans.

Essentials

This paragraph deals with *1. Antimicrobial therapy*, *2. Antiviral drugs*, *3. Antifungal drugs*, *4. Anti-protozoan therapy* and *5. Antihelmintic drugs*.

1. Antimicrobial therapy

Antimicrobial agents attack young, growing cells. An increasing number of microbes develop resistance for multiple drugs. The practical cause is an increasing use of antibiotics for minor infections and blind inappropriate therapy. Resistance to an antibiotic develops by mutation. Bacteria that can produce penicillinase are resistant to some of the penicillin molecules.

Antibiotics have toxic effects, especially when used in overdose or in the presence of other disease such as hepatic or renal failure. Aminoglycosides are both oto- and nephrotoxic.

Antibiotics can lead to secondary infections with fungi or with Clostridium difficile.

Serious bacterial infections and infections in immunodeficient patients are commonly treated with bactericidal drugs (aminoglycosides, cephalosporins, and penicillin), although there is no evidence that these drugs are more effective than bacteriostatic drugs. Such serious infections (eg, endocarditis, meningitis, TB, septicaemia) are preferably treated with combinations of synergistic antibiotics. - On the other hand, penicillin inactivates aminoglycosides.

β-Lactam drugs. The general structures of penicillin and cephalosporins include a b-lactam square. Lactam is a cyclic amide, and here the amine group is located on the second carbon from the carbonyl (Fig. 33-1). The small square has a low stability, and benzylpenicillin (penicillin G) undergoes hydrolysis rapidly by the acid gastric juice when given orally. Phenoxymethyl-penicillin (penicillin V) is more stable, and a sufficient fraction is absorbed following oral administration. *Ampicillin* is produced by introduction of an amino group into the phenyl radical of benzylpenicillin, which makes it active against both Gram-negative and Gram-positive bacteria. Further modification of the acyl side chain results in *cloxacillin,* which is insensitive to penicillinase, and thus preferable in treating infections with penicillinase-producing bacteria.

The advantages of *cephalosporins* over penicillin are their resistance to penicillinase and their wider antimicrobial spectrum (both Gram-positive and Gram-negative bacteria).

Fig. 33-1: Penicillin is build up by acyl side-chains linked to β-lactam and a thiazolidine ring. Cephalosporins have a dihydrothiazine ring instead.

The b-lactam drugs interfere with *bacterial cell wall* development. Both gram-positive and gram-negative bacteria have a rigid cell wall composed of a matrix of peptidoglycan strands. Gram-positive bacteria do not have a lipopolysaccharide coat in contrast to the gram-negative bacteria (Fig. 33-1). The two groups of b-lactam antibiotics bind to and inactivate specific *penicillin binding proteins.* These molecules are peptidases involved in the final cross-link of the rigid peptidoglycan matrix. The β-lactam antibiotics block this final cross-linking reaction and thus the development of a normal bacterial cell wall. Bacteria that can produce β-lactamase (penicillinase) are resistant to several penicillins.

Inhibitors of ribosomal actions. Aminoglycosides, chloramphenicol, clindamycin, erythromycin, fusidic acid, mupirocin, spectinomycin, tetracycline all bind to bacterial ribosomes and interfere with protein synthesis.

Aminoglycosides are compounds of amino sugars derived from Streptomyces. Streptomycin is seldomly used due to oto-and nephrotoxicity. Neomycin is used for local infections. Gentamicin and tobramycin are bactericidal for many Gram-negative microbes. Netilmicin and amikacin are resistant to aminoglycoside-inactivating enzymes.

Chloramphenicol is a nitrobenzene-containing antibiotic indicated for the treatment of *plague* (Yersinia pestis) and *enteric fevers* (Salmonella typhi murium, Salmonella paratyphi). Chloramphenicol binds to ribosomes, and prevents the addition of new amino acids to growing peptides. The most serious toxic effect of chloramphenicol is irreversible bone marrow depression.

Erythromycin and other macrolides (azithromycin, clarithromycin) are used against infections with Bordetella pertussis (whooping cough), Campylobacter, Clamydia, Coxiella, Legionella, Listeria, Mycoplasma and Toxoplasma gondii. Macrolides bind to and prevent translocation on bacterial ribosomes but not on mammalian ribosomes.

Fusidic acid has a bile acid structure and inhibits Gram-positive Cocci including penicillinase-producing Staphylococcus aureus in combination with other antibiotics. Fusidic acid inhibits translocation on the ribosomes and is hereby bactericidal for many bacteria.

Tetracycline contains four-ring structure and act by preventing attachment of t-RNA to the bacterial mRMA-ribosomal complex. Tetracycline inhibits a wide spectrum of aerobic and anaerobic Gram-positive and Gram- negative bacteria, Clamydia, Mycoplasma and rickettsia.

Nitroimidazoles. Nitroimidazoles, such as metronidazole, inactivate anaerobic bacteria and certain protozoa by breaking up the DNA of the microbes. Indications are amoebiasis, giardiasis, trichomoniasis and anaerobic bacterial infections.

Quinolones (nalidixic acid, cinoxacin, and fluoro-based ciprofloxacin, lomefloxacin, ofloxacin, and norfloxacin) inhibit bacterial DNA gyrase, which normally maintains the helical structure of DNA. These antibiotics are used for infections with resistant microbes.

Sulphonamides contain an essential para-amino group and inhibit folic acid synthesis in the microbes. These drugs are used for urinary tract infections and for inflammatory bowel disease.

2. Antiviral drugs

Idoxuridine and *vidarabine* are used as early as possible against herpes simplex virus and against varicella zoster virus (herpes zoster and severe varicella or chickenpox).

Acyclovir inhibits viral DNA synthesis and is effective against the same virus disorders.

Ribavirin is used against Lassa fever (see later).

Azido-deoxy-thymidine inhibits HIV reverse transcriptase and thus impairs viral replication. This substance is used for HIV patients, although there is danger of bone marrow depression.

Interferon is produced by the T-lymphocytes during virus infections. Interferon is used for hepatitis B and C.

3. Antifungal drugs

The target of antifungal drugs is young, growing cells.

Polyenes such as *Amphotericin B* and *nystatin* bind to sterols in the fungal membranes and increase the permeability, so K^+ and Mg^{2+} leaks out of the cell. Hereby, the fungi are killed.

Mammalian membranes have a lower affinity for polyenes than fungal membranes containing ergosterol. Since amphotericin B is nephrotoxic, a combination therapy with *flucytosine* is used for serious systemic fungal infections. Flucytosine inhibits fungal DNA synthesis.

Griseofulvin is used for long lasting fungal infections in skin and nails, because it is concentrated in keratin. Griseofulvin enters susceptible fungi and inhibits their mitosis.

Imidazoles inhibit a membrane ATPase and incorporation of ergosterol in the cell membrane of susceptible fungi. This results in K^+-efflux from the cell.

4.Anti- protozoan therapy

Leishmaniasis is treated with meglumine antimoniate and Na-stibogluconate. These substances inhibit phosphofructokinase and Krebs cycle enzymes in the Leishmania organism.

Malaria is caused by infection with the Plasmodium protozoa, which is present in the human host in the blood (erythrocytic) or the tissue (extra-erythrocytic stage). The erythrocytic stage is treated with aminoquinolines (eg, chloroquine) or analogues. Their antimalarial action is unclear, but they seem to be toxic to both the red cell and the plasmodia membranes. The extra-erythocytic stage is treated with primaquine.

The *diaminopyrimidines*, pyrimethamine and trimethoprim, inhibit dihydrofolate reductase in malarial parasites. Diaminopyrimidines and sulphonamides act synergistic to prevent the formation of schizonts in erythrocytes and hepatocytes.

African trypanosomiasis is treated with tryparsamide or melarsoprol. These arsenicals inhibit sulphydryl groups on carbohydrate enzymes in the parasite, whereby the production of ATP decreases.

Chagas disease is treated with nifurtimox, which produces free radicals. The free radicals react with oxygen to form superoxide anion, hydroxyl-free radical, and hydrogen peroxide. The trypanosomes do not contain enzymes able to inactivate these reactive substances, so they cause peroxidation of lipids and ribonucleic acids (DNA, RNA).

Giardiasis and *trichomoniasis* are treated with quinacrine, which prevent parasite replication by interference with DNA.

5. Anti-helminthic drugs

Helminths (worms) cause medical problems world-wide, because they are carried by parasitized humans all over.

The helminths are divided into three groups: *5a. Roundworms (ie, nematodes), 5b. Flukes (ie, trematodes, flatworms), and 5c. Tapeworms (ie, cestodes, flatworms).*

The existence of the worm in the host depends on its movements in order to reach an optimal location for food, on glucose and other nutrients, and on intact cytoplasmic microtubules.

5a. *Roundworm disorders* are treated with *mebendazole* or with *thiabendazole*. These substances bind to the cytoplasmic microtubules of the roundworm and inhibit glucose transport and protein secretion.

5b. *Flukes* are sensitive to *praziquantel*, which attacks the surface barrier (tegumentum) and the muscles of the worm. Praziquantel destroys the surface of the worm so host antibodies can bind to the worm antigens. Hereby, host complement and leucocytes are attracted to the site and a large Ca^{2+}-influx into the worm cells is triggered. The worm muscles enter a spastic contraction and the worm is killed by killer cells.

5c. *Tapeworms* are also killed by praziquantel, but as an alternative they are sensitive to *niclosamide*. Niclosamide binds to the mitochondria of both host and parasite and block ATP

formation. Gut-dwelling *tapeworms* absorb niclosamide but not gut-dwelling *roundworms*. Therefore, tapeworms exposed to niclosamide cannot produce ATP, their muscles are ineffective, they loose their grip in the intestinal mucosa and they are expelled.

Pathophysiology

This section deals with *1. Bacterial infections*, *2. Viral infections*, *3. Fungal infections*, *4. Protozoan infections*, *and* *5. Helminthic infections.*

1. Bacterial infections

Gram-positive Cocci, gram-negative cocci, gram-positive bacilli, gram-negative bacilli, actinomycetes, mycobacterium, spirochaetes, rickettsia, and Clamydia cause bacterial infections.

1A. Gram-positive cocci

Gram-positive cocci include staphylococci and streptococci.

Staphylococci present in the microscope like clusters or grapes (Fig. 33-2). Staphylococci are normally found on the surface of the skin and at all natural openings of the human body.

Skin infections with *Staphylococcus aureus* include cellulitis, furuncles, carbuncles, impetigo and *staphylococcal toxin* scarlet fever. Typical is also parotitis and osteomyelitis. Lung and airway infections may result in abscesses also in the lung. Endocarditis and pericarditis are serious just as meningitis and brain abscesses.

A special syndrome occurs in young females using high-absorbency polyacrylate tampons. There is rapidly developing fever, erythema, vomiting, diarrhoea, myalgia and shock (ie, *toxic shock syndrome* caused by staphylococcal toxin-1). Staphylococcal food poisoning is caused by heat-stable enterotoxin (A-E) from replicating staphylococci in manufactured food (cans, milk etc). There is diarrhoea and the patient vomits persistently.

Staphylococci outside hospitals are usually penicillin sensitive, and penicillin is preferred for infections with these organisms. Staphylococcal infections acquired in hospitals are often penicillin resistant. Accordingly, the patient is treated with fusidic acid in synergy with other antibiotics. Fusidic acid inhibits translocation on the ribosomes and is hereby bactericidal for many bacteria.

Staphylococcus aureus and other species resistant to aminoglycosides and also to all b-lactam drugs are now spread throughout the world. These infections are preferably treated with the lipophilic cell wall inhibitor, teichoplamin, and with quinolones

Fig. 33-2: Staphylococci and streptococci in pus under the microscope.

Streptococcal infections

Streptococci are round Gram-positive bacteria often arranged like a pearl chain in the infected cells (Fig. 33-2). Streptococcus pyogenes (group A b-haemolytic streptococci) cause almost all human infections such as classical scarlet fever and erysipelas.

Persons deprived of neutralizing antitoxin-antibodies to *erythrogenic toxin* develop classical *scarlet fever*, when infected with streptococci. Scarlet fever is usually a mild childhood disease with an incubation period of 2-4 days. Fever, headache and vomiting are typical clinical features. On the second day a rash occurs on the skin covered by a T-shirt. After 5 days of rash the skin desquamates. Typical is circumoral palor and strawberry tongue. Otitis media or even a retropharyngeal abscess complicates scarlet fever, which can be life threatening. *Phenoxymethylpenicillin* is given orally for 10 days or erythromycin in penicillin-allergic individuals.

Group A streptococci causes erysipelas. *Erysipelas* is a skin infection with red swollen skin. Sepsis can be terminal, so immediate penicillin therapy is imperative.

1B. Gram-negative diplococci

Gram-negative diplococci are unmoveable bacteria, also called *Neisseria*, N (detected in 1879 by Albert Neisser). Neisseria gonorrhoea and N. meningitides are pathogenic to humans, and N. catarrhalis is present in the nasopharynx of all healthy persons. The diplococci are formed like a coffee bean (Fig. 33-3).

Fig. 33-3: Neisseria gonorrhoea (N.G.) and Neisseria meningitides. β-Lactam drugs (penicillins and cephalosporins) block the development of the normal meningococcal or gonococcal cell wall. Aminoglycosides bind to bacterial ribosomes and interfere with protein synthesis. Quinolones inhibit bacterial DNA gyrase.

Gonorrhoea. The incubation period is usually less than a week. The Neisseria gonorrhoea has a special affinity towards the epithelium of the genital tract and other natural openings. Quite a few patients - especially females - experience no symptoms. Males develop urethritis with discharge and dysuria (painful urination). Complications in the males include infections of the epididymis, the testes or the prostate. Females are frequently infected in the endocervical channel, the vagina and in the rectum as proctitis just as homosexual males. Ascending infection in females leads to salpingitis and a high incidence of infertility. Arthritis is a late complication. Direct microscopy reveals the diagnosis, or culture of the fastidious bacteria on special media may be necessary. Uncomplicated cases are cured by 3 g of the amino-penicillin, amoxillin, with probenecid to delay the renal secretion. Patients with allergy or penicillin-resistant cocci require the quinolone, ciprofloxacin. The effects of antibiotics are shown in Fig. 33-3.

Meningococcal infection

The meningococcus reaches the blood through the naso-pharynx. The disease is either a sepsis with or without meningitis. Acute meningococcal sepsis, where death ensues within hours, sometimes occurs epidemically.

Meningococci are found in the blood or CSF. Cephalosporins, such as cefotaxime, or benzylpenicillin are effective when given immediately on suspicion. Vaccination is recommendable when travelling.

1C.Gram-positive bacilli

The most important infections are diphtheria, tetanus, botulism, gas gangrene, listeria, anthrax and pseudomembraneous colitis.

Diphteria is caused by airborne spread of Corynebacterium diphteriae. This is a club-shaped bacillus, which produces *Diphteria toxin* when exposed to bacteriophag B. Diphteria toxin has two subunits. Subunit A is toxic and subunit B transfers A to toxin-receptors on peripheral nerves and on the myocardium. Following a week incubation there is nasal discharge with crusts around the nares. A white membrane of fibrin, cells and bacilli characterizes the tonsillar and pharyngeal infection. Typical is also the bull-neck of swollen lymph glands. Laryngeal infection is life threatening due to airway obstruction with dyspnoea and cyanosis. Later myocarditis and palatal paralysis occur, followed by paresthesia, cranial nerve palsies and eventually encephalitis. Rapid administration of antitoxin i.v. is the only specific therapy of the toxin effects. Penicillin is given in order to kill the corynebacteria. Active immunisation in childhood is important.

Tetanus

Even a trivial wound contaminated with Clostridium tetani can kill the un-immunised victim within 24 hours.*Tetanospasmin* is an extremely potent neurotoxin. Muscle spasms occur in the

masseter muscle (lockjaw or trismus) and in the mimic facial muscles (grinning expression or risus sardonicus). Light or noise triggers spasms. Spasms of the back muscles result in opisthotonus (ie, arching of the body by spasms of the neck and back muscles). Death ensues from laryngeal spasms with respiratory failure or from cardiac arrest. Human anti-tetanus immunoglobulins are given i.m. and penicillins given i.v.

Botulism

Clostridium botulinum spores are thermo-stable, proliferate and produce neurotoxins pathogenic to humans (A, B and E) in canned foods. Botulinus toxins are thermo-labile and the most potent neurotoxins known. They block the neuromuscular endplate effectively. Neurological symptoms and signs occur within 24 hours after ingestion of contaminated food. Laryngeal spasm, strabismus, generalized paralysis including respiratory paralysis occurs, and the mortality is up to 70% of all hospitalised patients. Guanidine hydrochloride reverses the neuromuscular blockade, and assisted ventilation is often necessary.

Gas gangrene

Clostridium perfringens causes most case of gas gangrene, which occurs in lacerated wounds such as gun wounds. The muscle tissue is oedematous and slowly filled with gas, which is felt as crepitus. There is a tachycardia and the patient dies in shock or from renal or hepatic failure. Debridement and antibiotic combination therapy is imperative. See also hyperbaric oxygenation in Chapter 19.

Listeria

Listeria monocytogenes causes listeriosis with septicaemia, abortions and meningitis. Listeriosis is treated with ampicillin and Gentamicin.

Anthrax

Bacillus anthracis is spread from infected animals to man. There are erythematous skin lesions, retrosternal pains, pleural effusions, haematemesis and bloody diarrhoea. Death is often caused by respiratory failure. Serious cases are treated with penicillin i.v.

Pseudomembraneous colitis

Clostridium difficile produces toxins A and B following termination of antibiotic therapy. There is diarrhoea, fever and abdominal pain. The diagnosis is confirmed by the presence of toxin in stool specimens. Suspected antibiotics are of course seponated. Metronidazole or vancomycin is used for 10-14 days.

1D. Gram-negative bacilli

Brucellosis is caused by Brucella abortus Bang, Brucella militensis and Brucella Suis. These microbes are Gram-negative coccobacilli. Brucella is spread by intake of raw milk from infected cows and goats. In the body Brucella spread along the lymphatic system from where they reach the blood and finally the reticulo-endothelial system. The patient experiences oscillating fever, myalgia and lympadenopathy. Blood culture establishes the diagnosis. Tetracycline combined with rifampicillin for 6 weeks is effective.

Bordetella infections include *pertussis (Bordetella pertussis) and parapertussis*

(Bordetella parapertussis). Pertussis or *whooping cough* is spread by droplets and contagious. The first week is the catarrhal stage, where the patient feels as with a common cold with light fever, and coughing triggered by vagal signals. The paroxysmal stage is characterized by attacks of typical whooping cough - often just following change of thoracic position. A whoop is a prolonged, wheezing inspiration caused by airways obstructed with oedema and mucus. The cough attack terminates in vomiting and expectoration of a viscid mucus. The total white blood count is dominated by lymphocytes. Complications include atelectasis, pneumonia,

facial cyanosis, herniation, conjunctival petechiae and nose bleeding. Parapertussis is milder. The diagnosis is confirmed by culture of nasopharyngeal swabs. Early administration of erythromycin is beneficial.

Escherichia coli infections include travellers diarrhoea, children's diarrhoea, and diarrhoea in developing countries in general. The coli bacterium is a large, unmoveable Gram-negative stick (Fig. 33- 4) first described by Escherich. Serious infectious cholera-like diseases and bloody diarrhoea are also seen. The quinolone, ciprofloxacin, is effective, and in severe cases the aminoglycoside, Gentamicin, is used.

Fig. 33-4: Klebsiella pneumoniae and Eschericia coli are closely related. β-Lactam drugs (cephalosporins) block the development of the normal Klebsiella pneumoniae cell wall. Aminoglycosides bind to the ribosomes of coli bacteria and interfere with protein synthesis. Quinolones inhibit colic DNA gyrase and interfere with the normal helical structure of its DNA.

Klebsiella pneumoniae (Friedlander) is a short, unmoveable Gram-negative stick (Fig. 33-4) causing institution-acquired pneumonia in immuno-deficient patients. Cephalosporins are required, but the mortality is high.

The effects of antibiotics are shown in Fig. 33-4.

Haemophilus influenzae infections include meningitis, pneumonia, otitis media, endocarditis, pericarditis and milder infections. The diagnosis is confirmed by culture of Haemophilus influenzae Pfeiffer, which is a Gram-negative aerobic bacillus (Fig. 33-5). Serious infections must be treated immediately with cephalosporins (oral cefaclor). Chloramphenicol therapy is dangerous due to granulocytopenia or agranulocytosis.

Cholera. Cholera is a disease caused by an enterotoxin produced by the gram-negative bacillus, *vibrio cholerae*. If this bacillus is allowed to proliferate within the lumen of the small intestine, it causes profuse watery diarrhoea - up to 24 l per 24 hours - dehydration and circulatory shock. Fulminant cholera can kill the patient within a day.

The cholera enterotoxin contains A and B subunits. Subunit A is an enzyme, which enters the enterocyte and irreversibly activates adenylcyclase, whereby a cascade of reactions is triggered. Subunit B links the enterotoxin to the brush borders of the enterocytes, so the effect is persistent. As cAMP is activated in the cell it activates an *electrogenic chloride-channel* in the brush border membrane. This causes an enormous secretion of *NaCl and water* into the lumen of the small intestine. Although the colonic lavage (reabsorption capacity) is normally extremely large (up to 4200 ml reabsorbed fluid per day), this is not enough to prevent the massive diarrhoea (*ricewater stool*).

Acute dehydration with loss of base leads to metabolic acidosis and hypovolaemic shock.

Therapy is immediate rehydration and tetracycline. Tetracycline acts by preventing attachment of t-RNA to the mRMA-ribosomal complex of the cholera bacillus.

Fig. 33-5: Haemophilus influenzae (Pfeiffer) and Treponema pallidum. β-Lactam drugs (penicillins and cephalosporins) block the development of the normal cell wall of the Pfeiffer bacillus. Penicillin for syphilis interferes with spirochaetal cell wall development.

Salmonellosis includes typhoid fever (Salmonella typhi), paratyphoid fever (Salmonella paratyphi), and enterocolitis (Salmonella enteritidis, Salmonella typhi-murium). Salmonellae proliferate in Peyers plaques of the small intestine, and finally reach the reticuloendothelial system via lymph and blood. The quinolone, ciprofloxacin, is used as the preferred treatment.

Shigellosis or *bacillary dysentery* is caused by a series of Gram-negative bacilli. Dysentery is the name of severe bloody diarrhoea with mucus and cramping abdominal pain. Complications include dysentery-arthritis, widespread infections and life-threatening shock. Sulphonamides or the quinolone, ciprofloxacin, are used.

Campylobacter infection also causes bloody diarrhoea. Quinolone treatment is a frequent choice.

Helicobacter infection

Helicobacter pylori grow and proliferate in the mucous layer of the gastric and duodenal epithelium. This Gram-negative microbe is involved in the pathogenesis of peptic ulcer disease. Clarithromycin is a macrolide that binds to and prevents translocation on Helicobacter pylori- ribosomes, which is an effective basic therapy of peptic ulcer.

Yersinia infections include *plague* (Yersinia pestis), *enterocolitis* (Yersinia enterocolitica), and Yersinia pseudotuberculosis.

Plague is spread from woodland rodents to domestic rats. The rat flee bites humans. Bubonic plague occurs suddenly with high fever and the patient is confuse and may seem drunk due to toxaemia. The inguinal lymph nodes rapidly develop lymph-adenopathy or *buboes.* Pneumonic plague is fulminant pneumonia with bloody sputum and cardiac failure. Petechiae and severe cyanosis with a terminal outcome is responsible for the name *the Black Death.* Septicaemia plague is dominated by septicaemic shock and death occurs in days. The diagnosis is confirmed by demonstration of the bacilli in blood, sputum or lymph node aspirate. Rapid administration of antibiotics in synergism-combination is essential (streptomycin, tetracycline etc).

Entero-colitis or terminal ileitis caused by the Y. enterocolitica, and pseudotuberculosis (mesenteric lymphadenitis) is treated with tetracycline.

Tularaemia (Tulare is a town in California) is caused by Francisella tularensis, and spread from animals (in particular rabbits and squirrels) to man through flies, ticks and mosquitoes. There is an ulcer with lympadenopathy and sometimes pneumonia or septicaemia. The diagnosis is confirmed with an agglutination test. The aminoglycoside, gentamicin, is effective.

Legionnaire disease is caused by Legionella pneumophilia. The clinical picture is that of a severe pneumonia with coughing, fever, tachypnoea, cyanosis, confusion and diarrhoea. The diagnosis is confirmed by rapid antigen tests. Erythromycin and rifampicin (anti-mycobacterial drug that inhibits RNA synthesis and thus protein synthesis) are effective. Macrolides, such as erythromycin, bind to and prevent translocation on bacterial ribosomes but not on mammalian ribosomes.

1E. Actinomycetes

are branching bacteria (Gram-positive) that cause slowly and insidious developing infections with a typically chronic pattern. *Actinomycosis* and *nocardiosis* are rare disorders localised in lymph nodes, lung tissue, and coecum. Surgery and penicillin, tetracycline or the sulphonamide, sulphadiazine, are used.

1F. Mycobacteria are aerobic bacilli that grow slowly.

Mycobacterium tuberculosis is the cause of most cases of human tuberculosis. The primary infection includes the lungs, where granulomas are formed with caseation centrally. Although the primary granuloma heals there is often surviving tubercle bacilli in the granuloma, and later the bacilli spread locally and with the blood. The spread can attack all parts of the body - frequently bones and the kidneys. Daily intake of rifampicin 600 mg and isoniazid 300 mg for half a year is necessary in order to cure pulmonary tuberculosis.

Mycobacterium leprae causes leprosy or Hansens disease. This mycobacterium can oxidise 3,4-dihydroxy-phenylalanine to measurable pigments. Tuberculoid leprosy is localised, because the patient has maintained his cell-mediated immunity. The hypopigmented patches on

the skin are anaesthetic.*Lepromatous leprosy* is generalized because the patient has impaired cell-mediated immunity. The lesions affect the skin of the face, and limbs. Leprosy is treated with dapsone (a folate synthetase inhibitor) and clofazimine daily. Each month a dose of rifampicin is given.

Rifampicin (rifampin) binds to the mycobacterial RNA polymerase (less to mammalian RNA polymerase), and inhibits initiation of RNA synthesis. Isoniazid is bactericidal to growing mycobacteria - possibly by blocking the synthesis of mycolic acid in the bacterial wall.

1G. Spirochaetes

Spirochaetes include Treponema, Leptospira and Borrelia.

Syphilis (Lues) is caused by Treponema pallidum, which is mobile spirochaetes (Fig. 33-5).

Congenital syphilis is acquired transplacentally from mother to foetus and is apparent a few weeks after birth. The babies pass through three stages of the disease similar to those of the adults. *Adult syphilis* is acquired by intimate sexual contact.

Primary syphilis: Three weeks after exposure to the spirochaetes there is a papule at the epithelial lesion. The papule ulcerates and develops into a painless *chancre* with swelling of the regional lymph glands.

Secondary syphilis: Three months following the exposure the patient experiences fever, sore throat, lympadenopathy, rashes, condylomata and arthralgia. These symptoms and signs usually fade away over a few months without therapy.

Tertiary syphilis: Gumma in the skin, cardiovascular and neural damage (*neurosyphilis*) is typical findings years after the primary lesion. A *gumma* is a granulomatous mass often found to expand intracranially, where the pressure rises and focal disorders occur (epilepsy or hemiplegia).

The fluorescent treponema antibody absorption test is specific for *treponema* and remains positive for life. Tertiary syphilis, with demyelinisation of the dorsal roots of the spinal cord is known as *tabes dorsalis*. The syndrome includes lightening or knife-tap pain, ataxic gait, neuropathic joints, ptosis, optic atrophy, and Argyll-Robertson's light-stiff pupil (Chapter 7).

Early syphilis is treated with procaine penicillin in large daily doses for two weeks. *Neurosyphilis* is treated with parenteral penicillin with steroid cover for three weeks.

Yaws, Bejel, and *pinta* are caused by other treponema strains, but they produce the same late stages as seen in syphilis and they are all treated with the long-acting procaine penicillin.

Leptospirosis is caused by Leptospira *interrogans* ictero-haemorrhagiae. This disorder is also called *Weil's disease*.

Borrelia recurrentis causes *relapsing fever*.

Borrelosis or *Lyme disease* (first described in the city Lime in Connecticut) is caused by Borrelia burgdorferi. The disease is transmitted by infected tics. Borrelosis produces the same three stages as seen in syphilis. The first stage occurs within a week with skin lesions. Cardiac or neurological findings and IgM antibodies in the blood plasma characterize the second stage after months. Arthritis and IgG antibodies in the blood or cerebrospinal fluid characterize the third stage after years. Prompt treatment with large doses of benzathine penicillin or tetracycline for weeks may shorten the duration. Later stages must be treated with intravenous cephalosporins.

1 H. rickettsia

Rickettsias are *extremely small bacteria* that are transferred to humans by rat fleas, human lice, mites and tics - often by scratching. They cause different types of Typhus fever

1 I. Chlamydiae

Chlamydiae are intracellular bacteria infecting one of five persons. Chlamydia trachomatis causes *trachoma*, which is a common cause of blindness that can be avoided by tetracycline therapy. Other strains cause genital infections including lymphogranuloma venereum. Chlamydia psittaci causes **p***sittacosis* or *ornithosis*, which is spread from infected birds. Chlamydia pneumoniae causes *pneumonia,* which is treated with tetracycline.

2. Viral infections

Viruses are small infectious agents, which contain either DNA or RNA, and only survive intracellularly. The general structure is a nucleic acid core and an antigen protein shell that is specific for each virus.

The DNA viruses comprise *adenovirus* (acute pharyngitis, laryngitis or croup, mesenteric lymphadenitis), *herpes viruses* (cytomegalovirus infection, herpes simplex, varicella-zoster, Epstein-Barr virus infection, roseola infantum), *papova-viruses* (genital warts, carcinoma of the cervix), *parvo-virus* (slapped-cheek disease or erythema infectiosum, arthropathy, chronic infectious anaemia), and *pox virus* (smallpox or variola).

The RNA viruses comprise *picorna-viruses* (poliovirus, Coxsackievirus, Hepatitis A, Echovirus, Enterovirus and Rhinovirus with common cold), *reoviruses* (reo-and rota-virus causing childhood diarrhoea), *togaviruses* (alphaviruses, flaviviruses, and rubella virus causing epidemics of fever, yellow fever, dengue haemorrhagic fever, Japanese encephalitis and rubella), *orthomyxo-virus* (influenza A, B, C), *paramyxovirus* (measles and mumps), *rhabdovirus* (rabies), *retrovirus* (HIV), and *arena-virus* (Lassa fever in Lassa, Nigeria).

HIV and AIDS

HIV means *Human Immunodeficiency Virus*. HIV is the cause of AIDS (ie, *Acquired Immune Deficiency Syndrome*). HIV triggers a progressive and irreversible depletion of T-helper lymphocytes (Chapter 32).

The lack of immune defence in HIV-patients make them easy victims to opportunistic infections and cancer, which often occur simultaneously.

The transmission pathway is sexual contact - both homo- and hetero-sexual - or through parenteral exposure to blood.

Opportunistic infections in AIDS are caused by bacteria & mycobacteria (Legionella, Listeria, Salmonella, Shigella, TB), viruses (Cytomegalovirus, Herpes Simplex, Varicella-Zoster), fungi (Candida, Histoplasma, Cytococcus), and protozoa (Entamoeba Histolytica, Pneumocyctis Carinii, Toxoplasma).

The asymptomatic HIV patient is typically attacked by pneumonia (Pneumocystis Carinii), which must be treated carefully. Alternatively it is terminal. There is oral candidiasis, recurrent diarrhoea, and progressive demyelinisation of CNS (ie, leuco-encephalopathy).

Viral hepatitis is an infection of the hepatocytes causing cell destruction (necrosis) and inflammatory reactions. The five hepatitis viruses are marked A, B, C, D, and E. Also yellow fever virus, Epstein-Barr virus (infectious mononucleosis), cytomegalovirus, rare enteroviruses, herpes simplex virus, rubella virus, and Ebola-virus can cause hepatitis.

Hepatitis A virus (HAV) is a small picornavirus, which is only replicated in hepatocytes, excreted with the bile and found in the faeces of infected patients. HAV is evidenced in the faeces for 3 weeks just before the onset of the jaundice and a week after.

The virus is directly *cytotoxic* but probably also acts by immuno-mediation such as helping natural killer cells to kill hepatocytes. Following a relatively short incubation period, the

patient develops increasing fatigue, lack of appetite, vomiting, diarrhoea and fever. Abdominal pain may simulate acute abdomen, but soon the jaundice is diagnostic. When the icteric phase begins there is often subjective improvement. The urine becomes dark with green foam by bilirubin, and the stools are pale owing to intrahepatic cholestasis. The liver damage is shown by a rise in *aspartate serum amino-transferase* (AST) and 5 weeks from the exposure there is a rise in *IgM anti-HAV* in the blood. The latter is an indicator of acute infection, and soon disappears. As the patient recovers more and more *IgG anti-HAV* appears and persists for life. Complete recovery is followed by lifelong immunity.

Hepatitis B virus (HBV) is a hepatotropic virus, which was the first member of the group, termed hepadnavirus. HBV probably affects the hepatocyte by immunological processes without being directly cytotoxic. HBV consists of a shell and a core. The *shell* expresses an antigen called *Hepatitis B surface Antigen* (HBsAg), which is secreted into the blood plasma from the infected hepatocyte. The *core* of the virus contains double-stranded DNA, DNA polymerase, and immunogenic material termed *core Antigen* (HBcAg) and a degradation product of HBcAg called *e Antigen* (HBeAg). Many cases are asymptomatic, other cases show a clinical course like the HAV infection, and still others show a fulminant pattern or end in chronic hepatitis. Complete recovery is typical, but all recovered patients are potential carriers. HBV is spread by blood or blood products and is also found in semen and saliva. The intravenous route includes blood transfusion, contaminated needles, or sexual contact.

HBsAg appears in the blood even before the onset of symptoms. Typically, the recovery is rapid and HBsAg disappears again, but now antibodies to HBsAg (IgM anti-HBs) appear. These *IgM anti-HBs molecules* provide lifelong immunity, and are lifelong markers of previous HBV infection. Also *IgM anti-HBc,* the antibody to HBcAg, is a useful marker of previous HBV infection, probably without protective functions. HBeAg appears in the blood while the infection is severe, and declines rapidly in recovery. If HBeAg remains in the blood it usually expresses fulminant viral replication. *IgM anti-HBe* appears as the *HBeAg* disappears from the blood, and their presence indicates a relative improvement of hepatocyte function.

Precipitation of immune complexes (HBsAg- IgM anti-HBs) sometimes causes *serum sickness*, with glomerulonephritis, polyarthritis, pancreatitis, and urticarial rashes.

Chronic hepatitis is hepatic failure with sustained antigenaemia (ie, *HBsAg*) for more than 6 months. *Chronic hepatitis B patients* carry a significantly increased risk of *liver cirrhosis* and *hepatocellular carcinoma.*

Hepatitis C virus (HCV) is a 60 nm large RNA flavovirus, which is responsible for most cases of post-transfusion hepatitis. The clinical course is discrete or asymptomatic, but later many patients develop chronic hepatitis. HCV antigens produce antibodies, which are found in the blood plasma (*IgM anti-HC*). Some cases are treated with *interferon-α.*

Hepatitis D virus (HDV) is an RNA virus frequently affecting drug abusers. In the blood HDV is coated with *HBsAg,* and HDV infection is thus associated exclusively with HBV co-infection or with HBsAg-positive patients. The clinical course is like acute HBV infection, but sometimes fulminant or fatal.

Diagnosis is confirmed by a blood sample showing *IgM anti-D* together with *IgM anti-HBc*.

Hepatitis E virus (HEV) is an RNA calcivirus, which causes a water-borne hepatitis. Epidemics have occurred among risk groups for HBV infection and in developing countries. Good hygiene is essential, and boiling of water for at least 10 min is necessary in areas with contaminated water.

The therapy of viral hepatitis is symptomatic. Prevention by avoidance of risk factors and immunisation is the only rationale. Chronic HBV carriers can be treated with *interferon-α,* which may reduce the risk of later development of hepatic cirrhosis and hepatocellular carcinoma.

3. Fungal infections

Yeast reproduces by budding, whereas *moulds* grow by branching hyphae. Histoplasmosis, cryptococosis, blastomycosis, and candidiasis are all systemic infections, which can be treated with the polyene, amphotericin, which kills young growing cells.

Histoplasmosis and *cryptococcosis* are world-wide mycoses caused by Histoplasma capsulatum and Cryptococcus neoformans (yeast fungi). The reservoir for the spores is bird and bat droppings and wet soil. The clinical features of pulmonary mycosis are like pulmonary tuberculosis with a tendency for fistulae. Cryptococcosis usually presents as meningitis. *Candidiasis* is caused by Candida albicans - the most common fungus in humans mainly occurring as vaginal infections and oral thrust.

Blastomycosis is primarily a skin disease caused by blastomyces dermatitidis. Pulmonary lesions look like TB or malignancy on X-rays.

Dermatophytoses affect the skin, hair and nails (Trichophyton, Microsporum, and Epidermophyton).

4. Protozoan infections

Different types of *Leishmania* cause *leishmaniasis*. The protozoa are transferred via the sandfly. It invades the reticuloendothelial cells, and replicate before they spread to cause bone marrow hyperplasy, hepatomegaly, lympadenopathy, and splenomegaly. Visceral Leishmaniasis is called Kala-azar.

Malaria (meaning *bad air*) in 1897 was found to be caused by the parasite, plasmodium, found in the stomach of mosquitoes. Four species of Plasmodium exist: The malignant, virulent P. falciparum, and the three less virulent P.malariae, P. ovale and P. vivax. Malaria is said to have spoiled civilisations throughout history.

The parasites are transmitted by female Anopheles-mosquitoes, by infected drug syringes or by blood transfusion. As the mosquito bites, the parasite enters a red cell and multiplicate. Phagocytosis of parasitized red cells leads to reticuloendothelial hyperplasia with hepato-splenomegaly. P. Falciparum can invade red cells at any age and is the most malignant type, because of red cell haemolysis.

A new mosquito sucks blood-containing parasites from the patient. The parasites multiply in the stomach and salivary glands of the mosquito. The next mosquito bite of a new victim is the start of a new life cycle for the parasite. The incubation period is around 2 weeks.

Cerebral malaria is characterised by attacks of extremely high body core temperature, convulsions, coma, and possibly death. The attacks occur each 2. or 3. day when the parasites synchronously leave the red cells and spread into the blood. Blackwater fever is named after the black urine caused by intravascular haemolysis. Encephalopathy, congestive heart failure, pulmonary oedema, emboli, and renal failure may develop into a terminal crisis.

Elimination of mosquitoes has failed to prevent malaria. In several regions of Asia, the parasite is resistant to all traditional antimalarial drugs, because the parasites can change their coating, so the host immuno-defence system does not work.

Toxoplasmosis is caused by infection with the intracellular Toxoplasma gondii. There are many mammals functioning as intermediate hosts but the final host is the cat. Infection occurs by ingestion of food contaminated by Toxoplasma gondii cysts. Sulphadiazine and the diaminopyrimidine, pyrimethamine, for one month is effective as a synergistic therapy.

Trypanosomiasis is found in Africa (sleeping sickness) and in South and Central America (Chagas disease). The African Trypanosoma brucei is transmitted by both male and female tsetse flies, whereas various insects transmit the American Trypanosoma Cruzi. First there is a

tender nodule at the site of the bite. Then the protozoa invade the lymphatic system and finally the blood. In *sleeping sickness* the protozoa reach all organs of the body but the CNS is particularly occupied. A meningo-encephalomyelitis develops, where the patient becomes apathetic and sleepy. In *Chagas disease* there is regional lympadenopathy, fever, hepatomegaly, and myocarditis. Chronic Chagas disease is an autoimmune disorder caused by T-cells and antibodies against vital organs (cardiac failure, emboli).

5. Helminthic infections

The helminths are divided into three groups: *5a. Roundworms (ie nematodes), 5b. Flukes (ie trematodes, flatworms), and 5c. Tapeworms (ie cestodes, flatworms).*

5a. Nematodes

or roundworms are divided into *intestinal* (*Ascaris lumbricoides, Enterobius vermicularis or pinworm, Trichinella spiralis, Tricuris trichiura or whipworm, and the hookworm or Anchylostoma duodenale/Necator americanus),* and *extraintestinal roundworms (filariae or Wuchereria bancrofti/Brugia malayi, Loa loa or eyeworm, Onchocerca volvulus, and Dracunculus mediensis or guinea worm).*

5a.a. Ingestion of infective eggs is the usual cause of infection with *intestinal roundworms*. Identifying eggs in the faeces makes the diagnosis.

Ascaris lumbricoides causes *ascariasis,* where adult roundworms attach to the small intestine, and pregnant worms discharge eggs that are expelled with the faeces. In moist soil the eggs become infective within a month. When ingested in food or water, the eggs hatch, and the larvae enter the portal blood. The larvae pass with the blood through the liver and heart, reach the lungs and develop into 3.th stage larvae. They ascend up the trachea, down through the alimentary tract and stay in the small intestine, where they mature. Massive infections cause intestinal obstruction, appendicitis and malnutrition. Eggs can also be deposited in other tissues.

Enterobius vermicularis (pinworm) causes *enterobiasis*, where the worms live in the colon. Pregnant worms migrate through the anus and deposit eggs in the perianal region, where the victim experiences anal pruritus. Humans ingest the infective eggs, they hatch in the intestine and the larvae mature in the colon.

Trichinella spiralis

Trichinella spiralis is transmitted, when we are eating undercooked or raw meat (pig) containing encysted larvae. During digestion the encysted larvae are liberated and mature into adult worms that attach to the mucosa of the small intestine. The worms then discharge larvae that reach the blood stream and penetrate the vessel walls to be encysted in striated muscle. This may cause muscle pain, oedema, eosinophilia and fever. The larvae also migrate to the heart, the lungs and the nervous system.

Tricuris trichiura (whipworm)

The whipworm infects humans when ingesting their eggs in food and water. The eggs hatch larvae in the gut and they mature into adults in the colon, where they thread their whip end into the mucosa. Massive infection causes abdominal pain including acute abdomen with appendicitis or diarrhoea with bloody stools. The loss of blood leads to anaemia.

Hookworm (Anchylostoma duodenale/Necator americanus)

The hookworm larvae survive for months in soil protected from direct sunlight. The larvae penetrate the skin of humans and enter the blood. In the lungs they migrate to the alveoli, ascend to the epiglottis, and are swallowed. In the intestine they attach to the mucosa with their toothlike hooks and eat the villi. The bleeding leads to anaemia, and the skin and lung

lesions result in dermatitis and pneumonitis.

5a.b. The extraintestinal roundworms penetrate the skin. The diagnosis is established by microfilariae in the blood or by filariae in biopsies.

Filariae (Wuchereria bancrofti/Brugia malayi)

These parasites cause *filariasis*. Both filariae live in and block lymph vessels near the testes. Lots of different mosquitoes are intermediate hosts and humans are the final host. Symptoms are due to blockade and dilatation of lymphatic vessels to the testis, the epididymis and the spermatic cord. The tissue is inflamed with accumulation of lymphocytes, eosinophils and plasma cells. Pregnant filariae discharge microfilariae to the blood, where they may cause thrombo-embolism and create pulmonary infiltrates. The lymphoedema of the extremities results in elephant legs and is called *elephantiasis*.

Loa Loa (eyeworm)

The Loa Loa worm causes the African loiasis. Mango flies transmit the infection, and human are the definitive host. The worm migrates through the skin or the conjunctiva of the eye. Pregnant worms discharge microfilariae into the blood. Dermal swellings of worms surrounded by inflammation are typical findings. Worms can be extracted beneath the conjunctiva.

Onchocerca volvulus

This worm causes *onchocerciasis*, which is transmitted by blackflies and man is the ultimate host. The adult worm is coiled up in the subcutaneous tissue and form inflammatory nodules. Pregnant worms discharge large amounts of microfilariae, which migrate, to the eyes, lymph nodes and other organs. In the eye the microfilariae cause inflammation and blindness, and in the inguinal lymph nodes they cause *genital elephantiasis*.

Dracunculus mediensis (guinea worm)

The guinea worm is transmitted with water contaminated with an intermediate host, a crustacean of the type *Cyclops.* The adult worm lives in the subcutaneous tissue and cause urticaria with blisters. The blisters burst in contact with water and a multitude of larvae are discharged from the pregnant worm into the water.

Treatment of intestinal and extraintestinal roundworms is performed with antibiotics (mebendazole, thiabendazole), that inhibit glucose transport and protein secretion in the worms.

5b. Flukes (trematodes)

are found in the blood, the intestine or in the lungs.

The most important flukes are the blood-dwelling Schistosoma haematobium, S. japonicum and S. mansoni. The intermediate hosts are fresh-water snails.

Schistosomiasis (Bilharziasis) presents with swimmers itch at the site of invasion of the parasite. The blood to the heart, lungs and liver transports the cerariae. Eggs surviving in the mesenteric venules, the liver, and the urinary bladder cause most of the clinical manifestations. Generalized allergy develops with urticaria, asthma, eosinophilia, myalgia and fever. The patient eventually develops pneumonia, hepatosplenomegaly and lympadenopathy.

The life cycle of the parasite is complicated.

The adult worms reside in pairs with the female lying in the gynecophoric canal of the male. The pair copulates and produces several hundred eggs daily in humans. Within the egg a larval form, miracidium, develops. The schistosome eggs must leave the body to complete the life cycle of the parasite, and they penetrate the intestinal wall or the bladder wall to be expelled with the faeces or the urine. On contact with fresh water the miracidiae hatch from the eggs

and find a snail, which they penetrate. In the snail they multiply enormously, and thousands of infective cerariae are expelled daily. During swimming or contact with water, humans are infected by cerariae penetrating the skin or mucous membranes.

Flukes are treated with praziquantel, which destroys the surface of the worm, after which the worm is killed by host killer cells (see above).

5c. Tapeworms (Cestoda)

Many tapeworm infections are asymptomatic. Taenia saginata (beef tapeworm), T. solium (pork tapeworm), Diphyllobotrium latum (fish tapeworm), and Hymenolepsis nana (dwarf tapeworm) all live within the human intestinal tract. When human ingests the T. solium eggs from human faeces and become infected with cysticerci, a serious condition - called *cysticercosis* - is in progress. The eggs release larvae, which enter the blood, remain in different tissues, and develop into an intermediate larval form, the cysticercus that can survive indefinitely. The cysticercus grows and compresses the surrounding tissues. In the brain the cysticerci may cause hemiplegia or epilepsy, and in the eye they may cause blindness

Tapeworms are also killed by praziquantel (see above).

6. Sexually related infections

Gonorrhoea, Clamydia, Herpes simplex, Trichomoniasis, Candidiasis, hepatitis, HIV and AIDS, syphilis, yaws, Bejel and pinta are sexually related infections already described above

This paragraph deals with Chancroid, Lymphogranuloma venereum, Granuloma inguinale, Scabies, and Pediculosis pubis.

Chancroid or soft chancre (ulcus molle) is a venereal disease caused by a short bacillus, Haemophilus Ducreyi. The ulcer is soft and extremely tender in contrast to the hard syphilitic chancre. The bleeding ulcers may transfer HIV infection. The bacillus grows on special culture media, but it is fastidious. Cephalosporins and quinolones may prove effective.

Lymphogranuloma venereum is caused by Chlamydia trachomatis (besides trachoma, see above). The ulcer is painless, and the inguinal lymph nodes grow. Oxytetracycline is necessary.

Granuloma inguinale is caused by Calymmato-bacterium granulomatis, which is identified with the microscope. Erythromycin or tetracycline is effective.

The mite, Sarcoptes scabiei, causes scabies. A skin lotion with Malathion is applied. Malathion is a widely used pesticide that works as an anti-choline-esterase and kills the eggs of mites and lice.

Pediculosis pubis or pubic lice is a venereal disease, where lice are found in the pubic hair or in all hairy areas. The eggs are killed with Malathion, which is safer than other pesticides.

Genital ulcer diseases associate with HIV infection, and their cure is important in the prophylaxis against spread of AIDS.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. Prostaglandins are important for the inflammatory reaction, and they are products of cyclo-oxygenase activity. ASA and corticosteroids inhibit the synthesis.

B. The maximal reabsorption of NaCl and water occurs in the distal tubule of the nephron, and

is driven by the Na^+-K^+-pump.

C. Japanese encephalitis is caused by a flavivirus and is transferred by mosquitoes.

D. Acyclovir inhibits viral DNA synthesis, and is effective enough to kill HIV.

E. Influenza is not a clinical entity caused by influenza virus A, B and C, only. Infection with other viruses can produce the same clinical picture.

Case History A

A male comatous patient is brought to the city hospital in Patuakhali (at the river, Ganges). The patient is 1.80 m tall, but is obviously dehydrated. His normal body weight is 75 kg, and his weight at admittance is 60.

The normal extended extracellular fluid volume is 20% of the body weight (Chapter 17).

There is a severe rice-water diarrhoea of more than 1 litre in the first hour. Motile vibrios are demonstrated in the microscope by dark-field illumination, and the motility is reduced with vibrio cholerae antiserum. The haematocrit is 60%, serum Na^+ is 120 mM and the Base Excess is -15 mM. The pH_a is 7.17 and P_{aCO2} is 33 mmHg (4.4 kPa).

The patient is immediately treated with intravenous infusion of the WHO diarrhoea treatment solution. This is 68 mmol NaCl (4 g), 13 mmol KCl (1 g), 50 mmol glucose(9 g), and 80 mmol sodium acetate (6.5 g) dissolved in 1 litre of water.

The rehydration has a dramatic effect following 2 litres of infusion. The patient is awake and complains of thirst. The Base Excess is now -8 mM, the pH_a is 7.33 and P_{aCO2} is 31 mmHg (4.14 kPa).

The patient is also given tetracycline for 3 days, which cure the condition, and he is released from hospital after 4 days. The last two days he only receives an oral glucose-electrolyte solution.

1. *Describe the pathophysiology of cholera.*
2. *Describe and calculate the normal extended ECV at the start of the disease.*
3. *What is the fractional loss of extended ECV per hour?*
4. *Is rehydration important?*
5. *What is the effect of tetracycline?*
6. *Calculate the number of base equivalents missing in the normal extended ECV. The approximate amount was included in the first 2 litres of infusion. Why was this amount insufficient (change of base excess from -15 to -8 instead of zero)?*

Case History B

A 22-year old male returns from Africa, where he has been travelling for 6 weeks. Two weeks later he is not at work, and he does not answer the telephone. A friend goes to his flat, where newspapers are piled in front of the door, and there is no reply. The police and a key expert are called, and when entering the rooms they find the young man in coma. Immediate admission to the emergency department is arranged. Here the core temperature is recorded to 40.8^{o}C. The hospital doctor institutes a series of tests and calls the patients GP in order to get

relevant information. He is informed that the patient before going to Africa was vaccinated against hepatitis A, typhoid, tetanus, and meningococcal meningitis. The patient was also instructed about malaria prophylaxis and advised to take 300 mg chloroquine each Sunday during the travel and continue for 7 weeks after returning home.

1. What is the probable diagnosis?

2. How are the diagnosis confirmed?

3. What are the proper treatment?

Case History C

A male refugee from ex-Yugoslavia , 36 years old (weight 69 kg, height 1.82 m), is examined by a doctor at the refugee camp, because of coughing and shortness of breath. He has lost 7 kg in weight recently. There has been night sweats and sparse mucous blood-stained sputum. During auscultation a few crackles are heard over both lungs. The patient has never smoked. The patient lives in close quarters with many people, and he is admitted to hospital for further examination. The peak expiratory flow is measured to 510 l min^{-1}*. A chest X-ray shows patchy shadows in the upper zones of both lung fields. Two of the shadows are calcified. Sputum is stained with the Ziel-Nielsen stain, but no acid and alcohol-fast bacilli are found. Gastric juice is aspirated at 3 occasions. The gastric juice is cultured on Løwenstein-Jensen medium for 4 weeks, where characteristic acid and alcohol-fast bacilli are found after Ziel-Nielsen staining.*

1. What is the diagnosis?

2. What are the treatment?

3. Does this patient benefit from steroid therapy?

4. What are the necessary screening procedures for the persons living with the patient before he was sent to hospital?

Try to solve the problems before looking up the answers .

Highlights

- *Antibiotics have toxic effects, especially when used in overdose or in the presence of other disease such as hepatic or renal failure. Aminoglycosides are both oto- and nephro-toxic.*
- *Antibiotics can lead to secondary infections with fungi or with Clostridium difficile.*
- *Serious bacterial infections (eg, endocarditis, meningitis, TB, septicaemia) and infections in immunodeficient patients are commonly treated with bactericidal drugs (aminoglycosides, cephalosporins, and penicillins). Such serious infections are preferably treated with combinations of synergistic antibiotics.*
- *The advantages of cephalosporins over penicillins are their resistance to penicillinase and their wider antimicrobial spectrum (both Gram-positive and Gram-negative bacteria).*
- *Aminoglycosides, chloramphenicol, clindamycin, erythromycin, fusidic acid, mupirocin, spectinomycin, tetracycline all bind to bacterial ribosomes and interfere with protein synthesis.*
- *Fusidic acid has a bile acid structure and inhibits Gram-positive cocci including*

penicillinase-producing Staphylococcus aureus in combination with other antibiotics.

- *Idoxuridine and vidarabine is used as early as possible against herpes simplex virus and against varicella zoster virus (herpes zoster and severe varicella or chickenpox).*
- *Acyclovir inhibits viral DNA synthesis and is effective against the same virus disorders.*
- *Azidodeoxythymidine inhibits HIV reverse transcriptase and thus impairs viral replication. This substance is used for HIV patients, although there is danger of bone marrow depression.*
- *Interferon is produced by the T-lymphocytes during virus infections. Interferon is used for hepatitis B and C.*
- *Leishmaniasis is treated with meglumine antimoniate and Na-stibogluconate. These substances inhibit phosphofructokinase and Krebs cycle enzymes in the Leishmania organism.*
- *Malaria is caused by infection with the Plasmodium protozoa, which is present in the human host in the blood (erythrocytic) or the tissue (extraerythrocytic stage. The erythrocytic stage is treated with aminoquinolines (eg chloroquine) or analogues.*
- *Chagas disease is treated with nifurtimox, which produces free radicals. The free radicals react with oxygen to form superoxide anion, hydroxyl-free radical, and hydrogen peroxide. The trypanosomes do not contain enzymes able to inactivate these reactive substances, so they cause peroxidation of lipids and ribonucleic acids (DNA, RNA).*
- *Roundworm disorders are treated with mebendazole or with thiabendazole. These substances bind to the cytoplasmic microtubules of the roundworm and inhibit glucose transport and protein secretion.*
- *Flukes are sensitive to praziquantel, which attacks the surface barrier (tegumentum) and the muscles of the worm. Praziquantel destroys the surface of the worm so host antibodies can bind to the worm antigens.*
- *Tapeworms are also killed by praziquantel, but as an alternative the tapeworms are sensitive to niclosamide. Niclosamide binds to the mitochondria of both host and parasite and block ATP formation.*
- *Cholera is a disease caused by an enterotoxin produced by the gram-negative bacillus, vibrio cholerae. If this bacillus is allowed to proliferate within the lumen of the small intestine, it causes profuse watery diarrhoea - up to 24 l per 24 hours - dehydration and circulatory shock. Fulminant cholera can kill the patient within a day.*
- *Plague is spread from woodland rodents to domestic rats. The rat flee bites humans. Bubonic plague occurs suddenly with high fever and the patient is confuse and may seem drunk due to toxaemia. The inguinal lymph nodes rapidly develop lympadenopathy or buboes. Pneumonic plague is fulminant pneumonia with bloody sputum and cardiac failure.*
- *Syphilis (Lues) is caused by Treponema pallidum, which is a mobile spirochaete. Congenital syphilis is acquired transplacentally from mother to foetus and is apparent a few weeks after birth. The babies pass through three stages of the disease similar to those of the adults. Adult syphilis is acquired by intimate sexual contact.*

- *Flukes (trematodes) are found in the blood, the intestine or in the lungs. The most important flukes are the blood-dwelling Schistosoma haematobium, S. japonicum and S. mansoni. The intermediate hosts are fresh-water snails. Schistosomiasis (Bilharziasis) presents with swimmers itch at the site of invasion of the parasite.*

- *Taenia saginata (beef tapeworm), T. solium (pork tapeworm), Diphyllobotrium latum (fish tapeworm), and Hymenolepsis nana (dwarf tapeworm) all live within the human intestinal tract. When humans ingest the T. solium eggs from human faeces and become infected with cysticerci, a serious condition - called cysticercosis - is in progress.*

Further Reading

Clinical infectious Diseases. Monthly journal published by the Infectious Diseases Soc., University of Chicago Press, Journals Division, PO Box 37005, 5720 South Woodlawn, Chicago IL 60637, USA.

Mims C, Playfair J, Wakelin D, and R Williams. *Medical Microbiology*. Mosby, London, 1998.

Return to top

Return to Content

Symbols And Units

1. Force is measured in Newton (N). *One Newton* (kgm s^{-2}) is the force required to accelerate a mass of 1 kg with an acceleration of 1 m s^{-2}. The acceleration due to gravity is generally accepted as g or G = 9.8067 or 9.807 m s^{-2}.

2. Joule established already in 1848 that *mechanical work* and *heat ener*
used unit of energy is the calorie (cal), which is the energy, required to raise the temperature of 1 gram (g) of water from 14.5 to 15.5°C. Work is force times distance, and it is measured as Newton-meter or Joule (J). The *Joules equivalent* has been determined to be 4.187 J cal^{-1}.

3. Finally, *work-rate* or *power* is calculated as work per second (s). The power unit 1 W equals 1 J s^{-1}.

4. *Pressure* is measured as force per area unit that is in *N m^{-2}* or *Pascal*.

In the gravity field of the earth G or g equals 9.807 m s^{-2}. Blood and sea water has a relative density of 1033 kg m^{-3}. A 10 m high sea water column resting on one square m, corresponds to the following pressure: (10 m × 1033 kg m^{-3} × 9.807 m s^{-2}) =

101 306.3 (kg m s^{-2}) m^{-2}. This is *101 306.3 N m^{-2}* or *101.3 kPa* (= 1 atmosphere). The classical concept is that 1 atmosphere equals 760 mmHg. Accordingly, 1 Torr or 1 mmHg equals (101 306.3 Pa/760 =) *133.3 Pa*. In this book pressures are given in Pa (or kPa) together with mmHg.

5. *Concentration* is mass per volume unit. Squared brackets around a substance or **C** denote concentration. The international unit is mM = mmol l^{-1} = mol m^{-3}.

6. A prefix scale for different units is used as follows: milli = m = 10^{-3}; micro = m = 10^{-6}; nano = n = 10^{-9}; pico = p = 10^{-12}; femto = f = 10^{-15}.

International Symbols

(Fed.Proc. 9: 602-605, 1950).

This is a precise short-cut for intellectual transfer used by all physiologists.

A *dash* next to any symbol ($\bar{}$) indicates a mean value. A *dot* next to any symbol (·) denotes a time derivative. *Small letters* in a suffix denote gas dissolved in blood, whereas *large letters* denote gas in air. The symbol is often the first letter in the English word.

A:		
a	=	Solubility: The Bunsen solubility coefficient (ml STPD per ml fluid per 760 mmHg)
A	=	Alveolar gas
AA	=	arachidonic acid
Ach	=	acetylcholine
ACTH	=	adrenocorticotropic hormone
Ad	=	adrenaline

ADH	=	antidiuretic hormone
ADP	=	adenine diphosphate
AIDS	=	acquired immunodeficiency syndrome
AMP	=	adenine monophosphate
AMPA	=	special glutamate receptors
ANF	=	atrial natriuretic factor
ANH/ANP	=	atrial natriuretic hormone/peptide
AP	=	action potential
AR	=	absolute refractory period
ASA	=	acetylsalicylic acid
ATP	=	adenine triphosphate
ATPS	=	ambient temperature, pressure, saturated with water vapour
AV node	=	atrioventricular node

B:		
BB	=	buffer base
BD	=	base deficit
BE	=	base excess
BMR	=	basal metabolic rate
BSA	=	body surface area
BTPS	=	body temperature and ambient pressure, saturated with water vapour

C:		
C	=	concentration of gas in blood. Squared brackets around a substance also denote concentration
Cal	=	calorie
C_{v_CO2}	=	concentration of CO_2 in mixed venous blood
CA	=	carbonanhydrase
cAMP	=	cyclic adenine monophosphate
CBF	=	cerebral bloodflow
CBG	=	corticosteroid binding globulin
CCh	=	carbacholine
CCK	=	cholecystokinin
cGMP	=	cyclic guanosine monophosphate

CNS	=	central nervous system
CSF	=	cerebrospinal fluid
COLD	=	chronic obstructive lung disease
COMT	=	catechol-O-methyl transferase
C peptide	=	connecting peptide
CRH	=	corticotropin releasing hormone
CVP	=	central venous pressure
D:		
D	=	diffusion capacity
Da	=	Daltons (MW units)
DAG	=	diacylglycerol
1, 25-D_3	=	1,25-dihydroxy-cholecalciferol
25-OH-D	=	25-hydroxy-cholecalciferol
DIT	=	di-iodine-thyronin
DM	=	Diabetes mellitus
DMNV	=	dorsal motor nucleus of the vagus
DMPP	=	dimethylphenylpiperazine
DNA	=	deoxyribonucleic acid
DOPA	=	dihydroxy-phenylalanine
2,3-DPG	=	diphosphoglycerate
DPPC	=	dipalmitoyl phosphatidylcholine
E:		
E	=	expiration
E_{net}	=	mechanical net-efficiency of external work
EAA	=	excitatory amino acids
ECG	=	electrocardiogram
ECF	=	extracellular fluid
ECV	=	extracellular fluid volume
EDIP	=	end-diastolic intraventricular pressure
EDRF	=	endothelium-derived relaxing factor
EDTA	=	ethylene-diamine-tetra-acetate

EEG	=	electroencephalogram
EF	=	excretion fraction
EGF	=	epidermal growth factor
e.p.	=	equilibrium potential
EPSP	=	excitatory postsynaptic potential
ER	=	endoplasmic reticulum
ERBF	=	effective renal blood low
ERPF	=	effective renal plasma flow
ERV	=	expiratory reserve volume
ESV	=	end systolic volume
F:		
F	=	fraction of gas in dry air or force
f	=	respiratory frequency (breath/min)
FABP	=	fatty acid binding protein
FAD	=	flavine adenine dinocleotide
$FADH_2$	=	flavine adenine dinucleotide (reduced)
FFA	=	free fatty acids
FGF	=	fibroblast growth factor
FRC	=	functional residual capacity (= RV + ERV)
FSH	=	follicle stimulating hormone
FU	=	Flow units in ml of blood (100 g tissue)$^{-1}$ min^{-1}
G:		
G	=	Gibbs energy (free, chemical energy)
GABA	=	gamma-aminobutyric acid
GFF	=	glomerular filtration fraction
GFR	=	glomerular filtration rate (normal 118-120 ml min^{-1})
GH	=	growth hormone
GHIH	=	growth hormone inhibiting hormone
GHRH	=	growth hormone releasing hormone
GIP	=	gastric inhibitory peptide or glucose-dependent insulin-releasing peptide
GLP	=	glucagon-like peptide

GnRH	=	gonadotropin releasing hormone
GLUT	=	glucose transporter
GRP	=	gastrin releasing peptide
GTP	=	guanosine triphosphate
H:		
H	=	heat content (enthalpy; all energy when the pressure-volume work is zero)
Hb	=	haemoglobin (haemoglobin F = foetal haemoglobin)
HBF	=	hepatic blood flow
hCG	=	human chorionic gonadotropin
HDL	=	high density lipoprotein
HGF	=	hepatocytic growth factor
HGH	=	human growth hormone
HIP	=	hydrostatic indifference point
HIV	=	human immunodeficiency virus
hPL	=	human placental Lactogen
HPLC	=	high pressure liquid chromatography
HSS	=	hepatocyte stimulating substance
I:		
I	=	inspired gas
ICSH	=	interstitial cell stimulating hormone
ICV	=	intracellular fluid volume
IDDM	=	insulin-dependent diabetes mellitus
IDL	=	intermediate density lipoprotein
IGF	=	insulin-like growth factor
IGF-BP	=	IGF-binding protein
IP_3	=	inositol triphosphate
IRV	=	inspiratory reserve volume
ISF	=	interstitial fluid (tissue fluid)
Iso	=	isoprenaline
ISS	=	interpreted signal strength
i.v.	=	intravenous
J:		

J	=	flux of a substance (mol min^{-1}) through an area unit
J	=	Joule
JG	=	juxtaglomerular
K:		
K	=	Kelvin degrees of temperature
L:		
LAT	=	lactic acid threshold
LBNP	=	lower-body-negative-pressure
LES	=	lower oesophageal sphincter
LH	=	luteinizing hormone
LHRH	=	luteinizing hormone releasing hormone
LPL	=	lipoprotein-lipase
LDL	=	low density lipoprotein
LTH	=	prolactin
LVET	=	left ventricular ejection time
M:		
MAO	=	monoamine oxidase
MAP	=	mean arterial pressure/mean aortic pressure
MeCH	=	metacholine
MEOS	=	microsomal ethanol oxidation system
MG	=	monoglycerides
2MG	=	2-monoglyceride
MIH	=	Muller inhibiting hormone
MIT	=	mono-iodine-thyronin
mM	=	mmol l^{-1}
MR	=	metabolic rate
MSH	=	melanocytic stimulating hormone
MW	=	molecular weight (in Daltons)
N:		
N	=	Newton
NA	=	noradrenaline

NAD	=	nicotinamide adenine dinucleotide
$NADH_2$	=	nicotinamide adenine dinucleotide (reduced)
NANC	=	non-adrenergic, non-cholinergic
NBB	=	normal buffer base/neutral brush border
NGF	=	nerve growth factor
NIDDM	=	non-insulin-dependent diabetes mellitus
NMDA	=	*N*-methyl-D aspartate
NOS	=	nitric oxide synthase
NSAID	=	non-steroid anti-inflammatory drug

P:		
P	=	partial pressure of gas in air or blood
PAH	=	para-amino hippuric acid
PCV	=	packed cell volume
PDE	=	phosphodiesterase
PDGF	=	platelet derived growth factor
PEF	=	peak expiratory flow
PG	=	prostaglandins
PG_2	=	prostacyclin
PGE_2	=	prostaglandin E_2
PIF	=	prolactin inhibiting factor
PIP_2	=	phosphatidyl-inositol diphosphate
P_B	=	barometric pressure
$P_{c'CO2}$	=	partial pressure of CO_2 in end-capillary blood
P_{IO2}	=	partial pressure of O_2 in inspired air in trachea
P_{aO2}	=	partial pressure of O_2 in arterial blood
POMC	=	pro-opiomelanocortin
PP	=	pancreatic polypeptide/ pulse pressure amplitude
PRL	=	prolactin
PRU	=	pressure resistance unit
PTH	=	parathyroid hormone
PVR	=	pulmonary vascular resistance

Q:		
$\dot{Q}$	=	Cardiac output (l min^{-1})
QRS	=	the ventricle complex of the ECG
R:		
R	=	ventilatory exchange ratio (pulmonic)
R	=	Gas constant
RAS	=	reticular activating system
RBF	=	Renal bloodflow
RC	=	respiratory controller/ respiratory centres
REM	=	rapid eye movements
RES	=	reticulo-endothelial system
RIA	=	radio-immuno assay
RMP	=	resting membrane potential
RNA	=	ribonucleic acid
RPF	=	renal plasma flow
RPM	=	revolutions per minute
RQ	=	respiratory quotient (metabolic)
RR	=	relative refractory period
RV	=	residual volume
S:		
S	=	entropy (the tendency to spread in a maximum space)
S	=	saturation degree
SA	=	specific activity
SAmode	=	sinoatrial node
SB	=	standard bicarbonate concentration
SBE	=	standard base excess
SDA	=	specific dynamic activity
SR	=	sarcoplasmic reticulum
SS	=	steady state/ stimulus strength
STPD	=	standard temperature and pressure, dry (0^{o}C, 760 mmHg)
STN	=	solitary tract nucleus

sv	=	stroke volume

T:		
T	=	*t*ension (force)
T	=	temperature
T_3	=	Tri-iodo-thyronine
T_4	=	tetra-iodo-thyronine
TBA	=	thyroxine-binding albumin
TBG	=	thyroxine-binding globulin
TBPA	=	thyroxine-binding prealbumin
TBV	=	total blood volume
TCA	=	tri-carboxylic acid
TEV	=	total erythrocyte volume
TFGF	=	transforming growth factor
TG	=	triglycerides
TGF	=	tubuloglomerular feedback
TH	=	total haemoglobin content
TLC	=	total lung capacity (=RV+VC)
TP	=	threshold potential
TPVR	=	total peripheral vascular resistance
TRH	=	thyrotropin-releasing hormone
tRNA	=	transfer RNA
TSH	=	thyroid-stimulating hormone
TV	=	tidal volume
TxA2	=	thromboxane A2

V:		
v dash	=	linear mean velocity
$V\dot{}$	=	volume velocity of gas
V	=	volume
$V\dot{}_A$	=	expired alveolar ventilation (l min^{-1})
VC	=	vital capacity (=IRV+TV+ERV)
V_D	=	dead volume

W	=	Watts (J s^{-1})
W	=	external work (with pressure-volume work zero)

Nutritive Equivalents And Enthalpy

Nutritive equivalents for oxygen are:

Carbohydrate 37 mmol oxygen g^{-1}, fat 91 mmol oxygen g^{-1} , and protein 43 mmol oxygen g^{-1}. On a *mixed diet 20 kJ of energy* is transferred per litre STPD of oxygen used; the RQ is 0.8.

Nutritive equivalents for carbon dioxide are:

Carbohydrate 37 mmol g^{-1}, fats 64 mmol g^{-1}, and protein 34 mmol g^{-1}.

Metabolic enthalpies (heat energy liberated in the body per g combusted nutrient) in kJ g^{-1} substance: Protein 17, fat 39 and carbohydrate 17.5.

Essential Atomic And Molecular Weights

These are given in g mol^{-1} (or Daltons, Da) throughout the text. Calcium 40; Carbon 12; Glucose 180; Helium 4; Hydrogen 1; Nitrogen 14; Oxygen 16; PAH 194.2; Phosphorus 31; Potassium 39; Sodium 23; Xenon 131.

Physical Constants And Conversion Factors

Acceleration due to gravity (standard 1 G): 9.81 m/s^2.

Avogadro's constant: 6.02 10^{23} molecules mol^{-1} .

Diffusion coefficients for most molecules: 10^{-10} m^2 s^{-1} per molecule.

Energy (J = N m = Volts Coulomb): 1 cal = 4.187 J.

Farad = Coulomb/Volts.

Faraday's constant: 96 487 (10^4) Coulomb/mol monovalent ion.

Molar gas constant (R): 8.31 J mol^{-1} per degree Kelvin (K).

Specific heat capacity of the human body: 3.47 kJ kg^{-1} $^oC^{-1}$.

Energy transfer by evaporation of 1 kg of water at the usual skin temperature: 2436 kJ.

Pressure (Pascal = Pa = N m^{-2}): 1 mmHg = 1 Torr = 133.3 Pa.

Surface tension of body warm water: 0.07 N m^{-1}.

Temperature conversion between degrees of Fahrenheit (^oF) and degrees of Celsius (^oC): (^oF) = 9/5 (^oC) + 32.

Calculated Partial Pressures

The partial pressures of respiratory gasses are calculated in the alveoli and in the surrounding air of a healthy person, resting at sea level (101.3 kPa = 760 mmHg or Torr = 1 atmosphere).

The water vapour tension in a fluid (air or liquid) of the temperature 310 K (37^{o}C) is *6.27 kPa* or *47 mmHg.* At 293 K (20^{o}C) the tension is 2.4 kPa or 18 mmHg. The alveolar gas fractions are: F_{AO2} = 0.15, and F_{ACO2} = 0.056. The composition of atmospheric air is: F_{IO2} = 0.2093 and F_{ICO2} = 0.0003.

$P_{O2} = F_{O2}$ (101.3 - 6.27) kPa.

P_{AO2} = 13.3 kPa (100 mmHg); P_{aO2} = 12.7 kPa (95 mmHg); P_{vO2} = 6 kPa (45 mmHg).

P_{ACO2} = 5.3 kPa (40 mmHg) ; P_{aCO2} = 5.3 kPa (40 mmHg); P_{vCO2} = 6.1 kPa (46 mmHg).

Return to content

Go to Content

Selected books

33. Paulev, P.-E. Problemer og fortolkninger i medicinsk fysiologi. *Lærebog*. Pp. 1-200. FADL's forlag, København, Odense, Århus, 1976. Supplement. Pp.1-67.

36. Paulev, P.-E. Medicinsk fysiologi. Spørgsmål og svar. Bd. 2 og 3. *Lærebog. Pp. 201-1169. (Problemer og fortolkning i medicinsk fysiologi udgør bd. 1).* FADL's forlag, København, Odense, Aarhus, 1977.

132. Paulev, P.-E. *Questions and Answers in Medical Physiology.* W.B. Saunders Co Ltd, London, 1996. ISBN 0-7020-2043-5.

140. Paulev, P.-E. *Medical physiology and Pathophysiology.* Internet. 1999-2000. ISBN 87-984078-0-5

Return to home page

Go to content

Selected curriculum vitae

Postgraduate teaching:

- Organiser of "Second Annual Scientific Meeting", EUBS, Copenhagen, British Bahamas, Stockholm, and San Diego.
- Organiser of *Hyperbaric Physiology And Diving Medicine, Copenhagen.*
- Administrator or lecturer at conferences concerning circulatory- respiratory- exercise- physiology at University of Lund, State University of New York, McGuill University, Yamagata and Chiba Universities, Japan and at Polish Academy of Sciences.
- Administrator or lecturer at conferences concerning sport medicine, exercise medicine, diving medicine, hypertension and other risk factors, clinical physiology in Denmark.

Present research:

- Neural and endocrine factors controlling respiration (with Polish Academy of Sciences)
- The carotid body in the control of cardiopulmonary function during exercise
- The diving and survival reflex

Education:

- 1962 : MD at University of Aarhus, Denmark
- 1962-1966: Junior scientist at Institute of Physiology, University of Aarhus
- 1966-1971: Scientist at Department of Medical Physiology, University of Copenhagen
- 1971 - : Professor (Afdelingsleder, lektor), University of Copenhagen

Research visits at foreign universities:

- **USA** and **Canada:** Duke University - University of Rochester - State University of New York at Buffalo - University of Philadelphia - McGuill University, Canada - University of California (UCLA), Los Angeles.
- **Germany:** Freie Universitet, Berlin.
- **Sweden:** Karolinska Institutet, Stockholm, and Fysiologisk Institut, Lund
- **England:** University of Oxford
- **Japan:** Chiba and Yamagata Universities.
- **Poland:** Polish Academy of Sciences, Warszawa

Scientific societies:

- European Undersea Biomedical Society (EUBS): President London 1972, Stockholm, 1973. Past president, Marseilles, 1974.
- Member of the EU commission *Medical Group, Mines Safety and Health Commission*, Luxembourg.

Editorial work:

- European editor of *Undersea Biomedical Research*
- Referee for *Acta Physiologica Scandinavica, Pflügers Archive - Journal of Applied Physiology - Canadian Journal of Zoology - Scandinavian Journal of Clinical Laboratoy Investigation - Undersea Biomedical Research - Scandinavian Journal of Pharmacology & Toxicology - European Journal of Applied Physiology & Occupational Physiology - Japanese Journal of Physiology - Respiratory Physiology.*
- Evaluation of dissertations at University of Rochester, University of Lund, Sweden, Chiba University, Tokyo, Universities of Aarhus and Odense, Denmark.

Postgraduate courses

- Physical chemistry, mathematics and thermodynamics.
- Electronic data
- University pedagogic
- Isotope technique
- Indicator technique
- Myocardial contractility
- Mathematics and Thermodynamics
- Medical research methodology
- Statistics
- Work toxicology (Helsinki)
- Computer technique and programming RECKU/UNI-C, University of Copenhagen

Pregraduate teaching:

Organisor of the *pregraduate teaching administration*, for medical physiology (1971-1973).

- Censor in *medical physics*, University of Copenhagen
- Censor in *fluid-mechanics and medico-technique*, The Technical University of Denmark (DTU), Copenhagen.

Guest Scientists
Professor J.Honda, Chiba University, Japan
Professor M. Pokorski, Polish Academy of Sciences.Warszawa

Return to home page
Go to Content

Go to Content

Selected bibliography

Of the 134 publications in the Bibliography:
90 are printed in international, highly recommended journals with referees,
24 are of pathophysiological or of scientific, post-graduate character in journals with one editor and referee,
7 are pathophysiological for the public in journals without referee,
12 are books (6 textbooks, 3 popular science physiology books, and 3 monographies.

1. Paulev, P.-E. Pressure tests on submariners, divers and hospital personel. *J. of the Royal Naval Medical Service.* Spring 1965.

2. Paulev, P.-E. Decompression sickness following repeated breath-hold dives. *J.Appl.Physiol.,* 20: 1028-1031, 1965.

3. Paulev, P.-E. Decompression sickness following repeated breath-hold dives. In: *Physiology of Breath-hold Diving and the Ama of Japan*, edited by: H.Rahn and T. Yokoyama, Washington D.C.: National Academy of Sciences -National Research Council, Publ. 1341, p.221-226. 1965.

5. Paulev, P.-E. & Inge Dahn Limb blood flow during breath-holding. *Acta physiol. scand.*, Suppl. 277, 68: 37, 1966.

7. Paulev, P.-E. & N. Næraa Hypoxia and carbon dioxide retention following breath-hold diving. *J.Appl.Physiol.*, 22: 436- 440, 1967.

8. Paulev, P.-E. Nitrogen tissue tensions following repeated breath- hold dives. *J. Appl. Physiol*., 22: 714-718, 1967.

9. Paulev, P.-E. Cardiac rhythm during breath-holding and water immersion in man. *Acta Physiol. Scand.*, 73: 139-150, 1968.

11. Paulev, P.-E. & H.Wetterquist Cardiac output during voluntary apnea in man. *Scand. J. Clin. Lab. Invest.*, 22:115-23, 1968.

12. Paulev, P.-E. Disputats, Københavns Universitet, 1969. Respiratory and Cardiovascular Effects of Breath-Holding. *Acta Physiol. Scand.*, Suppl. 324, 1969.

17. Paulev, P.-E. Respiratory and cardiac responses to exercise in man. *J. Appl. Physiol.*, 30: 165-172, 1971

20. Paulev, P.-E. & H.G. Hansen The cardiac response to apnea and water immersion during exercise in man. *J. Appl. Physiol.*, 33 (2): 193-198, 1972.

23. Paulev, P.-E. & Else Stibolt Jensen Cardiac Rate and Ventilatory Volume Rate Reactions to a Muscle Contraction. *J. Appl Physiol.*, 34: 578-583, 1973.

34. Døssing, M., N. Gerdes & Paulev, P.-E. Local blood flow, pressure and resistance measured with strain gauge plethysmograph in limbs with constricting burns before and following treatment with incisional decompression. *Burns* 2:226- 237, 1976.

58. Nørregaard-Hansen, K.J. Bjerre-Knudsen, U. Brodthagen, R.Jordal & Paulev, P.-E. Muscle Cell Leakage Due to Long Distance Training. *Eur. J. Appl. Physiol*. 48:177-188, 1982.

61. Paulev, P.-E., Kroppen - sådan fungerer den. *Bogan* Pp. 1-148. København, 1983.

ISBN 87-87533-89-8.

63. Paulev, P.-E., Jordal, R. & Pedersen, N. Strandberg. Dermal excretion of iron in intensely training athletes. *Clin. Chem. Acta.* 127:19-27. 1983.

66. Paulev, P.-E. Exercise and risk factors for arteriosclerosis in 42 married couples followed over 4 years. *J.Chronic Diseases* Maj, 37(7):547-553, 1984.

68. Paulev, P.-E.,Jordal R., Kristensen O., Ladefoged J. Therapeutic effect of exercise on hypertension. *Europ. J.Appl. Physiol.* 53: 180-185, 1984.

76. Brodthagen U.A., Nørregaard Hansen K., Bjerre Knudsen J., Jordal R., Kristensen O., Paulev, P.-E. Red cell 2,3-DPG, ATP, and mean cell volume in highly trained athletes. Effect of long term submaximal exercise. *Europ. J. Appl. Physiol.* 53: 334-338, 1985.

78. Kruse P., Ladefoged J., Nielsen U., Paulev, P.-E., Sørensen J.P. Beta-blockade used in precision sports. The effect on pistol shooting performance. *J. Appl. Physiol.* 61(2):417-420, 1986.

86. Paulev, P.-E., Honda, Y., Sakakibara, Y., Morikawa, T., Tanaka, Y., Nakamura, W., Nakazono, Y., & Miyamoto, Y. Respiratory and cardiac response to dynamic exercise in man. *Jpn. J. Physiol.* 38:375-386, 1988.

88. Paulev, P.-E.,Honda, Y., Sakakibara, Y., Morikawa, T., Tanaka,Y. and Nakamura, W. Brady- and tachycardia in light of the Valsalva and the Mueller maneuver (Apnea). *Jpn. J. Physiol.* 38:507-517, 1988.

93. Ahn, B., Nishibayashi, Y., Okita, S., Masuda, A., Takaishi, S., Paulev, P.-E. and Honda, Y. Heart rate response to breath holding with supramaximal exercise. *Europ. J. Appl. Physiol. and Occup, Physiol.* 59: 146-151, 1989.

98. Ahn, B., Sakakibara, Y., Paulev, P.-E., Masuda, A., Nishibayashi, Y., Nakamura, W. and Honda, Y. Circulatory and respiratory responses to lower body negative pressure in man. *Jpn. J. Physiol.* 39: 919-930, 1989.

101. Paulev, P-E., Pokorski, M., Honda, Y., Morikawa, T., Sakakibara, Y. and Tanaka, Y. Cardiac output and heart rate in man during swimming while breath-holding. *Jpn. J.Physiol.* 40:117-125, 1990.

105. Mussell, M.J., Miyamoto, Y., Paulev, P.-E., Nakazono, Y. And Sugawara, T. A constant flux of CO2 injected into the airways mimics metabolic CO2 in exercise. *Jpn. J. Physiol.* 40 (6):877-891, 1990.

106. Paulev, P.-E., Pokorski, M., Honda, Y., Ahn, B., Masuda, A., Kobayashi, T., Nishibayashi, Y., Sakakibara, T., Tanaka, Y. and Nakamura, W. Facial cold receptors and the survival reflex diving bradycardia in man. *Jpn. J. Physiol.* 40 (5):701-712, 1990.

107. Pokorski, M., Masuda, A., Paulev, P.-E., Sakakibara, Y., Ahn, B., Takaishi, S., Nishibayashi, Y. and Honda, Y. Ventilatory and cardiovascular responses to hypoxic and hyperoxic static handgrip exercise in man. *Resp. Physiol.* 81:189-202, 1990.

114. Paulev, P.-E., Pokorski, M., Masuda, Y.Sakakibara, Y., Honda, Y. Cardiorespiratory reactions to static, isometric exercise in Man. *Jpn. J. Physiol.* 41 (5): 785-795, 1991.

130. Paulev, Poul-Erik, Janos Porszasz, William W. Stringer, and Karlman Wasserman. Lactate accumulation linked to mitochondrial purine degradation in muscle cells during exercise. *IUPS Meetings. Session 353, Workshop 97, 107-108,* Glasgow, 1.-6- August, 1993.

131. Pokorski, M., P.-E. Paulev, M. Szereda-Przestaszewska. Endogenous benzodiazepine system and regulation of respiration in the cat. *Respiration Physiology* 0:1-13, 1994.

132. Paulev, P.-E. Questions and Answers in Medical Physiology. *W.B. Saunders Co Ltd*, London, 1996. ISBN 0-7020-2043-5.

133. Pokorski, M., P.-E. Paulev, and L. Faff. Cannabinoids in the regulation of respiration. XXI Congress of the Polish Physiological Society. *J Physiol and Pharmacology*. Suppl. 1, vol 50: 121, 1999.

134. Paulev, Poul-Erik. Medical Physiology and Pathophysiology. Essentials and Clinical Problems. Online book. *Copenhagen Medical Publishers*, København 1999-2000 ISBN 87-984078-05

Return to homepage

Go to content

Chapter 1.
Cells And Action Potentials

Study Objectives

- To *define* activity, activity coefficient, concentration, diffusion, flux, molality, molarity, normality, osmolality, osmosis, pressure and radioactivity.

- To *describe* the diffusion potential, the equilibrium potential, facilitated diffusion, the Donnan-effect, the resting membrane potential, the action potential, and membrane transport including that of glucose.
- To *calculate* the equilibrium potential, osmotic pressure and other variables from relevant variables given.
- To *draw* the action potential curve.
- To *explain* the colloid osmotic pressure of plasma, hyponatraemia, overbreathing, radioactive decay and the elimination rate constant, the nerve conduction, signal transduction, the function of the action potential, the Na^+-K^+ -pump and the transport proteins.
- To *use* the ideal gas law and the above concepts in problem solving and case histories.

Principles

- *The ideal gas law relates the pressure P, the volume V, and the number of mol of the gas n, to the Kelvin temperature T: ($P \times V = n*R*T$). At standard temperature, pressure, dry (STPD) the volume occupied by 1 mol of any ideal gas is 22.4 litre (l). - By analogy in an ideal solution, one mol of solute will exert an osmotic pressure of one atmosphere, if it is dissolved in 22.4 l of water. Van´t Hoff´s law for ideal solutions is an equivalent to the ideal gas law.*

Definitions

Absolute temperature (T) is measured in Kelvin or K from the *absolute zero point* (-273 ^oC).

Activity is a corrected concentration measure of any species (ie, the free concentration multiplied by the activity coefficient). The activity is often measured with ion selective electrodes. - In diluted solutions - below 10^{-3} molar - there is no correction for uncharged molecules.

Activity coefficient is the fraction of the free ions, which is chemically active. - For sodium the activity coefficient is 0.75 in many biological solutions.

Action Potential (AP) is an *all-or-none electrical signal*, which appears as a positive wave when recording internally. The AP is conducted with the same shape and size along the whole length of a muscle cell or a nerve fibre.

Amphipathic molecules contain both a polar and a non-polar region.

The **membrane potential difference** is defined as the potential inside the cell minus the potential outside the cell – the difference is transiently reversed during an AP.

Becquerel (1 Bq) is the preferred unit for disintegration rate of radioactive decay, namely *one radioactive disintegration per s*. Disintegration rates were previously expressed in Curies (Ci), in honour of Marie Curie, who discovered radium.

Concentration (**C** or brackets around a substance $[Na^+]$) is the mass or moles (mol) per unit of fluid volume.

Diffusion is a net transport of atoms or molecules caused by their random thermal motion in an attempt to equalise concentration differences (DC).

The **Donnan effect** is the *extra osmotic pressure of protein solutions* caused by impermeable protein molecules resulting in uneven distribution of small, permeant cations and anions (in blood plasma).

The elimination rate constant (k) is the fraction of the total amount of a given substance in the distribution volume of the body eliminated per time unit. Elimination with a constant rate is exponential. The *half-life* for a substance eliminated exponentially is equal to 0.693 k^{-1}. This is just a simple mathematical deduction.

Flux (*J*) is the amount of a substance transported along a pressure gradient through an area unit (A is measured in m^2) of a membrane in moles per second (s). *Convective flux* is the net amount of molecules transported through A per time unit (mol s^{-1} m^{-2}), caused by a pressure gradient and *fluid* (liquid or air) volume transport.

An **ideal semipermeable membrane** is permeable to water only, but impermeable to all solutes. Most real *semipermeable membranes* are permeable to water and to low molecular substances (crystalloids), but not to macromolecular substances (colloids such as proteins).

Molar concentration (molarity) is the number of moles of a substance totally dissolved per litre (l) of solution - often given in mmol per l or mM. One mol of a substance is the amount of that substance containing Avogadros number, $6.022 * 10^{23}$ molecules per mol.

Molality is the number of mol totally dissolved substance per kg of solvent, frequently water. One *equivalent* is the molar mass of all the ions that contain $6.022 * 10^{23}$ single charges or *valences* when fully dissociated.

Motility is the reciprocal resistance of a molecule towards movements.

Normality of a solution is the number of equivalents per l (Eq l^{-1}).

Osmolality is a measure of the osmotic active particles in one kg of water. Plasma-osmolality is given in Osmol per kg of water. Water occupies 93% of plasma in healthy persons.

Osmolarity is the number of osmotically active particles dissolved in a litre of solution.

A **permeable membrane** allows the passage of all dissolved substances and the *solvent* (mainly water).

A **selectively permeable membrane** is permeable to a particular compound (sucrose, Na^+, Ca^{2+}, anions only or to cations only).

- **Pressure** (***P***) is measured as force per area unit - that is in Newton per square m or *Pascal.*
- **Osmosis** is transport of solvent molecules (mainly water) through a semipermeable membrane. **Osmotic pressure** (π) is the hydrostatic pressure, that must be applied to the side of a rigid ideal semipermeable membrane with higher solute concentration in order to stop the water flux, so that the net water flux is zero.
- **Radioactivity.** Some nuclei are *unstable* or *radioactive*, because they release certain particles such as helium nuclei or electrons. Other radioactive substances emmanate gamma-rays with an extremely short wavelength. All radioactive decay processes follow an exponential pattern. If N_o is the initial number of unstable nuclei, the number of nuclei remaining after a time t (N) is given by $N = N_o * e^{-kt}$, where k is a constant characteristic of each nuclide, called the *disintegration constant*. This is the *law of radioactive decay* (Eq. 1-6).
- **Volume** (***V***) in litres (l). *Standard temperature, pressure, dry* (STPD) is an abbreviation for a volume at standard temperature of 273 K, standard pressure of 101.3 kPa or 760 mmHg, and dry air.

Essentials

Three topics are treated here: 1. Transport through membranes, 2. Resting membrane potentials, and 3. Action potentials.

1. Transport through membranes

Membrane transport refers to solute and solvent transfer across both cell membranes, epithelial and capillary membranes.

1a. Membranes

Biological membranes are composed of phospholipids stabilised by hydrophobic interactions into bilayers (Fig. 1-1). The membranes contain approximately 50% lipids and 50% proteins.

Fig. 1-1: Model of a cell membrane built by phospholipids separating receptors, channels, proteins (Pr^-), glycoproteins (receptors, antigens etc) and glycolipids.

Phospholipids are *amphipathic*. One region is polar consisting of charged choline, etanolamine and phosphate head-groups (bullets in Fig. 1-1). The other region is non-polar, consisting of tails of fatty acyl chains (Fig. 1-1). The non-polar regions tend to avoid contact with water by self-association. Any other arrangement with disruption of *hydrogen bonds* (between O and H atoms) of water has a high energy cost. Integral proteins are deeply imbedded in the membrane, and the model shows 3 protein molecules spanning the membrane (ie, transmembrane proteins). Surface proteins are not shown. The proteins carry receptors to which transmitter substance bind. Carbohydrate chains are shown forming glycolipids with antigenic or receptor function, or glycoproteins with other receptor functions.

The *molar concentration* (molarity) is the number of mol totally dissolved substance per litre (l) of solution - often given in mmol per l or mM. Ions in plasma are conventionally measured in mM with flame-photometry by the ability to absorb monochromatic light. Na^+ is mainly dissolved in the water phase of plasma (93%). The $[Na^+]$ in plasma is therefore smaller than the *Na^+-activity* which is recorded in the water phase alone with ion selective electrodes. For conventional reasons ion selective electrodes are calibrated to match the well-known flame photometry values, although the *activity* is the biological important variable. Molality, normality, and flux are described above in *Definitions.*

Mechanical, electrical, thermal, or *gravitational forces* drive *migration* of molecules. These forces move the molecules passively in a direction determined by the vector of the force.

Diffusive flux (J^{dif}) is the movement of molecules by diffusion caused by a concentration gradient (dC/dx) in the direction x. The *diffusion coefficient* (D) is a proportionality constant that relates flux to the concentration gradient (dC/dx). *Einstein* defined D as (k×T×B), where T is absolute temperature and B is *motility* of molecules. Motility is the reciprocal resistance towards movements (velocity/N or $m*s^{-1}*N^{-1}$). The concept (k*T) is the thermal, molecular energy. A molecule diffuses from higher to lower concentration that is *down its concentration gradient.* Accordingly, dC/dx has a negative slope, when molecules diffuse in the direction x. Then, it is easy to calculate the diffusive flux (per m^2) according to Eq. 1-2. This relationship was first recognized as early as in 1855 by the anatomist and physiologist Fick, and it has since been named after him: *Fick's first law* of diffusion. The flux by simple passive diffusion is directly proportional to the concentration of dissolved molecules (Fig. 1-2).

Fig. 1-2: Two types of passive molecular transport: Simple diffusion and the much larger facilitated diffusion. C is concentration.

Einstein's relation states that for average molecules in biological media, the mean displacement squared, $(dx)^2$, is equal to 2 multiplied by D and by the time (t) elapsed, since the molecules started to diffuse (see Eq. 1-3). For molecules with $D = 10^{-9}\ m^2\ s^{-1}$, the time required to diffuse 1 mm is 0.5 milli-second (ms). To diffuse 10 and 100 mm, the time required increases 100-fold each time: 50 ms and 5000 ms.

Facilitated diffusion takes place through *transport proteins* not linked directly to metabolic energy processes (Fig. 1-2). Facilitated diffusion shows saturation or Michaelis-Menten kinetics, because the number of transport proteins is limited. The *saturation kinetics* is different from the energy limitation in primary active transport. *Amino acids, glucose, galactose and other monosaccharides* cross many cell membranes by facilitated diffusion.

An ideal *semipermeable membrane* is permeable to water only, but impermeable to all solutes. Most real *semipermeable membranes* are permeable to water and to low molecular substances (crystalloids), but not to macromolecular substances (colloids such as proteins).

+ 2+ - -

An ideal *ion-selective membrane* is permeable to anions only or to cations only (Na , Ca , or to Cl and NO_3). Ion-selective membranes are used in ion-selective electrodes to measure the activity of selective ions in plasma water (Box 1-1).

Box 1-1: Ion-selective membranes	
Selectivity:	**Either anions or cations**
Specificity:	**Cation membranes distinguish between Na^+, Ca^{2+}, and K^+.**
	Anion membranes distinguish between Cl^- and NO_3^-.

The relation between the potential difference measured with an ion-selective electrode and the activity (ionised or fully dissociated form) for a certain ion (eg, K^+) is given by the Nernst equation (Eq. 1-5).

1b. Osmosis and osmotic pressure

Osmosis is transport of solvent molecules (mainly water) through a semipermeable membrane. The water flows from a compartment of high water concentration (or low solute concentration) to one of low water concentration (or high solute concentration). The greater difference between the solute concentration of the two compartments, the more is water unevenly distributed between the two compartments. Water diffuses down its chemical potential gradient into the compartment with higher solute concentration, causing the chemical potential gradient to be reduced until solute equilibrium is reached.

Osmotic pressure is the hydrostatic pressure, that must be applied to the side of an ideal semipermeable membrane with higher solute concentration in order to stop the water flux, so that the net water flux is zero.

The *colligative properties of water* are strictly related to the solvent or water concentration alone. Water molecules are bound together by hydrogen bonds in clusters of several hundred molecules, forming a structure looking almost like crystals. Sites between the clusters, where the distance between water molecules are larger than elsewhere are called *bubble nuclei* because these sites seem to initiate formation of gas bubbles in decompression sickness. These sites are also likely locations for substances dissolved in water. With decreasing water concentration, the water vapour tension, and the freezing point is reduced, whereas the boiling point, and the osmotic pressure of the solution is increased as compared with pure water. The size of the osmotic pressure of a solution depends of the number of dissolved particles per volume unit.

The osmotic pressure (π) depends on the absolute temperature (T Kelvin or K) and on the number of dissolved particles per volume unit (N/V equal to the molar fraction).

This relationship was first recognized by *van't Hoff* and applies to ideal solutions only. Real physiological solutions, such as the cytosolic phase and extracellular fluid, differ from the ideal solutions, which are very dilute.

A correction factor called the *osmotic coefficient* (f) corrects for these differences in osmolality. For physiologic electrolytes it is 0.92 - 0.96, and for carbohydrates it is 1.01.

A solution has the *ideal osmolarity one*, when it contains ($6.022* 10^{23}$) osmotically active particles per l. Diluted solutions have an *osmolar concentration* or *osmolarity* (Osmol per l) numerically equal to the *sum total* molarity of all dissolved particles (mol per l). In biological solutions the molarity is different from osmolar concentration. The number of Osmol per l is: (f N/V). - The corrected van't Hoff law is developed in Eq. 1-4.

Osmolality is simply the number of mol per kg of water in the fluid frequently given as mOsmol kg^{-1}. Fully dissociated molecules have twice the osmolality of undissociated molecules. Plasma- osmolality can be approximated by the calculation expressed in Eq. 1-7. Plasma-osmolality is measured by freeze point depression or by boiling point

increase. The osmolality of the ICV is approximately 290 mOsmol kg^{-1}, which is simplified to 300 mOsmol kg^{-1} in Fig. 1-4. The osmolality of the extracellular fluid must be the same, since cell membranes are not rigid, so they cannot carry any essential pressure gradient. The total number of mOsmol in the ICV and ECV of a standard person is thus 8400 and 4200, respectively (Fig. 1-4).

The *colloid osmotic pressure* is equal in magnitude to the hydrostatic pressure, which must be applied at the luminal side of the capillary barrier, in order to stop net transport of water caused by uncharged colloids in the blood plasma. *Colloids* are mainly *plasma proteins*.

The *osmotic pressure* is equal in magnitude of a certain hydrostatic pressure. This pressure column must be applied to the solution to restore the free energy or *chemical potential* of its water to that of pure water. The tendency of water to pass a membrane depends on its chemical potential (ie, vapour pressure). The chemical potential of water decreases with solutes present and with decreasing temperature.

Uniformly distributed substances in diluted solutions behave like gas molecules at atmospheric pressure (atm). The osmotic pressure can be expressed as $P_{osmot} = C \times RT$, which is the equivalent of the ideal gas equation ($P \times V = nRT$). Here C is the concentration of dissolved solutes, and the derivation of the relationship is based on the chemical potential of water. R is the gas constant ($= 0.082\ l \times atm \times Osmol^{-1} \times K^{-1}$). At *standard temperature, pressure, dry* (STPD) the volume occupied by 1 mol of any ideal gas is *22.4 l.* Thus, STPD is an abbreviation for a volume at standard temperature of 273 K, standard pressure of 101.3 kPa or 760 mmHg, and dry air. In an ideal solution, one mol of solute will by analogy be dissolved in 22.4 l of water, and will exert an osmotic pressure of one atmosphere.

In biological solutions at 310 K, such as an ultrafiltrate of plasma (interstitial fluid, ISF), with an osmolality of 0.300 Osmol per kg water, the osmotic pressure must be:

> $P_{osmot} = 0.3\ (Osmol\ kg^{-1}) \times 0.082\ (kg \times atm \times Osmol^{-1} \times K^{-1}) \times 310\ (K) = 7.63$ atm (=773 kPa) or the pressure exerted by a column of water *76 m high*.

Only *net gradients* across endothelial and plasma membranes are important, and they depend upon *protein concentration gradients*. This is because all the electrolytes (crystalloids) have diffused to equilibrium across the capillary endothelial membrane, whereas proteins (colloids) cannot.

The average colloid osmotic pressure (π_{coll}) of plasma is approximately *3.6 kPa* (27 mmHg). The dissolved proteins have a molality of 1 mmol per kg water, and a net average of 17 negative charges per molecule (ie, 1 mmol kg^{-1} or 17 mEq kg^{-1}). Milli-Equivalents abbreviates *mEq.* The proteins are directly responsible for *2.4 kPa* (18 mmHg). The remaining *1.2 kPa* (9 mmHg) of the colloid osmotic pressure is due to the unequal distribution of permeable ions, the *Gibbs-Donnan law* or the *Donnan effect* (see below).

1c. The Donnan effect across the capillary membrane

Let us consider a closed system with two compartments separated by a rigid membrane that is permeable to water and to small ions. In the presence of solutions with different NaCl concentrations, water and ions permeate rapidly in both directions across the membrane. Electrical neutrality in each of the two solutions requires that the simultaneous movement of Cl^- match any net movement of Na^+, so the equivalents of anions and cations are the same. The number of times the two ions collide with one side of the membrane is *proportional* to the product of their concentrations: $[Na^+] \times [Cl^-]$. At equilibrium the fluxes of NaCl in each direction are identical, and ultimately the concentrations are the same all over.

Let us now add protein to one compartment ($compartment_p$ modelling streaming plasma), which is separated by a membrane (modelling the capillary endothelial membrane) from the other compartment ($compartment_{ISF}$ modelling interstitial or tissue fluid). The model still contains only Na^+ and Cl^- that can cross the membrane. At equilibrium the product of concentrations of the two ions on either side of the membrane must be equal, and the transmembrane

potential corresponds to the equilibrium potential of the small permeant ions (the Nernst equation, Eq 1- 5). Transforming the Nernst equation reveals that the concentration product of any pair of diffusible ions is identical on either side of the membrane at equilibrium:

$$[Na^+]_{ISF} \times [Cl^-]_{ISF} = [Na^+]_p \times [Cl^-]_p.$$

On the plasma side, which contains *impermeant anions* (negative proteinates), the concentration of permeant anion (Cl^- is the model) must always be less than on the interstitial fluid side. The concentration of permeant cation (Na^+ is the model) must always be greater than in the ISF.

The sum of permeant anion and cation concentrations in plasma is always greater than the sum of the same anion and cation concentrations in ISF:

$$[Na^+]_{ISF} = [Cl^-]_{ISF} \; ; \; [Na^+]_p + [Cl^-]_p > [Na^+]_{ISF} + [Cl^-]_{ISF}$$

This is a simple mathematical argument: The sum of unequal sides of a rectangle is greater than the sum of the sides of the square with the same area. It explains why the osmotic pressure in plasma exceeds that of the tissue fluid. This is not due to the plasma proteins alone, but is also due to the *higher* concentration of small, permeant ions in the plasma.

The *Donnan effect* is the *extra osmotic pressure* of protein solutions caused by the uneven distribution of small, permeable cations and anions. The Donnan effect causes a 5% and 10% concentration difference across the capillary barrier between the plasma and ultrafiltrate concentrations of monovalent and divalent ions, respectively.

In the above equations Na^+ and Cl^- are model ions for all the cations and anions. In our body other anions and cations are present, the Na^+ and Cl^- concentrations are not alike, and the *capillary membrane* is far from rigid. Nevertheless, the Donnan equilibrium implies an accumulation of charges on the side with the negatively charged proteins. This potential difference across the *capillary membrane* is termed the *Donnan potential* **–** a potential, which is developed across cell membranes without a sodium-potassium pump.

The Donnan factor at the capillary membrane is 0.95, so a plasma- $[Na^{+}]$ of 150 mmol measured in each kg of plasma-water is in equilibrium with 142 mM in the interstitial fluid between the cells.

Strictly speaking, there is no such thing as a cell with rigid cell walls in the animal kingdom, so the Donnan effect is theoretically unfounded in animal cell membranes – except at the capillary barrier.

1d. The Na^+-K^+-pump.

The Na^+-K^+-pump is a *transmembrane protein* in the cell membrane (Fig. 1-3). The pump contains a channel, which consists of two double subunits: 2 a - and 2 b - subunits. The catalytic subunit (a) is an Na^+-K^+-activated ATPase of 112 000 Dalton, and the b -*subunit* is a glycoprotein of 35 000 D.

Fig. 1-3: The Na^+-K^+-pump consists of 2 α - and 2 β - subunits (Pi = Phosphate).

The pump is a *primary active transporter*, because it uses the cellular energy of the terminal phosphate bond of ATP (Fig. 1-3). The Na^+-K^+-pump transports 3 Na^+ out of the cell and 2 K^+ into the cell for each ATP hydrolysed. This is a net movement of positive ions out of the cell, and therefore called an *electrogenic transport*. The constant influx of Na^+ is shown as well as the leakage of K^+- and Cl^-. In a steady state the net transport of each ion across the resting membrane is zero.

The Na^+-K^+-pump is located in the *basolateral* exit-membrane of the epithelial cell (Fig. 1-3). The primary active ion-transport provides metabolic energy for the secondary water absorption through the luminal membrane. Hereby, the active pump in the exit-membrane drives the luminal transport across the entry membrane. This transport of NaCl

and water is surprisingly nearly isotonic. The bulk flow can take place against a large osmotic gradient, and increases in diluted solutions. The *entry membrane* is often highly permeable to water.

The Na^+-K^+-pump builds up a high cellular electrochemical gradient for K^+ and indirectly for Cl^- (Fig. 1-3). The water outflux is coupled to the outward transport of K^+ and Cl^-. The interstitial fluid receives ions and glucose, causing its osmolarity to increase. The osmotic force causes water to enter the interstitial fluid via the cell membranes and the gaps between the cells (tight junctions). This in turn causes the hydrostatic pressure in the interstitial fluid to rise. The hydrostatic force transfers the *bulk of water, ions and molecules* through the thin-walled, tubular capillaries to the blood. When excess of water (solvent) passes through tight junctions, they lose part of their *tightness* and the solvent water drags many Na^+/Cl^--ions out (*solvent drag).*

In a healthy standard person nutrients and oxygen are transported into the cell interior from the extracellular fluid through the cell membrane (Fig. 1-4). The intracellular fluid volume, ICV, is 26-28 l. The *extracellular fluid* volume (ie, ECV of 14 l) consists of the circulating blood plasma (3-3.5 l) and the *interstitial fluid* (ISF) with a volume of 10.5-11 l in the spaces between cells. Total body water (here 42 l) accounts for 60% of body weight. The body is cleared of 24 mol of carbon dioxide by the lungs in 24 hours and of other substances by the kidneys (Fig. 1-4). A yellow tube on the diagram symbolises the gastrointestinal channel, where nutrient molecules are absorbed and waste products are eliminated through the liver bile.

Fig. 1-4: Salt- and water- transport through a cell membrane separating the intracellular and extracellular compartment.

The diagram also shows the Na^+-K^+-pump together with leakage of K^+, Cl- and water (Fig. 1-4). The net transport of each substance is zero in the steady state.

The Na^+-K^+-pump is responsible for maintaining the high intracellular [K^+] and the low intracellular [Na^+]. The energy of the terminal phosphate bond of ATP is used to actively extrude Na^+ and pump K^+ into the cell.

Jens Christian Skou of Denmark initiated the study of the Na^+-K^+-pump already in the 1950ties and received the Nobel Prize for his contribution to basic chemistry and physiology in 1997.

The membrane also contains many K^+- and Cl^--channels, through which the two ions *leak* through the cell membrane.

Intestinal and kidney tubule cells transport substrates, such as glucose and amino acids, in a *substrate-Na^+ cotransport* in the luminal membrane, linked to the Na^+-K^+-pump of the basolateral membrane. This is called a *secondary* active transport of substrate. Such a transport is powered by an actively established gradient (ie, the Na^+-gradient)

The many ion-transporting ATPases form classes or families showing amino acid sequence homology.

1e. Glucose transport proteins (GLUTs) and insulin receptors

A family of homologous carrier proteins that are coded by distinct genes mediates glucose-transport. The transport proteins (GLUTs) show a marked tissue-specificity, which reflects differing transport needs of various tissues. This is facilitated transport (Fig. 1-2).

Five human *glucose-transporters* are cloned and identified (GLUT 1-5). The GLUT 1 resides in placenta, brain, perineural sheaths, red cells, adipose and muscle tissues. GLUT 2 is found in the liver, pancreatic b-cells, proximal renal tubule cells, and the basolateral membranes of small intestinal cells. GLUT 3 is ubiquitously distributed, found predominantly in the brain and in lower concentrations in fat, kidney, liver and muscle tissues. GLUT 4 is confined to tissues with insulin-responsive glucose uptake (muscle, heart and fat stores). GLUT 5 is found in the luminal membrane of small intestinal cells, and also in brain, muscle and adipose tissues. Some of these transporters also allow fructose and galactose to pass.

Fig. 1-5: Insulin, insulin receptors, with D-glucose transport proteins (GLUTs) and their translocation.

In *adipocytes and muscle cells,* glucose transport is profoundly influenced by insulin (Fig. 1-5).

1. As insulin binds to its large T-shaped *insulin receptor*, many intracellular vesicles are stimulated.
2. They contain a high number of membrane penetrating GLUTs, which translocate from the intracellular pool towards the cell membrane.
3. When these vesicles - rich in glucose transporters - fuse with the cell membrane, the number of glucose transporters increases substantially, thereby increasing D-glucose uptake up to ten times.
4. As the insulin-receptor complex dissociates, the GLUTs translocate again to the intracellular stores in the vesicles (Fig. 1-5).
5. The glucose transport ceases.

The *insulin receptor* is a glycoprotein found in the cell membranes. The T-shaped receptor protruding from the cell membrane contains 1370 amino acids forming two a- and two b-sub-units. The two a-subunits are entirely extracellular, whereas the two b-subunits span the membrane. Insulin binding on a-subunits stimulates a *protein kinase* on the intracellular part of the receptor to phosphorylate tyrosine residues on the b-subunit and on endogenous proteins. The exact molecular mechanism linking the receptor kinase activity to changes in cellular enzyme activity and transport processes remain uncertain; but it is shown that the *kinase activity* is essential for signal transduction.

2. Resting membrane potentials

A *membrane potential difference* is conventionally defined as the intracellular (j^{i}) minus the extracellular (j^{o}) electrical potential. The ion concentrations (activities) inside the cell and outside the cell are called C^{i} and C^{o}, respectively.

When a microelectrode penetrates a membrane, it records a negative potential with respect to an external reference electrode caused by different permeability of anions and cations. This is the *resting membrane potential* (RMP values in Box 1-2). The resting membrane potential is an essential mechanism in storing and processing information in neurons and other cells.

Concentration gradients across cell membranes are present for several ions, whereby they diffuse from one location to another. The ion with the highest permeability and concentration gradient, such as the potassium ion, establishes a membrane potential. This potential enhances or inhibits the flux of other ions and the ultimate situation is an electroneutral flux.

The chloride ion diffuses extremely rapidly, but otherwise positive ions (cations) diffuse more rapidly than negative ions (anions) through a membrane. However, as an example the permeability for Na^{+} is low (0.2 nm $*s^{-1}$) compared to that of K^{+} (5-40 nm/s) in neurons.

The *equilibrium potential* for a certain diffusible ion across a membrane that has a concentration gradient over the membrane, is precisely that *membrane potential difference*, which opposes the flux due to the concentration gradient so that the net transport of the ion concerned is zero. The equilibrium potential is simply calculated by balancing the diffusion potential of the ion with the opposing electrical force. The electrical force working on the ion is proportional to the electromotive force of the field. As a consequence, the total driving force on the ion and its *diffusion flux* is *zero*. Nernst introduced this equilibrium potential shortly before year 1900. The *Nernst equation* for the equilibrium potential of Na^{+} across a selective permeable membrane at 310 K is found by insertion of the ion activities (concentrations) inside and outside the cell (Eq. 1-5).

Box 1-2: Resting membrane potentials (RMP) and equilibrium potentials (V_{Eq}) in different cells.

	RMP (mV)	**V_{Eq} (mV)**

Resting skeletal muscle	- 80	- 80 for Cl^-
and myocardial cells	- 90	- 94 for K^+
		+ 60 for Na^+
		+ 130 for Ca^{2+}
Smooth muscle cells	- 40 to –60 (oscillations)	Variable
Neurons	- 70	As above

In skeletal muscle cells the resting membrane potential is -80 mV, and the *equilibrium potential* of Na^+ is +60 mV. Hence, the electrical driving force is: (-80 - (+60)) = -140 mVolts. Accordingly, there is a net passive influx of Na^+ into these cells down an electrochemical gradient (Box 1-2). The net influx is small because the resting Na^+ penetration is almost exactly balanced by active extrusion.

The resting membrane potential (RMP) is calculated from the *Millman equation* (Eq. 1-8). The RMP is mainly a diffusion potential (see above).

Neurons typically have four structures: The cell body, dendrites, axon and axon terminals (Fig. 1-6). Dendrites are elaborate branching processes that arise from the cell body, and they are pathways for incoming signals from other neurons to the cell body. Integration of incoming signals occurs mainly in the *axon hillock*. This is the part of the cell body, which gives rise to an elongated tube called an axon, a fibre that can be up to 1.2 m long.

Fig. 1-6: The neuron with cell body, dendrites, axon and axon terminals.

Near its termination each axon divides into fine branches, each of which ends in an *axon terminal* (ie, synaptic button or Bolton terminal). The axon terminals contain mitochondria and synaptic vesicles filled with neurotransmitter. These presynaptic structures are the sites where electrical signals are converted into chemical messages for transmission to nearby neurons. Unipolar neurons only have a single major process extending from the cell body. Bipolar and multipolar neurons have two or more major processes arising from the cell body. Most neurons have only one axon, a few more than one and some neurons function without an axon. Their location, structure and functional properties (Box 3-1) classify neurons. Communication from an axon to a dendrite is called *axodendritic*, from a dendrite to another is termed *dendro-dendritic*, from a dendrite to an axon is called *dendro-axonal*, from a dendrite to the soma is called *dendro-somatic*, and between two axons is referred to as *axo-axonal*.

Neuronal membranes are composed of lipid bilayers stabilised by hydrophobic interactions, and thus function as barriers to free diffusion for water-soluble molecules. The ability of the neuronal membrane to control the movement and concentration of charged particles generates ion gradients with a charge difference across the membrane. The potential difference across the resting membrane is called the *resting membrane potential* (RMP).

Ion channels and gates are classified according to the gating stimulus to which they respond. Voltage-gated ion channels are located along the axon of a neuron and responsible for the action potential. These ion channels are sensitive to local anaesthetics. Voltage-gated Na^+- K^+- and Ca^{2+}-channels contain membrane spanning helices – often with amino acid sequence homology. Action potentials in cardiac muscle cells have a plateau phase, where Ca^{2+} enters the cytosol via slow Ca^{2+}-channels. This Ca^{2+}-entry plays an important role in excitation-contraction coupling.

Ligand-gated ion channels are responsive to particular neurotransmitters. Ligand-gated ion channels open in response to substances such as acetylcholine. These channels are permeable to small cations (often unselectively: Na^+- K^+- NH_4^+- and Ca^{2+}). These channels are involved in generation of the postsynaptic potential and the endplate potential.

NaCl is found in high concentration outside the neurons, whereas [K^+] is high inside the cell. These ion gradients

maintain a constant leakage of NaCl into the cell, and a leakage of K^+ out. The gradients are maintained by the Na^+-K^+-pump, which is thus controlling the resting membrane potential. *Cl^--ions* distributes passively across most neuronal membranes and contributes little to the resting membrane potential, but they are important for the modulation of incoming signals. At rest, many K^+ -channels are open and K^+ moves down its concentration gradient out of the cell, whereby the inside becomes negatively charged (until it is difficult for K^+ to leave the cell, and the K^+-outflux slows down). The RMP approaches the equilibrium potential for K^+.

3. The action potential

Neurons can carry electrical signals along their whole length without any loss of signal strength. This electrical signal is an all-or-none phenomenon, termed the *action potential***.** The incoming signals to dendrites and cell bodies consist of small, graded changes (ie, small *synaptic potentials*) in the resting membrane potential caused by the actions of neurotransmitters and modulators. *Synaptic potentials* are spatially and temporally summated in the axon hillock of the cell body. The synaptic potential is graded according to the stimulus and shows decrement conduction in that its size decreases with increasing distance (wave length several mm). The local synaptic potential cannot in itself initiate an action potential. When the strength of the summated synaptic potentials is sufficient to reduce the resting membrane potential at the axon hillock below the *threshold* it opens Na^+ -channels. These Na^+ -channels are voltage-gated, because the change in voltage opens or closes a gate over each pore. The Na^+ -channels is usually closed at conditions with a normal resting membrane potential. When the Na^+ -channels open and allow Na^+ to flow into the cell down its concentration gradient, the influx itself depolarizes the neuron further, whereby more voltage-gated Na^+ -channels open. A propagating action potential (approaches the equilibrium potential for Na^+) in the axon is generated with a *positive* voltage overshoot simultaneously with the peak membrane conductance to Na^+ ($g_{Na}{}^+$ in Fig. 1-7). This is followed by the *repolarisation phase* (conductance for Na^+ goes down and up for K^+), when the potential returns toward the resting membrane potential. The potential may overshoot the resting value, causing a transient hyperpolarization known as the *hyperpolarising afterpotential* (Fig. 1-7) close to the equilibrium potential for K^+.

Sustained depolarization inactivates the voltage-gated Na^+ -channels, and shuts off the Na^+ -influx. Opening of voltage-gated K^+ -channels allows an increased outflux of K^+ to counterbalance the influx of Na^+. The membrane conductance to K^+ increases *slowly*, and reaches a peak in the repolarization phase ($g_K{}^+$ in Fig. 1-7). This K^+ -outflux causes the neuronal membrane potential to return to its normal resting value, when the Na^+ -channel is inactivated. The signal conduction is unidirectional, because newly opened Na^+ -channels become refractory for a time, when they are inactivated. As these areas are blocked for further depolarization for a time, the depolarization can proceed only in the forward direction towards resting Na^+ -channels.

Fig. 1-7: Transmembrane potentials and Na^+-K^+- conductance (flux) in a neuron.

The action potential is an *all-or-none electrical signal*, which appears as a positive wave when recording internally. The action potential is conducted with the same shape and size along the whole length of a muscle cell or a nerve fibre.

The refractory periods

During the early part of the action potential the cell membrane is completely refractory. A new stimulus, regardless of its size, cannot evoke an action potential. Almost all Na^+-channels are inactivated, and will not reopen until the cell membrane is repolarized. This is the *absolute* refractory period covering most of the peak and lasting until well into the repolarizing phase (ARP in Fig. 1-7).

During the hyperpolarizing afterpotential, a *suprathreshold stimulus* is able to trigger a new AP, albeit of smaller amplitude than the first action potential. This period is called the *relative* refractory period (RRP in Fig. 1-7). The cell membrane is relatively refractory, because some Na^+-channels are voltage-inactivated and at the same time K^+-conductance is increased.

Nerve conduction

The lipophilic core of the cell membrane is an electrical insulator, but the salt solutions of the cytoplasm and the extracellular fluid act as conductors of electrical current. Opening of many voltage-gated Na^+-channels, whereby the Na^+-conductance is increased about 104-fold, so the membrane is instantly depolarised, causes the action potential. The action potential essentially spread by alterations of the voltage-gated Na^+-channels.

Depolarization spreads along the membrane of excitable cells by local currents flowing to the adjacent segments of the membrane. This is shown in Fig. 1-8A. The phenomenon is called the *local response* or *electrotonic conduction*. The depolarization decreases mono-exponentially from the excitation site. Na^+-channels will be recruited in all areas of the membrane, where the threshold potential is exceeded. The Na^+-channels behind the peak of the action potential are refractory. This explains why an action potential travels in both directions, when it is evoked in the middle of a nerve.

Fig. 1-8: **Spread of the action potential along an unmyelinated (A) and a myelinated (B) axon. The refractory channels prevent the action potential from proceeding in more than one direction. The action potential (wavelength in cm) essentially jumps from node to node or over several nodes facilitating high-speed conduction.**

The *myelin sheath* consists of 20-300 layers of insulator substance produced by Schwann cells wrapping round the axon. The *nodes of Ranvier* are the lateral spaces (1 mm wide) between adjacent Schwann cells, which stretch 1-2 mm.

The effects of this arrangement are as follows:

Very little current is lost through the electrical insulation of the *myelin sheath*. Thus, the *electrotonic conduction* is rapid with only a small decrement in amplitude. The electrotonic conduction is virtually instantaneous. Because of the insulation the depolarization can spreadmuch faster.

Saltatory or *leaping* conduction occurs, because the action potential is generated only at the nodes (Fig. 1-8). The cell membrane below the myelin sheaths has hardly any Na^+- channels and is therefore inexcitable. Saltatory conduction is up to 50 times faster than the conduction through the fastest unmyelinated axons. The action potential can also jump over a number of nodes to that farthest away, because the action potential wavelength is several m.

The Na^+-channels there are activated by the *electrotonic conduction*.

Since the ionic currents are restricted to the nodes of Ranvier in the myelinated axons, this minimises disturbances in the Na^+- and K^+-gradients, that are restored by an active process in which the Na^+-K^+-pump, driven by ATP, pumps Na^+ out and K^+ into the cell. The main energy cost is to restore the Na^+- K^+- balance.

Myelination of the nerve fibre thus reduces energy cost of maintaining the resting membrane potential following an action potential.

Typical values for normal ion concentrations (intracellularly and extracellularly) are given in Box 1-3.

Box 1-3: Normal ion concentrations in muscle cells and in plasma.

Intracellular osmolality mmol * (kg of water)$^{-1}$		**Plasma concentration ranges mmol * (l of plasma)$^{-1}$**
Na^+	10	135 - 146
K^+	155	3.5 - 5.0
Ca^{2+}	0.0001	1 – 1.2

Mg^{2+}	12	0.7 – 1.1
Cl^-	5	95 – 106
HCO_3^-	10	22 - 27

The total calcium concentration in plasma is 2.2 – 2.7 mM, but only 45% is ionized (Chapter 30).

Pathophysiology
This paragraph deals with simple conditions, where the student needs no prior knowledge, so only hyponatraemia, pseudo-hyponatraemia and overbreathing are described.

1. Hyponatraemia is defined as a plasma-[Na^+] below 135 mM. This is a common condition caused by a high water intake (*water intoxication*), reduced water excretion in kidney disease, salt loss or other causes described in Chapter 24.

Hyponatraemia must be distinguished from the rare condition *pseudo-hyponatraemia.* Spuriously low [Na^+] are measured in plasma (Na^+ is predominantly confined to the water phase), simply because its concentration is expressed per l of plasma. Normally, 93% of plasma is water, and the non-water fraction is 7% (mainly proteins). In cases with too much lipid, protein, glucose, urea, or alcohol in the blood plasma (ie, hyperlipidaemia, hyperproteinaemia, hyperglucosaemia, uraemia, alcoholaemia etc), the normal [Na^+] is reduced by dilution with the increased non-water fraction. Thus, the *calculated* plasma-osmolality (Eq. 1-7) is less than the *freeze-point-measured* osmolality. This discrepancy is called the *osmolality-gap.*

There is no need for treatment with salt-solutions in such spurious conditions. Pseudo-hyponatraemia also occurs, when a blood sample is taken from an arm vein, where a glucose solution is infused. The plasma-[Na^+] obtained from such a blood sample is low.

2. Overbreathing is also called *hyperventilation*. Overbreathing is frequently caused by *panic attacks* (ie, hyperventilation tetany). The high ventilation washes out too much carbon dioxide (CO_2)/carbonic acid, whereby the CO_2 tension of the arterial blood decreases simultaneously with an increase in its pH. This is an alkalosis (arterial pH above 7.45). Alkalosis dissociates proteins by mass action and form Ca^{2+}-proteinate (Fig. 17-9). The falling extracellular concentration of free Ca^{2+}-ions reduces the threshold and opens Na^+-channels in neurons, muscle cells and myocardium. The resulting reduction in membrane potential increases the excitability of the tissues, which causes continuous muscular contractions (ie, *tetanic cramps*).

Equations

- **Uniformly distributed substances** in diluted solutions behave like gas molecules at atmospheric pressure (atm). The osmotic pressure can be expressed as:

$$\text{Eq. 1-1: } P_{osmot} = C * R*T,$$

which is the equivalent of the *ideal gas equation* ($P = n/V* R*T$). C is the concentration of dissolved solute. R is the gas constant (= 0.082 l atm mol^{-1} K^{-1}).

- **Fick's first law of diffusion** deals with the diffusive flux per m^2 :

$$\text{Eq. 1-2: } J^{dif} = -D* dC/dx.$$

The dimension of D is found by dimension analysis of the equation for J^{dif}:

$$(J^{dif} \text{ moles } s^{-1}\ m^{-2}) = D \times (dC/dx \text{ moles } m^{-3}\ m^{-1}).$$

Accordingly, D has the dimension: $m^2\ s^{-1}$. D is small, when the molecules are large and when the surrounding medium is viscous. The *permeability coefficient* (ie, permeability) for a membrane is the flux (mol $m^{-2}\ s^{-1}$) divided by the concentration (mol m^{-3}) for a given substance and has the unit m s^{-1}.

Einstein's relation states that the displacement squared, $(dx)^2$, is equal to 2 multiplied by D and by the time (t) elapsed, since the molecule started to diffuse:

$$\text{Eq. 1-3: } (dx)^2 = 2 * D * t.$$

The corrected van't Hoff law:

$$\text{Eq. 1-4: } \pi = T \times R \times f \times N/V \text{ or } \pi = T \times R \times DC$$

where R is the ideal gas constant (0.082 l × atm ×mol^{-1} × K^{-1} or 8.31 J $(K\ mole)^{-1}$), and DC is the *concentration gradient*. This is the law for ideal or extremely dilute solutions.

A **membrane potential difference** is conventionally defined as the *intracellular* (j^{i}) minus the extracellular (j^{o}) electrical potential. The *ion activities* (concentrations multiplied by the activity coefficient) inside the cell and outside the cell are here called C^i and C^o, respectively. The *Nernst equation* for the equilibrium potential of Na^+ across an ion-selective membrane at 310 K reads:

$$\text{Eq. 1-5: } j^{i} - j^{o} = (\mathbf{R}\ T/z\ F) \ln(C^{o}_{Na^+}/C^{i}_{Na^+}) \text{ Volts (V)}$$

$$V_{EqNa^+} = 61.5 \log (C^{o}_{Na^+}/C^{i}_{Na^+}) \text{ mV}.$$

In the equation above **R** is the *ideal gas constant* (8.31 J (K× mole)$^{-1}$), T is the absolute temperature, z is valence of the ion with sign (z= +1), and F is the Faraday constant (96 500 coulombs per equivalent). The activity coefficient for sodium is 0.75 and used to convert concentration to activity.

The law of radioactive decay: If N_o is the initial number of unstable nuclei, the number of nuclei remaining after a time t (N) is given by:

$$\text{Eq. 1-6: } N = N_o * e^{-k\ t}.$$

where k is a constant characteristic of each nucleide, called the *disintegration constant*.

Plasma-osmolality is calculated as follows:

$$\text{Eq. 1-7: Plasma-osmolality} = (2 * [\ Na^{+}]\ ) + [\ \text{glucose}] + [\ \text{urea}]$$

Normally, the plasma-$[\ Na^{+}]$ is 140 mmol in 1 litre of water, and both plasma-[glucose] and plasma-[urea] are around 5 mmol per l of water. The *plasma-osmolality* is given in mOsmol per kg of water. One l of water is approx. equal to 1 kg of water.

The **Millman equation.** A convenient version to calculate the resting membrane potential (RMP) at body temperature is:

$$\text{Eq. 1-8: RMP} = (g_{K^+} \times Eq_{K^+} + g_{Na^+} \times Eq_{Na^+} + g_{Cl^-} \times Eq_{Cl^-})/(g_{K^+} + g_{Na^+} + g_{Cl^-})$$

This equation shows that the RMP is determined by the conductance (g) of the membrane to K^+, Na^+ and Cl^-, and by their equilibrium potentials.

Self-assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

A. Positive ions (cations) diffuse more rapidly through a membrane than negative ions (anions).
B. The Na^+-K^+ -pump located in the cell membrane, is responsible for maintaining the high intracellular [K^+] and the low intracellular [Na^+].
C. The permeability for Na^+ in cell membranes is high compared to that of K^+ and Cl^-.
D. In skeletal muscle cells the resting membrane potential is -80 mV, and the equilibrium potential of Na^+ is +60 mV.

A membrane potential difference is conventionally defined as the extracellular minus the intracellular electrical potential.

II. Each of the following five statements have False/True options:

A. The local, subthreshold response is graded according to the stimulus.
B. Accommodation is a progressive decrease in firing frequency despite maintained depolarization.
C. Voltage inactivation of Na^+-channels is involved in the accommodation and in the refractory periods.
D. During the early part of the action potential the cell membrane is relatively refractory.

Voltage-gated Na^+-, K^+- and Ca^{2+} -channels are comprised of subunits with membrane spanning domains.

III. Each of the following five statements have False/True options:

A. Saltatory conduction is up to 50 times faster than the conduction through the fastest unmyelinated axons.
B. Saltatory conduction occurs because the cell membrane beneath the myelin sheaths has a high density of Na^+-channels.
C. At standard temperature, pressure, dry (STPD) the volume occupied by 1 mol of any ideal gas is 22.4 l.
D. Elimination with decreasing rate is exponential.

Facilitated diffusion shows saturation kinetics, and takes place through transport proteins not linked to the metabolic energy processes.

Each of the following five statements have False/True options:

A. Osmolarity is the number of osmotically active particles in each l of solution.

B. Facilitated diffusion does not show saturation kinetics but moves solutes up-hill.

   A family of homologous carrier proteins that are coded by distinct genes mediates facilitated diffusion.

C. Glucose-transport. The transport proteins (GLUTs) show a marked tissue-specificity which reflects differing transport needs of various tissues.
D. The relation between the potential difference measured with an ion-selective electrode and the activity is given by the Nernst equation.
E. The Donnan effect is the extra colloid osmotic pressure of protein solutions caused by uneven distribution of small, diffusible cations and anions.

Try to solve the problems before looking up the answers.

Highlights

The ideal gas law relates the pressure P, the volume V, and the number of mol of the gas n, and the Kelvin temperature T: ($P \times V = nRT$).

- *One mol is the amount of a given substance containing Avogadro's number. The volume occupied by any ideal gas is 22.4 l at STPD.*
- *Fick's first law of diffusion relates the diffusive flux per m^2 to the concentration gradient.*
- *The Donnan effect of plasma is the extra osmotic pressure of protein solutions caused by uneven distribution of small, permeable cations and anions.*
- *The membrane potential difference is conventionally defined as the intracellular minus the extracellular electrical potential.*
- *The action potential is an all-or-none electrical signal, which appears as a positive wave when recording internally during activity in a neuron or a muscle cell.*
- *Saltatory or leaping conduction occurs, because the action potential is generated only at the nodes. The cell membrane below the myelin sheaths has hardly any Na^+- channels and is therefore inexcitable. Saltatory conduction is up to 50 times faster than the conduction through the fastest unmyelinated axons.*
- *In cases with too much lipid, protein, glucose, urea, or alcohol in the blood plasma (ie hyperlipidaemia, hyperproteinaemia, hyperglucosaemia, uraemia, and alcoholaemia) etc, the normal plasma-$[Na^+]$ is reduced by dilution with the increased non-water fraction.*
- *Overbreathing is caused by panic attacks. The high ventilation washes out the carbon dioxide (CO_2)/carbonic acid, whereby the CO_2 tension of the arterial blood decreases simultaneously with an increase in its pH. This is an alkalosis. Alkalosis dissociates proteins and form Ca^{2+}-proteinate. The falling extracellular concentration of free Ca^{2+}-ions opens Na^+-channels in neurons, muscle cells and myocardium. The resulting reduction in membrane potential increases the excitability of the tissues, which causes continuous muscular contractions (ie, tetanic cramps).*

Further Reading

Alberts B, Brady D, Lewis J, Raff M, Roberts K, and JD Watson. Molecular biology of the cell. Garland Publ Inc, NY & London, 1994.

Apps DK, BB Cohen, and CM Steel. "Biochemistry." Bailliere Tindall, London, 1994

Model Of A Cell Membrane

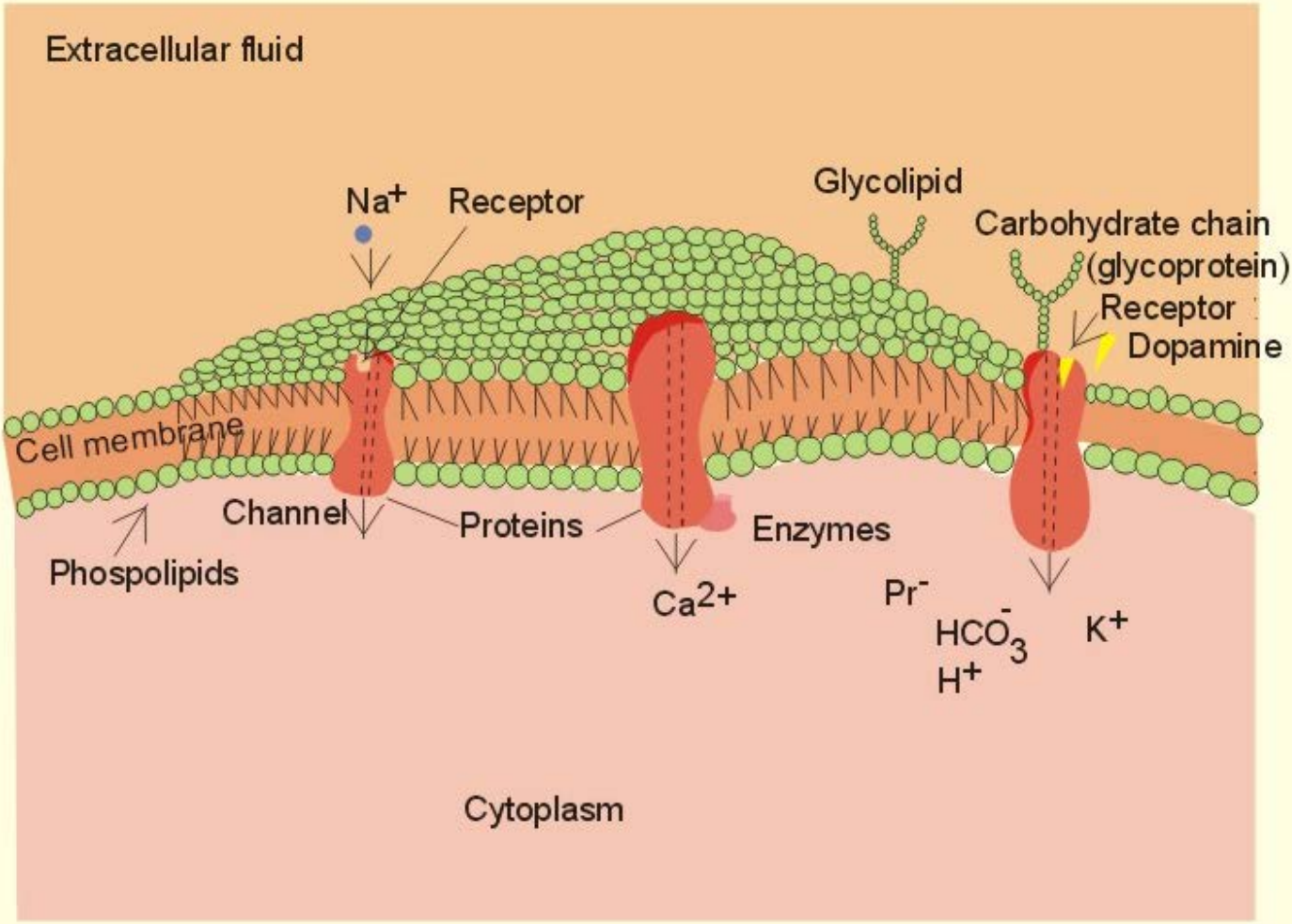


Fig. 1-1

KMc

Two Types of Passive Transport

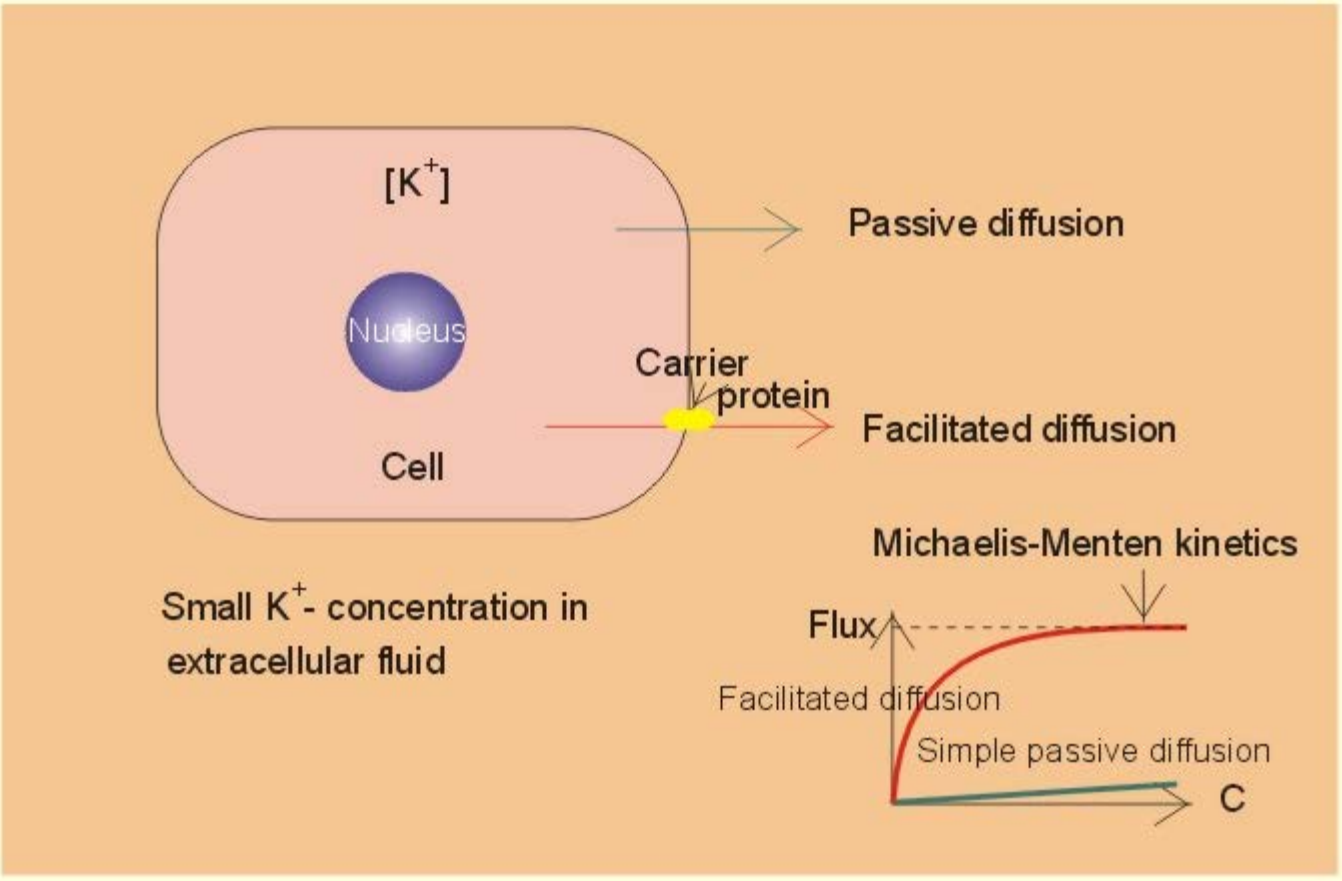
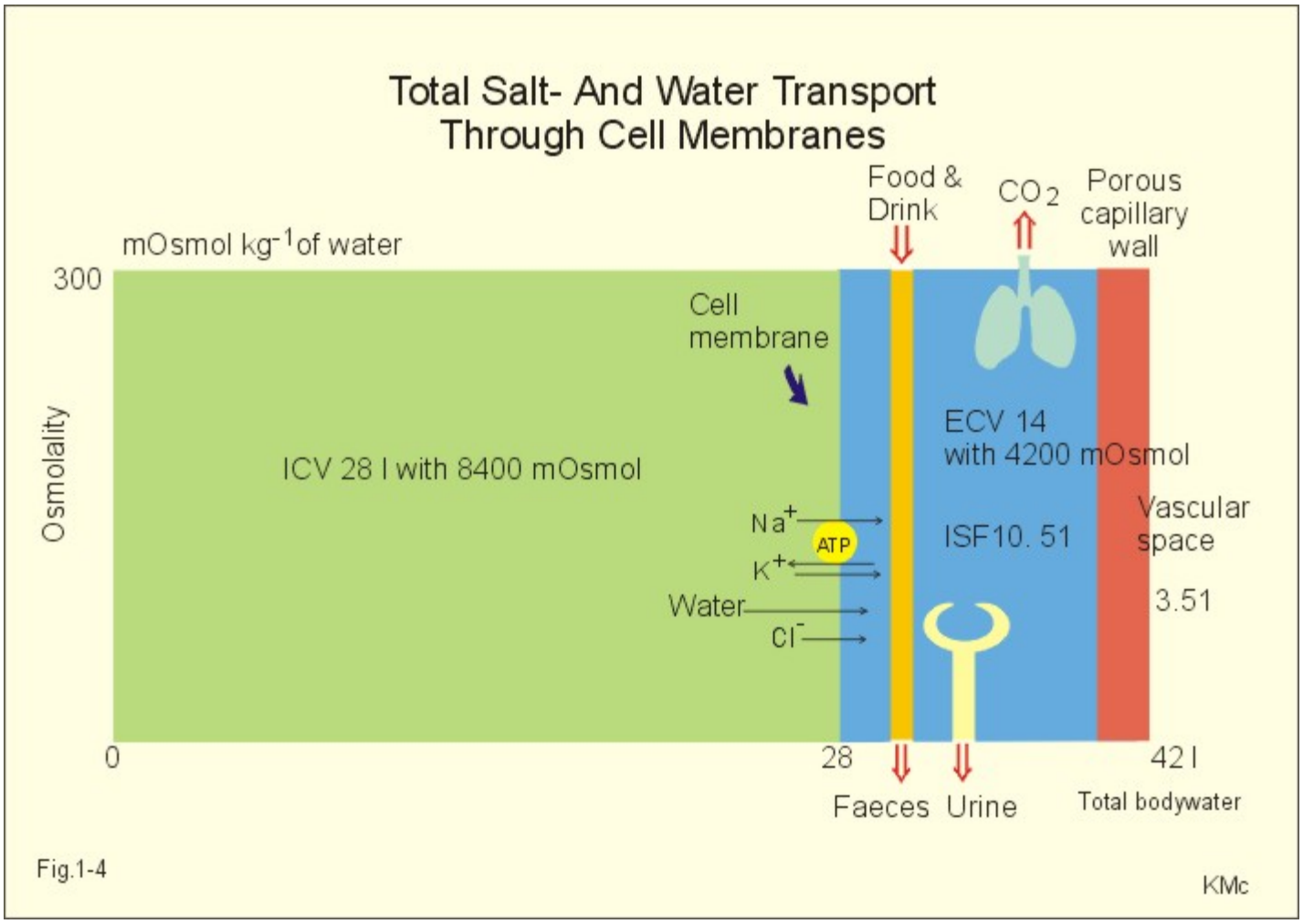
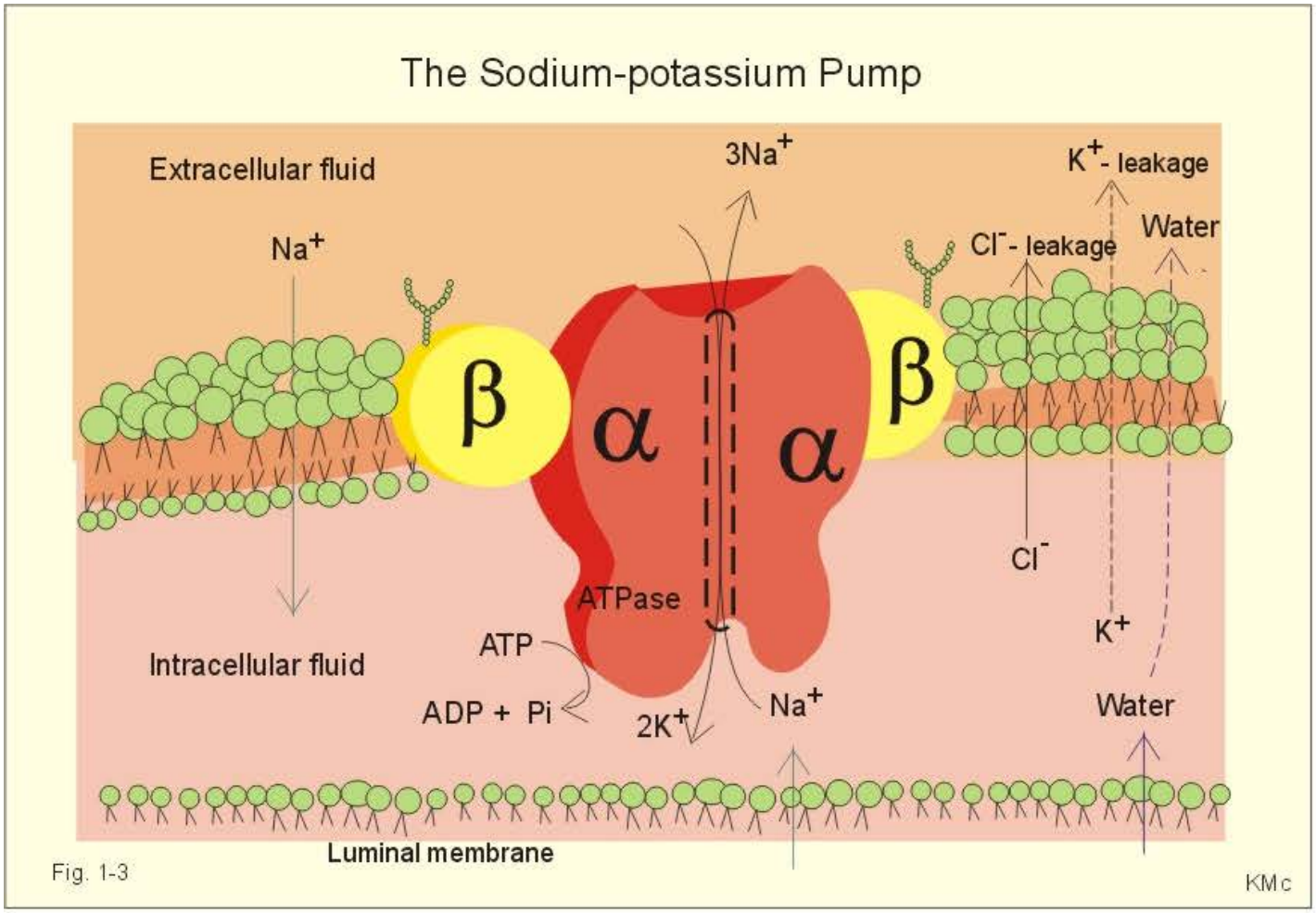


Fig. 1-2

KMc


Total Salt- And Water Transport
Through Cell Membranes
Food &
Drink
CO_2
Porous
capillary
wall
mOsmol kg^{-1} of water
300
Osmolality
Cell
membrane
ICV 28 l with 8400 mOsmol
ECV 14
with 4200 mOsmol
ISF10. 51
Vascular
space
3.51
Na^+
ATP
K^+
Water
Cl^-
0
28
42 l
Faeces
Urine
Total bodywater
Fig.1-4
KMc


The Sodium-potassium Pump
Extracellular fluid
3Na+
K+- leakage
Water
Na+
Cl- - leakage
β
α
α
β
Cl-
ATPase
K+
Intracellular fluid
ATP
ADP + Pi
2K+
Na+
Water
Luminal membrane
Fig. 1-3
KMc

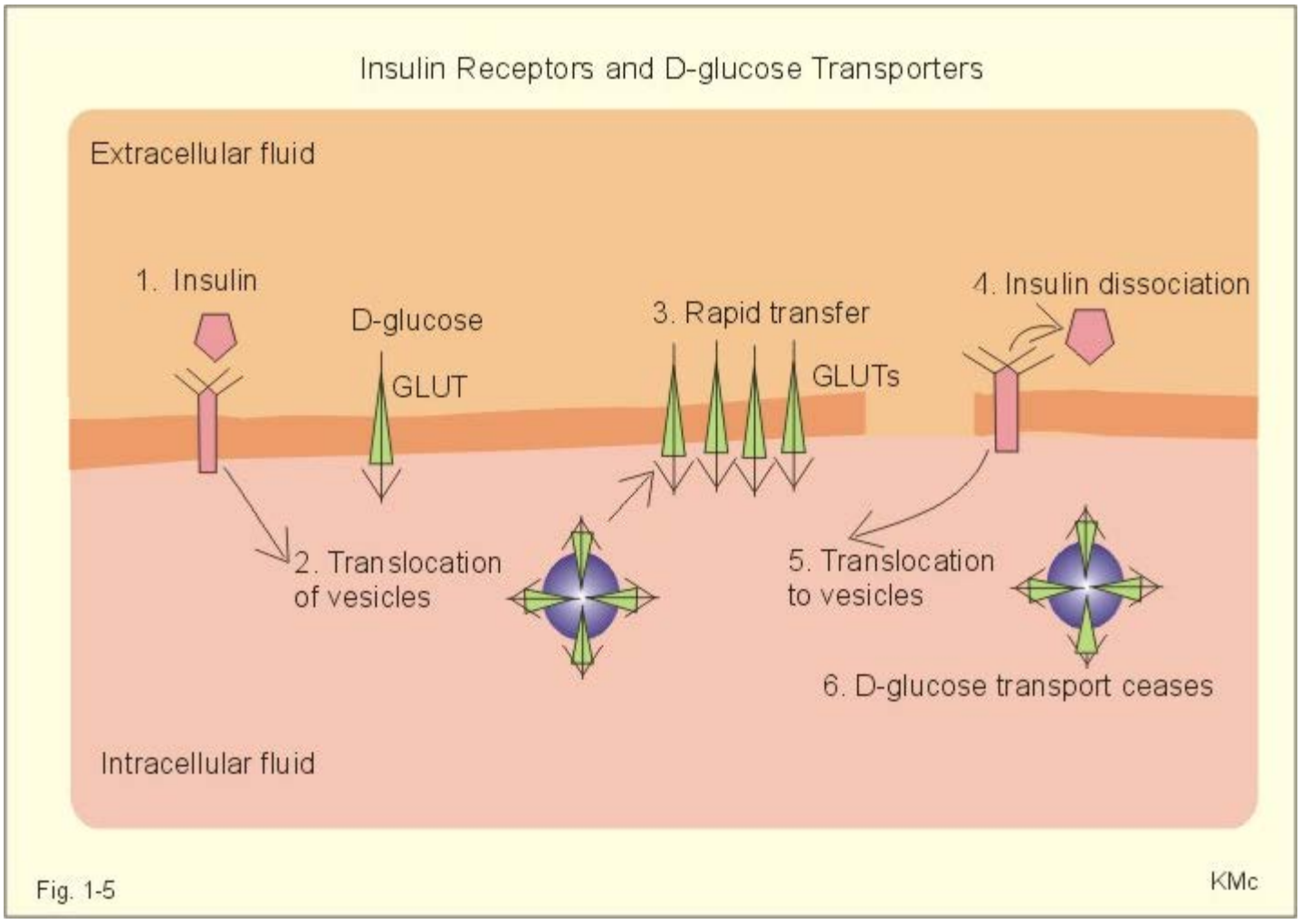
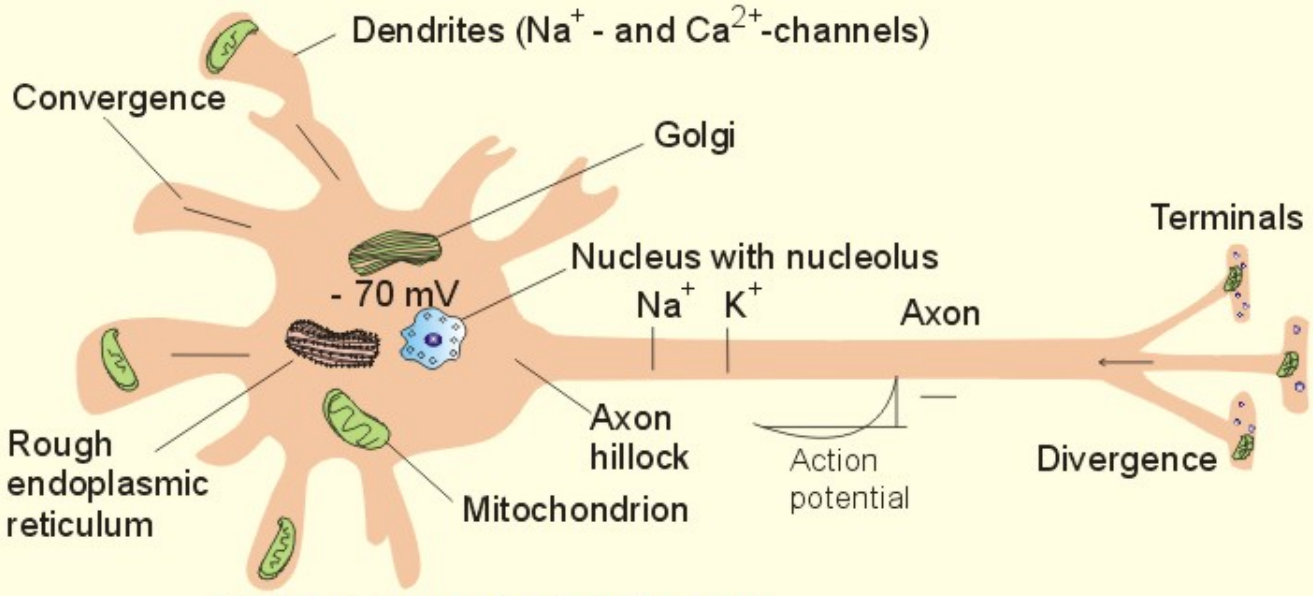

Insulin Receptors and D-glucose Transporters
Extracellular fluid
1. Insulin
D-glucose
GLUT
3. Rapid transfer
GLUTs
4. Insulin dissociation
2. Translocation
of vesicles
5. Translocation
to vesicles
6. D-glucose transport ceases
Intracellular fluid

Fig. 1-5

KMc

The Neuron With Its Four Structures



Cell body: Ligand-gated channels

Axon: Voltage-gated channels

Fig.1-6

KMc

SECTION II.

The Nervous System

My colleagues Jørn Hounsgaard and Poul Dyhre-Poulsen have contributed importantly to this section.

The nervous system is the essential control system for the human body. It consists of a central nervous system (CNS) and a peripheral nervous system. All information originates in sensory receptors and, enters the CNS via the peripheral nervous system. The CNS controls all the activities of the human body ranging from contractions of striated and smooth muscles to the exocrine and endocrine secretions.

The nervous system is an extremely rapid signal transduction system and the most important communication network in our body. The main integrative functions allow selection and processing of incoming signals to produce an appropriate response.

The nervous system includes sensory receptors that detect events in the body as well as in the outer world. Signals or action potentials from the sense organs travel through peripheral, afferent nerves to the CNS, where they are processed. The CNS controls the various activities of the body by motor mechanisms that generate movements and glandular secretions through efferent nerves. The afferent and efferent nerve fibres distributed throughout the body form the peripheral nervous system that is subdivided into a somatic and an autonomic part.

Neurons are highly specialised cells that are excitatory, inhibitory and sometimes neurosecretory. Neurons receive and transmit signals (action potentials) to other neurons or effectors. Neuronal networks account for information in a memory, evaluation of available knowledge, decision making, and transmission of response signals to appropriate effectors. The human nervous system contains about 10^{12} neurons forming at least 10^{15} synapses.

Frequently used abbreviations in this section are CSF for cerebrospinal fluid, EAA for excitatory amino acids, ECF for extracellular fluid, ECV for extracellular fluid volume, EEG for electroencephalogram, and REM for rapid eye movements. A complete list of abbreviations is present in Chapter 35.

Chapter 3.

The Somatosensory System And Disorders

Study Objectives

- To *define* adaptation, adequate stimulus, coding, sensory receptors including taste and smell, molecular receptors, receptor potential, stimulus transfer, types of sensory nerve fibres, conduction velocity, and threshold stimulus.
- To *describe* skin receptors, articular receptors, nociceptors and central pathways, the effect of chordotomy, thalamic surgery, and prefrontal lobotomy.
- To *draw* Hills force-velocity curve and the voltage-duration curve for nervous stimulation.
- To *calculate* one variable from relevant information's given.
- To *explain* cortical somatotropic and columnar organisation, the control of taste and smell, the control of nociceptive transmission (gatecontrol), central analysis, central pain, headache, referred pain, allodynia, causalgia, hyperalgesia, trigeminal neuralgia, thalamic syndrome, phantom limb pain, hyperalgesia, and Brown-Sequards syndrome.
- To *use* the concepts in problem solving and case histories.

Principles

- *Critical empiricism. In brain research any scientific observation presupposes a theory that can be falsified. Theories that fail to be falsified in repeated scientific projects are temporarily acceptable. This philosophy is generally applicable.*
- *Sherrington's integration law. The integrative action of the nervous system unifies separate organs to form an individual personality.*

Definitions

Adaptation or *accommodation* of sensory receptors refers to a progressive decrease in firing frequency despite maintained depolarisation.

Adequate stimulus refers to the stimulus, for which the receptor has a lower energy threshold than for other stimuli - ie, the stimulus to which the receptor is most sensitive.

AMPA is an abbreviation of *alpha-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid*. This is an excitatory amino acid (EAA) and one of the EAA-gated ion channels carries AMPA-receptors with high affinity for a subclass of glutamate receptors.

Causalgia means hyperalgesia (hypersensitivity to pain) elicited by a cold stimulus.

Coding: In neurons the stimulus intensity is coded by the frequency of action potentials.

Dermatomes: Every spinal dorsal root is destined to a segment of the skin called a dermatome.

Endogenous opioids are substances in the CNS with opiate-like effects.

GABA is the abbreviation for gamma-amino butyric acid – the most common inhibitory transmitter in the brain.

The **gate-control hypothesis** of pain states that pain transmission is suppressed by innocious signals in thick myelinated afferents (group II), whereas the pain sensation is enhanced by signals in thin afferents.

Headache is pain on the surface of the head, which is actually due to anomalies in intracranial or extracranial structures.

Hyperalgesia means hypersensitivity to pain.

NMDA is the abbreviation for N-methyl-D-aspartate.

Nociceptors or nocireceptors are responsive to stimuli that potentially cause injury.

Plasma membrane receptors consist of a protein or glycoprotein molecule, an ion channel pore or a specific enzyme (G-protein).

Receptor potential: The stimulation of a receptor elicits a *generator potential* that is graded according to the stimulus strength. When the stimulus is strong enough to reach the threshold, an action potential is fired. Sensory receptors translate stimulus intensity to impulse activity in sensory afferents.

Sensory receptors are either *neurons* in the case of vision, smell and cutaneous senses, or *modified epithelial cells* in the case of auditory, vestibular and taste senses.

Threshold stimulus refers to the weakest stimulus to which the receptor will react.

Trigeminal neuralgia (French: *Tic douloureux*) is a condition with daily paroxysms of violent pain in part of the trigeminal area lasting only some seconds. The pain is provoked by eating, washing the face or by cold in the face.

Essentials

This paragraph deals with *1. Sensory receptors and nerve fibres, 2. Blood-brain barrier and CSF, 3. Regeneration of nervous tissue, 4. Sensory pathways, 5. Central opiate receptors, 6. Taste and smell.*

1. Sensory receptors and nerve fibres

Sensory receptors are either *neurons* in the case of vision, smell and cutaneous senses, or *modified epithelial cells* in the case of vision, auditory, vestibular, smell and taste senses. The special sensory receptors for vision, hearing and balance are described in Chapter 5.

Some sensory receptors have characteristics similar to the well-known *plasma membrane receptors*. Plasma membrane receptors consist of a protein or glycoprotein molecule, an ion channel or a specific enzyme (G-protein).

The stimulation of a receptor elicits a receptor potential (*generator potential*) that is graded continuously with stimulus intensity. When the stimulus is strong enough to reach the threshold, action potentials (APs) are fired. In neurons the stimulus intensity is coded by the frequency of action potentials.

Sensory receptor systems are *biological transducers* with a dynamic range of *up to* 10^{12} in the most sensitive organ, the *ear*. The *threshold* is the reciprocal of the sensitivity. The *threshold* is the weakest stimulus to which the receptor will react. The *sensitivity* of a sensory receptor is greater the smaller its threshold stimulus is.

The Pacinian corpuscles in the skin are gigantic receptors (almost 1 mm long) consisting of concentric layers like onion-scales in the microscope. There is a single axon in the axis of each corpuscle (Fig. 3-1). Stimulus intensity is coded by the single axon. Compression deforms the axon and depolarises the membrane by opening of Na^+-channels.

Fig. 3-1: The Pacinian corpuscle (1 mm long) is a vibration detector.

The depolarisation generates a graded *receptor potential* forcing a current towards the first node of Ranvier with maintained stimulus. The receptor potential rapidly decreases (rapid adaptation), because the adequate stimulus is alterations in the deformity rate (vibrations in the range 150-300 Hz). At the first node of Ranvier, a propagating *action potential* along the axon is released, provided the generator potential is sufficiently large (Fig. 3-1). The Pacinian corpuscle is located in the deeper layers of the skin and connective tissue. The afferent fibres are thick (Type Ab or II), and they lead the signals to synapses in nucleus gracilis and cuneatus of the spinal cord.

Steven proposed the *power law* that is given as Eq. 3-1. The *interpreted stimulus strength* (ISS) is equal to a constant (k) multiplied by the *actual* stimulus strength (SS) raised to the power n. The situation n equal to 1 describes a linear relation between stimulus and resulting activity in the conducting neuron (Fig. 3-2).

Perception of *taste, heat and angular acceleration* are described by *power functions* or *transfer functions* with n just above 1, whereas *hearing and smell* are described by functions with n lower than 1. The only sensation with a particularly large value of n (about 3) is *pain,* which is why pain is felt so severe with increasing stimulus intensity!

Fig. 3-2: A graphical description of the power law.

However, the power law and other types of curve fitting with transfer functions have hardly improved our understanding of sensory modalities.

Both *conduction velocity* and *size* is used in classification of nerve fibres. The fibres are divided into types A, B and C, based on the three main *conduction velocities* shown in the record of the *compound action potential* from a mixed nerve. Type A, B and C refer to phases in the combined action potential. Type A fibres are the fast conducting myelinated fibres (thick fibres subdivided into a , b , g , and d), type B are preganglionic sympathetic fibres, and type C are the small, unmyelinated fibres. Another classification is based on the *thickness of the axons* (I-IV). The size classification became necessary, when Aa - fibres were separated into two subgroups: Ia and Ib.

In Box 3-1, the velocity classification (A-C) is written first, and the size classification (I-IV) is given in parentheses.

Box 3-1: Classification of nerve fibres

Fibre type	Function	Axon diameter m m	Conduction
		/ Myelin + -	velocity, m per s
Aα (I)	motor a - fibres	9-18/+	70-120
	spindle afferents (Ia)		
	tendon organs (Ib)		
Aβ (II)	touch and pressure	5-12/+	30-75
Aγ (II)	motor to muscle spindles	3-6/+	18-36
Aδ (III)	pain, pressure, temperature	1-5/-	4-30
B (III)	preganglionic	3/-	3-12
C (IV)	pain, touch, heat	1/-	1-2

The Aa (I) fibres are *motor* α *-fibres* and *proprioceptors* from the annulospiral endings of muscle spindles (Ia) and from Golgi tendon organs (Ib).

The Ab (II) fibres conduct discrete touch and fine pressure signals from cutaneous tactile receptors.

The Ag (II) fibres are motor fibres to muscle spindles. They have their origin in the spinal cord.

The Ad (III) fibres transfer pain sensations, decline in skin temperature as well as crude, passive touch and deep pressure.

- The B fibres (III) are autonomic preganglionic fibres.
- The C fibres (IV) are unmyelinated and lead *pain, touch* and *signals* from heat receptors from the skin. The C fibres have no myelin sheath.

Sensory receptors in the nervous system are classified as *exteroceptors* (located on the body surface), *proprioceptors* (located in muscles, tendons and joint capsules), *interoceptors* (located in the viscera), and *telereceptors* (stimulated by events far from the person).

Cutaneous receptors are exteroceptors (Fig. 3-3). Pacinian and Meissner corpuscles are rapidly adapting (dynamic) touch velocity detectors in glabrous skin. In hairy skin, hair-follicle receptors are velocity detectors (they adapt rapidly). Meissner corpuscles are located in the papillae of the hairless skin such as fingertips, lips and clitoris. Merkels discs and Ruffini-end organs are slowly adapting (static) touch intensity detectors both in hairless and hairy skin. Merkels discs are found in elevated dome corpuscles in hairy skin (up to 50 Merkel discs in a corpuscle of 0.5 mm in diameter). These so-called *Iggo dome receptors* are extremely sensitive and transmit touch signals to a single nerve fibre. In this way, even weak tactile stimuli can create sensation in the CNS.

Fig. 3-3: Cross section through an area of hairless and a hairy skin showing 6 types of sensory mechanoreceptors. All the mechanoreceptors are supplied by afferent nerve fibres of group II, except the free nerve endings (group IV).

The free nerve endings, with unmyelinated afferents (group IV) conduct signals with low velocity (Box 3-1), and register passive touch, such as, slow strokes with a piece of cotton.

Thermoreceptors are also exteroreceptors. We have cold-receptors just below the skin surface (200 nm deep). Cold receptors respond to changes in temperature. Heat receptors are also located in the skin. The location of certain heat and cold points in the skin is determined by bringing a thin hot or cold object in contact with the skin. *Thermoreceptors* react to temperature changes. *Cold receptors* and *heat receptors* in the skin are located close to the surface. Both types of receptors are also located in the deep tissue and in the CNS. Both types of receptors *discharge spontaneously* at normal temperature, *dynamically* when skin temperature is changing rapidly, and *adapt slowly*.

Proprioreceptors, located in muscles, joints and joint capsules, are mechanoreceptors (muscle spindles, Golgi-receptors, Pacinian and Ruffini corpuscles, and free nerve endings). The Ruffini mechanoreceptors are also called *joint receptors*, because they are located in ligaments, tendons and *articular capsules*. They provide information for the CNS concerning articular movements, movement velocity and joint position. Joint receptors of the proximal joints are particularly sensitive. The static and dynamic receptors inform the CNS about the position and movement of the joint, respectively. These receptors enable us to sense the position of the joint with great accuracy.

Accommodation of sensory receptors or *adaptation* is a progressive decrease in firing frequency despite maintained depolarization. The frequency of action potentials from stimulated receptors fall, although the stimulus is maintained at constant strength. Accommodation or adaptation occurs, when a proportion of the voltage-gated Na^+-channels is rapidly *inactivated* by depolarisation, which also opens K^+- channels. This makes the cell more refractory to stimulation. Accommodation can also be caused by a hyperpolarization induced by gradual activation of Ca^{2+} -dependent K^+ -channels.

Fig 3-4: Accommodation or adaptation curves from different sensory receptors.

Pain- and cold-receptors, Merkel discs and Ruffini-end organs adapt extremely slowly and incompletely (Fig. 3-4). Joint receptors, smell-taste-receptors, muscle spindles, carotid sinus- and pulmonary stretch receptors, and the optic nerve, all adapt somewhat better (Fig. 3-4).

Hair-follicle receptors, Meissner corpuscles and Pacinian corpuscles adapt rapidly, just as many free nerve endings (Fig. 3-4).

Nociceptors or *nocireceptors* (pain receptors) are responsive to stimuli that potentially cause injury. Nociceptors are free nerve endings of two types. The *fast adapting* Ad fibre mechanical nociceptors (group III) are high-threshold, finely myelinated afferents that originate superficially in the skin. The *slowly adapting* C-polymodal nociceptors (group IV) are unmyelinated afferent fibres that originate in the deeper cutaneous tissue, and respond to various mechanical, thermal and chemical stimuli (Fig. 3-5). In the spinal cord nociceptive afferents synapse with secondary neurons in lamina I and II. These sensory neurons ascent in the spinothalamic tracts.

The fast adapting pain through group III fibres is bearable (acute, sharp, stinging, somatic pain), compared to the slowly adapting unbearable pain (diffuse, burning, prolonged secondary, visceral pain) through group IV fibres.

Fig 3-5: Mechanical, polymodal and visceral nociceptors. A visceral pain afferent synapses in the spinal cord with the neuron of the lateral spinothalamic tract on which the cutaneous group IV pain afferent terminates.

When nociceptors become sensitised (ie, more responsive), their thresholds are reduced, thus causing *hyperalgesia* (ie, hypersensitivity to pain). Many substances such as bradykinin, histamine, leucotrienes, prostaglandins, serotonin, and K+ that are often released near damaged or dying cells sensitise nociceptors. K^+ *activates* the nociceptors. Substance P is also released from polymodal nociceptors through an axon reflex with antidromal signal transduction in afferent group IV fibres, causing hyperalgesia, vasodilatation and increased capillary permeability (Fig. 3-5). Glutamate may be co-released with *substance P* from the polymodal C-fibre terminals.

The *gate-control hypothesis* of pain states that pain transmission is suppressed by innocious signals in thick myelinated afferents (group II), whereas the pain sensation is enhanced by signals in thin afferents. Inhibitory interneurons of the lamina II in the dorsal horn of the spinal cord perform the gate-control through a special type of presynaptic inhibition called primary afferent depolarization (PAD), and the receptors on the cell body of the secondary neuron is the gate. The gate control hypothesis explains why innocuous signals, mediated by large myelinated afferents, can inhibit pain mediated by thin myelinated afferents.

The *adequate stimulus* is the stimulus, for which the receptor has a lower energy threshold than for other stimuli (ie, the stimulus to which the receptor is most sensitive). The adequate stimulus for pain receptors is mechanical deformation, extreme temperature or tissue damage. The *sense impression* depends on the site in the brain which receives the sensory signal (ie, *central analysis*) and on the receptor localisation (ie, *peripheral analysis*). This is how different neurons transmit different types of sensations, even though they may transmit the same electrical signals (see Chapter 8).

The CNS discards more than 99% of all incoming signals as irrelevant.

The visual system is an example of a *specific information line* for a certain modality of sensation. The neurons in the retina, the optic nerve, the lateral geniculate nucleus, and the visual cortex describe just such a dedicated neuronal pathway. The specific information line through which the signal is conducted determines the way in which a suprathreshold stimulus is perceived (eg, pressure applied on the eye will be perceived as light).

The auditory system also forms a *specific or labelled line* all the way from receptor to cortex. In all cases the specific region in the cerebral cortex, where the nerve fibre ends determines the modality of sensation.

Now, where is the sense interpretation localised and what is its intensity?

1. Coding in the sense organ is *peripheral analysis*, which is based on the peripheral location of the receptor. External energy is transformed to a *receptor potential* that triggers APs in afferent nerve fibres. Peripheral analysis depends upon the location and the special structure and sensitivity of the receptor. The pattern of firing of APs is the only possible variable for coding information in a single neuron. Examples of firing patterns are *on-off* patterns with mean frequencies, *off-on* patterns, *transient* patterns or *adaptation*, *long-lasting* patterns, firing with latency etc.
2. *Central location coding in the CNS* is termed *central analysis*, which is related to the sense impression.

2. Blood-brain barrier and CSF

The blood-brain barrier consists of *tight junctions* between the endothelial cells of the capillaries in the CNS and of neuroglia. This barrier only allows extremely small or hydrophobic molecules to pass into the brain. The cerebral

microcirculation consists of *strong arterioles* that can constrict to carry a high arterial pressure without brain oedema.

Many large molecules cannot pass from the blood to the cerebrospinal fluid (CSF) across the choroid plexus, a tight junction barrier that is called the *blood-cerebrospinal fluid barrier*.

The *blood-CSF barrier* of the choroid plexus allows some large molecules to pass from the blood to the CSF.

Fig. 3-6: The blood-brain and blood-CSF barriers showing the daily formation of 500 ml of CSF.

The blood-brain and the blood-CSF barriers exist in all areas of the brain, except in the so-called *circumventricular organs* (hypothalamus, the pineal gland, and the area postrema). These discrete organs have highly fenestrated capillaries that are easily penetrated by large and small molecules as well as ions. The circumventricular organs are located close to essential control centres in the hypothalamus and brain stem regions regulating respiration, blood glucose concentration, and extracellular fluid osmolality.

The two brain barriers are almost impermeable to large molecules such as plasma proteins, but highly permeable to CO_2, oxygen, water, alcohol, anaesthetics, hallucinogens, and other lipophilic substances. The blood-brain barrier is almost completely impermeable to water-soluble molecules, electrolytes such as H^+, whereas CO_2 passes through the barrier to the medullary chemoreceptors (Fig.16-3).

Humans produce 500 ml of CSF daily. The total CSF volume is only 1/3 of the daily production. Most of the 500 ml of CSF is produced in the *choroid plexuses* in the four brain ventricles, and the remaining is produced across the blood-brain barrier.

The ventricular system and the central spinal channel are covered with *ependyma*. The absorption of CSF takes place through the *arachnoidal granulations*, which protrude, into the *sinus sagittalis*. The rate of absorption is directly related to the pressure in the cranial cavity - in particular the CSF- pressure. Proteins can pass through large holes in the endothelial cells. The CSF is separated from the brain cells by the *thin* pia mater. Substances that enter the CSF can easily diffuse into the brain interstitial fluid. Drugs that cannot pass the *blood-brain barrier* can enter the brain through pia mater, when infused into the CSF (Fig. 3-6).

The CSF passes from the lateral ventricles (I and II) through the *foramen of Monroe* into the third ventricle (III), through the *aquaduct of Sylvius*, the fourth ventricle (IV), and out into the subarachnoid space through the *foramina Luschkae & Magendie* (Fig. 3-7).

Fig. 3-7: Anatomical structures involved in CSF-formation and absorption.

The normal CSF-pressure in a supine person is up to 10 mmHg (1.3 kPa) or 136 mm of water.

The secretion of fluid by the choroid plexus depends on the active Na^+-transport across the cells into the CSF. The electrical gradient pulls along Cl^-, and both ions drag water by osmosis. The CSF has lower [K^+], [glucose], and much lower [protein] than blood plasma, and higher concentrations of Na^+ and Cl^-. The production of CSF in the choroid plexuses is an active secretory process, and not directly dependent on the arterial blood pressure. The CSF is separated from the brain cells by the extremely thin *pia mater*. All natural substances that enter the CSF can easily diffuse into the brain extracellular fluid (Fig. 3-6).

CSF leaves the four ventricles through the roof of the 4th ventricle, traverses the subarachnoid space, and is reabsorbed into the blood of the venous sinuses via the arachnoidal villi. The *absorption* here is directly related to the CSF pressure in the cranial cavity. Large holes through the endothelial cells allow proteins to enter the blood.

3. Regeneration of nervous tissue

Severe injury to nervous tissue causes cell death. Neurons are postmitotic cells. For this reason lost neurons cannot be replaced.

There is, however, considerable capacity for regeneration of axons in the peripheral nervous system. Both growth and maintenance of axons require the *nerve growth factors* (NGF). NGF is an essential survival factor for neurons outside the CNS - in particular sensory neurons. NFG binds to receptors belonging to the insulin receptor family (*tyrosine kinase family*).

When a motor axon has been severed, the cell body undergoes **chromatolysis**. This is a neuronal reaction, where the rough endoplasmic reticulum (the *Nissl bodies*) becomes active. The Nissl bodies accumulate proteins required for repair of the axon. The *axonal reaction* is an attempt to repair the fibre by production of new protein structures that are transported along the axon. Therefore, proteins distend the rough endoplasmic reticulum. The axon and the myelin

sheath distal to the injury die and are phagocytized. The neuroglial Schwann cells that had formed the myelin remain alive. This is the so-called *wallerian degeneration* named after Waller.

The Schwann cells proliferate and form long rows along the pathway previously occupied by the dead axon. The severed axon regenerates along this pathway, and *growth cones* may eventually reinnervate the target organ.

Neurological injury probably involves excessive *glutamate receptor stimulation* as a common pathway.

Glutamate is the most important of the *excitatory amino acids* (EAAs) in the spinal cord and the brain. Glutamate stimulates the family of EAA-receptors including AMPA-, NMDA-and metabotropic receptors. NMDA means N-methyl-D-aspartate. - Effective glutamate antagonists are applied in clinical studies of pain.

The inhibitory amino acids, GABA and glycine, and the monamines and endogenous opiods inhibit the second-order neurons of the spinothalamic tract.

Fast axonal transport of organelles in the cytosol occurs as rapidly as 0.4 m per day. At this rate synaptic vesicles can travel along the motor axon from the spinal cord to a patient's foot within three days. Fast axonal transport of enzymes and organelles occurs on microtubuli in the axons, and is not interrupted by resting periods in cell compartments outside the transport system (Fig. 3-8). Oxidation of glucose in the mitochondria provides ATP for the Na^+-K^+-pump and for transport filaments and microtubules embedded in the axonal cytoplasm (Fig. 3-8).

Fig. 3-8: Axonal transport of vesicles, organelles and proteins by microtubuli.

Slow axonal transport occurs as diffusion of cytosolic proteins and organelles such as mitochondria. This transport occurs at a rate 100 times more slowly than fast axonal transport. Organelles or enzymes are stored in different cell compartments on their way or their direction of transport reverses.

Axonal transport can be *anterograde*, when it occurs in the direction from the soma to the axonal terminals. Axonal transport can also be *retrograde*, when it occurs in the opposite direction. Here vesicles are degraded by lysosomes, when returned to the soma. A typical example of slow transport is the transfer of the many mitochondria towards the terminal of an axon.

In the CNS, fast neurotransmission is *inhibitory* or *excitatory*. In the neuromuscular junction, *each signal* is always excitatory and sufficient to trigger a muscular contraction. In the neuromuscular junction, acetylcholine is the only neurotransmitter, whereas in the CNS there is a large variety of neurotransmitters (see Box 7-1 and 7-2).

The sensory system transmits signals from sensory nerve receptors in the body. The nerve receptors are located in the skin, muscles, tendons, joints and viscera. The signals are transferred to the CNS by a pathway of first, second, third, and higher-order neurons. The third and higher order neurons are located in the *thalamus* and the *cortex*. The cell body of the first order afferent neuron is located in the dorsal root or in the cranial nerve ganglia. The signals pass through the spinal cord, the brain stem, and the thalamus before reaching the cerebral cortex.

4. Sensory pathways

Several sensory tracts and pathways synapse in the *nuclei of the thalamus* (the spinothalamic tracts). The *somatosensory thalamus* is a relay station for most sensory modalities. The sensory inputs are processed in somatotopic areas of the thalamus, and are then transferred to appropriate cortical areas. The *somatotopic organisation* is maintained all the way to the cortex.

The *reticular activating system* (RAS) of the brainstem is involved in arousal acting in concert with the thalamus.

The *spinothalamic tract* conveys pain and temperature (lateral tract), and also crude passive touch (ventral tract). The first-order neurons are afferent Ad fibres (III) which have cell bodies in the spinal ganglia. Second-order neurons cross immediately to the opposite side of the spinal cord, and ascend in the lateral and ventral spinothalamic tract.

Fig. 3-9: The spinothalamic tracts and their sensory function.

Pain and temperature reach the thalamus in the lateral spinothalamic tract (in the lateral funiculus). The second-order axon terminates in the *somatosensory thalamus* (the ventral posterior lateral nucleus and the central lateral nucleus). The third-order neurons pass from the somatosensory thalamus via the *thalamocortical fasciculus* to the *somatosensory cortex* or the *primary sensory cortex* (somatic sensory area I, or area 1, 2, 3 in Fig. 4-2) with the *sensory homunculus.* Some third-order neurons also pass to the somatic sensory area II of both hemispheres.

Proprioception and active tactile signals are transmitted through sensory nerve fibres to the spinal cord. Primary afferent fibres ascend in the *dorsal columns* all the way to the medulla oblongata. These primary axons synapse with second-order neurons in the gracile and the cuneate nuclei. These second order neurons cross the midline in the

medulla, and ascend in the medial lemniscus to end in the somatosensory thalamus. The *medial lemniscus pathway* transmits proprioception and fine tactile senses.

The *spinothalamic tract* is the most important pathway for *pain*. The *second order neurons* of the spinal tracts have their cell bodies in the lamina I, II and V of the spinal cord. These cells receive excitatory signals from nociceptors in the skin, muscles and viscera. The action potentials from the nociceptors are conducted along the axon to the spinal cord and release neurotransmitters such as the excitatory amino acid, glutamate, and different neuropeptides. When these neurotransmitters bind to the receptors on the postsynaptic membrane of the secondary neurone, they increase the permeability to small ions, and excite secondary, postsynaptic neurons. The secondary neurons of the spinothalamic tract projects mainly to the *contralateral thalamus* by crossing over immediately through the anterior commissure to the opposite side of the spinal cord within the incoming segment.

5. Central opiate receptors

The *endogenous analgesia system* is a pain control system descending from brainstem to the spinal cord (Fig. 3-10).

As an example, this system may explain why a runner who twists his leg during a competition may finish the run before he really feels the pain. As soon as he has passed the goal and stop running the pain often becomes severe, and he cannot run at all.

The cell bodies of the neurons belonging to this system are located in the *periaqueductal grey area* of the midbrain, the *periventricular areas*, locus coeruleus, and the areas surrounding the *aqueduct of Sylvius* (Fig. 3-10). Signals from these cell bodies reach the medullary *raphe magnus* nucleus and the medullary *nucleus reticularis gigantocellularis* with nucleus reticularis *paragigantocellularis* lateralis. The nuclei transmit signals via the *descending pain-suppressing pathway* in the dorsolateral column to a *pain inhibitory complex*. Stimulation or increased tone of the analgesia system can suppress strong pain signals entering the spinal cord through the dorsal spinal horn. These regions contain *opioid receptors*. There are at least 4 types of *central opiate receptors* and their subtypes: **u** for morphine-like drugs, d and k for enkephalins, and the non-selective s -*receptors*.

Endogenous opioids are substances with opiate-like effects. These substances are naturally occurring in the nervous system (b -endorphin, met-enkephalin, leu-enkephalin, dynorphin and many others). Endogenous opioids are derivatives of *three* large protein molecules encoded by three different genes. These mother-molecules are *pro-opio-melanocortin* (POMC), *proenkephalin* and *prodynorphin.*

Fig. 3-10: Opiate receptors in the CNS. Enkephalin, b-endorphin and dynorphin (3 peptides) appear in normal human cerebrospinal fluid.

Enkephalins inhibit both type C and type Ad (III) pain fibres presynaptically in the dorsal horns. *Enkephalin* is the endogenous ligand for the d -opiate receptors. Dynorphin has much higher affinity than morphine and is only found in small quantities close to the dynorphinergic k -opiate receptors. b -*endorphin* is present in the hypothalamo-hypophysary system.

Presynaptically located opiates inhibit depolarization of nerve terminals and reduce synaptic transmission. The purpose of pain is to protect the body from further or imminent harm.

A special type of burning pain is provoked by noxious heat or by capsaicin (which contain a vanillyl-group) in chilli, paprica and pepper. These spices and heat stimuli seem to activate a vanniloid receptor subtype 1 in sensory nociceptors with terminals in the dorsal horn of the spinal tract. Activation opens Ca^{2+}-channels and the Ca^{2+}-influx is probably involved in the burning sensation.

Gyrus cinguli has the highest density of *central opiate receptors*. Pyramidal cells are contacted by *opiate secreting interneurons* that inhibit arriving pain signals.

6. Taste and smell

The sensations from the anterior 2/3 of the tongue travel with the trigeminal nerve fibres, through the *chorda tympani* into the facial nerve (VIIth), and eventually reach the *solitary tract* of the *brain stem*. Taste signals from the back of the tongue and surrounding tissues are transmitted through the glossopharyngeal nerve (IXth) into the tractus solitarius. All taste fibres synapse in the *nuclei of the solitary tract* and the axons of these neurons project to the thalamus. From the thalamus third-order neurons reach the lower part of the *primary sensory cortex* in the postcentral gyrus (somatosensory area I = area 1 in Fig. 4-2).

Fig. 3-11: Taste buds and taste pathways from the tongue.

Acids evoke *sourness*, because H^+ stimulates special *H^+-receptors* in the taste buds. *Saltiness* is produced by the anions of inorganic salts. The *Cl^--receptor* is particularly effective in registering saltiness. Our taste buds at the base of the tongue also have *bitter-receptors* stimulated by many long-chain organic compounds. Many alkaloids (quinine, caffeine, and nicotine) also taste bitter. *Sweet-receptors* are stimulated by sucrose, glucose, lactose, maltose, glycerol, alcohol, aldehyde, ketone, and organic chemicals.

In the *upper nasal cavity* the mucous membrane is yellow and termed the *olfactory membrane*. It contains 100 million bipolar neurons called *olfactory cells*. They contain hairs or *olfactory cilia*. The olfactory cells are *smell receptors*. They work as telereceptors, and the smell pathways do not include the thalamic relay station and a neocortical projection area. Instead, the olfactory cells pierce the cibriform plate and synapse in the olfactory bulb. The olfactory tract then transmits the olfactory signals to the olfactory cortex at the surface of the temporal lobe. In the *limbic system* (Fig. 4-3), olfactory information is correlated with feeding behaviour and emotional-motivational behaviour.

Fig. 3-12: The olfactory region, its receptors and pathways.

Pathophysiology

This paragraph deals with *1.The thalamic syndrome, 2. Brown-Sequards syndrome, 3. Special sensory pain disorders, 4. Taste and smell disorders*.

1. The thalamic syndrome

The thalamic syndrome is frequently caused by *thrombotic* blockade of bloodflow to the somatosensory thalamus. The destruction of thalamic neurons in one hemisphere leads to *ataxia* and *loss of sensations* from the opposite side of the body. After a few months different types of sensations return, but they are accompanied by pain.

2. The Brown-Sequards syndrome

The Brown-Sequards syndrome or paresis includes all effects of *transection* of only one half of the spinal cord at a certain level. All motor functions on the side of the lesion are blocked in the segments below the level (paresis, spasticity, and loss of vasoconstrictor tone). Sensations of pain and temperature from all lower dermatomes on the opposite side of the body are lost, because of transection of the contralateral spinothalamic tract. The only sensation left on the side of transection is crude touch, because it is transmitted in the opposite ventral spinothalamic tract. The total sensory loss is therefore termed *dissociated anaesthesia*.

3. Pathological pain

Hyperalgesia means *hypersensitivity to pain*. Hyperalgesia is caused by either hypersensitive pain receptors (sunburned skin), or by facilitated transmission. Facilitated transmission is due to abnormal stimulation of peripheral nerve fibres and neurons of the spinal cord or of the thalamus.

A special type of hyperalgesia is present when *herpes virus* infects one or more dorsal root ganglia. The virus excites the neurons and causes pain in the dermatomal segment subserving the ganglion. The segmental pain circles halfway around the truncus on the affected side. The virus is also transported by axonal flow to the cutaneous terminals, where it causes a characteristic rash confined to the dermatome. The disease is called *herpes zoster*.

Causalgia is *hyperalgesia and heat-cold-sensations* accompanied by sweat secretion in a region with nerve lesion. The hyperactivity in sympathetic efferent neurons and in nociceptive afferents running along arteries is unexplained.

Even the lightest touch at sensitised *trigger areas* release - within seconds - severe lancinating pain throughout the affected branch of the trigeminal nerve. The cause is unknown, and therapy is usually unsuccessful.

Referred pain and central pain

Pain that originates from deep organs is poorly localised and often referred coming from superficial structures. This may be explained by the fact that pain signals from viscera are transmitted through neurons in the CNS that also transmit pain signals from a specific area of the skin. Pain due to *myocardial ischaemia* (angina pectoris) is commonly described as pain originating from the inner side of the left arm, and termed *referred pain*.

Dermatome anaesthesia

The mammalian embryo is segmented into so-called *somites*, which are innervated by an adjacent part of the spinal cord. Every spinal dorsal root is destined to a segment of the skin called a *dermatome*. Also muscles (*myotomes*) bone (*sclerotomes*), and viscera are related to specific segments of the spinal cord or brain stem. The dermatomes are drawn

with sharp borders, which is unrealistic (Fig. 3-13). There is considerable overlap and several successive dorsal roots must be interrupted to produce *dermatome anaesthesia*. In case of serious spinal cord injury the dermatome map is useful for determination of the level and extent of the lesion.

Fig. 3-13: Dermatomes

Central pain is a sensation of pain in absence of peripheral nociceptive stimuli. Central pain is processed in the *cortical pain areas*, and caused by lesions along the nociceptive pathways (peripheral nerves, the spino-thalamo-cortical tract and the thalamus).

Amputation of a limb is sometimes followed by *phantom limb pain*. The patient suffers from severe pains, and the sensation is projected to the amputated limb. It is not known whether the mechanism is central or peripheral.

Trigeminal neuralgia (*tic douloureux*) is a condition with daily paroxysms of violent pain in part of the trigeminal area lasting only some seconds. The pain is provoked by eating, washing the face or by cold in the face. The mandibular and maxillary area is involved. The paroxysm finishes with saliva-tears- and sweat secretion. The disease is probably located to the trigeminal Gasserian ganglion, but the cause is usually unknown. Drugs or surgical procedures on the ganglion have variable effect.

Headache

Headache is caused by anomalies in intracranial or extracranial structures.

1. Intracranial headache is released from nociceptors in the meninges or in the arteries and veins at the base of the skull. The brain tissue itself hardly contains pain receptors. The intracranial types cover frontal, occipital, migraine, and meningeal psychogenic and pressure headache.

Nociceptors are stimulated by stretch (dura or tentorium), by dilatation (vessels), or by chemical means (eg. histamine, 5-hydroxytryptamine etc). The pain signals reach the CNS through the 5^{th} and 9^{th} cranial nerve and the cervical sensory fibres.

Stimulation of supratentorial nociceptors is referred to the frontal area via the 5^{th} cranial nerve as *frontal headache*.

Subtentorial nociceptors cause *occipital headache* through the 2^{nd} cervical nerve.

Migraine headache

Migraine or hemicrania means *unilateral headache*, which is frequently but not always present. Migraine is defined as *recurrent attacks* of headache associated with gastrointestinal and visual disorders. There is evidence for a genetic aetiology of migraine: an autosomal dominant inheritance with reduced penetrance (ie, expression only in permissive environments).

Classical migraine has prodromal symptoms (aura) with visual disturbances due to ischaemia in the retina. The onset is often in the eye region with spread towards the vertex or towards the other eye and typically accompanied with nausea, emesis, photophobia and *scintillating scotomata*. Sometimes sensorimotor abnormalities occur on one side of the body with ataxia, dysphasia and syncope.

Frequently migraine occurs without aura, which make the diagnosis difficult.

Migraine is of unknown origin, but it is sometimes associated with prolonged psychological stress. A hypothesis claims that prolonged emotional stress in sensitive individuals causes *reflex vasoconstriction* of intra- and extra-cranial arteries. The brain ischaemia explains the prodromal phenomena, and leads to accumulation of vasodilatating substances such as adenosine, ADP, NO etc. At the onset of the aura, the plasma concentration of 5-hydroxytryptamine rises, and it falls during the migraine attack.

After a brief period of aura the vessels dilatate, pulsate forcefully, and the walls become oedematous. These changes are believed to cause the *migraine headache*.

Food containing nitrites and tyramine may precipitate migraine attacks.

Severe attacks of migraine are treated with 5-hydroxytryptamine$_1$ agonists, such as sumatriptan, or with ergotamine tartrat.

Prophylaxis of migraine is carried out with *5-hydroxytryptamine antagonists* (pizotifen, methysergide) and b -*adrenergic blockers*.

Psychogenic headache varies in severity and location. This headache is generally accentuated by conflicts or by

anxiety with excessive sweating, tachycardia and hyperreflexia. This type of headache can be the first sign of depression, if the condition is worse in the morning following sleep disturbances.

Meningitis headache is accompanied by contraction of the neck muscles (stiff neck). The dura and the venous sinuses are inflamed, and the headache is severe.

Pressure headache. Intracranial mass lesions (tumours, abscess, bleeding, and traumata) are usually surrounded by brain oedema, whereby the basal vessels and the meninges are displaced and generate pain. This causes a special pressure headache, which is exaberated by supine rest, bending over, straining, sneezing and coughing. Any elevation of intracranial pressure induces this type of headache. Pressure headache is often accompanied by vomiting.

Subdural haematoma must be suspected after head trauma with pressure headache.

Suddenly occurring headache following a trauma may be caused by subarachnoid haematoma.

When combined with fever, neck stiffness, back stiffness, and vomiting the cause may also be meningitis, where a history of *sore throat* is frequently obtainable.

Fig. 3-14: The cutaneous innervation of the head and types of headache.

2. Extracranial headache is common and released from nociceptors in extracranial vessels, in the muscles of the head and neck or by inflamed mucous membranes of the sinuses, *sinus pain* (Fig. 3-14). The pain is felt directly over the frontal or the maxillary sinuses in the case of sinusitis.

Typical is the *muscle contraction headache* located in the frontal or occipital muscles. The frontal and/or occipital-nuchal muscles are tender. Frequently, both the occipital and the cervico-trapezial muscles are tense and tender with specific pain points (*loci dolendi*). Acupuncture or lazer therapy of these points is often effective. Treatment is difficult when depression or accident sequelae is the underlying cause.

3. Taste and smell disorders

The geniculate ganglion is a sensory ganglion for taste, which lies at the genu of the facial nerve. The nerve fibres join the facial nerve in the chorda tympani and carry taste from the anterior two-thirds of the tongue. Cranial lesions involving the petrous temporal bone cause *loss of taste* in this area together with an unpleasant loud distortion of noise called hyperacusis. Hyperacusis is due to paralysis of the stapedius muscle.

The sensory fibres of the glossopharyngeal nerve carry taste from the posterior third of the tongue. Cranial lesions involving the jugular foramen may damage the glossopharyngeal nerve often together with the vagus and the accessory nerve.

Loss of the ability to smell is called *anosmia*. Head injuries involving the cibriform plate or tumours damaging the sensory pathway may cause anosmia. Damage of the olfactory receptors in the nasal mucosa by upper respiratory infections may lead to anosmia.

Equations

- Stevens proposed the *power law* to account for the non-linearity of most physiological mechanisms. The interpreted stimulus strength (ISS) is equal to a constant (k) multiplied by the actual stimulus strength (SS) raised to the power n:
    - Eq. 3-1: $ISS = k*SS^n$.

In his original version, only the exponent **n** differed for each type of sensation. The equation can be modified by subtracting different constants from SS before raising it to the power **n**, or by changing the value of **k**.

Self-Assessment

Chapter 3. Multiple Choice Questions

I. Each of the following statements has True/False options:

- A. The somatosensory thalamus is a relay station for most sensory modalities.
- B. Glutamate is the main inhibitory transmitter in the CNS, whereas GABA is the dominant excitatory transmitter.
- C. Pain and temperature reach the thalamus through the lateral spinothalamic tract.

- D. Presynaptic transmission of opiates inhibits depolarization of nerve cell membranes.
- E. The adequate stimulus of the cutaneous mechanoreceptors is deformation of the receptor.

II. Each of the following statements has False/True options:

- A. The Nissl bodies are stacks of rough endoplasmic reticulum.
- B. Taste, heat and angular acceleration follow transfer functions, so the interpreted stimulus strength decreases with the rise in actual stimulus strength.
- C. Sensory receptor systems are biological transducers with a dynamic range up to 10^{12}.
- D. The B nerve fibres are autonomic preganglionic axons with a diameter less than 3 m m and a conduction velocity of 3-12 m per s.
- E. The CSF has higher concentrations of K^+, glucose, and protein than blood plasma, and lower concentrations of Na^+ and Cl^-.

Case History A

A female of 32 years is admitted to a neurosurgical ward with a discrete lesion in the spinal cord caused by a traffic accident. Her vital functions are unaffected. The most important signs are a complete lack of cutaneous temperature sensibility and pain sensibility in the left leg and the lower left side of the trunk below the umbilicus (the navel).

Where in her spinal cord is the lesion localised?

Try to solve the problems before looking up the answers

Highlights

- *The nervous system is a rapid signal transduction system and the main communication network in our body. The integrative functions allow selection and processing of incoming signals to produce an appropriate response.*
- *The nervous system includes sensory receptors that detect events in the body as well as in the outer world. Several sensory tracts and pathways synapse in the nuclei of the thalamus (the spinothalamic tract). The somatosensory thalamus is a relay station for many sensory modalities.*
- *Neuroglia is supportive cells that sheath and protect neurons. Myelinated axons propagate APs up to 50 times faster than unmyelinated with the same diameter. Neuroglia also eliminates transmitters more rapidly from the synapse. The neuroglia constitutes half of the brain volume, and there are about 10^{12} to 10^{13} glial cells in the human brain.*
- *Sensory receptors in the nervous system are classified as exteroceptors (located on the body surface), proprioceptors (located in muscles, tendons and joint capsules), interoceptors (located in the viscera), and telereceptors (stimulated by events far from the person).*
- *Sensory receptors are either neurons in the case of vision, smell and cutaneous senses, or modified epithelial cells in the case of auditory, vestibular, smell and taste senses.*
- *C fibres (IV) are unmyelinated and lead pain, touch and heat signals from the skin.*
- *The gate-control hypothesis of pain states that pain transmission is suppressed by innocious signals in thick myelinated afferents (group II), whereas the pain sensation is enhanced by signals in thin afferents. Inhibitory interneurons in the dorsal horn of the spinal cord perform the gate-control through a special type of presynaptic inhibition called primary afferent depolarization (PAD), and the receptors on the cell body of the secondary neuron is the gate.*
- *All taste fibres synapse in the nuclei of the solitary tract and the axons of these neurons project to the thalamus. From the thalamus third-order neurons reach the lower part of the primary sensory cortex in the postcentral gyrus (somatosensory area I).*
- *Dermatome anaesthesia. The mammalian embryo is segmented into so-called somites, which are innervated by an adjacent part of the spinal cord. Every spinal dorsal root is destined to a segment of the skin called a dermatome. Also muscles (myotomes) bone (sclerotomes), and viscera are related to specific segments of the*

spinal cord or brain stem.

- *Loss of the ability to smell is called anosmia. Head injuries involving the cibriform plate or tumours damaging the sensory pathway may cause anosmia. Damage of the olfactory receptors in the nasal mucosa by upper respiratory infections may lead to anosmia.*
- *The thalamic syndrome is frequently caused by thrombotic blockade of bloodflow to the somatosensory thalamus. The destruction of thalamic neurons in one hemisphere leads to ataxia and loss of sensations from the opposite side of the body. After a few months some of sensations return, but they are often accompanied by pain.*
- *Meningitis headache is accompanied by contraction of the neck muscles (stiff neck). The dura and the venous sinuses are inflamed, and the headache is severe.*
- *Migraine headache begins with prodromal nausea and vision disturbances, often occurring about one hour prior to the headache. The pain is located on one side of the head in classical cases.*

Further Reading

- Caterina, MJ, MA Schumacher, M Tominaga, TA Rosen, JD Levine, and D Julius. The capsacain receptor: A heat-activated ion channel in the pain pathway. Nature 389: 816-24, 1997.
- Kandel, E.R., J.M. Schwartz and Jessop. "Principles of neural science." New York: *Elsevier Science Publ. Co.*, 1991.
- Russell, M.B. and J. Olesen. "The genetics of migraine without aura and migraine with aura." *Cephalalgia* 13 (4): 245-8, 1993.

Return to content

Return to chapter 3

Neuronal Transmembrane Potentials

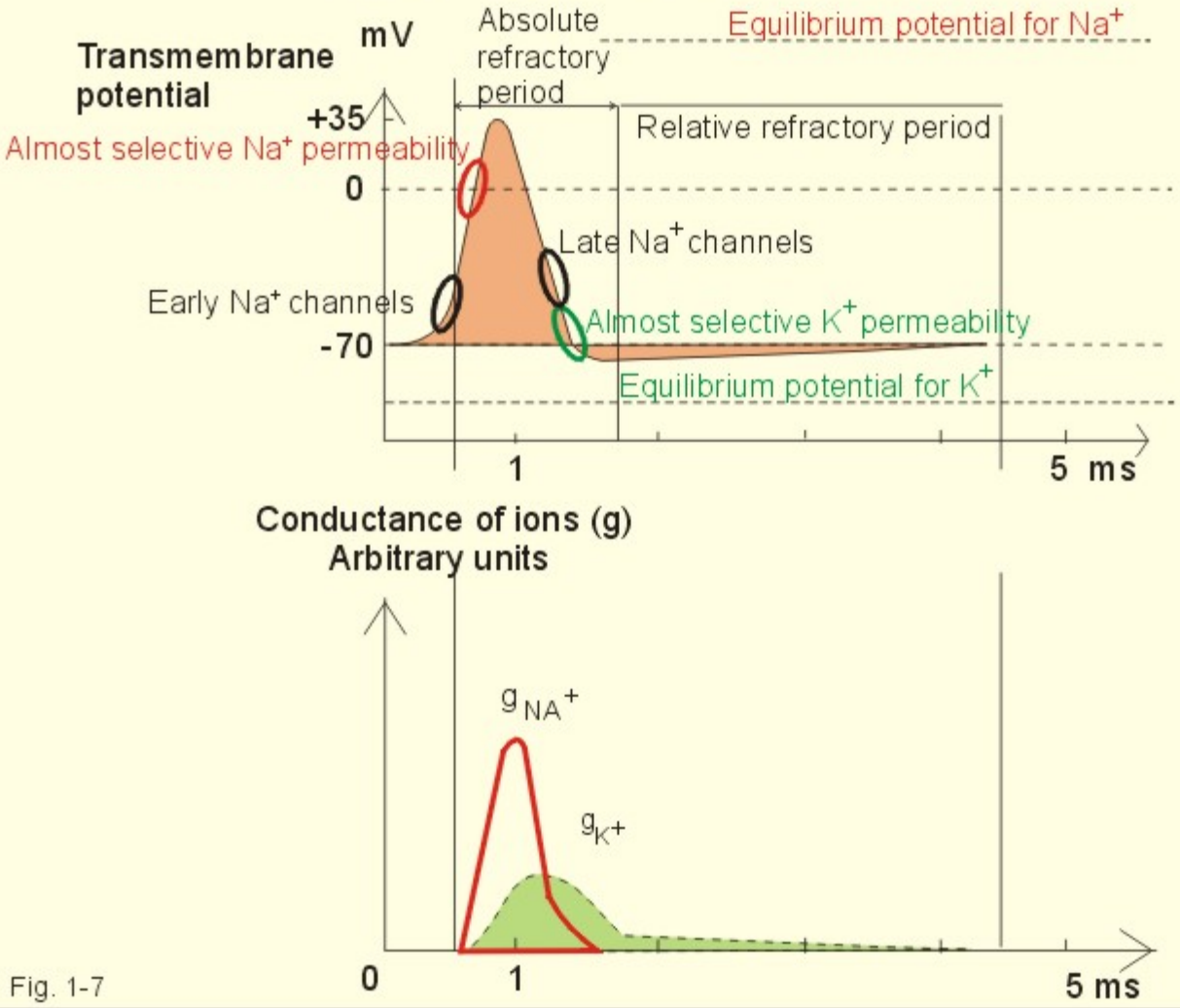


Fig. 1-7

KMc

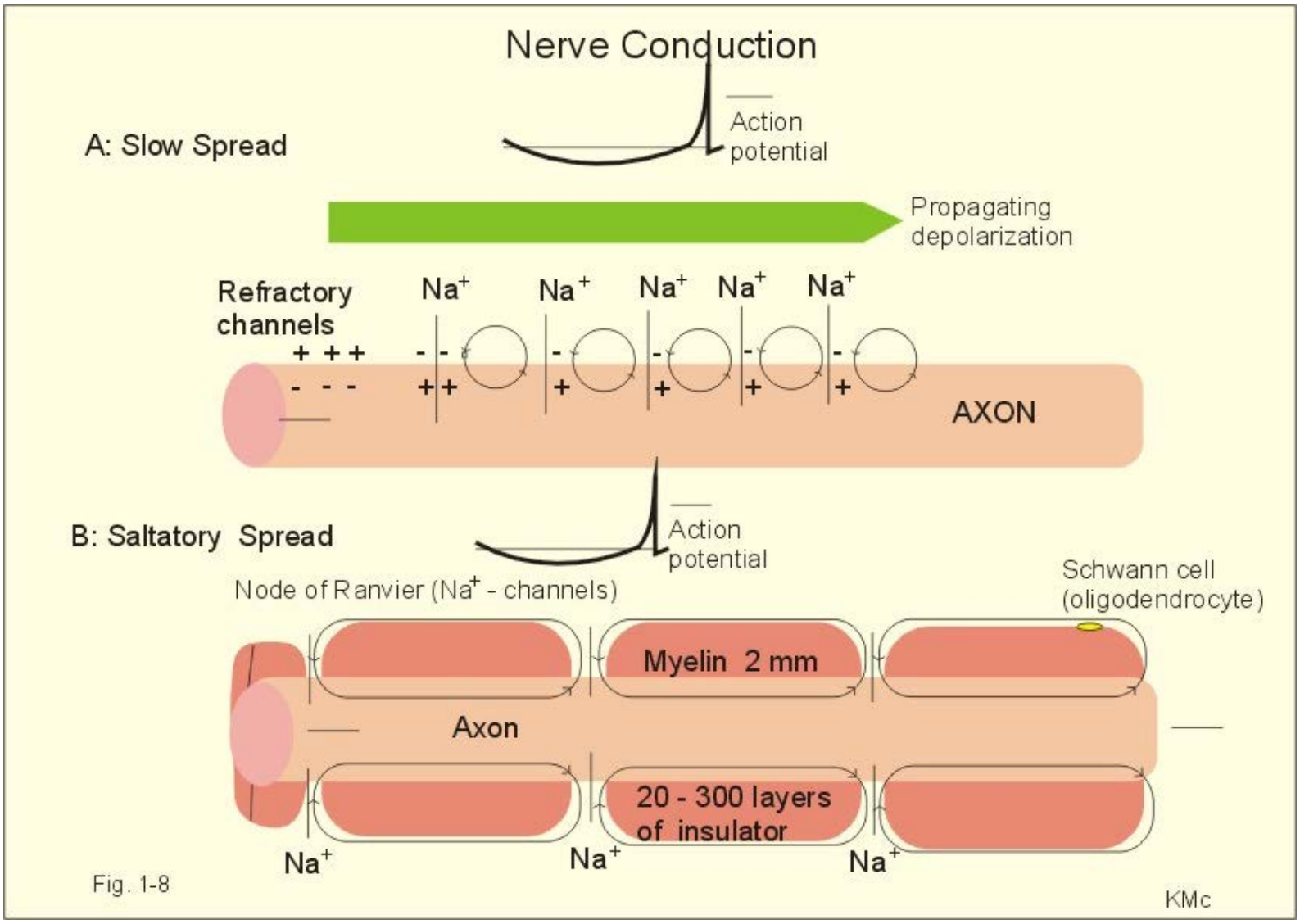

Nerve Conduction
A: Slow Spread
Action potential
Propagating depolarization
Refractory channels
Na+
AXON
B: Saltatory Spread
Action potential
Node of Ranvier (Na+ - channels)
Schwann cell (oligodendrocyte)
Myelin 2 mm
Axon
20 - 300 layers of insulator
Na+
KMc

Fig. 1-8

Chapter 30.

Other Hormones And Disorders

Study Objectives

- To *define* or *describe* concepts such as catecholamines, nanism, pituitary dwarfism, pseudo-hypo-para-thyroidism, somatomedins, somatostatin, and somatotropin.
- To *describe* the regulation of extracellular and intracellular Ca^{2+}, the bone structure, remodelling and the function of osteoblasts, osteoclasts and osteocytes. To describe the biosynthesis of calcitonin, parathyroid hormone, vitamin D, mineralocorticoids, glucocorticoids, adrenaline and noradrenaline. To describe osteoporosis, osteomalacia, and rachitis. To describe the Ca^{2+} and phosphate balance in healthy persons.
- To *explain* the effects of calcitonin, parathyroid hormone, vitamin D, mineralocorticoids, glucocorticoids, androgens and oestrogens from the adrenal cortex, and adrenaline/noradrenaline. To explain phenomena such as phaeochromocytoma, shock, adrenogenital syndrome, virilization and pubertas praecox.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The sympatho-adrenergic system give rise to the fright, flight or fight reactions that are all acute stress situations.*
- *The adrenal cortex is concerned with the carbohydrate metabolism and with the electrolyte balance.*
- *The parathyroid hormone and vitamin D maintain normal levels of calcium and phosphate in the body.*

Definitions

- **Calcium concentration** (total) in plasma: 2.2-2.7 mM. Half of the total calcium is ionized.
- **Catecholamines** are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The important catecholamines in humans are adrenaline, noradrenaline and dopamine.
- **Cushing's disease** is hypercorticism (increased glucocorticoid production) caused by a pituitary basophilic adenoma.
- **Cushing's syndrome** refers to the consequences of increased plasma glucocorticoid concentration from any source.
- **Growth hormone releasing hormone** (GHRH) is a peptide produced by GHRH cells in the hypothalamus. GHRH stimulates the release of growth hormone (Somatotropin) from the adenohypophysis.
- **Hypercalcaemia** is an increase in the plasma concentration of ionized calcium. The condition is characterised by neurological signs (hyporeflexia, lethargy or coma). When Ca^{2+} is deposited in the inner medulla of the kidney, ADH resistance develops with polyuria and polydipsia.
- **Hypocalcaemia** is a decrease in the plasma concentration of ionized Ca^{2+} (cramps, tetany, twitching and tingling of the fingers). Hypocalcaemia is caused by renal failure or by thyroidectomy and parathyroidectomy.
- **Hyperparathyroidism.** Primary hyperparathyroidism is caused by parathyroid adenomas or by hyperplasia. Secondary hyperparathyroidism is compensatory hypertrophia due to hypocalcaemia.

- **Hypoparathyroidism.** The idiopathic form is a rare autoimmune destruction of the parathyroid (with hypocalcaemia).
- **Infantile nanism** is a term used to indicate that the somatic age of a child is below its chronological age.
- **Nanism** or *dwarf growth* expresses that the height of an adult is below a certain limit. In the Anglo-Saxon countries the limit is 1.40 m for females and 1.50 m for males.
- **Pituitary dwarfism** refers to low pituitary secretion of GH or insensitive target organ receptors (African pygmies).
- **Pseudo-hypo-para-thyroidism** is used for conditions, where target-organs (bone, kidneys, and gut) are *resistant to PTH*. The plasma- [Ca^{2+}] is low, plasma-[phosphate] is high, and basic phosphatase activity is high.
- **Somatomedins** are small peptides (5 kDa) produced in the liver, promote bone growth and protein synthesis. Somatomedin C is also called insulin-like growth-factor -1 (IGF-1). The hepatic production of somatomedins is stimulated by growth hormone. Somatomedins stimulate the hypothalamic secretion of somatostatin.
- **Somatotropin or growth hormone** (GH) from the anterior pituitary stimulates body growth in humans. GHRH stimulates and somatostatin inhibits the release of GH. Growth hormone stimulates the hepatic production of Somatomedin C (IGF-1).
- **Somatostatin** or growth hormone inhibiting hormone (GHIH) is a cyclic peptide, which is produced in the pancreatic islets and in the CNS, where it is co-transmitter with noradrenaline. GHIH inhibits the release of growth hormone from the adenohypophysis and is therefore also called somatotropin releasing inhibiting factor, SRIF.

Essentials

This paragraph deals with 1. Growth hormone, 2. Parathyroid hormone. Calcium and phosphate homeostasis, 3. Adrenal corticoids, 4. Adrenal catecholamines, and 5. Vitamin D and bone minerals.

1. Growth hormone (GH)

Growth hormone is the main stimulator of body growth in humans. Growth hormone is produced in the pituitary, stimulated by hypothalamic growth hormone releasing hormone (GHRH) and inhibited by hypothalamic *somatostatin* (Fig. 30-1). Pituitary growth hormone stimulates the hepatic production of Somatomedin C (IGF-1) and its specific *IGF-binding protein 1* (IGF-BP1). Actually, it is somatomedin C that stimulates body growth - both bone and muscle growth.

Fig. 30-1: Pituitary GH and hepatic IGF-1 secretion in the hypothalamic-pituitary feedback system.

Growth hormone releasing hormone (GHRH) is produced by GHRH cells in the hypothalamus and reaches the adenohypophysis via the portal system. GHRH stimulates the release of growth hormone (GH) from the adenohypophysis (Fig. 30-1).

Growth hormone and insulin are important anabolic hormones in humans, although growth hormone has many insulin-antagonistic effects. Growth hormone is released as *hunger spikes* during the day and during sleep.

Growth hormone (GH) produced in the placenta differs from the *pituitary GH* by a few Amino acid residues. *Placental GH* suppresses release of maternal, pituitary GH during Pregnancy. Placental GH stimulates maternal metabolism and foetal cell proliferation and Hypertrophia.

Foetal thyroid hormones stimulate brain development, and *foetal insulin* stimulates foetal Growth, cellular glucose uptake and glucose utilisation. Paracrine and autocrine growth Factors are also important for foetal growth: Insulin-like growth factor-II (IGF-II), nerve Growth factors (NGF), epidermal growth factor (EGF), and platelet derived growth factor (PDGF).

The *pituitary GH* has an endocrine effect on the production of the growth factor *Somatomedin C* (IGF-1) and its specific binding protein in the liver. Hepatic IGF-1 circulates in plasma bound to *IGF-binding protein 1* (IGF-BP1).

The binding proteins are important, not only as a vehicle in plasma, but also for the final binding of the hormone to its cell membrane receptor. *Hepatic IGF-I* stimulates bone formation, and GH stimulates the precondrocytes directly. *Locally released IGF-I* stimulates the condrocytes in healthy states, but it also has a regenerative function in damaged tissues.

Pituitary GH also stimulates the production of receptors for other growth factors. GH serves to commit a precursor cell to a specific pathway, and IGF-I enhances its growth and replication. Somato
the somatotropic cells of the adenohypophysis. Somatomedins also stimulate the hypothalamic secretion of *somatostatin* (GHIH or somatotropin releasing inhibiting factor, SRIF).

Tissue specific growth factors - just like those important for foetal growth - are produced in damaged tissues, where they are important for regeneration. Nervous tissues produce NGF, epidermal tissues produce EGF, thrombocytes produce PDGF, and fibroblasts produce both *fibroblast growth factor* (FGF) and *transforming growth factors* (TGF-a, and TGF-b). Hepatocyte proliferation after liver damage is induced by both *hepatocyte growth factor* (HGF) and by *hepatocyte stimulating substance* (HSS). Growth factors are used in regenerative therapy.

Fig. 30-2: 68-1: Growth curves in boys - a similar pattern is found in girls.

The relative large growth of the testes at puberty shows the importance of *sex hormones* for early puberty growth (Fig. 30-2). In general, the growth spurt in early puberty is caused by increased production of *sex steroids* (stimulated by luteinizing hormone and by follicle stimulating hormone, FSH), because the sex steroids stimulate hypothalamus to release *more growth hormone.*

Growth hormone and insulin are probably the most important *anabolic hormones* in the body - more effective than oestradiol and testosterone; however, GH has many insulin-antagonistic effects. Human GH stimulates protein synthesis, mitosis in cells, chondrogenesis, ossifica
anaerobic breakdown of glycogen). GH stimulates hepatic glucose production (glycogenolysis) but not its gluconeogenesis. GH enhances RNA synthesis, accelerates glucose uptake and antagonise the lipolytic effect of adrenaline. - GH is produced in spikes during the day (hunger spikes) and during deep sleep (EEG stage III and IV), and the half-life of GH is 20 min.

Children with *pan-hypopituitarism* become *infantile dwarfs* and later turn into *juvenile dwarfs*, as they do not develop sexually.

Children with *hyperpituitarism* become giantesses or giants. The clinical diagnosis is *gigantismus.*

Hypothyroid children (cretins) also become dwarfs. The *thyroid hormones* are necessary or *permissive* for normal growth and development. *Cretins* are *mentally retarded, hypothyroid dwarfs.*

Sex steroids in high concentrations close the epiphyseal lines. If this occurs early in life, the child also becomes an infantile dwarf, but such an individual is often sexually active. The clinical diagnosis is *precocious puberty* or *pubertas praecox.*

Without proper *insulin treatment*, children with diabetes (mellitus) become dwarfs (diabetic, infantile nanismus), because of the intracellular hypoglycaemia. Poorly regulated diabetics also develop *osteopenia* (bone decalcification).

Genetic factors are essential for optimal growth and development, as shown by the strong correlation between the height of the parents and the final height of the child. Tall people are high before the puberty growth spurt, which is more or less the same for tall and shorter people. - Persons with only one sex chromosome (X, 0) show retarded growth from birth, whereas persons with an extra sex chromosome (XXY; XYY) become tall.

Optimal growth depends upon *optimal nutrition* (essential amino acids, vitamins, minerals, and fatty acids) and optimal Neuroendocrine, metabolic control.

Optimal growth also depends upon an *optimal health.* Most disease states and all immuno-defence threatening treatments retard growth or imply weight loss. Such disorders cause catabolic hormones to dominate and induce anorexia.

2. Parathyroid hormone (PTH). Calcium and phosphate homeostasis.

We overlooked the existence of the parathyroid glands until the consequence of their surgical removal was realised.

Without the parathyroid, a person develops *tetanic cramps* due to a fall in the concentration of ionised calcium ([Ca^{2+}]) in the blood plasma (explained in Chapter 17). - Already in 1909, MacCallum treated this condition successfully with calcium salts.

The *chief cells* (C cells) of the parathyroid glands produce *pre-proparathyroid hormone*, which is cleaved in their endoplasmic reticulum to *proparathyroid hormone*. This in turn is cleaved in the Golgi apparatus to *parathyroid hormone* (PTH). PTH is a single chain peptide with a molecular weight of 9500 Da.

PTH is water-soluble and binds to membrane receptors on the surface of the target cells. Thus the PTH action is dependent on second messengers. Both intracellular Ca^{2+} and cyclic adenosine monophosphate (cAMP) are used.

Humans carry a single PTH gene on *chromosome 11*. The chief cells also produce a *parathyroid secretory protein* of unknown function. In man, the four parathyroid glands are located just behind the thyroid gland. Ectopic tissue sometimes develops in the mediastinum or in the neck.

PTH binds to *membrane receptors* on all target cells. The major effects of PTH are on *three target organs:*

1. Bone: PTH accelerates the removal of Ca^{2+} and phosphate from bones (osteolysis by surface osteocytes, resorption of bone by osteoclasts). PTH stimulates the osteolysis by surface osteocytes causing the release of Ca^{2+} for rapid equilibrium with the ECV. After 12 hours, the delayed effect of PTH, which stimulates the osteoclasts to reabsorb mineralised bone, sets in.

2.Kidney: PTH reduces the reabsorption of Ca^{2+} and phosphate from the proximal tubules, and increases the reabsorption of Ca^{2+} from the distal tubules, frequently resulting in an increased net loss of Ca^{2+} in the urine. PTH binds to the basolateral membrane of the tubule cell, and stimulates cAMP, which in turn diffuses through the cell to the luminal membrane. Here cAMP activates a Ca^{2+} -reabsorption port.

The glomerular filtration of Ca^{2+} is easy to calculate, since approximately half the total plasma concentration is free and filterable (2.5/2 mM). Since 0.94 parts of plasma is water, the [Ca^{2+}] in the ultrafiltrate is (1.25/0.94) 1.33 mM. A person with an average *glomerular filtration rate* of 0.125 l per min will produce a 24-hour ultrafiltrate of 180 litre.

Thus, a total Ca^{2+} flux of (1.33×180 =) 239 mmol or almost *10 000 mg daily,* will pass the glomerular *ultrafilter*.

Fortunately, almost all Ca^{2+} is reabsorbed in the kidney tubules (about 67% is reabsorbed in the proximal, and the reabsorption of the balance in the distal tubules is regulated by PTH). We only excrete 100 to 200 mg daily (or 2.5 to 5 mmol daily) in the urine and 50 mg or 1.1 mmol through the skin. This is a daily maximum of *250 mg* (or 6 mmol) *Ca^{2+} excretion from the body*.

Another drastic action of PTH is on the proximal tubules, where PTH *inhibits* phosphate reabsorption so efficiently that its excretion in the urine increases within 5 minutes.

3.Gut: The PTH action on the gut is indirect. PTH stimulates the renal production of biologically active vitamin-D (1,25-dihydroxy-vitamin D), which stimulates the active absorption of Ca^{2+} and phosphate across the gut mucosa (Fig. 30-3), and potentiates the action of PTH on bone resorption.

Fig. 30-3: The normal calcium transfer (all numbers are in mg per day or in [mmol per day]). The balance is 250 mg per day.

The end result of these three actions is an increase in plasma [Ca^{2+}] and a decrease in plasma [phosphate]. Simultaneously, the bone resorption activity is illustrated by a high basic phosphatase concentration. *Cystic areas*, corresponding to periostal reabsorption, are visible on bone radiographs (osteitis fibrosa cystica). Hypercalcaemia predisposes to kidney stones and to *metastatic calcification* in synovial membranes and meninges, in the lungs, kidneys, pancreas and elsewhere. Certain tumours produce *PTH-related protein* with sequence homology to PTH. Perhaps this substance is related to the *hypercalcaemia of malignant processes* (see ectopic tumours).

The daily need of phosphate is 18 mmol just as the daily need of Ca^{2+} .The human body contains 25 mol of phosphorous (31 Da) as phosphate. Out of the 25 mol, 20 mol is located in bones and about 4 mol in muscle and other

soft tissues (Fig. 30-4). A daily food intake of 46 mmol of phosphate (equal to the Ca^{2+} intake) with a secretion of 3 mmol and an intestinal absorption of 39 mmol, leaves (10 + 36) mmol daily for faecal and renal excretion in order to be in balance. The *normal* plasma concentration of phosphate is 1 mM. This implies a daily phosphate filtration rate (net-flux) of 180 mmol (1 mM*180 l of plasma) - at a glomerular filtration rate of 0.125 l per min. Phosphate is a threshold substance with a tubular reabsorption capacity (T_{max}) of 0.1 mmol per min; this means that 80% of the filtered load is reabsorbed. The *maximal* daily filtration- reabsorption- and excretion flux is calculated in Fig. 30-4.

Most of the reabsorption of phosphate takes place in the proximal tubules, where PTH inhibits phosphate reabsorption. Renal control maintains the phosphate concentration in blood plasma. The T_{max} for phosphate is up regulated by high phosphate intake and down regulated by low phosphate intake in the food. The daily exchange of phosphate from bone is 8 mmol and from soft tissue 16 mmol (Fig. 30-4). – Phosphate depletion may leads to muscle weakness (cardiac and skeletal muscles) and pathological bone formation.

Fig. 30-4: The daily phosphate balance.

The binding of PTH activates adenylcyclase and this raises the [cAMP], which interacts with protein kinase A. The enzyme then catalyses the *phosphorylation of effector proteins.* PTH is secreted in response to hypocalcaemia, in particular, a low [Ca^{2+}] (or low [Mg^{2+}]). The major effects of PTH are to increase plasma [Ca^{2+}] and decrease plasma [phos
reactions and important for the neuromuscular transmission.

3. Adrenal corticoids

The adrenal cortex of the adult human has three layers: The *outer* zona glomeru
narrow, the *middle* zona fasciculata is wide, and the *inner* zona reticu
The adrenal cortex is of mesodermal origin.

Steroids

The human adrenal cortex produces three types of steroids: Glucocorticoids, mineralocorticoids, and a minimal amount of sex steroids (androgens and oestrogens). All steroids are lipid-soluble and easily cross the lipid membrane. All steroids represent chemical modifications of four-ring struc
phenantren.

Steroids bind to specific *cytosol-receptor proteins* that are then translocated to the cell nucleus. Here they reversibly bind to DNA.

The synthesis of *glucocorticoids* (cortisol and corticosterone) occurs in the *zona fasciculata* with a small contribution from zona reticularis. The *mineralocorticoid*, aldosterone, is produced in no other region of the cortex than *zona glomerulosa.* The synthesis of *sex hormones* (androgens and oestrogens) occurs mainly in *zona reticularis.* The precursor for these hormones is *cholesterol* absorbed from the blood HDL and LDL fractions by the cortex cells. Most of the synthetic reactions involve *mixed oxygenases* (belonging to the cytochrome P-450 enzymes) localized in the endoplasmatic reticulum and in the mitochondria.

During ACTH stimulation the size and number of cells in the zona fasciculata and zona reticularis increase, mainly because the cortisol and the sex hormone production increase. The mitochondria, central ribosomes, vesicular cristae and endoplasmic reticulum grow in these cells. ACTH activates all steps in corticosteoid hormone synthesis.

A *microsomal desmolase* removes C_{20-21} from the precursors, pregnenolone and progesterone (C_{21} steroids). The residues are dehydroepi-androsterone and androstenedione (C_{19}). These androgens are weak and are converted to a more potent form, testosterone, in peripheral tissues.

In the zona reticularis, testosterone is converted to oestradiol (C_{18}), due to removal of a CH_3 group by *aromatase.* The gene for the human androgen receptors is found on the *X chromosome*. The receptor protein has a molecular weight of 98 kDa and is found both in the cytosol and the nucleus.

CRH & ACTH

The hypothalamic *Corticotropin Releasing Hormone (CRH)* stimulates the secretion of ACTH. Stress stimulates not only the sympatho-adrenergic system with catecholamine release, but also the CRH/ACTH release with increased

secretion of neuroregulatory peptides and cortisol. This is important, since small amounts of *cortisol* have *permissive effects* on catecholamines, while inhibiting TSH. Stress also releases *growth hormone* (GH that stimulates glycogenolysis/glycolysis) and *prolactin*, both from the acidophilic cells in the adenohypophysis. Prolactin released by stress is possibly mediated by *hypothalamic histaminergic neurons.*

ACTH binds to the cells of zona fasciculata and activates adenylcyclase, which results in a rise in the cAMP level. The major effect of ACTH - through increased cAMP level - is to stimulate the conversion of cholesterol to *pregnenolone* by desmolase. This is the rate-limiting step in the production of cortical steroids! Plasma levels of *cortisol* (hydrocortisone), *adrenal androgens* and their precursors rise within three minutes of intravenous ACTH injection. Cortisol inhibits the release of CRH and ACTH by *negative feedback.*

Glucocorticoids

Kendall isolated cortisone, and Hench & Reichstein directed the first administration of cortisone and ACTH to patients with rheumatoid arthritis. The result was a dramatic improvement. They shared the Nobel Prize in 1950. The serious side effects were recognized later.

The gene for the *human glucocorticoid receptor* is present in *chromosome 5*. The receptor protein (94 kDa) is found in both the cytosol and the nucleus.The effects of a glucocorticoid such as cortisol are physiologic or pharmacological dependent upon the dose.

Cortisol is the main endogenous *glucocorticoid* synthesized in the adrenal cortex.

Cortisol is essential for life and acts permissively to facilitate the mobilisation of fuels.

Cortisol defends the body against hypoglycaemia evoked by insulin.

Cortisol stimulates hepatic glucose production (gluconeogenesis).

Cortisol augments the glucagon stimulation of glycogenolysis.

Cortisol inhibits the glucose uptake in target cells (GLUT 4 in muscle cells, heart cells and adipocytes).

Cortisol is diabetogenic.

Cortisol is lipolytic and acts permissively on adrenaline, GH and other lipolytic substances to mobilize triglycerides. Cortisol also suppresses CRH release and an excess of cortisol may lead to truncal obesity.

Cortisol induces leptin synthesis in fat cells. This synthesis limits the appetite by negative feedback.

Cortisol maintains the contractility of striated and cardiac muscle (the Na^+- K^+-pump, b-adrenergic receptors etc). Excess cortisol increases muscle metabolism and reduces muscle mass and strength. Cortisol decreases the ratio of insulin-sensitive slow oxidative type I muscle fibres to the fast glycolytis type II-B muscle fibres.

Cortisol reduces the differentiation to active osteoblasts and the synthesis of collagen in bone matrix and connective tissue. Cortisol antagonises the action of 1,25-dihydroxy-cholecalciferol and thus the absorption of Ca^{2+} from the gut – cortisol excess leads to osteoporosis.

Cortisol is permissive to the vasoconstriction of catecholamines and angiotensin II.

Cortisol increases renal bloodflow and GFR.

Cortisol is important for the development of CNS, skin, gut, bone marrow and lungs. Glucocorticoids stimulate erythropoiesis.

Cortisol inhibits processes involved in inflammation, infections, tissue damage, and immune system reactions. Glucocorticoids inhibit most responses mediated by leucotrienes, NO, PAF, prostaglandins, and thromboxanes. Therapeutic doses of glucocorticoids are used to treat more diseases than any other group of drugs, because the hormones act anti-inflammatory and anti-allergic. The negative effects are delayed healing of wounds and increased gluconeogenesis with destruction of tissue proteins.

In plasma, most of the cortisol binds to *transcortin*, to corticosteroid binding globulin (CBG), or to albumin and only 5% is free. The protein binding means that cortisol has a long plasma-$T_{½}$ of more than 70 min. Oestrogens stimulate

the production of cortisol binding proteins. Patients with liver and kidney diseases do not produce enough. – Transcortin-cortisol complexes activate adenylyl cyclase, and the complex liberates free colesterol at sites of inflammation.

A low concentration of *hormone specific globulin* implies a low concentration of the *bound* hormone fraction, but the *free* hormone concentration can still be the same.

Cortisol is converted into tetrahydrocortisol, cortisone or to 17-ketosteroids. Cortisone is a biologic analog to cortisol. Exogenous cortisone is an important source of cortisol in many tissues due to the presence of the enzyme, 11beta-hydroxy-dehydrogenase.

Conjugation with glucuronic acid or sulphate forms water-soluble products. Only minimal amounts of the daily cortisol secretion is excreted in the urine. The fraction excreted in the bile is released to the enterohepatic circuit.

Half of the daily cortisol secretion is excreted in the urine as 17-hydroxycorticoids.

The plasma cortisol concentration is increased by anxiety, burns, fever, exercise, hypoglycaemia, pain, severe physical or psychiatric disease, stress and surgery.

The plasma cortisol concentration decreases promptly following administration of synthetic glucocorticoids (dexamethasone), which suppress ACTH secretion by negative feedback.

Mineralocorticoids

Aldosterone is the major *mineralocorticoid* with corticosterone contributing only little. Aldosterone maintains ECV by conserving body Na^+. When body Na^+ is depleted, the fall in ECV and plasma volume reduces renal bloodflow and pressure. A reduction in circulating fluid volume recorded in the kidneys releases aldosterone.

Aldosterone promotes the reabsorption of Na^+ and increases the secretion of K^+ and H^+ in the distal tubules and the cortical collecting ducts. A rise in plasma-[K^+] from normal or a fall in plasma-[Na^+] also releases aldosterone by direct action on zona glomerulosa.

The *renin-angiotensin-aldosterone cascade* (Chapter 24) controls the adrenal aldosterone secretion, not the ACTH.

Only a small fraction of the aldosterone binds to proteins. This makes its half-life *short* (20 min). Aldosterone is reduced to *tetra-hydro-aldosterone* in the liver and is conjugated with glucuronic acid.

Aldosterone like all other steroids is lipophilic, so that it passes easily through the cell membrane. The human mineralocorticoid receptor is found primarily in the cytoplasm, and its gene is present on *chromosome 4.* This receptor has been cloned revealing a molecular weight of 107 kDa. When aldosterone is bound to the receptor, the receptor reveals a *DNA-binding site.*

The receptor-aldosterone complex translocates from the cytosol to the nucleus, where it binds to chromatin by means of the DNA-binding site. This process activates the transcription of the specific gene producing mRNA, which is then translated into specific proteins. Following complete transmission, the *receptor-aldosterone complex* dissociates from chromatin and each other. The receptor returns to the cytosol with a *hidden DNA-binding site.*

By this process aldosterone promotes the synthesis of new proteins, that may stimulate the Na^+-K^+-pump or facilitate Na^+ entry into the tubular cell through integral Na^+ -channel proteins in the luminal membrane.

Increased Na^+-reabsorption from the tubules due to aldosterone causes a simultaneous reabsorption of water and thus an increased ECV, with increased arterial blood pressure. The pressure rise leads to increased GFR. The rapid filtration flow overrides the high reabsorptive effect of aldosterone, which down-regulates the size of ECV. – Beta-adrenergic blockers lower the arterial pressure in part by reducing the Na^+ -retention caused by the renin cascade. Inhibitors of angiotensin converting enzyme or angiotensin II receptor blockers also lower the arterial blood pressure. Antagonists of aldosterone at the renal distal tubule cells directly prevent Na^+ -reabsorption and reduce hypertension.

The ACTH effect is only a moderate stimulation in acute situations. The action of *angiotensin II* on aldosterone secretion is much more important.

The adenohypophysis probably produces an *aldosterone-stimulating factor*. Dopamine (Prolactin inhibiting hormone, PIH) is an *aldosterone inhibitor*.

Adrenal sex corticoids are mainly weak adrenal androgens produced largely in zona reticularis. By conversion to testosterone and dihydrotestosterone in peripheral tissues, the adrenal androgens play a physiological role. Also a small oestrogen production is present in healthy persons. In females, adrenal androgens produce testosterone for normal pubic and axillary hair development and erythrogenesis.

4. Adrenal catecholamines

The *medulla* is a *modified sympathetic ganglion* derived from neuroectodermal cells. The development of the sympathoadrenergic nervous system is stimulated by neural growth factors.

The adrenal medulla is the source of the circulating catecholamine, adrenaline. The medulla also secretes small amounts of nor-adrenaline, normally a neurotransmitter.

The *sympathetic system* consists of short preganglionic fibres, which have tropic centers in the lateral horn of the spinal cord from the first thoracic segment to the third lumbar segment.

The myelinated nerve fibres form ramus communicans albus and pass to the paravertebral or *lateral* sympathetic ganglion chain. Here most of the fibres end on the postganglionic cell bodies. Some preganglionic fibres pass the sympathetic chain and reach *collateral ganglia* such as the *solar plexus*, and the *ganglion mesentericus superior* and *inferior*.

In these lateral and collateral relay stations each preganglionic fibre is in contact with many cell bodies. These cells have long, postganglionic, unmyelinated fibres serving the different organs.

The preganglionic fibres to the adrenal medulla pass all the way to the chromaffin cells in the adrenal medulla. These *chromaffin cells* (high affinity for chromium cell stains) are embryological analogues to postganglionic fibres and the synapse is cholinergic.

Synthesis of catecholamines

Catecholamines are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The catecholamines are synthesises within the chromaffin cell.

The dietary amino acid *tyrosine* is absorbed as tyrosine from the gut (Fig. 30-5). Tyrosine is also produced in the liver from dietary phenylalanine.

The L-tyrosine is hydroxylated by *tyrosine hydroxylase* in the cytosol of the chromaffin cell, and the product is L-dihydroxy-phenyl-alanine (Dopa, Fig. 30-5). Tyrosine hydroxy
production. Dopa is decarboxylated to dopamine, catalysed by Dopa-decarboxylase. Dopamine is produced in the cytosol, and then taken up by the chromaffin granules, which contain the enzyme, dopamine b-hydroxylase. The enzyme catalyses the formation of nor-adrenaline from dopamine. In most granules noradrenaline diffuses back into the cytosol, where it is N-methylated by the enzyme, phenyl-ethanolamine N-methyl-transferase to form adrenaline. Adrenaline is then stored in the granules as the most important adrenomedullary hormone. The storage process requires ATP, and catecholamines are stored with ATP in the granules. - In the newborn the primary end product in the medulla is nor-adrenaline, but with advancing age there is a dramatic rise in the adrenaline content. This conversion depends upon cortisol.

The three important catecholamines in humans are *dopamine*, *epinephrine* or *adrenaline* (Ad), and *norepinephrine* or *noradrenaline* (NA in Fig. 30-5).

Dopa is converted to *melanin* in the melanocytes of the skin, stimulated by *melanocytic stimulating hormone* (MSH in Fig. 30-5).

Fig. 30-5: Synthesis of catecholamines within the chromaffin cell. Tyrosine hydroxylase and dopamine β-hydroxylase is activated by sympathetic stimulation.

Catecholamines are formed in *adrenergic, noradrenergic and dopaminergic neurons.*

The neurons of substantia nigra produce dopamine, which by axonal transport reaches the striatum (corpus striatum in

Fig. 30-5). Here, dopamine is stored in the granules of the terminals - ready for liberation in the synapses.

In the sympathetic ganglion cells, dopamine is just an intermediary step, which is oxidised (dopamine b-hydroxylase) to form noradrenaline.

Euler identified noradrenaline as the chemical transmitter from adrenergic nerves, and Axelrod demonstrated the reuptake of the transmitter after its release from nerve terminals. In 1970 they shared the Nobel Prize with Katz, who explained the role of calcium in synaptic transmission.

Actions of catecholamines

Catecholamines act on adrenergic membrane receptors designated a_1-, a_2-, b_1- b_2- and b_3-receptors. The a-receptors are located on effector cells that are highly sensitive to noradrenaline, less so to adrenaline, but much less so to isoprenaline (a synthetic catecholamine). The b-receptors are located on effector cells that are most sensitive to isoprenaline, but less so to adrenaline (and noradrenaline). The a_1-receptor acts via calcium and protein kinase C. The a_2-receptor acts through G-protein inhibition of membrane adenylyl cyclase. The b-receptors are transmembrane glycoproteins, and they use cAMP as second messenger. Catecholamines prevent hypoglycaemia and restore glucose delivery to the CNS. Catecholamines increase glucose production by gluconeogenesis in the liver, and stimulate glycogenolysis in liver and muscle cells. In the absence of glucagon, adrenaline is important for recovery from hypoglycaemia. Adrenaline inhibits the insulin-mediated glucose uptake by muscle and fat cells. Catecholamines stimulate glucagon secretion and inhibit insulin secretion.

Adrenaline activates lipase in fat cells and thus increases FFA release, b-oxidation of FFA in muscle and liver and ketogenesis.

Adrenaline increases basal metabolic rate, nonshivering thermogenesis, and diet-induced thermogenesis during cold exposure. Catecholamines stimulate proton transfer into the mitochondria and uncouple ATP synthesis from oxygen utilisation.

Catecholamines increase the heart rate, contractile force and the cardiac output by stimulation of the *adrenergic β_1-receptors* in the myocardium. Noradrenergic nerve fibres innervate vessels all over the body, and this system usually has some tonic, vasoconstrictor activity. The a_1-receptors are located on the surface of vascular smooth muscles. Adrenaline constricts the arterioles in the cutaneous, renal and splancnic beds.

Catecholamines dilatate the bronchial airways by stimulation of their adrenergic b_2-receptors. They increase both tidal volume and respiratory frequency. The result is an increased ventilation with an increased cardiac output.

Catecholamines relax the *smooth muscles of the digestive tract* (b_2-receptors), but *contract the sphincters* just like the sympathetic nerve system.

Finally, adrenaline stimulates the *ascending reticular system* (ie, the reticular activating system or RAS) in the brain stem, keeping us alert and causing *arousal reactions* with desynchronisation of the EEG.

Stress responses

Stress comprises severe emotional and physical burdens (fear, pain, hypoxia, hypothermia, hypoglycaemia, hypotension etc).

The adrenal medulla functions in concert with the adrenal cortex during stress and during immune system disorders. The two systems provide useful local cytokines and avoid systemic damages from accumulated cytokines.

Stress activates the CRH-ACTH-corticoid axis and adrenergic neurons in the hypothalamus with connections to the sympathetic nervous system. ACTH is released whereby plasma cortisol is elevated; the adrenergic stimulus elevates plasma adrenaline. These hormones increase glucose production.

Noradrenaline and CRH induce a state of arousal and aggressiveness. CRH inhibits sexual activity and feeding behaviour.

The sympatho-adrenergic system gives rise to the *fright, flight or fight-reactions*, which are all *acute* stress situations. During stress adrenaline induces hyperglycaemia and ketoacidosis (a diabetogenic hormone). Prolonged *stress* liberates

ACTH via hypothalamic signals. ACTH stimulates the glucocorticoid secretion through cAMP. Small amounts of glucocorticoids are *permissive* for the actions of catecholamines.

Acute stress activates the splancnic nerves and liberates large amounts of adrenaline from the medulla. Diabetics, who are developing acute hypoglycaemia, secrete large amounts of catecholamines.

In the endangered individual catecholamines dilatate bronchioles, inhibit gastrointestinal motility, and dilatate pupils to improve distant vision.

Exercise

Catecholamines support the sympathetic system in modifying the circulation during exercise. During *exercise* catecholamines promote the use of muscle glycogen, the gluconeogenesis from lactate, and the mobilisation of FFA as alternative fuel.

During exercise the blood is directed to the working muscles from the other parts. The resistance vessels of the striated muscles in hunting predators (and perhaps in humans) are also innervated by another system. This is the *cholinergic, sympathetic vasodilator system.* It is capable of a rapid and appropriate bloodflow response during hunting. The fall in the *α-adrenergic tone* in the muscular arterioles is probably the most important exercise response in humans.

Metabolism of catecholamines

Plasma catecholamines are rapidly removed from the blood and have half-lives in plasma of 20-60 s. This is the combined result of *rapid uptake by tissues* and *inactivation* in the liver and vessel walls.

Enzymes inactivate catecholamines by methylation or by oxidation. *Catechol-O-methyl-transferase* (COMT) on the surface of the target cells catalyses methylation to metanephrine, normetanephrine.

Monoamine oxidase (MAO) in the mitochondria catalyses oxidative removal of the amino group with formation of vanillylmandelic acid (VMA) and methoxy-hydroxy-phenyl-glycol (MOPG). The final products in the urine areVMA and MOPG. The normal excretion of VMA and MOPG in the daily urine averages 4 and 2 mg.

5. Vitamin D and bone minerals

Bones are compared to iron mixed with concrete. The *elasticity* is due to the collagen (iron) network, and the *stiffness* (concrete) is due to calcium salts (Ca - hydroxyapatite complex). - Bone tissue consists of two compartments, bone cells (metabolically active) and the bone matrix (metabolically inert extracellular compartment). The organic part of the bone matrix consists of *collagen fibres* and *ground substance,* and the inorganic part consists of *calcium-phosphate-hydroxyapatite.*

The atomic weight of calcium is 40, and we consume 1000 mg (25 mmol) per day. However, we absorb only 400 mg in total. The *net-absorption* is only *250 mg,* since we secrete 150 mg daily to the intestines. Thus, 750 mg (19 mmol) must be excreted in the faeces every day (Fig 30-3). The net-absorption is saturable (an active process), since it depends on available Ca^{2+} *-binding protein* in the brush border and in the cytosol of the enterocyte. The synthesis of this protein in the mucosal cells, and thus intestinal Ca^{2+} -absorption, is induced by active vitamin D and by the parathyroid hormone. Steroid hormones like vitamin D, exerts their major effects after binding to specific nuclear receptors and then stimulating the synthesis of mRNA that codes for *cytosolic* Ca^{2+} *-binding protein.* The basolateral membrane contains two transporters of Ca^{2+}: A Na^{+}/Ca^{2+} *exchanger,* which is more effective at high intracellular [Ca^{2+}], and a Ca^{2+}*-ATPase,* which is the major mechanism at low levels of intracellular [Ca^{2+}]. The duodenum-jejunum can concentrate Ca^{2+} against a tenfold concentration gradient.

The amount of phosphate absorbed and its concentration in plasma is determined by the amount available in the diet, but the active transport is somewhat dependent on vitamin D.

High plasma [Ca^{2+}] and [phosphate] promote bone formation in children. The children increase the precipitation of Ca-hydroxy-apatite in their bone matrix.

Intracellular phosphate is an essential component of nucleic acids, high-energy molecules, cofactors, regulatory phospho-proteins, and glycolytic intermediates.

The [total calcium] in the blood plasma of healthy persons is normally *2.5 mM* or 100 mg per l. Half the calcium binds

to plasma proteins and 5% binds to a citrate-phospha
ionised calcium ([Ca^{2+}]); this part has critical roles in neuromuscular function and coagulation. If the plasma [Ca^{2+}] falls to half its normal level, the body develops increased neural excitability with *tetanic cramps* (see below). On the other hand, an increase in [Ca^{2+}] leads to *calcification* of soft tissue with heart, kidney and intestinal diseases. The [Ca^{2+}] is the one variable regulated by parathyroid hormone (PTH) - see Fig. 30-3.

The normal adult contains 27 500 mmol or 1100 g of calcium and 600 g of phosphate. The major part exists in the bones (1080 g fixed as hydroxyapatite and only 4 g or 100 mmol as *exchangeable bone calcium*). We exchange this 4 g of calcium five (5) times a day. The flux between the bones and the ECV (16 l) is approximately 500 mg or 13 mmol daily. This fast exchange regulates [Ca^{2+}] in plasma. But we also have a small amount of daily bone formation and resorption (*bone remodulation*).

In our soft tissue cells 99% of all calcium is complexed with phosphate in the *mitochondria* (10^{-5} mM), or bound to membranes and the endoplasmic reticulum. The cytosol contains an extraordinarily low [Ca^{2+}]: 10^{-7} mol per l. The intracellular [Ca^{2+}] is of extreme physiologi
a *secondary messenger* for peptide hormones. In general Ca^{2+} is important to almost all biological systems: action and membrane potentials, blood clotting, cell division, cellular secretion, contraction, cytoskeletal rearrangements, and motility.

Both calcium and phosphate ions are in equilibrium with calcium-proteinate:

$$\text{Protein} + Ca^{2+} \Rightarrow 2\,H^{+} + \text{Calcium-proteinate}$$

A hyperventilating patient eliminates carbonic acid, liberates H^+ and reduces [Ca^{2+}] in plasma, because more Ca^{2+} binds to protein. Low [Ca^{2+}] in plasma leads to increased excitability of neuromuscular tissue and to tetany (tetanic muscle cramps). The treatment is simply to reinhale the exhaled air (rebreathing) resulting in carbonic acid accumulation (Chapter 17). – As seen from the reaction above, alkalosis reduces and acidosis increases [Ca^{2+}] in plasma.

Fig. 30-6: Vitamin D and the Ca^{2+} balance.

Vitamin D_3 (*cholecalciferol*) from the skin, and vitamin D_2 (*ergocalciferol*) from plant diet are concentrated in the liver. Both are produced in the skin by ultraviolet irradiation (Fig. 30-6). In the liver microsomal oxidase simply hydroxylates the molecules to the weakly active *25-hydroxy-cholecalciferol* (25-OH-D). Lack of vitamin D leads to insufficient bone formation, because the osteoid matrix does not calcify. This disease is called rickets (rachitis) in children and osteomalacia in adults.

This substance is transferred with the blood to the kidneys, where it is further hydroxylated at the C-1 position to the most potent form *1,25-dihydroxy- cholecalciferol* (Fig. 30-6). This is what makes vitamin D a potent *steroid hormone*. Vitamin D is stored in muscle and adipose tissue and circulates in plasma bound to *vitamin D-binding protein*.

Vitamin D_2 is derived from vegetable ergosterol and its chemical name is *ergocalciferol.* Vitamin D_3 or *cholecalciferol* is produced in the skin by ultraviolet irradiation. The two molecules have equal biological capacity in humans. Vitamin D must be hydroxylated to the weakly active *25-hydroxy-cholecalciferol* in the liver. This substance is transferred with the blood to the liver, where it is further hydroxylated at the C-1 position to the most potent form *1, 25-dihydroxy-cholecalciferol.* This is exactly what is necessary to make vitamin D a potent *steroid hormone.* Vitamin D circulates in plasma bound to *vitamin D-binding protein* and is stored in liver, muscle and adipose tissue.

Vitamin D deficiency causes rickets in children and *osteomalacia* in adults. The deficiency is caused by insufficient diet, inadequate absorption as in fat malabsorption, insufficient synthesis in the skin, or abnormal conversion to its potent form in the liver and kidneys.

Rickets (rachitis) is extremely rare in the rich part of the world. The epiphyseal plate of the growing skeleton is inefficiently mineralised and considerably thickened. Rachitic children have a *flat skull* with prominent frontal bones. The sternum protrudes as *pigeon breast* or *pectus carinatum*. The costal cartilage is enlarged forming the *rachitic rosary.* The limbs are shortened and deformed.

Osteomalacia or *soft bones* are inadequate mineralisation of the organic bone matrix. Patients are asymptomatic or they suffer from diffuse pains. Bone X-rays are *osteopenic.*

High plasma concentrations of parathyroid hormone, phosphate, and oestrogens stimulate the renal hydroxylation to active vitamin D. The active vitamin D-metabolite, the 1,25-dihydroxy-cholecalciferol, stimulates the transport of Ca^{2+} across the cell and mitochondrial membranes, and has the following two effects: 1. Enhanced effect on gut absorption of Ca^{2+}, so that plasma-[Ca^{2+}] increases, 2. Enhanced effect of PTH on bones. In vitamin D deficiency the low plasma-[Ca^{2+}] stimulates the parathyroid glands, which leads to secondary hyper-parathyroidism.

Ca^{2+} loss continues for months following space flight. Stimulation of bone deposition requires the stress of muscular activity as when a person is working against a gravity field. Appropriate exercise during space missions reduces the total Ca^{2+} loss.

Puberty growth is essentially stimulated by the increased production of sex hormones, because the sex steroids stimulate hypothalamus to secrete more GH. Testosterone and oestradiol are generally important anabolic hormones.

GH and insulin are also important anabolic hormones in humans, although GH has many insulin-antagonistic effects. GH is released as *hunger spikes* during the day and during sleep.

Primary hormones for bone remodelling are *GH, PTH, calcitonin and 1,25-D_3*.

Secondary hormones for bone remodelling are glucocorticoids, sex hormones, thyroid hormones, prostaglandins (PGE_2), insulin and IGF-I to III (somatomedins).

Bone remodelling is a balance between bone formation by osteoblasts and bone resorption by osteoclasts and mononuclear cells. This balance involves the following hormones besides GH:

1. 1,25-Dihydro-cholecalciferol (1,25-D_3) is a D-vitamin, but also a steroid hormo
   when stimulated by PTH. The steroid hormone increases the calcium absorption from the gut and mobilises Ca^{2+} from the bones. This increases the [Ca^{2+}] and [phosphate] in plasma. The 1,25-D_3 also increases the renal reabsorption of Ca^{2+} indirectly.
2. PTH also increases the plasma [Ca^{2+}]. PTH mobilises Ca^{2+} from the bones and increases the reabsorption of Ca^{2+} in the distal, renal tubule cells while inhibiting phosphate reabsorption in the proximal tubules of the kidneys.
3. Calcitonin from the thyroid inhibits bone resorption by blocking the PTH receptors on the osteoclasts. Thus plasma [Ca^{2+}] and [phosphate] fall.

Cyclic treatment with calcitonin or combined treatment with 1,25-D_3 and PTH *improve bone formation* in *osteoporosis* (ie, bone atrophia involving both minerals and matrix). Postmenopausal osteoporosis is treated successfully with calcitonin. Calcitonin is important in *bone remodelling* and in *treatment of osteoporosis*.

A parallel to the bone formation by osteoblasts is the dentin formation by odontoblasts in our teeth. Dentists stimulate dentin formation with $Ca(OH)_2$, when treating caries.

Pathophysiology

This paragraph deals with: *1. Acromegaly and gigantism, 2. Nanism, 3. Hyperparathyroidism, 4. Hypoparathyroidism, 5. Hyperaldosteronism (Conn´s disease), 6. Cushings disease and syndrome, 7. Congenital adrenal hyperplasia, 8. Primary hypoadrenalism, 9. Secondary hypoadrenalism, 10. Phaeochromocytoma*.

1. Acromegaly and gigantism

Pituitary acidophilic adenomas that secrete excess growth hormone (GH) causes gigantism in children and acromegaly in adults. Rare cases are caused by GHRH excess release from the hypothalamus. Since pituitary acidophilic adenomas often contain somatotropic as well as mammotropic cells, the combined adenomas secrete an excess of both GH and prolactin (causing galactorrhoea in males).

GH excess before the long bone epiphyses have closed, results in *gigantism*. The arms and legs are long, and the patients are tall. Giants may suffer from deficient sexual development, if the gonadotropic pituitary function also suffers.

The hypophyseal tumour is demonstrated by CT scans or by magnetic resonance imaging.

GH stimulates both soft-tissue and skeletal growth. In adults only bones in the cranium, the face, and the hands and feet respond the GH excess. Acromegaly often develops around the age of 30 years in both females and males.

Fig. 30-7: Clinical features of acromegaly.

The calvarium, the maxilla and the mandible grow (Fig. 30-7). The teeth are separated by spaces and may fall out. The hands and feet are large and *spade-shaped.* This is why the adult disorder is called *acromegaly* (from Greek: extremity-great). Visual field defects are present when the enlarged hypophysis damages the crossed fibres in chiasma opticum, resulting in *bilateral, heteronymous hemianopsia* (Fig. 5-6). The GH concentration in plasma is only increased during stress and as occasional peak values, but IGF-1 is increased, and *impaired glucose tolerance* is diagnostic for a secondary, diabetic condition. There is often polyuria and glucosuria (Fig. 30-7).

Therapy in cases with growing and pressure-damaging tumours is trans-sphenoidal or trans-frontal surgery with post-operative radiotherapy. Medical treatment with *octreotide*, an analogue to somatostatin (GHIH), is beneficial in some cases.

2. Nanism

Nanism or dwarf growth expresses that the height of an adult is below a certain limit. In the Anglo-Saxon countries the limit is 1.40 m for females and 1.50 m for males. Infantile nanism is a term used to indicate that the somatic age of a child is below its chronological age.

Low pituitary secretion of GH or insensitive target organ receptor (African pygmies') result in *pituitary dwarfism.* Many cases of dwarfism are idiopathic (of unknown origin). An *achondroplastic dwarf* has short limbs and a large head with a flat nose, because of abnormal growth of cartilage and bone. The abnormal gene is autosomal and dominant.

3. Hyperparathyroidism

Primary parathyroid hyperfunction is almost inevitably due to *parathyroid adenomas* or *hyperplasia* that secrete an excess of *parathyroid hormone*, PTH. Ectopic tumours have been found in the mediastinum and elsewhere. Excessive secretion of PTH leads to: bone resorption, high [Ca^{2+}] in plasma, high Ca^{2+} -excretion in the kidneys with renal stone formation, bone lesions, and metastatic calcification (Fig. 30-8).

The action of PTH on bone results in osteitis fibrosa cystica, there is *hypercalcaemia* with low plasma phosphate, metabolic acidosis and raised PTH levels in plasma.

Secondary hyperparathyroidism is an adequate physiological response to hypocalcaemia by any cause such as renal failure or malabsorption.

Hypercalcaemia has many other causes: hypervitaminosis D, excessive Ca^{2+} -intake, leukaemia, myelomata, sarcoidosis and cancer with bone metastases or malignant tumours producing PTH.

Fig. 30-8: Clinical manifestations of hyperparathyroidism.

4. Hypoparathyroidism

Hypoparathyroidism is characterised by *hypocalcaemia,* but most cases of hypocalcaemia are caused by renal failure with phosphate retention, vitamin D deficiency or resistance to vitamin D, resistance to PTH, pregnancy and lactation, or calcitonin administration.

Most cases of hypoparathyroidism are *iatrogenous* (ie, caused by doctors during thyroidectomy or parathyroidectomy). After seemingly successful surgery there is a dramatic fall in plasma [Ca^{2+}] and a rise in [phosphate], leading to *hypocalcaemia cramps.*

The patient is in tetany with paresthesia, convulsions, and laryngeal stridor. Without therapy death ensues.

Primary, idiopathic hypoparathyroidism is an extremely rare *autoimmune disease* often found in combination with other autoimmune disorders.

Pseudo-hypo-para-thyroidism is used for conditions, where target-organs (bone, kidneys, and gut) are *resistant to PTH.* The plasma- [Ca^{2+}] is low, plasma-[phosphate] is high, and basic phosphatase activity is high. A *PTH injection* does not increase renal phosphate excretion, because the kidneys are resistant.

Hypoparathyroidism is treated with Ca^{2+} -infusion and the continuation is a lifelong treatment with PTH and vitamin D.

5. Hyperaldosteronism (Conn´s disease)

Patients with adenoma or hyperplasia of the zona glomerulosa secrete large amounts of aldosterone - they suffer from *primary hyperaldosteronism*. Here, aldosterone works directly on the highly Na^+ -loaded distal tubules, so the patient becomes severely K^+-depleted. The renal retention of salt and water leads to *hypertension*. Loss of K^+ and H^+ leads to *hypokalaemia* with muscle weakness and cardiac arrhythmias and to metabolic alkalosis with hypocalcaemia and tetany.

Plasma aldosterone is elevated although the patient has a high intake of NaCl. Removal of the neoplasm cures Conn´s disease.

Secondary hyperaldosteronism. A patient with serious congestive heart failure may develop *secondary hyperaldosteronism* and become *K^+-depleted*. At rest the GFR, the renal bloodflow, and the cardiac output are close to normal. The cardiac patient has an abnormally high Na^+ retention and the accompanying high water retention that increases the ECV. This leads to oedema and increased venous return to the heart with a small rise in cardiac output. During *exercise* the rise in cardiac output is insufficient. The insufficient rise in bloodflow and blood pressure elicits (via the baroreceptors) a marked vasoconstriction of the renal circulation. Thus the RBF must decline. Other vascular beds also constrict markedly during exercise. The key factor to the *abnormal Na^+ retention* is the *reduction in RBF*. The major Na^+ reabsorption normally takes place in the proximal tubules. Therefore the distal tubular fluid contains a small load of Na^+, allowing only a small K^+ secretion here. This is in contrast to the patient with Conn's disease.

6. Cushings disease and syndrome

Cushing's disease is a pituitary disorder with increased secretion of pituitary ACTH.

Cushing's syndrome is used as a common term for all clinical cases of abnormally high glucocorticoid concentration, [cortisol], in the blood plasma.

All clinical cases are divided into two groups: ACTH-dependent Cushing and non-ACTH-dependent Cushing.

1. *ACTH-dependent Cushing* is caused either by pituitary disorder (Cushing's disease, see above) or by an ectopic ACTH-producing tumour. This group of cases is characterized by a high ACTH concentration in the plasma.
2. *Non-ACTH-dependent Cushing*. Most clinical cases are caused by glucocorticoid administration for long periods, but also adrenal tumours (adenomas and carcinomas) produce excess glucocorticoid. The ACTH secretion is suppressed.

Clinical manifestations

The high [cortisol] in the blood has two major effects on the body:

1. *Salt and water-retention* with renal loss of K^+ results in a plethoric *tomato face* or *moon face* (Fig. 30-9). The fluid-retention eventually leads to *cardiac hypertrophy* due to prolonged *hypertension.* There is often *peripheral oedema,* if not eliminated by polyuria due to the *glucocorticoid-induced diabetes*. This type of diabetes is typically resistant even to large doses of insulin.
2. *Catabolism* causes muscle wasting, fat accumulation, osteoporosis with kyphosis, buffalo hump, and fractures. The skin is thin with ulcers and red stirrer, and there is poor wound healing. There is impaired fibrocyte formation and capillary resistance.

Fig. 30-9: Manifestations of Cushing's syndrome.

In *Cushings disease* there is a *pituitary* tumour with enlarged sella turcica or pituitary hyperplasia. The pituitary corticotrophins produce excessive amounts of ACTH, which leads to *bilateral* adrenal cortical *hyperplasia.*

In other cases of *Cushing's syndrome* the primary cause is an *adrenal* tumour or *adrenal* hyperplasia. There is a *hypersecretion* in cortical zona reticularis and fasciculata.

The patient may suffer from *muscular and bone atrophy* due to the *catabolic* effect of the cortisol surplus. The bone atrophy leads to osteoporosis, vertebral fractures and necrosis of the hips. There is a *delayed healing* of lesions and wounds due to the slow fibrocyte formation. The skin is thin with purple striae and fragile capillaries. Sodium and water accumulate in the body, whereas potassium is lost. The nitrogen balance is negative. The Cushing patient is often depressed and suffers from insomnia.

The arms and legs are thin and weak due to the protein catabolic effect, but fat collects as truncal obesity and in the back and neck regions (*bulls neck* and *buffalo hump*). The fluid accumulation leads to a special look called *moon face* or *tomato face.*

The high blood [cortisol] also increases the blood [glucose]. A *diabetic condition,* which is resistant to even large doses of insulin, develops. The patient suffers from *hypertension* and rapidly develops *arteriosclerosis*. Female patients complain of amenorrhoea and poor libido.

The primary therapy is to reduce plasma [cortisol] to approximately 350 nM also as a preparation for surgery. In cases of *Cushing's disease* the tumour is typically removed trans-sphenoidally, and adrenal adenomas are resected.

The best assessment of increased ACTH production is by measuring the 24-hour urinary cortisol excretion. *Suppression* of ACTH and cortisol is performed with a standard dose of synthetic Glucocorticoid (*dexamethasone*), which is potent enough to suppress the ACTH production, without influencing the plasma [cortisol].

The principle of the *suppression test* is to *challenge the hypothalamic-pituitary feedback system.* If the system is intact a fall in blood [cortisol] is expected, and if not (hypothalamic-pituitary Cushing) the high [cortisol] is maintained.

In healthy persons such a standard dose will *suppress ACTH secretion,* and lead to a reduced blood [cortisol]. Such a normal observation is also typical for patients with a *primary adrenocortical hypertrophy*. Dexamethazone does not suppress the high cortisol level of patients with a *damaged feedback system* producing excess ACTH from a pituitary tumour.

Cortisol excess has been treated with ketoconazole, which inhibits several steps in the corticosteroid synthesis.

Differential diagnosis:

Alcoholic patients with so-called *Pseudo-Cushing* look exactly like *Cushing's syndrome,* but the glucocorticoid concentration in the plasma is within normal limits. The cause is unknown. Cases with plasma [cortisol] in the upper end of the normal scale can be explained in the following way. Their progressive hepatic failure probably impairs the normal hepatic destruction of glucocorticoids, whereby the plasma [cortisol] is rising. Such patients should be controlled, because the liver destruction is likely to proceed, and they become more and more Cushingoid with continued alcohol abuse.

7. Congenital adrenal hyperplasia

This is an autosomal, recessive enzyme deficiency blocking steroid synthetic pathway of the adrenal cortex. The most frequent of these rare genetic disorders is the *21-hydroxylase deficiency* on *chromosome 6*. Lack of hydroxylase blocks the conversion of 17-hydroxyprogesterone into *11- deoxycortisol* and further to *cortisol*. Instead androstenedione and testosterone is produced and virilization is an inevitable result. The impaired cortisol release from the adrenal increases the pituitary ACTH secretion by negative feedback and more precursor molecules are accumulated (ie, hydroxyprogesterone, androstenedione and testosterone).

The clinical features are those of an *adrenogenital syndrome.* Newborn girls may show clitoral hyperplasia and labioscrotal fusion, and later some boys develop *precocious puberty*. Varying degrees of virilism in females, *birdied ladies* in circus, or perhaps only hirsutism before puberty occur. The lack of aldosterone causes salt-and water-loss, which may be life threatening.

A defect in the *11-hydroxylase enzyme gene* also leads to overproduction of androgens and 11-deoxycorticosterone from accumulated precursors. The later has a massive mineralocorticoid effect.

The therapy is exact replacement of glucocorticoid activity and any other inefficiency.

8. Primary hypoadrenalism

Primary hypoadrenalism is a complete destruction of the adrenal cortex, and thus it impairs all three lines of steroid production (ie, glucocorticoids, mineralocorticoids and sex hormones). Primary hypoadrenalism occurs in an acute and a chronic form.

Acute hypoadrenalism (ie, *adrenal crisis* or Waterhouse-Friderichsens syndrome) is a fulminant infection with shock and massive bleeding all over. The large internal bleedings are visible at the skin as areas of purpura. The patient dies within a few hours, if not treated urgently. *Hydrocortisone* (100 mg) is given intramuscularly. Antibiotics and cortisol is administered.

Chronic hypoadrenalism (ie, Addison's disease) is hypocorticism due to destruction with atrophy of the entire adrenal cortex by autoimmune processes, malignancy, infarction or infection. Most cases develop *organ-specific autoantibodies;* these cases are associated with many other autoimmune disorders (eg, diabetes mellitus, hypoparathyroidism, pernicious anaemia, vitiligo and thyroiditis).

Addison's disease is a life-threatening condition with loss of Na^+ and thus also of ECV. Symptoms and signs include: reduced blood pressure, reduced blood [glucose], tiredness and skin pigmentation caused by MSH, whose level is increased concurrent with the *overproduction of ACTH*, due to the decreased negative feedback. Severe hypotension may develop into cardiovascular collapse, which is called an *Addisonian crisis*.

The clinical manifestations are consequences of the impaired secretion of the three hormone groups (Fig. 30-10).

Fig. 30-10: Clinical manifestations of primary, chronic adrenocortical insufficiency (Addison).

The low cortisol increases the hypothalamic CRH secretion through feedback and thereby increases ACTH secretion. Fragments of the ACTH molecule form MSH, which is responsible for the skin hyperpigmentation.

When Addison's disease is suspected a *single* plasma ACTH level is often diagnostic, as it will be elevated if hypocorticism originates in the *adrenal gland* (primary) and normal or low in *hypothalamic-pituitary hypocorticism* (secondary). - Differentiation between the hypothalamic and the pituitary hypocorticism can be accomplished by *injection of CRH*, which stimulates the pituitary directly. - *Insulin-induced hypoglycaemia* normally recruits the combined hypothalamic-pituitary axis. Hypoglycaemia is a prolonged stress, which will normally increase ACTH and glucocorticoid secretion. If there is adrenal gland insufficiency then this increase in glucocorticoid secretion is not observed. *Hypothalamic and pituitary failure* is detected when a dose of CRH restores glucocorticoid release. - *A single ACTH-injection* along with measurement of plasma [cortisol] is a simple screening procedure for adequate endogenous ACTH secretion. Synthetic ACTH is injected intravenously. If plasma [cortisol] is normal 30 minutes later, the disease is not a *primary* adrenocortical atrophia. - *The metyrapone test* is used when patients are suspected to have Addison's hypocorticism, but have reacted normally to the ACTH-test. Metyrapone inhibits the last step in cortisol synthesis, so 11-dehydrocortisol is accumulated and an acute cortisol deficiency is created. Plasma-[Cortisol] and -[11-deoxycortisol] is measured two mornings in succession, and the *adrenal 11-hydroxy*
500 mg metyrapone every second hour during 24 hours. The plasma- [cortisol] must fall while -[11-deoxycortisol] accumulates. The fall in cortisol biosynthesis increases the CRH and ACTH secretion by negative feedback, provided the hypothalamic-pituitary system is *intact*. Failure to respond to metyrapone suggests a hypothalamic-pituitary defect that has abolished this feedback.

Substitution therapy with gluco- and mineralo-corticoids is necessary for the rest of the patient's life.

9. Secondary hypoadrenalism

This is usually caused by prolonged corticoid therapy of chronic disorders such as rheumatoid arthritis or COLD. A rare form of secondary hypoadrenalism is *panhypopituitarism*.

10. Phaeochromocytoma

Rarely some patients have attacks of *severe hypertension*., which are caused by a *medullary tumour* (a

phaeochromocytoma of chromaffine cells), which liberates large amounts of catecholamines.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

A. Atropine blocks muscarinic receptors, and d-tubocurarine blocks nicotinic receptors.

B. b-Receptors are blocked by propranolol.

C. Catecholamines have a half-life in plasma of more than 20 min.

D. The three important catecholamines in humans are adrenaline, noradrenaline, and DOPA.

E. b_1-receptors are stimulated by adrenaline and located in the myocardium.

II. Each of the following five answers have False/True options:

ACTH stimulates the liberation of**:**

A**.** Hydrocortisone or cortisol

B. Adrenaline

C. Aldosterone

D. Adrenal androgens

E. Dopamine.

Case History A

A man, 49 years of age (height 1.86 m; weight 62 kg) is in hospital due to the following symptoms and signs. He is nervous and has a diffuse struma. A characteristic, blowing sound is heard from the thyroid gland with a stethoscope. The blood pressure is 145/70 mm Hg. An attack of cardiac arrhythmia is recorded with an ECG. A P-wave frequency above 400 per min is present during the attack. The concentration of thyroxine in blood serum is 180 nM. The distribution volume for radioactive thyroxine is 8 l. The elimination rate of this thyroxine is 14% of the total content per 24 hours.

1. *Calculate the serum [thyroxine] in μg per l.*
2. *How much thyroxine is eliminated per 24 hours expressed in μg daily?*
3. *Calculate the thyroid plasma clearance of iodide, when the concentration of free iodide in plasma is 4 μg per l.*
4. *Explain the condition of this patient***.**

Case History B

A 22-year-old medical student is treated with an intravenous dose of PTH. This changes his renal excretion flux of two substances and their plasma concentrations.

1. *Describe the alterations and explain the mechanisms.*
2. *What is the diagnosis?*

3. *Describe the most likely symptoms and signs of this patient before treatment.*

Case History C

A female (52 years of age; height 1.68 m; weight 62 kg) is in hospital due to her third attack of kidney stone pains. The first routine examination with arterial blood analysis reveals the following. Her blood pH is 7.21 and her plasma [Ca^{2+}] is 2 mmol per l (mM) in ionised form. Her ionised [Ca^{2+}] constitute 62% of the total calcium concentration in plasma. The inorganic phosphate concentration (total) is 0.84 mmol per l of plasma. The patient excretes 2-3 l of urine per day. pK_2 = 6.8 for H_3PO_4.

1. *Calculate the fractions of primary ($H_2PO_4^-$) and secondary (HPO_4^{--}) phosphate in her plasma.*

   Compare the results to the normal mean value of 1 mM of inorganic phosphate with 20% primary and 80% secondary phosphate at pH = 7.40.

2. *Calculate the total plasma [calcium] of this patient and compare the result to the normal mean value of 2.5 mM.*

3. *Why is the ionised calcium fraction much higher than normal (0.45)?*

4. *Could this condition be the result of a classical endocrine disease?*

5. Why did this patient develop kidney stones? Was her diuresis normal? If not explain why.

Try to solve the problems before looking up the answers.

Highlights

- *Growth hormone (GH) produced in the placenta differs from the pituitary GH by a few amino acid residues. Placental GH suppresses release of maternal, pituitary GH during pregnancy. Placental GH stimulates maternal metabolism and foetal cell proliferation and hypertrophia.*
- *Foetal thyroid hormones stimulate brain development, and foetal insulin stimulates foetal growth, cellular glucose uptake and glucose utilisation. Paracrine and autocrine growth factors are also important for foetal growth: Insulin-like growth factor-II (IGF-II), nerve growth factors (NGF), epidermal growth factor (EGF), and platelet derived growth factor (PDGF).*
- *Growth hormone and insulin are the most important anabolic hormones in the human body.*
- *PTH binds to membrane receptors on target cells in bone, kidney and gut. The PTH actions result in hypercalcaemia and hypophosphataemia. The bone resorption is illustrated by a high basic phosphatase concentration in blood.*
- *Cortisol stimulates hepatic glucose production, both glycogenolysis and gluconeogenesis, lipolysis, formation of FFA and of ketone bodies. Cortisol inhibits the glucose uptake in target cells (GLUT 4 in muscle cells, heart cells and adipocytes).*
- *Therapeutic doses of glucocorticoids are used for a multitude of diseases such as inflammations, allergy, malignancy and aplastic anaemia. The negative effects are delayed healing of wounds and increased gluconeogenesis with destruction of tissue proteins.*
- *Aldosterone is the major mineralocorticoid with corticosterone contributing only little. Aldosterone promotes the reabsorption of Na^+ and increases the secretion of K^+ and H^+ in the distal tubular system (ie the cortical collecting ducts and the connecting segment). A rise in serum - K^+ from normal or a fall in serum- Na^+ releases*

[ ] [ ]

aldosterone. The renin-angiotensin-aldosterone cascade controls the adrenal aldosterone secretion, not ACTH.

- *Sex steroids are mainly weak adrenal androgens, which are metabolised to testosterone and dihydrotestosterone. Also a small oestrogen production is present in healthy persons.*
- *Catecholamines increase the heart rate and the cardiac output by stimulation of the adrenergic β_1-receptors in the myocardium.*
- *Noradrenergic nerve fibres innervate vessels all over the body, and this system usually has some tonic, vasoconstrictor activity. The α_1-receptors are located on the surface of vascular smooth muscles.*
- *Catecholamines dilatate the bronchial airways by stimulation of their adrenergic β_2-receptors. Catecholamines increase both tidal volume and respiratory frequency. The result is an increase in ventilation together with an increase in cardiac output.*
- *Catecholamines relax the smooth muscles of the digestive tract (β_2-receptors), but contract the sphincters just like the sympathetic nerve system.*
- *Catecholamines stimulate metabolism (T_3). Adrenaline stimulates hepatic glycogenolysis and lipolysis in adipose tissue. Adrenaline increases the plasma concentrations of glucose, FFA and ketoacids.*
- *Adrenaline stimulates the ascending reticular system (ie, the reticular activating system or RAS) in the brain stem, keeping us alert and causing arousal reactions with desynchronisation of the EEG.*
- *Pituitary acidophilic adenomas that secrete excess growth hormone (GH) causes gigantism in children and acromegaly in adults. Rare cases are caused by GHRH excess release from the hypothalamus. Since pituitary acidophilic adenomas often contain both somatotropic and mammotropic cells, the combined adenomas secrete an excess of both GH and prolactin (causing galactorrhoea in males).*
- *Primary parathyroid hyperfunction is almost inevitably due to parathyroid adenomas or hyperplasia that secrete an excess of parathyroid hormone, PTH. Ectopic tumours have been found in the mediastinum and elsewhere. Excessive secretion of PTH leads to: bone resorption, high [Ca^{2+}] in plasma, high Ca^{2+} -excretion in the kidneys with renal stone formation, bone lesions, and metastatic calcification.*
- *Chronic hypoadrenalism (ie, Addison's disease) is hypocorticism due to destruction with atrophy of the entire adrenal cortex by autoimmune processes, malignancy, infarction or infection. Most cases develop organ-specific autoantibodies; these cases are associated with many other autoimmune disorders (eg diabetes mellitus, hypoparathyroidism, pernicious anaemia, vitiligo and thyroiditis).*

Further Reading

Costanzo, L S. Regulation of calcium and phosphate homeostasis. *Am. J. Physiol.* 275 (*Adv. Physiol. Educ*.): S206-S216, 1998.

Änggård, E. "Nitric oxide: mediator, murderer, and medicine." *Lancet* 343: 1199-1206, 1994.

Tonegawa, S. "The molecules of the immune system," *Sci. Am.* 253: 122-131, 1985.

Return to top

Return to Content

Chapter 24.

Body Fluids And Regulation

Study Objectives

- To *define* the concepts: Dehydration, hyponatraemia, intracellular fluid volume (ICV), extracellular fluid volume (ECV), interstitial fluid (ISF), overhydration, oxidation water, radioactivity, specific activity, and total body water.
- To *describe* the daily water balance, the K^+- and Na^+-balance, sweat secretion, the ionic composition in blood plasma, the water content of fat- and muscle- tissue and the daily water transfer across the gastro-intestinal mucosa. To describe the osmotic pressure in the body fluids, the measurement of fluid compartments by indicator dilution, the measurement of total body-K^+ and -Na^+ and the related dynamic pools.
- To *draw* models of the body fluid compartments.
- To *explain* the influence of age, sex and weight on the size of the total body water and its phases. To explain disorders with increased or reduced extracellular fluid volume and shock.
- To *apply* and *use* the above concepts in problem solving and in case histories.

Principles

- *The law of conservation of matter states that mass or energy can neither be created nor destroyed (the principle of mass balance). The principle is here used to measure physiological fluid compartments and the body content of ions.*

Definitions

- **Concentration:** The concentration of a solute is the amount of solute in a given fluid volume.
- **Dehydration** is a clinical condition with an abnormal reduction of one or more of the major fluid compartments (ie, total body water with shrinkage of blood volume or ISF).
- **Dextrans** are polysaccharides of high molecular weight.
- **Intracellular fluid volume** (ICV) refers to the volume of fluid inside all cells. This volume normally contains 26-28 litre (l) out of the total 42 l of water in a 70-kg person. - One litre of water equals one kg of water.
- **Extracellular fluid volume** (ECV) refers to the interstitial and the plasma volume. The ECV contains the remaining water (14-16 kg) with most of the water in tissue fluid (ISF) and about 3 kg of water in plasma. - Interstitial fluid (ISF) is the tissue fluid between the cells in the extravascular space.
- **Hyperkalaemia** refers to a clinical condition with plasma-[K^+] above 5 mM (mmol/l of plasma).
- **Hypokalaemia** refers to a clinical condition with plasma-[K^+] below 3.5 mM.
- **Hypernatraemia** refers to a clinical condition with plasma-[Na^+] above 145 mM.
- **Hyponatraemia** refers to a clinical condition with plasma-[Na^+] below 135 mM.
- **Oedema** refers to a clinical condition with an abnormal accumulation of tissue fluid or interstitial fluid.

- **Osmolality** is a measure of the osmotic active particles in one kg of water. Plasma-osmolality is given in Osmol per kg of water. Water occupies 93-94% of plasma in healthy persons. Plasma osmolality is normally maintained constant by the antidiuretic hormone feedback system.

- **Overhydration** refers to a clinical condition with an abnormal increase in total body water resulting in an increased ECV and thus salt accumulation.

- **Oxidation water** or *metabolic water* (oxidative phosphorylation) refers to the daily water production by combustion of food - normally 300-400 g of water daily in an adult.

- **Radioactivity** is measured as the number of radioactive disintegrations per s (in **Becquerel** or Bq per l). *One disintegration per s* equals *one Bq*.

- **Total body water** is destributed between two compartments separated by the cell membrane: The intracellular and the extracellular fluid.

Essentials

This paragraph deals with *1. The three major fluid compartments, 2. Water balance, 3. Body potassium, 4. Body sodium, 5. The indicator dilution principle, 6. The renin-angiotensin-aldosterone cascade, 7. Output contol, 8. Regulation of renal water excretion, and 9. Regulation of renal sodium excretion.*

Read first about the nephron (paragraph 1 of Chapter 25).

1. The three major fluid compartments

The three major body fluid compartments are the intracellular fluid volume (ICV), the interstitial fluid volume (ISV) and the vascular space (Chapter 1, Fig.1-4). *Water permeable* membranes separate the three compartments, so that they contain almost the same number of osmotically active particles per kg. The three compartments have the same concentration expressed as mOsmol per kg of water or the same freeze-point depression. They are said to be *isosmolal*, because they have the *same osmolality*.

The so-called *lean body mass*, which means a body stripped of fat, contains 0.69 parts of water (69%) of the total body weight in all persons. - Such high values are observed in the newborn and in extremely fit athletes with minimal body fat. Babies have a tenfold higher water turnover per kg of body weight than adults do.

As an average females have a low body water percentage compared to males. Such differences show *sex dependency*, but the important factor is the relative content of body fat, since fat tissue contains significantly less water (only 10%) than muscle and other tissues (70%). This is why the relative water content depends upon the relative fat content.

The average for most healthy persons is 60% of the body weight. Sedentary, overweight persons contain only 50-55 % water dependent on the body fat content.

The relative content of *body fat* rises with increasing age and body weight, and the relative mass of muscle tissue becomes less. Consequently, the body water fraction falls with increasing body weight and age. Aging implies loss of cells, but the ECV is remarkably constant through life and under disease conditions.

Each body (weight 70 kg) contains 4 mol of both sodium and potassium (ie, the total ion pool). A minor fraction of the potassium is radioactive. The calcium and magnesium content is 25 and 1 mol, respectively.

In the renal tubule cells the epithelium is a single layer of cells, joined by junctional complexes near their luminal border (Fig. 25-7). Solutes can traverse the epithelium through transcellular or paracellular pathways. Virtually every cell membrane in the body contains the Na^+- K^+-pump, which maintains the *low* intracellular Na^+-concentration and develops the negative, intracellular voltage. In the renal tubule cells the Na^+- K^+-pump, is located in the basolateral membrane. Read more about the nephron in Chapter 25 and about hormonal control later in paragraph 8 and 11 of this Chapter.

Unfortunately, the simple laws of dilute solutions are unprecise at physiological concentrations. Rough estimates are

based on the assumptions that extracellular sodium is associated with monovalent anions and that deviations in osmolality are twice the deviation in plasma sodium concentration.

ICV: The dominating intracellular solute is potassium (K^+), balanced by phosphate and anionic protein, whilst the dominating extracellular solute is NaCl. All compartments have almost the same osmolality 300 mOsmol* kg^{-1} of water. The thin cell membrane - or the endothelial barrier between ISF and plasma in the vascular phase - cannot carry any important hydrostatic gradient. Water passes freely between the extra- and intra-cellular compartment, as osmotic forces govern its distribution and the membranes are water permeable.

Fig. 24-1: The daily water transfer across the gastrointestinal barrier in a healthy standard person.

The ICV comprises 26-28 kg out of the total 42-kg of water in a 70-kg person (Fig. 1-4).

ECV: The ECV compartment comprises the remaining water (14-16 kg) with most of the water in *tissue fluid (interstitial fluid* or *ISF)* and 3 kg of water in plasma (Chapter 1, Fig. 1-4). The size of the ECV compartment is proportional to the total body Na^+. Changes in plasma osmolality indicate problems in water balance.

A [Na^+] in ECV of 150 mmol per kg of plasma water corresponds to a total osmolality of 300 mOsmol per kg.

Alterations in plasma-[Na^+] (osmolality) will be followed by similar changes of the ECV osmolality, because the permeability of of the capillary barrier for Na^+ and water is almost equal.

The daily water transfer across the gastrointestinal tract amounts to approximately 9 l in each direction (Fig. 24-1).

2. Water balance

A healthy person on a mixed diet in a temperate climate receives 1000 ml with the food and drinks 1200 ml daily. Balance is maintained as long as the water loss is the same (Fig. 24-2).

Fig. 24-2: The daily water balance in a 70-kg healthy person on a mixed diet. The apparent imbalance between input (2200 ml) and output (2500 ml) is covered by 300 ml of metabolic water.

Water is lost in the urine (1500 ml), in the stools (100 ml), in sweat and evaporation from the respiratory tract (900 ml) as a typical example.

The total loss of water is 2500 ml, and this corresponds perfectly to the intake plus a normal production of 300 ml of *metabolic water* per 24 hours (Fig. 24-2).

3. Body potassium

The daily dietary intake of potassium varies with the amount of fruit and vegetables consumed (75-150 mmol K^+daily).

More than 90% of the body potassium is located intracellularly. Only a few percent of the K^+ in the body pool are found outside the cells and subject to control (Fig. 24-3). The main renal K^+-reabsorption is passive and paracellular through *tight junctions* of the proximal tubules. Moreover K^+-excretion can vary over a wide range from almost complete reabsorption of filtered K^+ to urinary excretion rates in excess of filtered load (ie, net secretion of K^+).

The Na^+-K^+-pump located in the cell membrane, maintains the *high* intracellular [K^+] and the *low* intracellular [Na^+]. The energy of the terminal phosphate bond of ATP is used to actively extrude Na^+ and pump K^+ into the cell. The membrane also contains many K^+ - and Cl^- -channels, through which the two ions leak out of the cell.

In myocardial cells, as in skeletal muscle and nerve cells, K^+ plays a major role in determining the resting membrane potential (RMP), and K^+ is important for optimal operation of enzymatic processes.

Under normal conditions, the RMP of the myocardial cell is determined by the dynamic balance between the membrane conductance to K^+ and to Na^+. As [K^+] is reduced during hypokalaemia, the membrane depolarises

out

causing voltage-dependent inactivation of K^+-channels and activation of Na^+-channels, allowing Na^+ to make a proportionally larger contribution to the RMP.

Fig. 24-3: The total body K^+-pool in a healthy person comprises 4000 mmol with more than 90% intracellularly. The normal ECG and the ECG of a patient with hyperkalaemia is shown to the right.

The K^+ -permeability is around 50 times larger than the Na^+ -permeability, so the RMP of normal myocardial cells (typically: -90 mV) almost equals the equilibrium potential for K^+ (-94 mV).

The excretion of K^+ by overload is almost entirely determined by the *extent* of distal tubular secretion in the *principal cells*. Any rise in serum [K^+] immediately results in a marked rise in K^+-secretion. This transport mechanism is controlled by *aldosterone* and by K^+. Aldosterone stimulates the secretion of K^+ and H^+ by the principal cells of the renal distal tubules and collecting ducts (Fig. 25-11). This is why chronic acidosis decreases and chronic alkalosis increases K^+-secretion. – Actually, acute acidosis may reduce K^+-secretion.

Of the consumed K^+, 75-150 mmol is daily absorbed in the intestine. Since 90% is excreted renally in a healthy person, there must be a minimum in a typical volume of 1500 ml of daily urine with a concentration of (75/1.5) = 50 mM. Normal urinary [K^+] is at least 30 mM. A high urinary [K^+] is indicative of a high *total body K^+* or a high intake of K^+.

The normal *excretion fraction* (Chapter 25) for K^+ is 0.10 (10% or 90 mmol of the 900 mmol in the daily filtrate) corresponding to the daily intake (Fig. 24-4). A K^+ -poor diet leads to hypokalaemia with less than 20 mmol K^+ in the daily urine. A K^+ -rich diet triggers a large secretion and a high excretion in the urine (Box 25-1). A low urinary [K^+] is indicative of a low total body K^+ or of *extracellular acidosis* with transfer of K^+ from the cells in exchange of H^+. A low [K^+] in the distal tubule cells reduces the K^+ -excretion**.**

The *normal plasma-[K^+]* level is dependent upon the exchange with the cells, the renal excretion rate, and the extrarenal losses through the gastrointestinal tract or through sweat.

Measurement of total and exchangeable body potassium

Our natural body potassium is ^{39}K, but we also contain traces of naturally occurring radioactivity (0.00012 or 0.012% is ^{40}K with a half-life of 1.3×10^9 years). When using this natural tracer, injection of radioactive tracer is avoided.

The person to be examined is placed in a sensitive whole body counter, and the *total activity* of the tracer ^{40}K in the body (S Bq) is measured.

Specific activity (SA) is the concentration of radioactive tracer in a fluid volume divided by the concentration of naturally occurring, non-radioactive mother-substance. The concentration of mother-substance is traditionally measured in mmol per l (mM). SA is equal to *radioactivity* (A) per non-radioactive mass unit, m (ie, A/m in Bq/mol). Following even distribution, the SA for a certain substance must be the same all over the body. SA is preferably measured in plasma (with scintillation counters or similar equipment).

Specific activity (SA) is here the number of Bq ^{40}K per mol of mother substance (^{39}K) in the whole body. We can calcu
0.00012. The *total body potassium* of a healthy person is 4000 mmol. The SA of ^{40}K implies a $^{40}K/^{39}K$ ratio of 0.48 mmol/4000 mmol (=0.00012).

An *exchangeable ion pool* in our body is the dynamic part of the total specific ion content. The remaining content is fixed as insoluble salts in the bones. The dynamic character implies the use of a dilution principle to measure such a pool.

In order to measure the *exchangeable* body potassium pool, a radioactive tracer is injected, such as ^{42}K with a physical half-life of 12 hours (12.4 hours) and urine is collected. The first urine sample is from the first 12 hours, and the second sample is covering 12 - 24 hours. The total tracer dose given must be adjusted for by the loss of tracer in the urine and by the radioactive decay during the first 12 hours mixing period. The two urine samples obtained are examined for tracer and for natural potassium. The tracer is assumed to distribute just as natural potassium after 12 - 24 hours. When the tracer is distributed evenly in the exchangeable body potassium, its SA must be the same in urine, plasma or elsewhere in the body. The exchangeable body potassium is calculated by Eq. 24-2 .

The specific activity for the tracer (SA Bq per mol) is known from the plasma measurements. In this way we measure the exchangeable body potassium. The normal values are *41 mmol* ^{39}K per kg body weight for females, and *46 mmol* per kg for males.

4. Body sodium (^{23}Na)

The *exchangeable* body sodium is easy to measure using the dilution principle and a minimum of equipment.

Our natural non-radioactive body sodium is ^{23}Na. We administer the radioactive tracer, ^{24}Na, with a physical half-life of 15 hours. We have to use a total period of 30 hours to secure even distribution in the ECV.

The total tracer dose given, must be adjusted for by the loss of tracer in the urine, and the radioactive decay of ^{24}Na (see the decay law in Chapter 1). The exchangeable body sodium is calculated by Eq. 24-2.

We know the specific activity for the tracer (SA Bq/mol) from the plasma measurements; therefore calculation of the exchangeable body ^{23}Na is easy.

The normal value for exchangeable body sodium is 40 mmol/kg of body weight. In a patient with a body weight of 75 kg the exchangeable sodium is (75×40) = *3000 mmol*. The non-exchangeable sodium is fixed in the bones.

The *total body sodium* is measured following discrete radiation with a method called *neutron activation analysis*. The whole body of the patient is exposed to radiation with neutrons. A small fraction of the natural ^{23}Na now becomes radioactive sodium (^{24}Na) by uptake of an extra neutron.

A sensitive *whole body counter* records the radiation from ^{24}Na. Now we can calculate the *total body sodium.*

Normally, the total body sodium is 1000 mmol larger than the *exchangeable* sodium due to the fixed sodium content of the bones (1000 + 3000 mmol = 4000 mmol ^{23}Na).

Fig. 24-4: Body fluid electrolytes. Water permeable membranes separate the three compartments, which contain almost the same number of osmotically active particles per kg.

The sum columns of electrolyte equivalents in muscle cells are essentially higher than the extracellular sum columns of equivalents, because cells contain proteins, Ca^{2+}, Mg^{2+} and other molecules with several charges per particle (Fig. 24-4).

The above columns show the ionic composition per kg of water, so we have 150 mmol of Na per kg of plasma water. Normally, one litre of plasma has a weight of 1.040 kg and contains 10% of dry material. Consequently, one litre of plasma contains 0.940 l of water, and the rest consists of plasma proteins and small ions. Thus the fraction of water in plasma (F_{water}) is typically 0.94.

5. The indicator dilution principle

Mass conservation is always the underlying principle. The amount of indicator n mol distributes in V litres of distribution volume.

We measure the concentration C_p in mM, following even distribution, and calculate V:

$$V = n/C_p.$$

Errors: Uneven distribution of indicator introduces a systematic error. - A non-representative concentration of

indicator in the plasma makes it insufficient to correct for plasma proteins alone. - Loss of indicator to other compartments is inevitable. - Elimination or synthesis of indicator in the body occurs as frequent errors. - The indicator may be toxic or in other ways change the size of the compartment to be measured.

Total body water, ECV, plasma volume, and the elimination rate constant are measured as follows:

5 a. Total body water

Total water is measured by the help of the dilution principle. *Tritium marked water* is a good tracer. The equilibrium period is 3-6 hours. n mol of indicator divided by C_p mmol of indicator per l is equal to the distribution volume (V) for the indicator.

Healthy adolescents and children have normal values around 60% of the body weight assuming one l of water to be equal to one kg. Adult males and females with a sedentary life style and larger fat fractions contain only 50% of water.

5 b. The extracellular fluid volume (ECV)

is measured by administration of a *priming dose* of inulin intravenously. Then inulin is infused to maintain a steady state with constancy of the plasma concentration of inulin (C_p).

The patient then urinates, and the infusion is stopped with collection of a plasma sample. For the next 10 hours the patient collects his urine, which makes it possible to measure all the body inulin present at the end of the infusion (n mol) assuming all inulin excreted.

Dividing n with C_p gives the volume of distribution (V) after correcting for the difference in protein concentration between plasma and ISF (Eq. 24-1).

Chromium-ethylene-diamine-tetra-acetate (^{51}Cr-EDTA) is a chelate with a structure that cannot enter into cells. The chelate molecule contains radioactive Cr, making it easy to measure. The ^{51}Cr-EDTA distributes and eliminates itself in the extracellular fluid volume (ECV) just as inulin and is therefore used to measure ECV. – For clearance measurements, we inject a single dose intravenously, and draw blood samples every hour for 5 hours. The clearance of ^{51}Cr-EDTA is independent of C_p and a good estimate of GFR just like the *inulin clearance*. Since the indicator is cleared from the ECV only, it is possible to measure its size. Such methods - including renal lithium reabsorption - are important during renal function studies. Normal values for ECV are approximately 20% of the body weight or 14-17 kg.

Chronically ill patients with debilitating diseases often maintain their ECV remarkably well in spite of marked reductions in the cell mass of their body.

5 c. The plasma volume

Also here, the dilution principle is used. The indicator for plasma volume can be Evans Blue (T_{1824}) that binds to circulating plasma albumin. A small dose of albumin, marked with radioactive iodine, is also a good indicator (iodine ^{131} has a physical half-life of 8 days).

The indicator concentration in plasma (C_p) is measured every 10-min for an hour after the administration, and the log of C_p is plotted with time. Extrapolation to the time zero determines the maximum concentration of indicator in plasma. This corrects for the biological loss, while the indicator distributes itself in the plasma phase. The tracer dose divided by C_p at time zero provides us with the *intravascular* plasma volume. Normal values for the plasma volume are close to 5% of the body weight.

In *diabetics* and *hypertensive* patients the tracer is lost more readily through their leaky capillaries to the interstitial fluid than in healthy persons (increased transcapillary escape).

6. The renin-angiotensin-aldosterone cascade

Macula densa is described in paragraph 9 of Chapter 25.

The most likely intrarenal trigger of the renin-angiotensin-aldosterone cascade is the **falling** NaCl concentration of the

reduced fluid flow at the macula densa in the distal renal tubules (Fig. 24-5).

The NaCl concentration at the macula densa falls, when we lose extracellular fluid, move into the upright position and when the blood pressure falls.

Renin is a proteinase that separates the decapeptide, angiotensin I, from the liver globulin, angiotensinogen.

When angiotensin I passes the lungs or the kidneys, a dipeptide is separated from the decapeptide by angiotensin converting enzyme (ACE). This process produces the octapeptide, angiotensin II.

Angiotensin II has multiple actions that minimize renal fluid and sodium losses and maintain arterial blood pressure.

1. Angiotensin II stimulates the aldosterone secretion by the adrenal cortex, and through this hormone it stimulates Na^+-reabsorption and K^+-(H^+)-secretion in the distal tubules (Fig. 24-5). - Angiotensin II is in itself a potent stimulator of tubular Na^+-reabsorption.
2. Angiotensin II inhibits further renin release by negative feedback.
3. Angiotensin II constricts arterioles all over the body including a strong constriction of the efferent and to some extent also the afferent arteriole. Hereby, the renal bloodflow (RBF) and to a lesser extent the glomerular filtration rate (GFR) is reduced.
4. Angiotensin II inhibits the absolute proximal tubular reabsorption – contributing to the reduction of GFR.
5. Angiotensin II enhances sympathetic nervous activity.

Fig. 24-5: The renin-angiotensin-aldosterone cascade.

Sympathetic stimulation of the renal nerves stimulates renin secretion directly via b-adrenergic receptors on the JG cells just as falling blood pressure in the preglomerular arterioles. - b-blocking drugs and angiotensin II inhibit the renin secretion (Fig 24-5).

The combined effects from the whole renin cascade is extracellular fluid homeostasis.

In contrast, exposure to stress and painful stimuli triggers the combined sympatho-adrenergic system with release of catecholamines, gluco- and mineralo-corticoids, and ACTH from the hypophysis. ACTH stimulates further the secretion of the glucocorticoid, cortisol, from the adrenal cortex.

7. Output control

The body uses *output control*, when it is overloaded with water or with sodium.

The most important osmotically active solute in ECV is NaCl, because it only passes into cells in small amounts. Urea, glucose and other molecules with modest concentration gradients are without importance, because they distribute almost evenly in the fluid compartments.

Healthy persons use two primary control systems: 1) The osmolality (osmol per kg of water) or ion concentration controls our elimination of water. 2) The change of blood volume (ECV) or pressure controls sodium excretion - not osmolality.

Only when the arterial blood pressure falls *drastically* the body will drop its protection of normal concentration. In such a disease state large amounts of ADH molecules are released in an attempt to improve the volume and blood pressure.

8. Regulation of renal water excretion

The primary control of the renal water excretion is *osmolality control* (Fig. 24-6). Since 2/3 of the body water normally is located within the cells, this is also an intracellular volume control.

Following *water deprivation* even an increase in plasma osmolality of only one per cent stimulates both the hypothalamic osmoreceptors and similar (angiotensin-II-sensitive) thirst receptors. Thirst may increase the water

intake of the individual and thus increase the ECV, with negative feedback to the thirst receptors.

Activation of the hypothalamic osmoreceptors and thirst receptors increases the hypothalamic neurosecretion to the neurohypophysis and releases antidiuretic hormone (ADH or vasopressin). Hyperosmolality elicits a linear increase in plasma ADH, which causes water retention (Fig. 24-6) until isosmolality is reached.

ADH increases the reabsorption of water from the fluid in the renal cortical and medullary collecting ducts. ADH binds to receptors on the basolateral surface of the tubule cells, where they liberate and accumulate cAMP. This messenger passes through intermediary steps across the cell to the luminal membrane, where the number of water channels (aquaporin 2) are increased. The luminal cell membrane is thus rendered water-permeable, which increases the renal water retention. The increased water reabsorption leads to a small, concentrated urine volume (antidiuresis), and a net gain of water that returns ECF osmolality towards normal. Initially, osmolality control overrides blood volume control.

Fig. 24-6: Primary osmolality control of the renal water excretion. ADH and thirst systems maintain osmolality and ICV within narrow limits.

Water overload decreases ECF osmolality and has the reverse effect, because the hypothalamic osmoreceptors suppress the ADH release, and the renal water excretion is increased already after 30 min (Fig. 24-6). When a person rapidly drinks one litre of water, the intestine absorbs water. Ions diffuse into the intestinal lumen and the blood osmolality falls causing a block of the ADH secretion (Fig. 24-6).

Pure water is distributed evenly in all three body fluid compartments – just like intravenous infusion of one litre of 5% glucose in water.

Intake of one l of isotonic saline implies ECV expansion, without dilution of body fluids. This expansion will not increase the urine volume much, so the increased ECV can be sustained for many hours. An intravenous infusion of one l of large dextran molecules (macrodex) stays mainly in the vascular space.

9. Regulation of renal sodium excretion

In healthy persons, changes of *blood volume* (or ECV) or *blood pressure* control sodium excretion (Fig. 24-7). The dominating cation of the ECV is Na^+ . The sodium intake is balanced by the sodium excretion as long as the thirst and other homeostatic systems are functional.

During conditions where sodium intake exceeds renal sodium excretion, total body sodium and ECV increase. Conversely, total body sodium and ECV decrease, when sodium intake is lower than renal sodium excretion. This is because volume-pressor-receptors detect the size of the circulating blood volume (ECV) or pressure, and effector mechanisms adjust the renal sodium excretion accordingly.

The volume-pressor-receptors are widely distributed. *Low-pressure receptors* are found in the pulmonary vessels and in the atria. An increased blood volume can also increase the arterial blood pressure and stimulate the well-known *high-pressure baroreceptors* in the carotid sinus and the aortic arch. Increased arterial pressure reduces sympathetic tone – also in the kidneys, whereas decreasing arterial pressure enhances sympathetic tone and renal salt retention. Arterial pressure receptors are also located in the renal *preglomerular* arterioles. Both stimuli in Fig. 24-7 release renin from macula densa, whereby angiotensin II and aldosterone is secreted (both sodium retaining hormones).

A decrease in circulating blood volume leads to a decrease in NaCl delivery to the macula densa and release of the renin cascade. Conversely, an increase in circulating blood volume with increased NaCl delivery to the macula densa suppresses renin release and increases sodium excretion (Fig. 24-7).

Fig. 24-7: Primary blood volume-pressure control of the renal Na^+ -excretion. The effective circulating blood volume is protected – also during shock (Na^+ -retention) and during hypertension (natriuresis).

Increased salt intake increases blood volume and leads to natriuresis, possibly augmented by release of ANP (see below), nitric oxide and other factors. The excretion of Na^+ depends upon several effector mechanisms out of which

three are classical:

The *first* factor is the *glomerular filtration rate* (GFR), which is responsible for the size of the filtered flux of Na^+ across the glomerular barrier in the kidneys. Renal prostaglandins, generated in response to angiotensin II, are involved in maintaining the filtered flux of Na^+.

The *second* factor is the *renin-angiotensin-aldosterone cascade* (Fig. 24-5).

The *third* factor consists of *peptides* with natriuretic effects. The most well-known peptide is called *atrial natriuretic peptide* (ANP) and originates from granules of the atrial myocytes. A low circulating blood volume with low atrial pressure increases renal sympathetic tone, reduces the stimulus of the low-pressure receptors in the atrial wall and thus the ANP secretion. Hereby, the natriuresis is reduced. - Renal natriuretic peptide or urodilatin from the distal tubule cells is related to ANP. Urodilatin has been isolated from human urine and contains four amino acids more than ANP.

An increase in effective circulating blood volume, increases atrial pressure, reduces sympathetic tone and releases ANP and urodilatin leading to increased natriuresis.

The main purpose of these mechanisms is to maintain an effective circulating blood volume by an increase or a decrease of the renal excretion of Na^+. Initially, osmolality control is dominating. Finally, after a dangerous reduction in blood volume, volume-pressure receptors override the hypothalamic osmoreceptors and stimulate the ADH release and thirst. In the terminal phase, the body protects effective circulating blood volume at the expense of ECF osmolality.

Pathophysiology

This paragraph deals with 1. Dehydration, 2. Overhydration, 3. Hyponatraemia, 4. Hypernatraemia, 5. Hypokalaemia, and 6. Hyperkalaemia.

1. Dehydration

Dehydration is an *abnormal reduction of the major fluid volumes (total body water* with shrinkage of ECV). When we lose more than 5% of the total body water it has clinical consequences. The condition is life threatening if the patient loses 20 %.

Accidents and surgery with a period of water deprivation, imply a rise in ECF osmolality and thus stimulation of both thirst and the hypothalamic osmoreceptors, whereby ADH is released. - Symptoms and signs of dehydration are *thirst*, dry *mucous membranes*, and decreased skin elasticity or *turgor* due to loss of ISF.

Loss of effective circulating blood volume implies a low blood pressure in both the venous and the arterial system. Loss of more than one litre of ECV causes *postural hypotension* with dizziness, confusion and cerebral failure. Empty veins and cold skin characterise the peripheral venoconstriction. Finally, there is extreme tachycardia, which turns into terminal bradycardia and an arterial blood pressure that approach zero.

Loss of salt and water frequently develops into *hypo-osmolal dehydration* (Fig. 24-8). This is because the thirst forces the patient to drink (salt free) water. Water dragged into the cells further reduces the hyposmolal ECV (Fig. 24-8). The small ECV elicits a hyperaldosteronism, which is called *secondary*, because it is not initiated as primary hypercorticism in the adrenal cortex. A precise compensation of the water loss results in *pure hyponatraemia*, where water eventually is drawn from ECV into the cells. The low [Na^+] around the swelling cells reduces the potential gradient across the cell membranes with increased neuromuscular irritability (muscular twitching) and cardiac arrhythmias.

Isosmolal dehydration is a proportional loss of water and solutes. There is no concentration gradient over the cell membranes, and the loss is mainly from ECV (Fig. 24-8).

Fig. 24-8: Dehydration (hyperosmolal, isosmolal and hyposmolal).

Hyperosmolal dehydration occurs in persons deprived of water. The hypero
dehydrates the cells (Fig. 24-8). This is intracellular dehydration.

The hyperosmolality liberates ADH to restrict the water loss. The patient excretes a very small urine volume.

Persons deprived of water at sea may drink seawater. Sea water is hypertonic saline and the victims die faster. When hypertonic saline reaches the ECV it aggravates the intracellular dehydration simultaneously with an extracellular overhydration. Intracellular dehydration leads to respiratory arrest and death of thirst.

2. Overhydration

Overhydration is an *abnormal increase of total body water* - in particular ECV, and thus salt accumulation. The increase in the interstitial fluid volume is called oedema. Overhydration frequently occurs among patients in fluid therapy (ie, overhydration of iatrogenous origin).

Increased salt intake by mouth is compensated by increased salt excretion by normal kidneys.

However, a large saline infusion (0.9% NaCl) will expand ECV and total body water (isosmolal overhydration in Fig. 24-9). Inappropriately large infusions of saline lead to iatrogenous hyperosmolal overhydration, if they lose more water than salt (Fig. 24-9).

Hyperosmolality drags water from the cells, so that the patient develops *intracel*
loss of consciousness and eventually respiratory arrest.

The patient with *hyposmolal overhydration* is typically in fluid treatment and develops muscle cramps and disorientation. The skin turgor is normal. A low serum - [Na^+] confirms the diagnosis. The water overload in ECV is dragged into the cells in hyposmolal overhydration until osmolality balance (Fig. 24-9).

In the brain and the muscles this *intracellular overhydration* causes headache, disorientation, increased spinal pressure, coma and muscle cramps. Both *hyposmolal* and hyperosmolal *intracellular overhydration conditions* are characterised by cerebral symptoms and signs.

Fig. 24-9: Overhydration (hyperosmolal, isosmolal, and hyposmolal).

Acute renal failure with decreased GFR reduces the flux of filtered NaCl (*first factor*) and thus the Na^+ -excretion.

Oedema is a clinical condition where the interstitial fluid volume (ISF) is abnormally large.

A *voluminous ISF* is usually due to *increased* hydrostatic venous pressure (heart insufficiency), or a *reduced* colloid osmotic pressure (hypoproteinaemia) as predicted from Starlings law for transcapillary transport.

Reduced protein synthesis (liver disease) and *abnor*
hypoproteinaemia. Thus protein-losing kidneys are involved.

Capillary damage (allergy, burns, inflammation etc) with increased capillary permeability *causes* local oedema. – Obstruction to lymphatic drainage can also cause oedema (scarring after radiation therapy, elephantiasis etc).

Cardiac insufficiency with increased venous pressure and oedema formation increases sympathetic tone and thus releases the *renin-angiotensin-aldosterone cascade* (Fig. 24-5) causing Na^+ -retention.

Hepatic cirrhosis activates the cascade in a similar way - possibly including the release of nitric *oxide*.

Hypoalbuminaemia reduces the colloid osmotic pressure of plasma, whereby water is distributed from the vascular space to the ISF. The fall in effective circulatory volume activates the renin cascade and leads to Na^+ -retention.

NSAIDs can activate the renin-angiotensin-aldosterone cascade, and the increased aldosterone leads to Na^+ -retention and overhydration.

Angiotensin II-receptor antagonists and *ACE-inhibitors* are utilized clinically to block the effects of angiotensin II in congestive heart failure, diabetes mellitus and hypertension. Blockade of the cascade reduces both preglomerular and postglomerular resistances.

The *supine position* at bed rest increases venous return. This implies an increased cardiac output (Starlings law), a reduced ANF secretion from the atrial walls and a reduced renin-angiotensin-aldosterone cascade. This is why bed rest is beneficial for disorders with salt accumulation.

3. Hyponatraemia

Hyponatraemia (ie, plasma-[Na^+] below 135 mM) is associated with dehydration, overhydration or normohydration (ie, a normal ECV and total body sodium content).

Hyponatraemia with *reduced* ECV (ie, salt-deficient hyponatraemia) is caused by a salt loss in excess of the high water loss (ie, hyposmolal dehydration in Fig 24-10). This is seen in any type of hypoadrenalism including the rare primary hypoadrenalism (Addison's disease).

In Addison's disease the entire adrenal cortex is destroyed by autoimmune reactions (80%) or by malignancy or infection. All three types of hormones are insufficiently produced (mineralocorticoids, glucocorticoids and sex hormones). The lack of aldosterone leads to Na^+ -excretion and K^+-retention with hyponatraemia combined with hyperkalaemia resulting in dehydration and hypotension.

Hyponatraemia is developed in the following way (Fig. 24-10):

1. The first step is the salt loss in excess of the water loss.

2. Since the ECF-[Na^+] is low, the ADH secretion is suppressed, and the water excretion is increased. Hereby, both the ISF and the vascular spaces are reduced often by more than 10%.

3. This is an adequate stimulus for the *volume-pressure receptors*, which override the osmoreceptors, whenever the effective circulatory volume is threatened.

Fig. 24-10: The three body fluid compartments in a patient with salt-deficient hyponatraemia.

The volume-pressure receptors stimulate both thirst and the release of ADH. The effective circulating volume is protected at the expense of osmolality! Still the blood pressure is falling, which impairs cerebral perfusion, causing confusion, headache and coma.

The hyponatraemia implies a reduced resting membrane potential and thus a low threshold for neuromuscular stimulation resulting in muscle cramps.

The large renal loss is seen with osmotic diuresis (hyperglycaemia and uraemia), excessive use of diuretics, renal tubular reabsorption defects, adreno-cortical insufficiency as aldosterone-antagonist-intoxication or other types of hypoaldosteronism.

The extra-renal loss is often large from excessive sweating, diarrhoea, haemorrhage, vomiting, loss with ascites or bronchial secretion, and transudation from cutaneous defects. Normal kidneys normally compensate extra-renal loss. The urinary excretion of salt and water falls in response to volume depletion, so the urine is concentrated - but with less than 10 mM Na^+.

Normal sweat is a *hypotonic* solution, because Na^+ is reabsorbed in the duct system. The [Na^+] can increase up to 80 mM with increasing sweat flow - due to the limited time for the aldosterone-controlled Na^+-reabsorption.

Increased salt intake by mouth or intravenously is required as a supplement to the treatment directed at the primary cause.

+

Low plasma- [Na] in a chronically salt-deficient patient suggests a high aldosterone secretion from the adrenal zona glomerulosa. Further administration of aldosterone therefore may not have any effect.

Hyponatraemia with *increased* ECV (water-excess hyponatraemia) is often caused by cardiac, hepatic, and renal insufficiency or by hypoalbuminaemia - see hyposmolal overhydration (Fig. 24-9).

Hyponatraemia with *normal* ECV is often caused by stress (surgery, psychogenic polydipsia), abnormally high ADH release (in the syndrome of inappropriate antidiuretic hormone secretion, and in vagal neuropathy), increased sensitivity to ADH by drugs such as chlorpropamide and tolbutamide, or by intake of ADH-like substances (oxytocin).

Pseudo-hyponatraemia is characterised by a spuriously low plasma value measured conventionally in the total volume of plasma, which includes an extra volume in cases with hyperlipidaemia or hyperproteinaemia etc. Plasma osmolality or plasma-Na^+ measured with ion selective electrodes is the choise and the direct read value is normal. This is because Na^+ is confined to the aqueous phase.

Treatment of *artefactual hyponatraemia* (taking blood from an extremity into which isotonic glucose is infused) is also unnecessary.

4. Hypernatraemia

The *normal* plasma-[Na^+] is *135-145 mM*, and values above 170 mM are rare. Excessive infusion of saline (0.9% NaCl or 154 mM) can lead to hypernatraemia. Such alarmingly high levels create an emergency situation, where glucose infusion is indicated initially in order to reduce the high level slowly. The increased plasma osmolality elicits a strong desire to drink.

The cause is sometimes *water deficit* due to pituitary diabetes insipidus, or to nephrogenic diabetes insipidus, where ingestion of nephrotoxic drugs have made the renal collecting ducts resistant to ADH. – Osmotic diuresis also causes water deficit with hypernatraemia just as excessive loss of water through the skin or lungs.

Primary hyperaldosteronism (Conns disease) and all types of secondary hyperaldosteronism also lead to hypernatraemia combined with hypokalaemia and enlarged blood volume.

Cerebral failure and convulsions are alarming signs, but there are no specific symptoms and signs of hypernatraemia.

Polyuria, polydipsia and thirst suggest diabetes. Diabetes mellitus is easy to diagnose, and diabetes insipidus shows a low urinary osmolality. Pituitary diabetes insipidus is treated with an analogue of ADH (desmopressin, with a low pressor-effect).

5. Hypokalaemia

The normal potassium ion concentration in blood plasma is 3.5-5 mM. Hypokalaemia is caused by renal or extra-renal K^+ -loss or by restricted intake.

Long standing use of diuretics without KCl compensation is a frequent cause of hypokalaemia.

Hyperaldosteronism (increased aldosterone secretion) is another cause.

Vomit fluid only contains 5-10 mM of K^+. Still, prolonged vomiting develops into hypokalaemia, because the Na^+-loss stimulates the aldosterone secretion, which increases $\mathbf{K^+}$-excretion in the kidneys.

Profuse diarrhoea causes marked hypokalaemia, also because the diarrhoea fluid contains up to 50 mM of $\mathbf{K^+}$.

Hypokalaemia is seen in cardiac patients receiving *digoxin* treatment. Digoxin toxicity is imminent, because digoxin firmly binds to myocardial cells in hypokalaemia. Treatment must be directed towards the underlying cause. Infusion of potassium -rich fluid is dangerous, because of the marginal distance to hyperkalaemia.

+

The reduced extracellular K hyperpolarises the cell membrane (increases the negativity of the voltage across the membrane). This reduces the excitability of neurons and muscle cells. Thus, hypokalaemia can result in muscle weakness and paresis. Hypokalaemia is associated with an increased frequency of cardiac arrhythmias with atrial and ventricular ectopic beats in particular in patients with cardiac disease . - Hypokalaemia inhibits release of adrenaline, aldosterone and insulin.

6. Hyperkalaemia

Acute hyperkalaemia (ie, plasma-[K^+] above 5 mM) is a normal condition following severe exercise, and normal kidneys easily eliminate K^+ .

In disease states the causes are *insufficient* renal excretion or *increased* release from damaged body cells as during long lasting hunger, exercise or in severe burns. A plasma- [K^+] above 7 mM is life threatening due to asystolic cardiac arrest.

Long term intake of b-blocking drugs, which inhibit the Na^+-K^+-pump, leads to hyperkalaemia that is accentuated by exercise.

Hyperkalaemia reduces the size of the resting membrane potential (reduces the negativity of the voltage), whereby the threshold for firing is approached in neurons and striated muscle cells. The increased excitability in hyperkalaemia results in muscle contractions, cramps followed by muscle weakness. Hyperkalaemia leads to decreased cardiac excitability, hypotension, bradycardia and eventual asystole. The ECG is characterised by increased duration of the QRS-complexes and *tented* T-waves due to abrupt Ca^{2+}-influx, contraction, and abrupt Ca^{2+} -binding (Fig. 24-3). Cardiac arrest occurs as ventricular fibrillation (the heart can never produce smooth tetanus) or as asystole.

Insulin is used to drive K^+ back into the cells - either by insulin infusion or by glucose infusion in order to release more insulin from the pancreatic islets. Usually, a combined glucose-insulin drop is applied.

Other hormones (adrenaline, aldosterone) also stimulate the Na^+-K^+-pump and thereby increase cellular K^+ -influx (Fig. 24-3)

Equations

- *The indicator dilution method*: The indicator n mmol distributes in V litres of distribu concentration C_p in mM, following even distribution, and calculate the volume, V:

  Eq. 24-1: $V = n/C_p$. (litre = mmol/(mmol/l)

- When the tracer is radioactive potassium and thus distributed evenly in the *exchangeable potassium pool,* its specific Activity (SA) must be the same in urine, plasma or elsewhere in the pool.

  Eq. 24-2: Exchangeable body potassium =

  (Injected - eliminated)/SA. (Mol = Bq/(Bq per mol)

We know the specific activity for the tracer (SA Bq per mol) from the plasma measurements. In this way we measure the exchangeable body potassium. The normal values are *41 mmol* ^{39}K per kg body weight for females, and *46 mmol* per kg for males.

- The following *concentrations* are found in normal plasma:

[Na^+] 135-145, [K^+] 3.5-5, [Cl^-] 96-106, [bicarbonate] 24, and total-[Ca^{2+}] 2.5 mM.

- The concentration of low molecular ions in the *ultrafiltrate* is affected by the Donnan effect (normally 5% for monovalent ions), and by the fractional content of water in plasma (0.94 normally):

  Eq. 24-3: [Low molecular ions] = Plasma conc. * Donnan factor/ 0.94.

$[Na^+]$ = 141× 0.95/0.94 = 143 mmol/l of ultrafiltrate. Based on the Donnan effect alone, this result is less than 141. The Donnan effect on monovalent cations is simply more than compensated by the protein volume effect or fractional content of water in plasma (0.94).

$[Cl^-]$ = 103×1.05/0.94 = 115 mmol/l of ultrafiltrate. Based on the Donnan effect this result should be greater than 103 and the protein volume effect contribute further. Such an ultrafiltrate is present in the kidneys and in ISF.

- The *extracellular fluid volume* (ECV) can be measured if all inulin molecules are collected in the urine over 10-15 hours after the inulin infusion stopped.

  Eq. 24-4: ECV = Amount of inulin excreted/(C_p/0.94).

  The inulin distribution volume is more or less identical to the ECV.

- Concentration of molecules in the filtrate are calculated as follows:

  Eq. 24-5: $C_{filtr} = C_p \times F_{free}/0.94$ (mmol per l of ultrafiltrate).

This value depends upon the fractional content of water in plasma (F_{water} = 0.94 l of water per l of plasma) and of the fraction of free, unbound molecules (F_{free}). For uncharged, free molecules like inulin F_{free} is 1, and for protein-bound molecules F_{free} is lower than 1

Self-Assessment

Multiple Choice Questions:

Each of the following statements has True/False options:

A. Hyponatraemia with normal ECV is often caused by stress, abnormally high ADH release, increased sensitivity to ADH by drugs, or by intake of ADH-like substances.

B. The total water content of a healthy person is 60%, and an extremely obese adult contains relatively more water.

C. Hyponatraemia is defined as a plasma-$[Na^+]$ below 145 mM.

D. A plasma- $[K^+]$ above 4.5 mM is life-threatening.

E. An infusion of one l of 5% glucose is distributed evenly into all three compartments just as pure water. An infusion of one l of saline remains mainly in the ECV, whereas an infusion of one l of macrodex stays mainly in the vascular space.

Case History A

A healthy male with a body weight of 70 kg has a normal extracellular osmolality (300 mOsmol kg^{-1}), and a normal ICV/ECV of 28/14 kg or l of water.

One day he is the victim of severe burns and he suffers a water loss of 2.5 l of water (the salt loss is covered).

1. Calculate the new ECV osmolality following the water loss.

2. Does this hyperosmolality have consequences?

3. Following total restitution of the water compartments the patient undergoes surgery with skin grafts. During the long procedure he receives sufficient water by glucose infusion, but he looses 900 mOsmol NaCl. Calculate the new osmolality.

4. Is it dangerous for a healthy individual to lose 6 kg of water without solutes?

5. Is it dangerous for a healthy individual to lose 6 kg of water as an isosmolal fluid from the ECV?

Case History B

A female patient (age 22 years; weight 71 kg) is in hospital suspect of potassium imbalance. She has taken diuretics for 2 years. She is tired and sleepy; her legs are paretic. The ECG shows prolongation of the Q-T interval, depression of the S-T segments and flattening of the T-waves. Her blood pH is 7.57 and the serum K^+ -concentration is 2.9 mM. One morning she receives an intravenous injection of a solution containing the radioactive isotope of potassium (555,000 Becquerel, Bq, of $^{42}K^+$ with a physical half-life of 12 hours). Following the injection her urine is collected in two periods (0-12 and 12 -24 hours). The first urine collection contained 40 mmol K^+ ($^{39}K^+$) and 4144 Bq $^{42}K^+$. The second urine specimen contained 40 mmol K^+ and 2220 Bq $^{42}K^+$. Both urine specimens were analysed for radioactivity exactly 24 hours after the injection, where the specific activity of her plasma was 55.5 Bq/mmol. The $^{42}K^+$, retained after the first 12 hours, distribute in her body just like all other exchangeable K^+. The body contains traces (0.012% of the total) of naturally occurring radioactivity (^{40}K) with a half-life of 1.3×10^9 years.

1. Calculate the exchangeable K^+ pool of her body after the 12-hour distribution period. - Is the result normal?

2. Calculate the elimination rate constant (k) for exchangeable K^+ in her body, and the biological half-life for this K^+ in hours. Calculate the ratio between the physical and the biological half-life of K^+.

3. What is the cause of her disease?

4. Describe the actions of diuretics.

5. Describe a method for measurement of her total body potassium.

Case History C

This case requires knowledge of the renal function (Chapter 25).

Two groups of substances are evenly distributed in the ECV of a healthy 25-year-old man. His weight is 70 kg, and his extracellular volume (ECV) is 14 L. Both groups of substances disappear solely by excretion through the kidneys. His GFR is 120 ml/min, and his renal plasma flow (RPF) is 700 ml/min.

1. Inulin is representative for one family of substances. Inulin is only ultrafiltered in the kidneys. What fraction (k_1) of the total amount of inulin in the body is maximally excreted in the urine per min?

2. The other substances are not only ultrafiltered, but they are also undergoing tubular secretion to such an extent that they totally disappear from the blood during the first passage. What is the elimination rate constant (k_2) for these substances?

Try to solve the problems before looking up the answers.

Highlights

- *Water permeable membranes separate the three body fluid compartments, so that they contain almost the same number of osmotically active particles (expressed as mOsmol per kg of water or the same freeze-point depression). The three compartments are the intracellular fluid volume (ICV), the interstitial fluid volume (ISV) and the vascular space.*

- *The sum columns of electrolyte equivalents in muscle cells are essentially higher than the extracellular sum columns, because cells contain proteins, Ca^{2+}, Mg^{2+} and other molecules with several charges per particle.*

- *Females contain less water as an average compared to males. Such differences show sex dependency, but the important factor is the fraction of body fat, since fat tissue contains significantly less water than other tissues (only 10%). Sedentary, overweight persons contain 50-55 % water dependent on the body fat content, and regardless of sex.*

- *Primary hyperaldosteronism (Conns hypercorticism disease) and all types of secondary hyperaldosteronism also lead to hypernatraemia combined with hypokalaemia and enlarged blood volume. Cerebral failure and convulsions are alarming signs, but there are no specific symptoms and signs of hypernatraemia.*

- *Polyuria, polydipsia and thirst suggest diabetes insipidus and low urinary osmolality is a clear indication. Pituitary diabetes insipidus is treated with an analogue of ADH (desmopressin, with a low pressor-effect).*

- *Regulation of K^+-balance: The daily intake of K^+ is matched by the renal K^+-excretion and our daily urine contains 2-5 g of K^+.*

- *Acid-base balance. The pH of the ICV and the ECV is maintained within narrow limits (many metabolic processes are sensitive to pH). The acid-base balance is accomplished by co-operative action of the kidneys and the lungs.*

- *Hypokalaemia reduces the excitability of neurons, muscle cells and the myocardial syncytium. Thus, hypokalaemia can result in muscle weakness, paresis, and cardiac arrhythmias with ectopic beats and cardiac arrest in diastole.*

- *Hyperkalaemia increases the excitability of neurons, muscle cells and the myocardium.*

- *Acute hyperkalaemia is a normal condition following severe exercise, and normal kidneys easily eliminate this. In disease states the causes of hyperkalaemia are insufficient renal excretion or increased release from the body cells as during long lasting hunger.*

- *A plasma- $[K^+]$ above 7 mM is life threatening due to ventricular fibrillation or cardiac arrest in systole. Tented T-waves and increased QRS-complexes characterise the ECG.*

Further Reading

Astrup, P, P. Bie, and H.C. Engell**.** 'Salt and water in culture and medicine.' *Munksgaard*, Copenhagen, 1993**.**

Knox, F. G. "Physiology of potassium balance." *Am.J. Physiol*. 275 (*Adv. Physiol. Educ*. 20): S142-S147, 1998.

Return to chapter 24

Return to Content

Respiratory Acidosis (↓ pH_a & ↑ P_{aCO2})

Hypoventilation

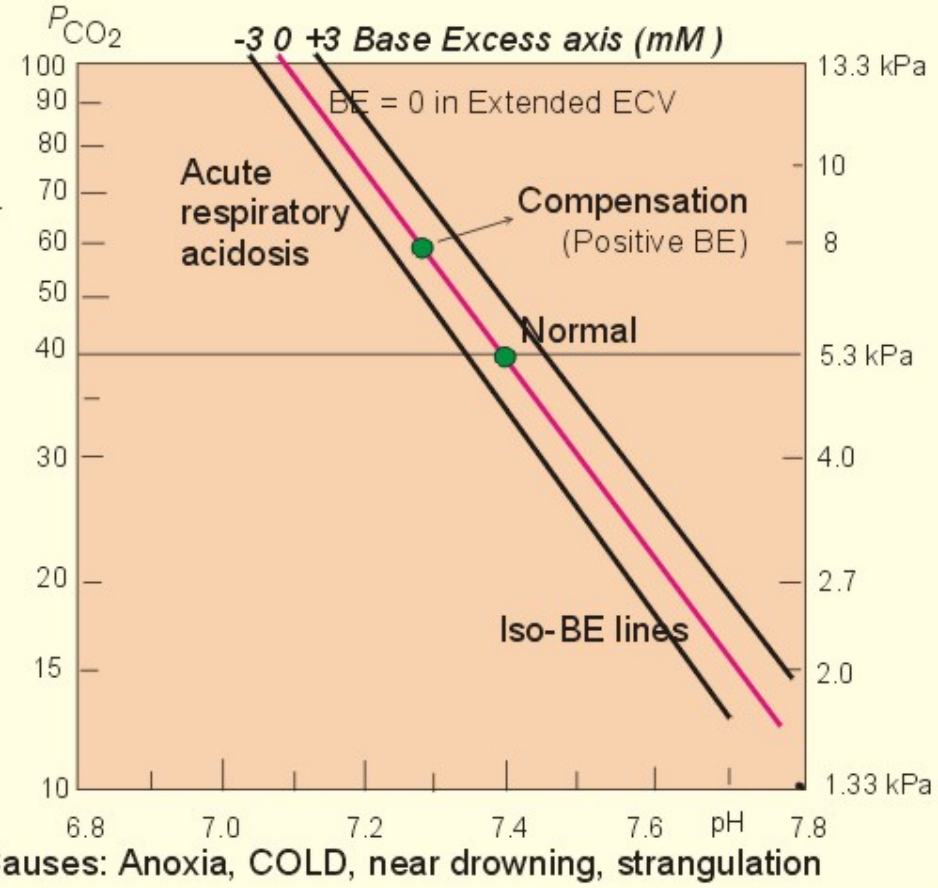
Increase

$CO_2 + H_2O \xrightarrow{*} H_2CO_3 \rightarrow H^+ + HCO_3^-$

$HB \rightarrow H^+ + B^-$

Rise in HCO_3^- = fall in B^-

BE = 0



Causes: Anoxia, COLD, near drowning, strangulation

Fig. 17-9

KMc

Answers

Chapter 1

Multiple Choice Questions

I. Answers **A, B** and D is true, whereas **C** and **E** are false statements.

II. Answers **A, B, C,** and **E** are true statements, whereas **D** is false.

III. Answers **A, C,** and **E** are true statements, whereas **B** and **D** are false

IV. Answers **A, C,** and **E** are true statements, whereas **B** and **D** are false.

Return to chapter 1

Return to content

Chapter 2
Muscle and Cells Disorders

Study Objectives

- To *define* the concepts gap junction, motor unit, synaptic & neuromuscular transfer, isometric and isotonic contraction, plasticity, post-synaptic potentials, and recruitment.
- To *describe* the electromyogram, three types of motor units and three types of muscle tissue (striated, smooth, and myocardial tissue), modulation of neurotransmission with facilitation, potentiation, neurotransmitters and receptors.
- To *explain* the function of the neuromuscular junction, the synapses, the neurotransmitters, and the control of the muscular force by frequency variation and recruitment. To explain disorders of the neuromuscular junction, the skeletal muscles, the smooth muscles and the myocardium.
- To *use* the above concepts in problem solving.

Principles

- *Waller's law of neuronal degeneration: When a motor axon has been severed, the rough endoplasmic reticulum accumulates proteins required for repair of the axon. The axon and the myelin sheath distal to the injury die and are phagocytized. The neuroglial Schwann cells remain alive, proliferate and form long rows along the pathway previously occupied by the dead axon. The severed axon regenerate along this pathway.*
- *Dale's law: A single neuron liberates only one neurotransmitter at all its synapses. Although the law is frequently valid, there are several exceptions, where two or more co-transmitters are released at all the synapses of a single neuron.*

Definitions

- **Excitatory postsynaptic potential** (EPSP) refers to a transientdepolarization of a neuron membrane. The combined effect of EPSPs from hundreds of presynaptic terminals can summate to evoke an action potential.
- **Gap junctions** are transmembrane protein pores between cells. The pores represent a low electrical resistance. Most *electrical synapses* contain many gap junctions allowing free passage of ions and small molecules in both directions when open.
- **Inhibitory postsynaptic potential** (IPSP) is a transient hyperpolarization of a neuron membrane. The negativity of the resting membrane potential increases (normally -70 mV) and summation of IPSPs may result in an effect.
- **Isometric contraction** is a muscular contraction at constant length.
- **Isotonic contraction** is a muscle contraction at constant tension (load).
- A **miniature endplate potential** is probably caused by the spontaneous release of a single acetylcholine vesicle into the synaptic cleft. This is called *quantal release*.
- **Motor unit** refers to one motor neuron and the group of muscle fibres it innervates. All muscle fibres belonging to a certain motor unit are of the same type.
- **Neurotransmission** refers to transfer of signals from one neuron to another mediated electrically or chemically.
- The **neuromuscular endplate** is the *contact zone* between the axons of motor neurons and striated muscle fibres. The acetylcholine containing vesicles of the axon terminals dock on the *release sites* of the presynaptic membrane with high affinity. The muscle cell membrane at the endplate is folded in *junctional crypts*. Nicotinic acetylcholine receptors are concentrated at the openings of these crypts.
- **Plasticity** refers to mechanical plasticity of smooth muscle tissue or to an amplification produced by synapses, which transmit better when frequently used.
- **Recruitment** refers to the increase in force and contraction velocity of a muscle by activation of more and more motor units.
- **Sarcomere** is a contractile unit of a muscle fibril containing the halves of two I-bands with the A-band in between (ie, the part of the fibril between two neighbour Z-lines).
- **Synaptic transfer** refers to the transmission of signals from one neuron to another, and the site of contact between the two neurons is called the *synapse*.

Essentials

This paragraph deals with 1. Neuromuscular junctions, 2. Synapses, 3. Skeletal muscles, 4. Smooth muscles and 5. Cardiac muscle tissue.

1. Neuromuscular junctions

The neuromuscular endplate is the *contact zone* between the axons of motor neurons and striated muscle fibres. Axon terminals have vesicles containing *acetylcholine* (Fig. 2 -1). The vesicles dock on the *active zones* or *release sites* of the presynaptic membrane with high affinity. The muscle cell membrane at the endplate is folded in junctional folds or *crypts* (Fig. 2-1). *Nicotinic acetylcholine receptors* (Chapter 6) are concentrated at the openings of these junctional crypts. The release sites are located directly over the acetylcholine receptors (Fig. 2-1). The postsynaptic membrane has *acetylcholinesterase* all over its surface.

The *nicotinic acetylcholine receptor* is related to a ligand (acetylcholine)-gated ion channel found not only in the neuromuscular junction, but also at all autonomic ganglia (Chapter 6) and in the central nervous system (CNS). The receptor is fixed into the postjunctional membrane, whereas *acetylcholinesterase* is loosely attached to its surface. The receptor has five integral protein subunits (2a , 1b , 1g , 1d), surrounding a *central ion channel pore* that is opened by the binding of 2 acetylcholine molecules to the 2 a - proteins (Fig. 2-1). Opening of the ion channel increases the conductance for small cations (Na^+ and K^+) across the postjunctional membrane, depolarising the membrane potential of the cell. These ion channels are not voltage-gated (not dependent on changes in membrane potential), like most cation channels in neurons, cardiac and skeletal muscle cell membranes.

Fig. 2-1: The neuromuscular junction and intracellular events. Acetylcholine = ACh. The ACh-receptor to the right is magnified.

The acetylcholine-vesicles are probably already stored close to the *release zones*, awaiting the release signal (Fig. 2-1). When the action potential (AP) reaches the axon terminals, the axon membrane is depolarised, and voltage-gated Ca^{2+}-channels are transiently activated. This causes Ca^{2+} to flow down its concentration gradient from the outside into the axon terminal. The influx of Ca^{2+} at the release zones causes the vesicles to fuse with the axon membrane, and empty acetylcholine into the 50 nm wide cleft by *exocytosis* (Fig. 2-1).

After crossing the synaptic cleft by diffusion, acetylcholine binds to its *receptor protein* on the muscle cell membrane. This binding complex opens the ion channel and increases the conductance for small cations across the muscle cell membrane. The influxes of Na^+ depolarise the *endplate* temporarily, the transient depolarization is termed the *endplate potential* (EPP). The EPP dies away when acetylcholine is hydrolysed to acetate and choline by the enzyme, *acetylcholinesterase*. The EPP has a large safety margin, as a *single action potential* in the motor axon will produce an EPP that always reaches the threshold potential in the muscle fibre.

Rapid contraction of the muscle fibre is achieved by propagation of the muscle action potential along the whole length of the muscle fibre membrane and into the small, transverse tubules, which penetrate all the way through the muscle fibre (T-tubules in Fig. 2-1).

The acetylcholine binding at the motor endplate increases endplate conductance and generates an action potential (AP) in all directions from the end plate (Fig. 2-1). The electrical excitation of the sarcolemma and the transverse tubules (T-tubules) during the AP triggers – by an unknown mechanism - the *sarcoplasmic reticulum* to release a pulse of Ca^{2+} (Fig. 2-1). The Ca^{2+}-channels opens transiently in the vicinity of each sarcomere (Fig. 2-1). The sarcoplasmic $[Ca^{2+}]$ increases from 10^{-7} to 10^{-6} M (which is the threshold). This Ca^{2+} diffuses to the adjacent myofilaments, where they bind strongly to troponin C on the active filament, and end the troponin-tropomyosin blockade. This enables cyclic crossbridges to work as long as the high $[Ca^{2+}]$ is maintained, whereby contraction occurs. A continually active Ca^{2+}-pump returns Ca^{2+} to the sarcoplasmic reticulum, and another Ca^{2+}-pump in the cell membrane also reduces sarcoplasmic $[Ca^{2+}]$. Then the thin filament is *off duty*, because Ca^{2+} is withdrawn from its troponin C, the troponin-tropomyosin-blockade is re-established and *relaxation* ensues. The terminal cisternae of the sarcoplasmic reticulum contain granules of calsequestrin, a protein that can bind Ca^{2+} and reduce the concentration gradient (Fig. 2-1).

Neurons with *motor* function have the ability to synthesise acetylcholine, because they contain *choline-acetyltransferase*. This enzyme catalyses the production of acetylcholine from acetyl-CoA and choline. Almost all cells produce *acetyl-CoA* and *choline*. Choline is also actively taken up from the extracellular fluid via a mechanism indirectly powered by the Na^+-K^+-pump. There is a 50% reuptake of choline from the synaptic cleft; hence some choline must be synthesized in the motor nerve.

The postjunctional membrane depolarizes spontaneously - resulting in so-called *miniature endplate potentials* (MEP-potentials). A miniature endplate potential is probably caused by the spontaneous release of a single vesicle into the cleft. This is called *quantal release*.

An endplate potential is prolonged when *cholinesterase-inhibitors* are present in the synaptic cleft. This is because these substances (eserine, edrophonium, malathion, parathion etc.) inhibits the enzyme and thereby protects acetylcholine from being hydrolysed by the enzyme. The life dangerous parathion poisoning is described in chapter 6. Under normal conditions, the endplate potential is terminated by the rapid hydrolysis of acetylcholine by acetyl-cholinesterase.

Acetylcholine is a transmitter in the CNS, in all motor neurons, in all preganglionic neurons of the autonomic nervous system and postganglionic parasympathetic fibres, and in a few postganglionic sympathetic fibres. The cholinergic receptor subtypes are shown in Table 6-2.

2. Synapses

Chemical synapses prevail in humans, but we also have electrical synapses in gap junctions.

A chemical synapse consists of a neuronal presynaptic terminal, a synaptic cleft and a subsynaptic (or postsynaptic) membrane with associated receptor proteins (Fig. 2-2). The chemical synapse is highly developed in the CNS. It conducts the signal one way only, and has a characteristic *synaptic delay*.

The *presynaptic axon terminal* typically broadens to form a *bouton terminaux* (presynaptic terminal).

Fig. 2-2: A synapse between a preganglionic and a postganglionic neuron.

1. The action potential, originating in the CNS, *depolarises* the axon membrane by selective influx of Na^+ , which has a large electrochemical gradient. Repolarization follows rapidly by selective K^+-efflux (Fig. 2-2).
2. When the action potential reaches the presynaptic membrane, Ca^{2+} enters the terminal through voltage-gated *Ca^{2+}-channels*.
3. Vesicles containing transmitter, fuse with the presynaptic membrane and release their contents of acetylcholine into the synaptic cleft (Ca^{2+}-induced exocytosis).
4. Transmitter molecules (acetylcholine, ACh) diffuse across the synaptic cleft and bind to specific receptors, which are located into the postsynaptic membrane (Fig. 2-2). This ligand binding elicits a transient opening of *pores***,** which are specifically permeable to small cations. The synaptic cleft of a chemical synapse is about 30 nm.
5. The ACh-receptor opens and allows influx of Na^+, whereby the membrane depolarizes and an action potential is generated which propagates along the length of the postganglionic axon (Fig. 2-2). This is an appropriate response of the postsynaptic cell to the received signal.
6. The effect is rapidly terminated by the highly specific enzyme acetylcholinesterase, which hydrolyses acetylcholine into two inactive products (acetic acid and choline).

Influx of Na^+ or efflux of K^+ through the pores of such receptors changes the postsynaptic membrane potential. If the presynaptic action potential (AP) results in a postsynaptic depolarization, the transient is called an *Excitatory Post-Synaptic Potential* (EPSP). If the AP results in a postsynaptic hyperpolarization, the transient is called an *Inhibitory Post-Synaptic Potential* (IPSP). Excitatory synapses often use *glutamate* as the transmitter. The pores are penetrated mainly by Na^+, which enters the cell, depolarizes the membrane, and produces an EPSP.

The axon hillock on the *cell body* has a high density of voltage-gated Na^+- and K^+-channels. The axon hillock probably integrates the many synaptic potentials, and from here the action potential is generated. The *dendrites* have voltage-gated channels for K^+ and for Ca^{2+}. Recent evidence suggests that dendrites also contain voltage-gated Na^+-channels, which are involved in *electrogenesis* (ie, movement of charge across the membrane).

Each neuron in the CNS is in contact with up to 10^5 presynaptic axon terminals. Synaptic inputs are integrated at the axon hillock by either *spatial* or *temporal* summation.

Spatial summation occurs when inputs from several axons arrive simultaneously at the same postsynaptic cell. Their postsynaptic potentials are additive. EPSPs summate and move the membrane potential closer to the threshold level for firing. Conversely, EPSPs and IPSPs cancel each other out.

Temporal summation occurs when successive APs in a presynaptic neuron follow in rapid succession, so that the postsynaptic responses overlap and summate. Summation is possible because the synaptic potential lasts longer than action potentials by a factor of 10-100 times.

Each individual synapse contains receptors, ion channels, and other key molecules, which are sensitive to the neurotransmitters released at the site. These specific protein molecules are involved in synaptic plasticity and summation.

Electrical synapses. A *gap junction* is a transmembrane pathway of low electrical resistance that connects the cytoplasm of adjacent cells. A gap junction allows the membrane potential of the adjacent cells to be *electrically coupled*. Gap junctions form *electrical synapses***,** which differ from chemical synapses in that transmission, is instantaneous.

An electrical synapse consists of several protein pores, which close in response to increased intracellular [Ca^{2+}] or [H^+] in a cell, thereby increasing their resistance. Open gap junctions exchange ions and small molecules up to a molecular weight of 1000 Dalton.

Gap junctions are found in simple reflex pathways, where rapid transfer of the electrical potential is essential, and between non-neural cells such as epithelial and myocardial cells, smooth muscle cells and hepatocytes.

Neurotransmitters are divided into classical, rapidly acting non-peptides (Box 7-1) and putative, slowly acting neuropeptides (Box 7-2) - all dealt with in Chapter 7.

Here is only described the function of GABA, neuropeptides and dopamine.

The major *inhibitory* transmitters are GABA (gamma-aminobutyric acid) in the brain and glycine in the spinal cord. Binding of GABA to the GABA-receptor opens the pore for Cl^- influx, whereby the subsynaptic cell membrane *hyperpolarises* (Fig. 2-3). The increase in Cl^- conductance stabilises the membrane potential and decreases the efficacy of excitatory transmission. The GABA-receptor pore is permeable to K^+ besides Cl^-. The GABA-receptor has a major inhibitory role in brain function and is the binding site for barbiturates (used as hypnotics in anaesthesia) and for benzodiazepines (used to relieve anxiety).

Fig. 2-3: A $GABA_A$ -receptor in an inhibitory synapse.

The $GABA_A$-receptor shown here is related to *sedation* and *mood*, whereas the $GABA_B$-receptor controls *spasticity* (Chapter 7). Picrotin blocks the GABA-channel.

Glutamate, *aspartate* and related acidic amino acids are the most important *excitatory transmitters* in the brain and spinal cord. Excitatory neurons possess *excitatory* amino acid (EAA) receptors. EAA receptors are a family of receptors with at least four different ions channels: The N-methyl-D-aspartate-receptor **(**NMDA), and three so-called

+

non-NMDA receptors - one of which is the *glutamate receptor*. The NMDA-receptor operates with K -efflux, while Na^+ and Ca^{2+} enters the subsynaptic neuron. Mg^{2+} and many antiepileptic drugs block the NMDA-receptor channel (Chapter 7). Opening of Na^+- and Ca^{2+}-channels, which allow an increased influx of Na^+ and Ca^{2+}, cause the membrane potential to approach the threshold level for excitation. Both a reduced Cl^--influx to the neuron and a reduced K^+-efflux move the membrane potential towards the threshold level and possible excitation. The NMDA-receptor has a separate glycine site.

Neuropeptides (Box 7-2) have slow excitatory or inhibitory transmitter actions. Peptides cannot be synthesized locally in the axon terminals, because they do not have ribosomes.

Fig. 2-4: Peptide neurotransmitters

Peptides are water soluble, and act as hormones by binding to specific cell-surface receptors. *Cell-surface receptors* are a family of guanosine triphosphate-binding proteins, so-called GTP-binding or *G-proteins,* which control and amplify the synthesis of second messengers. Cell-surface receptors for neurohormones can function as transport protein and possess enzyme activity (Fig. 2-4).

Neuropeptides are build by a sequence of amino acids. Neuropeptides are synthesized in the cell bodies of the neurons and transported to the terminal buttons by rapid axonal transport (Fig. 2-4). Some neuropeptides are released together with a non-peptide co-transmitter (Box 7-2).

Some neuropeptides are produced when a large *mother-peptide* is cleaved into several active neuropeptides. Neuropeptides are released from the nerve terminal near the surface of its target cell, and diffuse to the receptors of the target cell. Low concentrations of neuropeptides typically affect the membrane potential by changing the conductance of the target cell to small ions. The action of neuropeptides usually lasts longer than that of enzyme-inactivated transmitters. Following prolonged synaptic transmission, neuropeptides are deactivated by *proteolysis*.

Dopamine and other catecholamines derive from tyrosine via DOPA, which stands for the precursor 3,4-dihydroxy-phenylalanine. Dopamine is the actively accumulated into storage vesicles in the nerve endings together with noradrenaline and ATP. Dopamine activates both presynaptic and subsynaptic D_2-receptors (Fig. 2-5).

Fig. 2-5: Dopamine receptors and the interactions with noradrenaline (NA).

Noradrenaline can be oxidatively deaminated by monoamine oxidase (MAO) located on the external membrane of mitochondria (Fig. 2-5). The enzyme COMT (catechol-O-methyl transferase) can also methylate noradrenaline to nor-metanephrine. MAO and COMT are important in metabolising circulating catecholamines. Re-uptake of noradrenaline is the most important terminator of its actions.

Activation of both D_2-receptors opens K^+-channels and the increased outflux of K^+ hyperpolarizes the membrane. Blockage of the presynaptic D_2-receptors in substantia nigra with antipsychotic drugs reduces K^+-outflux and increases dopamine production and release.

Loss of *dopamine-containing neurons* in substantia nigra results in the lack of dopamine at the D_2-receptors of the striatal neurons. These neurons degenerate in Parkinson's disease causing *muscular rigidity* and *hand tremor* (Chapter 4).

3. Skeletal muscles

Skeletal or striated muscles are attached to a skeleton. Striated muscles are called *striated*, because they have a striking banding pattern. Microscopy with polarised light reveals *dark* (optically anisotropic) *striations* or *A bands* alternating with light or optically isotropic striations or *I bands*. Running along the axis of the muscle cell or *muscle fibre* is the *myofibril bundles* of *filaments* that are visible on electron micrographs. The A band contains the *thick filaments* of myosin, and the I band contains *thin filaments* of actin and tropomyosin (Fig. 2-6). The thin filaments are anchored to

a transverse structure termed the *Z disc* (Fig. 2-6). Each contractile unit contains the halves of two I-bands with the A-band in between. This unit is a *sarcomere*. Sarcomeres have a length of 2.3-2.5 mm between the two Z discs at rest. The central A band is a relatively isotropic substance - also termed the *H-band* - with an *M-line* of darkly stained proteins that link the thick filaments into a fixed position. Contraction takes place by sliding of the filaments.

The sliding of filaments against each other is called the *sliding filament hypothesis*, and since contraction works by cycling of millions of crossbridges, it is also called the *theory of crossbridge cycling*.

The thin filaments are 1-1.2 mm long and consist of small globular proteins that form two helic pearl strings. The *double helix of actin* is supported by a long, thin molecule of *tropomyosin* that is situated along the groove of the double strands of actin (Fig 2-6). Each tropomyosin molecule interacts with 7 actin molecules on each side. *Troponin* is composed of 3 subunits: Troponin-C binds Ca^{2+}, troponin-T reacts with tropomyosin, and troponin-I inhibits the actin-myosin-interaction, when Ca^{2+} is absent. *Dystrophin* is another normally occurring cytoskeletal muscle protein.

The thick filaments are 1.6 mm long, and consist of large myosin molecules. Myosin is a dimer of almost 500 kD. Each monomer consists of one heavy chain and two light chains. The *heavy chain* consists of a helical tail and a globular head (Fig. 2-6). The *light chains* are associated with the head of the heavy chain. Since myosin is a dimer, the *double-helix tail* must end in two *globular heads* (Fig. 2-6). The globular heads contain the ATPase activity and the actin-binding site. The *light chains* control the rate of cross bridge cycling.

Fig. 2-6: Thick and thin filaments. The crossbridge cycle.

The crossbridge cycle theory states that there are *multiple cycles* of *myosin-head* attachment and detachment to *actin* during a muscle contraction. When myosin binds to actin, an actinomyosin complex is formed - with an extremely active ATPase. The interaction between actin and myosin and the hydrolysis of ATP is the basic process that converts chemical energy into mechanical energy.

Each crossbridge consists of two heads. At rest the crossbridge from myosin is not attached to actin. The globular myosin heads are oriented perpendicular to the filament axis (Fig. 2-6), and they have a high standard affinity for actin.

1. Stimulation of a muscle liberates Ca^{2+} in the sarcoplasma, which removes the troponin-tropomyosin blockage of the actin, and actin can react with the binding sites on the globular heads. The crossbridge is now bound to the thin filaments (Fig. 2-6).
2. The binding accelerates the release of ADP and P_i from the actin-myosin complex, and the attached global heads change conformation by 45^o with respect to the filament axis. The head of the crossbridge drags the thick filament *10 nm* along towards the Z-disc or - at constant length - a proportional force is developed. Multiple repetitions of this short sliding process is necessary to result in an appreciable muscle shortening. In the absence of ATP, the crossbridge cycle stops here and the binding is immobile (*rigor link* and *rigor mortis*).
3. The following stage is the binding of ATP to the myosin heads, which weakens the binding to actin and disrupts the *rigor link.*
4. Then ATP is partially hydrolysed on the myosin head, and the resulting energy is stored in the perpendicular head, which has a renewed high standard affinity for actin. If Ca^{2+} is present, a new crossbridge cycle is initiated and may occur 100 times each s. With a cycle movement of 10 nm this is 1000 nm per s for each half of the sarcomere.

Fig. 2-7: Force-length diagram

Force is required to stretch a relaxed muscle, because muscle tissue is elastic, and the force increases with increasing muscle length (Fig. 2-7). The passive blue curve reflects the properties of the elastic, connective tissue, which becomes less compliant or stiffer with lengthening (Fig. 2-7).

A muscle contraction at constant length is termed *isometric*. Force is measured in Newton (N), and *one N* is the force

required to accelerate a mass of one kg with an acceleration of **one** m s^{-2}. In muscles, the traditional expression for *force* is stress or *tension* in N per cross-sectional area of the muscle (N m^{-2}), which is actually pressure (Pascal, Pa). Here, the ordinate is force expressed as a percentage of the maximal force (Fig. 2-7).

1. The length at which maximum active contractile force is developed is called L_o, corresponding to a sarcomere length of 2.15 m m (Fig. 2-7). L_o is the length of the muscle in the body when at rest. At this length there is a maximum number of active crossbridges (Fig. 2-7). When an isolated muscle in an *isometric force* or *stress-meter* is stimulated, the active muscle force decreases with the decrease in overlap between thin and thick filaments; at a sarcomere length of 3.65 m m the isometric force reaches zero (Fig. 2-7). The force is always proportional to the number of cycling cross-bridges interacting with the thin filament.

2. Force also declines at muscle lengths less than L_o (Fig. 2-7). Thin filaments overlapping, and thick filaments colliding against Z-discs cause this. The isometric force (stress) decreases as the sarcomere length is reduced, as shown with the sarcomere length of less than 2.15 m m (Fig. 2-7).

3. When the active muscle length is stretched beyond any overlapping between the thin and the thick filaments the muscle can only develop a force of zero (see the sarcomere length of 3.65 m m with an A-band of 1.6 m m in Fig. 2-7).

The lengths of the thick and thin filaments of human striated muscles are similar (1.6 and 1.2 m m, respectively). They generate *maximal tension forces* at L_o, corresponding to a sarcomere length of 2.2 m m, namely 300 kN per m^2 or kPa.

Muscle power or *work-rate* (Eq. 2-1) is the product of muscle force (N) and shortening velocity (m s^{-1}). The maximal work rate of human muscles is reached at a contraction velocity of 2.5 m s^{-1}. The maximal work-rate is thus (300 kPa *2.5 m s^{-1}) = 750 kW per square meter of cross sectional area.

Hill developed an equation for the shortening velocity of isotonic muscle contractions (Eq. 2-2). The equation is illustrated in Hills force-velocity diagram (Fig. 2-8).

The maximum force is developed at the initial length (Fig. 2-8 right: 18 g of load). At 18 g there is no shortening – the length is unchanged. Stimulation of the unloaded muscle results in maximum shortening velocity (100%). An unloaded crossbridge can cycle at maximal rate, indicated by maximal shortening velocity (Fig. 2-8 right).

The shortening velocity decreases rapidly as the afterload is increased describing a hyperbola (Fig. 2-8 right). With increasing loads the latency is increased and the shortening is reduced (4 and 9 g in Fig. 2-8 left). The latency depends on the length of thepreceding isometric phase. The maximal velocity of shortening is directly proportional to the *myosin ATPase activity*. We increase the velocity of muscle shortening under a given load by the recruitment of additional motor units.

The long human arm muscles shorten at a rate of 8 m per s. Muscles can bear a load of 1.6 times the maximal force before the crossbridges are broken, but under such extreme conditions the work-rate (power) of the muscle approach zero (no shortening in Fig. 2-8 left). This is also the case when a person attempts to lift a motor car - the speed of shortening is zero (isometric contraction). On the other hand, the speed at which a pocket thief operates is probably impressive, although the force is minimal.

Maximal work-rate occurs at a load of 1/3 of the maximal isometric force of the muscle. Here the contractile system has optimal efficiency in converting chemical energy into mechanical energy.

Fig. 2-8: Hill's force-velocity diagrams (right) and related shortening curves (left).

A further rise in filament velocity seems to reduce the potential for actin-myosin interaction. The *crossbridge cycling rate* falls as the load on the crossbridges increases (Fig. 2-8 right).

In a muscle, the force of contraction is graded by *increasing* the frequency of action potentials, and by *recruiting* more muscle cells. Prolonged crossbridge contraction results in physiological *tetanus*. This is a prolonged muscle contraction maintained by the prolonged Ca^{2+}-influx caused by repetitive stimulation.

Human skeletal muscles consist of three functional types of motor units. A motor unit is a motor neuron with the muscle fibres it innervates. All muscle fibres belonging to a motor unit are of the same type. The three types of muscle fibres are characterised in Box 2-1.

Box 2-1. Structural, functional and histochemical characteristics of twitch fibres.

Classification	***Red (I)***	***Red (IIA)***	***White (IIB)***
	Slow oxidative (SO)	FOG	FG
	Intermediate	Red	White
	Slow	FR	FF
	Slow-twitch	Fast-twitch red	Fast-twitch white
Myoglobin	High	High	Low
Oxidative enzymes	High	Intermediate	Low
Glycolytic activity	Low	Low	High
Glycogen	Low	High	Intermediate
Mitochondria	Intermediate	High	Low
Mitochond.ATPase	Intermediate	High	Low
Sarcoplasmic retic.	Intermediate	Dense	Dense
Fibre diameter	Small	Intermediate	Large
Contractions	Postural	Endurance	Powerful
Shortening velocity	Low (I)	Intermed. (IIA)	High (IIB)
Recruitment	First	Second	Last

Most human skeletal muscles are a mixture of all three types of motor units, although the proportions vary considerably.

Type I: The *slow* motor units contain *slow-oxidative* (SO) red *slow-twitch* fibres.They are adapted to continuous postural muscle activity. The fibres have many mitochondria and a high content of myoglobin (red fibres). They depend on aerobic metabolism and the glycogen content is high. Slow motor units have weak but long lasting contractions (slow reaction to a signal or twitch). The fibres are small and are first to be recruited. During light work these highly excitable motor units activate red fibres suited for prolonged activity or endurance activities. Endurance training increases the oxidative capacity of the activated motor units, whereas strength training increases cellular hypertrophy.

Type IIA: *Fast*-twitch, fatigue-resistant (FR) motor units have type IIA twitch fibres with a high or intermediate content of mitochondria, myoglobin, and glycogen. These fibres also rely upon oxidative metabolism (fast oxidative glycolytic = FOG) and have a high level of both oxidative and glycolytic metabolism. The motor units provide contractions of intermediate force and duration, and they resist fatigue. FOG fibres are of intermediate size, and they are recruited before the white fibres. This is in accordance with the *size recruitment principle*: Small or intermediate motor units are easier to activate by excitatory postsynaptic potentials (EPSPs) than large neurons.

Type IIB: Fast-twitch fatiguable (FF) motor units produce fast contractions (fast-twitch), and fatigue easily, as the name implies. Their *large white fibres*, with their dense sarcoplasmic reticuli, are adapted to activities requiring large forces with rapid control of contraction and relaxation. The fast-twitch white fibres (also called type IIB due to the highest shortening velocity) have few mitochondria, small amounts of myoglobin (white fibres), and depend on glycolysis (high anaerobic metabolism). They have only small amounts of glycogen (fast glycolytic = FG). The FF motor neuron is large, the axon is thick and it branches so greatly that the FF motor unit innervates more muscle fibres. This is why FF motor units are capable of powerful contractions. The cell body receives type Ia afferents. The FF units are recruited last and mainly during maximal efforts such as sprinting. The production of ATP by glycolysis matches the high rate of ATP consumption.

We have three major *metabolic sources* of ATP:

1. Phosphocreatine, which is an immediate energy source used for intense white fibre activity such as sprinting. Lohmann's creatine kinase catalyses the efficient reforming of ATP from ADP by the conversion of a small phosphocreatine pool to creatine. Following exercise the *oxygen debt* is repaid and the phosphocreatine pool is restored (Chapter 18).
2. The glycogen stores of the muscle produce ATP rapidly but inefficiently by *glycolysis*, with lactate as the end product.
3. Glucose, free fatty acids, triglycerides and amino acids in plasma are substrates for oxidative phosphorylation. This is a most efficient pathway and the slowest source of energy due to the many steps in the process (Chapter 20).

4. Smooth muscles

The same molecules as in striated muscle essentially cause contraction in smooth muscle, but the intracellular organisation and the dynamic characteristics are entirely different (Box 2-2).

Box 2-2: Characteristics of skeletal, cardiac and smooth muscle cells.

	Skeletal	***Cardiac***	***Smooth muscle***
Diameter (m m)	Up to 100	10	Up to 5
Length (m m)	200 000	50	Up to 200
T-tubules	Yes	Yes	No -Simple caveoli
Regular sarcomers	Distinct	Distinct	No -Look smooth
Regular Z-discs	Yes	Yes	No- but dense bodies
Regular myofibrils	Yes	Yes	Irregular myofibrils
Troponin	Yes	Yes	No
Sarcoplasmic	Yes	Yes	Simple reticulum

reticulum			
Gap junctions	No	Yes	Yes (single-unit)
Extracellular Ca^{2+}	No	Yes	Yes
Refractory period	Short	Long (300ms)	Long
Latency (ms)	10	10	200
Twitch (ms)	10-100	300	3000
Resting membrane pot.(mV)	-80	-90	-50
Force	High	High	Low maintained for days
Energy cost	300-fold	High	Low
Disorders	Atrophy	Cardiac	Asthma, hypertension

Smooth muscles are called so because they lack the distinct *sarcomeric bands* of *striated* muscles. Smooth muscle cells are spindle-shaped and line the hollow organs and the vascular system; the smooth muscle cells are extremely small (Box 2-2). Smooth muscle cells contain a few thick myosin-filaments, and many thin actin-filaments attached to *dense bodies* by a-actin (helical sarcomers). The cells are without *regular* sarcomers, Z disc's, myofibrils and T-tubules. Smooth muscle cells lack troponin. Dense bodies are analogous to Z disc's, and some dense areas are attached to the cell membrane. Smooth muscle cells do not contain a typical sarcoplasmic reticulum, which can store and release Ca^{2+}. Instead some fibres possess an analogous *simple reticular system* located near the *caveoli* of the cell membrane. Caveoli are small invaginations of the membrane, similar to the T-tubules of striated muscles. The more extensive the reticular system is in the smooth muscle fibre, the higher is its shortening velocity due to release of Ca^{2+} mediated by IP_3. Smooth muscle cells maintain large forces almost continually at extremely low energy costs.

The same tension or tone is maintained for days in smooth muscle organs (intestine, urinary bladder, gall bladder) and can be obtained in striated muscle at high energy cost (up to 300 times the smooth muscle rate of ATP consumption).

Smooth muscle cells are extremely sensitive to *extracellular* $[Ca^{2+}]$.

During an action potential the inward flux of ions is not Na^+, but Ca^{2+} through slow Ca^{2+}-channels. They open mainly in response to a ligand binding, but we have also voltage-dependent Ca^{2+}-channels.

The force-length relation is qualitative similar to that of striated muscles, so the *sliding-filament mechanism* is probably analogous (Fig. 2-6).

The smooth muscle mechanism is special, because stimulation results in a maintained isometric force with strongly reduced velocities. Smooth muscle contractions are extremely slow. Ca^{2+} probably regulates the number of active crossbridges in smooth muscle slowly and indirectly.

Smooth muscle cells contain some mitochondria, and they show a slow contraction pattern superimposed on the lasting tonus. Smooth muscle contractions typically last for 3 s, in contrast to striated muscle with total contraction periods of 10-100 ms. Since the energy demand in smooth muscle is extremely low, it is balanced by the *oxidative ATP synthesis*.

Smooth muscle cells do not have an oxygen debt as striated muscles do, although they produce *large amounts of lactate*. This is probably because the ATP-synthesising glycolytic mechanism is located in the cell membrane and is linked to the ATP-utilising Na^+-K^+-pump. Smooth muscle contains far fewer myosin filaments than striated muscle. The myosin crossbridge heads of smooth muscle contain an isoenzyme with much less ATPase activity than that of striated muscle. Ca^{2+}-entry through the cell membrane is much slower than internal release of Ca^{2+}.

A contracting smooth muscle fibre releases Ca^{2+} from two pools. The *large extracellular fluid pool* is essential. In the fibre that possesses a sarcoplasmic *reticulum* similar to the sarcoplasmic reticulum of striated muscle, there is a fast intracellular pool. The smooth muscle cell membrane contains a *$3Na^+$-$2K^+$-pump,* a delayed *K^+-channel,* a ligand-activated and a voltage-dependent *Ca^{2+}-channel,* a sarcolemmal *Ca^{2+}-pump,* and a Na^+- Ca^{2+}-exchanger (Fig. 2-9).

1. A stimulatory ligand is bound to membrane receptors for G-proteins and for ligand-gated Ca^{2+}-channels (Fig. 2-9). The major Ca^{2+}-influx takes place through the ligand-gated (noradrenaline) and the voltage-gated Ca^{2+}-channels. The Ca^{2+}-influx depolarizes their membrane, whereby Ca^{2+} further permeates the membranes. The depolarization by ligand binding thus indirectly opens the voltage-gated channels.
2. When a stimulus acts on reticular receptors via a G-protein it activates phospholipase C. *Phospholipase C* hydrolyses *phosphatidyl inositol diphosphate* (PIP_2) into IP_3 and diacyl-glycerol, DAG (Fig. 2-9).

Fig. 2-9: Contraction and relaxation in smooth muscle cells. The ligand is acetylcholine in visceral cells and noradrenaline, ATP and peptide hormones in vascular smooth muscle cells.

3. IP_3 is bound to a receptor on the simple sarcoplasmic reticulum and this second messenger binding elicits a controlled release of Ca^{2+} from the reticulum. Hereby, the sarcoplasmic [Ca^{2+}] rapidly increases above the threshold for contraction (0.1 m M).
4. The crossbridge cycling is regulated by a *myosin light chain kinase* (MLC kinase) dependent upon both Ca^{2+} and calmodulin. The phosphorylation of myosin to myosin-phosphate is drastically accentuated by the binding of 4 Ca^{2+}-calmodulin to MLC kinase forming a complex. The phosphorylated light chain myosin reacts with actin in the thin filaments and contracts. The rate of sliding and of ATP-splitting is up to 1000-fold slower than in striated muscles.
5. Ca^{2+} is actively pumped out of the cell by an ATP-demanding Ca^{2+}-pump and through a Na^+- Ca^{2+}-exchanger (antiport). The antiport uses the energy of the Na^+-gradient for influx. Reuptake into the poorly developed sarcoplasmic reticulum and the mitochondria is slow compared to cardiac and skeletal muscle tissue.
6. Below the Ca^{2+}-threshold the myosin light chains are dephosphorylated by myosin light chain phosphatase and the contractile structures relax.
7. The Na^+-K^+-gradient across the cell membrane is maintained by the Na^+-K^+-pump (Fig. 2-9).

When the high intracellular [Ca^{2+}] during an action potential is lowered again towards the resting level, the cell relaxes. This is accomplished by stimulation of the *sarcolemmal* Ca^{2+}-pump, and by blockade of both Ca^{2+}-input and Ca^{2+}-release.

Metarterioles and precapillary sphincters without nerve fibres can still respond to the needs of the tissue by the action of local tissue vasodilatators. The following factors cause smooth muscle relaxation, and therefore vasodilatation: Adenosine, NO, lack of oxygen, excess CO_2, increased [H^+], increased [K^+], diminished [Ca^{2+}], and increased [lactate].

Endothelial-derived relaxing factor (EDRF) is recently shown to be *nitric oxide* (NO). Activation of endothelial cells produces NO from arginine, and NO diffuses into the smooth muscle cells. NO stimulates directly the enzyme *guanylate-cyclase*, and by that intracellular [cGMP] elevates.

Circulating acetylcholine *contract***s** the arterial smooth muscles when bound to *cholinergic receptors*.

Smooth muscle cells grow (hypertrophies) as a response to the needs of the body, and they also retain the capacity to divide.

During *hypertension* the lamina media of the arterioles hypertrophies which increases the *total peripheral vascular resistance* in the systemic circulation. These topics are further developed in Chapter 9.

During *pregnancy* the (single-unit, see below) smooth muscles of the myometrium are quiescent and contain few *gap junctions* under the influence of progesterone. At term the myometrium grows and the number of gap junctions *increases*, due to the high oestrogen concentration. Now the myometrium is well prepared for the co-ordinated contractions during *parturition* (see Chapter 29).

Smooth muscle changes length without marked changes in tension. Initially, there is a high tension developed upon stretching; then the tension falls as the myosin and actin filaments are reorganised by slowly sliding against each other. A sudden expansion of the venous system with blood results in a sharp rise in pressure followed by a fall in pressure over minutes. The smooth muscle fibres in the walls of the venous system are highly compliant, because they have accepted a large blood volume without much rise in pressure (*delayed compliance*).

Smooth muscle cells are frequently involved targets in diseases such as hypertension, stroke,asthma, and many gastrointestinal diseases.Smooth muscle cells can be divided into multi-unit smooth muscle and single-unit smooth muscle.

Fig. 2-10: Contraction of multi-unit smooth muscle cells (vascular). A single contraction is elicited by an electrical stimulus and later acetylcholine elicits tetanus. Contraction of multi-unit smooth muscle is controlled by extrinsic innervation or by hormones. Mechanical contact junctions between the cells are not found.

1. In *multi-unit smooth muscle tissues* each cell operates entirely independent of other cells and the cell does not communicate with other muscle cells through gap junctions**.** The discrete cells are separated by a thin basement membrane and often innervated by a single neuron, and their main control is through nerve signals. Thousands of smooth muscle cells belonging to the multi-unit type join by the common innervation in a *functional syncytium***.** Multi-unit smooth muscle is found in the eye (the ciliary muscle and sphincters as the iris muscle of the eye), in large arteries, in the vas deferens, and in the piloerector muscles that cause erection of the hairs. These muscle cells are normally quiescent, insensitive to stretch and they are activated only through their autonomic nerves. Each muscle is composed of *multiple motor units*, hence the name: *multi-unit smooth muscles*. The nerve fibre branches on a bundle of smooth muscle fibres, and form *junctions* with varicosities filled with transmitters. These junctions are analogous to the neuromuscular junctions of striated muscles. The neurotransmitters are acetylcholine and noradrenaline. Multi-unit smooth muscles have developed a *contact junction* with shorter latency than the slowly operating *diffuse junctions* mainly found in the single-unit type.
2. *Single-unit smooth muscle cells* are arranged in bundles such as the arrangement in a viscera eg. intestine, uterus and ureter (Fig. 2-11). These smooth muscle cells commun-icate through hundreds of *gap junctions*, separating the cell membranes by only 2-3 nm, and from pacemaker tissue of variable location, action potentials are generated initiating a contraction of the muscle. In this respect single-unit cells resemble the cardiac muscle.

Fig. 2-11: Single-unit smooth muscle cells resemble cardiac muscle. Activity propagates from cell to cell through gap junctions forming an electrical syncytium. The dense bodies and dense areas contain alpha-actin.

Action potentials generated in one cell can activate adjacent cells by ionic currents spreading rapidly over the whole organ and securing a co-ordinated contraction as though the tissue were a *single unit* or a *syncytium*. These cells are characterized by their spontaneous motility and by their sensitivity to stretch. The spontaneous activity is usually modified by the autonomic nervous system. *Visceral smooth muscle* undergoing peristalsis, generates propagating action potentials from cell to cell.

Other cell-to-cell contacts are *desmosome's* and *intermediary junctions* subserving structural contact. These intermediary junctions transfer mechanical force from one smooth muscle cell to another on the plasma membrane,

causing the single-unit smooth muscle cell to function like a *stretch transducer*.

5. Cardiac muscle tissue

Myocardial cells are built of regular sarcomers just like the skeletal muscles, and they are contracting fast. Myocardial cells form an electrical syncytium in the same way as the single-unit smooth muscle cells. The characteristics of myocardial, skeletal and smooth muscle cells are presented in Box 2-2. Myocardial cells are mononuclear and the myoglobin, enzymatic and mitochondrial content are large just as the red fibres of skeletal muscles. The metabolism of myocardial cells is similar to that of red skeletal fibres, both being designed for endurance rather than speed and strength. The oxygen supply to the heart muscle must be maintained, if it is to synthesise ATP at a sufficient rate. Myocardial cells deprived of oxygen for 30-s cease to contract.

Myocardial cells most resembles smooth muscle in its auto-rhythmicity and syncytial function. Pacemaker cells in the sinus node determine the normal cardiac frequency, because they send out spontaneous action potentials along the conduction system of the heart with a higher frequency than any other cells in the heart. Vagal stimulation releases acetylcholine at the pacemaker cells. Acetylcholine increases the K^+-permeability, whereby K^+ leaves the cell and hyperpolarizes the cell membrane. This is why the pacemaker (cardiac) frequency is reduced by vagal nerve stimulation. Sympathetic stimulation or adrenaline reduces the K^+-permeability, so the depolarization is shortened, and the pacemaker frequency increased.

The prolonged action potential characteristic for myocardial cells is initiated by an abrupt Na^+-influx (phase 0) through *fast Na^+-channels* just as in the striated muscles. The AP plateau is due to a slow Na^+-Ca^{2+}-channel, which deliver Ca^{2+} for the contraction activation. The action potential releases Ca^{2+} from the sarcoplasmic reticulum to the sarcoplasma. The effect is distributed by the cardiac T-tubule system.

Cardiac contraction by crossbridge cycling depends on the presence of extracellular Ca^{2+} just as in smooth muscle tissue. Therefore, use of Ca^{2+}-antagonists reduces the contractile force of the heart, whereas drugs, which increase Ca^{2+}-permeability across the membrane, improve the contraction. In the heart, Ca^{2+}-influx tends to prolong the depolarization just as in smooth muscle cells. The cardiac glycoside, digoxin, selectively binds to and inhibits the sarcolemmal $3Na^+$-$2K^+$-pump, which leads to an increase in intracellular $[\ Na^+]$. Although the Na^+-efflux is inhibited, the redundancy of Na^+ affects the *Na^+-Ca^{2+}-exchanger* ($3\ Na^+$ out for one Ca^{2+} into the cell), leading to an increase in cellular $[\ Ca^{2+}]$ and in the force of contraction. This is the mechanism of the increase in contractile force by digitalis glycosides.

Pathophysiology

This paragraph deals with 1. Disorders of the neuromuscular junctions (myasthenia gravis), 2. Skeletal muscle disorders (dystrophia, dystonia, muscle injuries), 3. Smooth muscle disorders (asthma, hypertension etc) and 4. Myocardial disorders (coronary artery disease, arrhythmias, and chronic heart disease).

1. Disorder of the neuromuscular junction (Myasthenia gravis)

This serious disease is acquired, but the cause is unknown. The development of this *autoimmune disorder* may be related to other diseases. Rheumatoid arthritis treated with D-penicillamine has resulted in myasthenia gravis. More than 50% of the myasthenia patients have *thymic hyperplasia* and some patients have a real *thymoma*.

Many of these patients have an increased blood concentration of antibodies against their own *acetylcholine receptor protein*. There is a *decreased density* of receptor proteins on the postjunctional membrane. This was shown by the use of radioliganded toxins from poisonous snakes (which bind irreversibly to the acetylcholine receptor protein).

The patients are tired and the muscles are extremely weak. This is particularly so for the proximal limb muscles, the extraocular muscles and the neck muscles, whereby the patient has difficulties in lifting the head. Mastication and swallowing is a difficult process.

Fig. 2-12: Neuromuscular junction with antibodies and decreased density of acetylcholine-receptors in a patient with myasthenia gravis.

As just mentioned the blood of most patients with myasthenia gravis contains *autoantibodies* against *acetylcholine (ACh) receptor proteins* on the cell surfaces of the motor end plates etc. The autoantibody competes for the ACh receptor and inhibits synaptic transmission, so muscular contraction is greatly inhibited. Deposition of immune complexes eventually destroys the ACh-receptor protein.

Intravenous injection of an *anticholinesterase* improves the muscle strength immediately, but the beneficial effect is gone within 3 min.

Thymectomy improves the condition and the prognosis also in the group of patients without thymoma.

Oral anticholinesterase (such as pyridostigmine) has beneficial effect over 2-4 hours. They inhibit the enzyme acetylcholine-esterase, and thereby prolong the effect of naturally occurring acetylcholine on the receptors. In severe cases this treatment is inefficient, and immune-suppressants such as corticosteroids are sometimes favourable.

2. Skeletal muscle disorders

Muscular dystrophy is an inherited disorder of skeletal muscles. *Duchenne muscular dystrophy* is an X-linked recessive muscle disorder characterized by the absence of *dystrophin* in the striated muscles and in the myocardium. The locus is localised to the Xp21 region of the X chromosome. Dystrophin is a normally occurring cytoskeletal muscle protein. The patient is a boy, who has to climb up his legs in order to reach the erect posture. Typically, there is proximal weakness with compensatory *pseudohypertrophy* of the calves. There is no cure and the patient dies from myocardial damage.

Dystonias are *prolonged muscle contractions* leading to muscular spasms. There is a simultaneous action of opposing agonist and antagonist groups that produce abnormal postures. Dystonia is painful and particularly resistant to treatment.

Dystonia musculorum deformans begins in childhood with generalized spasms that affect gait and posture. In most cases the cause is a genetic defect.

Spasmodic torticollis causes the head to turn (torticollis) or change posture. Patients with a trigger zone on the jaw benefit from acupressure here.

Muscle injuries are dealt with in Chapter 18.

3. Smooth muscle disorders

The most important disorders are *asthma* (Chapter 14) and *systemic hypertension* (Chapter 12). Smooth muscles are also involved in a disorder of swallowing (*achalasia*), where the myenteric plexus and the lower oesophageal sphincter fail to respond with *receptive relaxation*, and the food accumulates in the oesophagus. Other disorders of the gastro-intestinal smooth muscles are also treated there.

4. Disorders of the myocardium

Coronary artery disorders (smooth muscles and myocardial disease) and congestive heart disease are treated in Chapter 10, and cardiac arrhythmias in Chapter 11.

Only direct therapeutic uses of the systems developed till now are described here.

Nitro-glycerine, nitroprusside and similar drugs relax smooth muscles by transfer of NO from endothelial cells. NO increases *intracellular* [cGMP] (Fig.11-1), which is the basis for the beneficial effect of the drugs on cardiac cramps. These second messengers activate protein kinases that phosphorylate effector proteins such as Ca^{2+}-pumps and K^+-channels. Such vasodilatators stimulate the *sarcoplasmic* Ca^{2+}*-pump*, inhibit Ca^{2+}-influx and stimulate K^+-efflux through the delayed K^+-channel (reduces the excitability). Hereby, the high intracellular [Ca^{2+}] during an action

potential is lowered towards the resting level (10^{-7} mM), and the smooth muscle cell relaxes producing vasodilatation.

Equations

- **Muscle power** (or work rate) equals the product of muscle force and shortening velocity

Eq. 2-1: Power (W) = Force (N) * Velocity (m s^{-1}).

- **Hill's equation.** The force-velocity curve is shown in Fig. 2-7. The curve fits Hill's equation:

Eq. 2-2: Initial shortening velocity (v) = (Po - P)*b/(P + a)

where P is the force or load acting on the muscle, Po is the maximal isometric force or load, **a** is a constant with the dimensions of a force, and b is a constant with the dimensions of velocity.

Self-assessment

Multiple Choice Questions

I. Each of the following five statements have False/True options:

- **A.** Motor neurons synthesise acetylcholine unrelated to their content of choline-acetyltransferase.
- **B.** There is a high density on the subsynaptic membrane of specific acetylcholine receptors.
- **C.** The receptor protein for acetylcholine contains a voltage-gated channel for cations.
- **D.** Binding of acetylcholine elicits a transient opening of ionophores, which are specifically permeable to small ions.
- **E.** Parkinson's disease is possibly caused by loss of dopamine containing neurons in the substantia nigra.

II. Each of the following five statements have False/True options:

A. Nitro-glycerine, nitroprusside and similar drugs contract smooth muscles by transfer of nitric oxide from endothelial cells.
B. All myasthenia patients have a thymoma.
C. During hypertension the lamina media of the arterioles hypertrophies which increases the total peripheral vascular resistance in the systemic circulation.
D. The nicotinic acetylcholine receptor is related to an acetylcholine-gated ion channel found not only in the neuromuscular junction, but also at all autonomic ganglia and in the central nervous system.
E. When the high intracellular [Ca^{2+}] during an action potential is lowered again towards the resting level, the cell contracts. This is accomplished by stimulation of the sarcolemmal Ca^{2+}-pump, and by blockade of both Ca^{2+}-input and Ca^{2+}-release.

Try to solve the problems before looking up the answers.

Highlights

- *Recruitment is the increase in force and contraction velocity of a muscle by activation of more and more motor units.*
- *Synaptic transfer refers to the transmission of signals from one neuron to another, and the site of contact between the two neurons is called the synapse.*
- *A chemical synapse consists of a neuronal presynaptic terminal, a synaptic cleft and a subsynaptic membrane with associated receptor proteins. The chemical synapse is highly developed in the CNS. It conducts the signal one way only, and has a characteristic synaptic delay.*

A gap junction or electrical synapse is a pathway of low electrical resistance that connects cytoplasm of adjacent cells. A junction couples adjacent cells electrically and thus allows synaptic transmission without delay.

- *Neurons with motor function have the ability to synthetize acetylcholine, because they contain choline-acetyltransferase.*
- *GABA (gamma-aminobutyric acid) in the brain and glycine in the spinal cord are inhibitory neurotransmitters. Binding of GABA to the GABA-receptor opens the pore for Cl^- influx, whereby the subsynaptic cell membrane hyperpolarizes. The GABA-receptor has a major inhibitory role in brain function and is the binding site for barbiturates (used in anaesthesia) and for benzodiazepines (used towards anxiety).*
- *Glutamate, aspartate and related acidic amino acids are the most important excitatory transmitters in the brain and spinal cord. Excitatory neurons possess excitatory amino acid (EAA) receptors. These EAA-mediated synapses predominate in the CNS.*
- *Each neuron in the CNS is in contact with up to 10^5 presynaptic axon terminals. Synaptic inputs are integrated by either spatial or temporal summation.*
- *Neuropeptides are built by a sequence of amino acids. Neuropeptides are synthesized in the cell bodies of the neurons and transported to the terminal buttons by rapid axonal transport.*
- *Loss of dopamine-containing neurons in substantia nigra results in lack of dopamine at the D_2-receptors of the striatal neurons. These neurons degenerate in Parkinson´s disease causing muscular rigidity and hand tremor.*
- *Blockade of the presynaptic D_2-receptors in substantia nigra with antipsychotic drugs reduces K^+-outflux and increases dopamine production and release.*
- *The crossbridge cycle theory states that there are multiple cycles of myosin-head attachment and detachment to actin during a muscle contraction. When myosin binds to actin, an actomyosin complex is formed - with an extremely active ATPase.*
- *Muscle power or work-rate is the product of muscle force (afterload in N) and shortening velocity (m s^{-1}). The maximal work rate of human muscles is reached at a contraction velocity of 2.5 m s^{-1}. The maximal work-rate is thus (300 kPa *2.5 m s^{-1}) = 750 kW per square meter of cross sectional area.*
- *Tetanus is a prolonged muscle contraction maintained by the prolonged Ca^{2+}-influx caused by a high stimulation frequency.*
- *Smooth muscle cells are frequently involved targets in diseases such as hypertension, stroke, asthma, and many gastrointestinal diseases.*
- *Smooth muscle cells maintain large forces almost continually at extremely low energy costs. The same tension or tone is maintained for days in smooth muscle organs (intestine, urinary bladder, and gall bladder).*
- *Myocardial cells form an electrical syncytium in the same way as the smooth muscle cells do.*
- *Myocardial cells deprived of oxygen for 30-s cease to contract.*
- *The most important smooth muscle disorders are asthma and hypertension.*
- *The most important myocardial disorders are coronary artery disease, arrhythmias, and chronic heart disease.*
- *Myasthenia gravis is a disorder of neuromuscular contraction. The patients frequently have an increased blood concentration of antibodies against their own acetylcholine receptor protein and thymic hyperplasia.*
- *Duchenne muscular dystrophy is an X-linked recessive muscle disorder characterized by the absence of dystrophin in the striated muscles and in the myocardium.*

Further Reading

- Kupfermann, I. "Functional studies of cotransmission." *Physiol. Rev.* 71: 683, 1991.
- Pollack, G.H. "Muscles and molecules: Uncovering the principles of biological motion." Seattle, Washington, 1990. *Ebner & Sons*.
- Alberts, B. et al. "Molecular biology of the cell." *Sec. Ed.*, 1989, Garland Publishing, Inc., New York & London.

[Return to top](#)

Return to Content

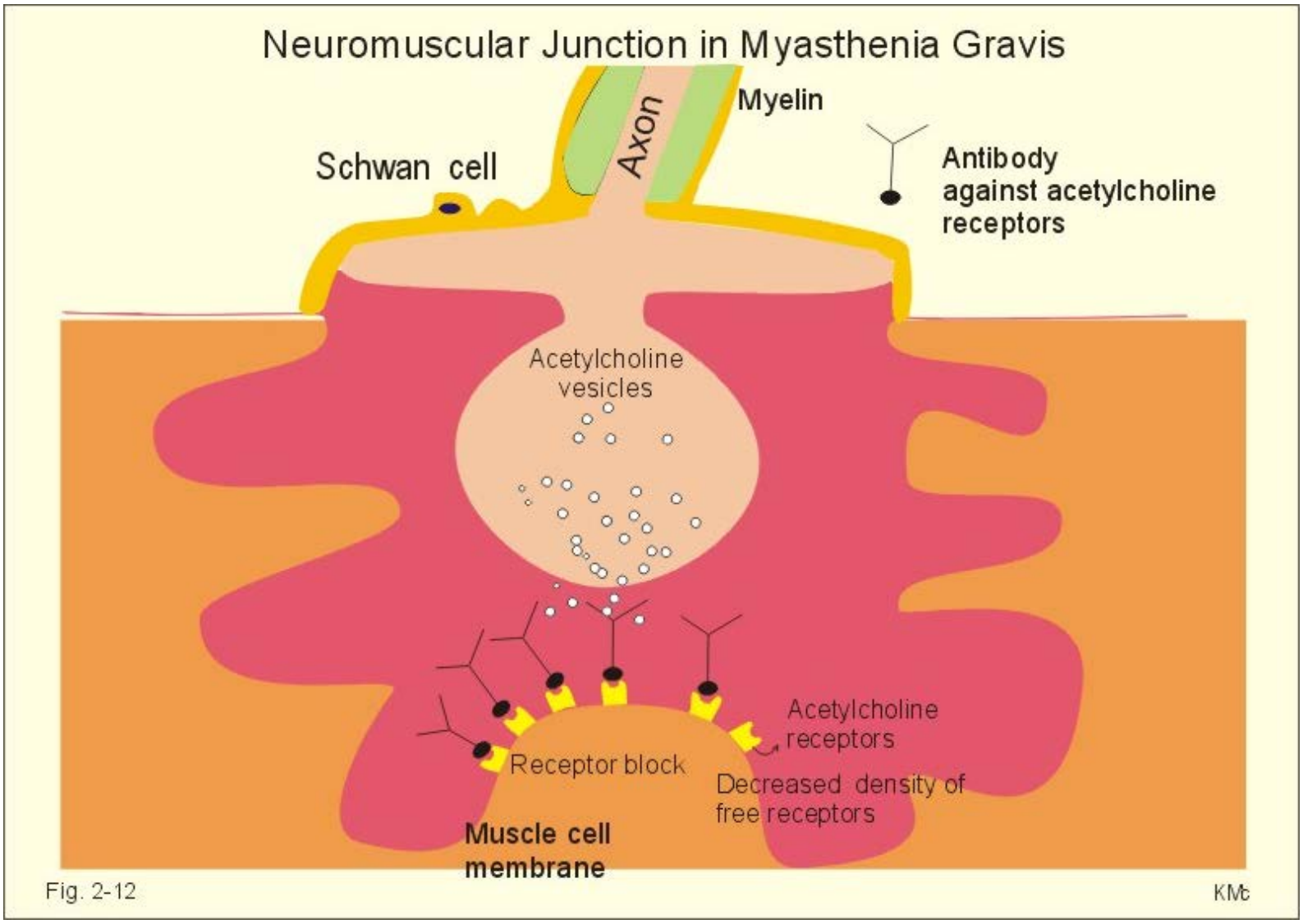

Neuromuscular Junction in Myasthenia Gravis
Myelin
Axon
Schwan cell
Antibody
against acetylcholine
receptors
Acetylcholine
vesicles
Acetylcholine
receptors
Receptor block
Decreased density of
free receptors
Muscle cell
membrane

Fig. 2-12

KMc

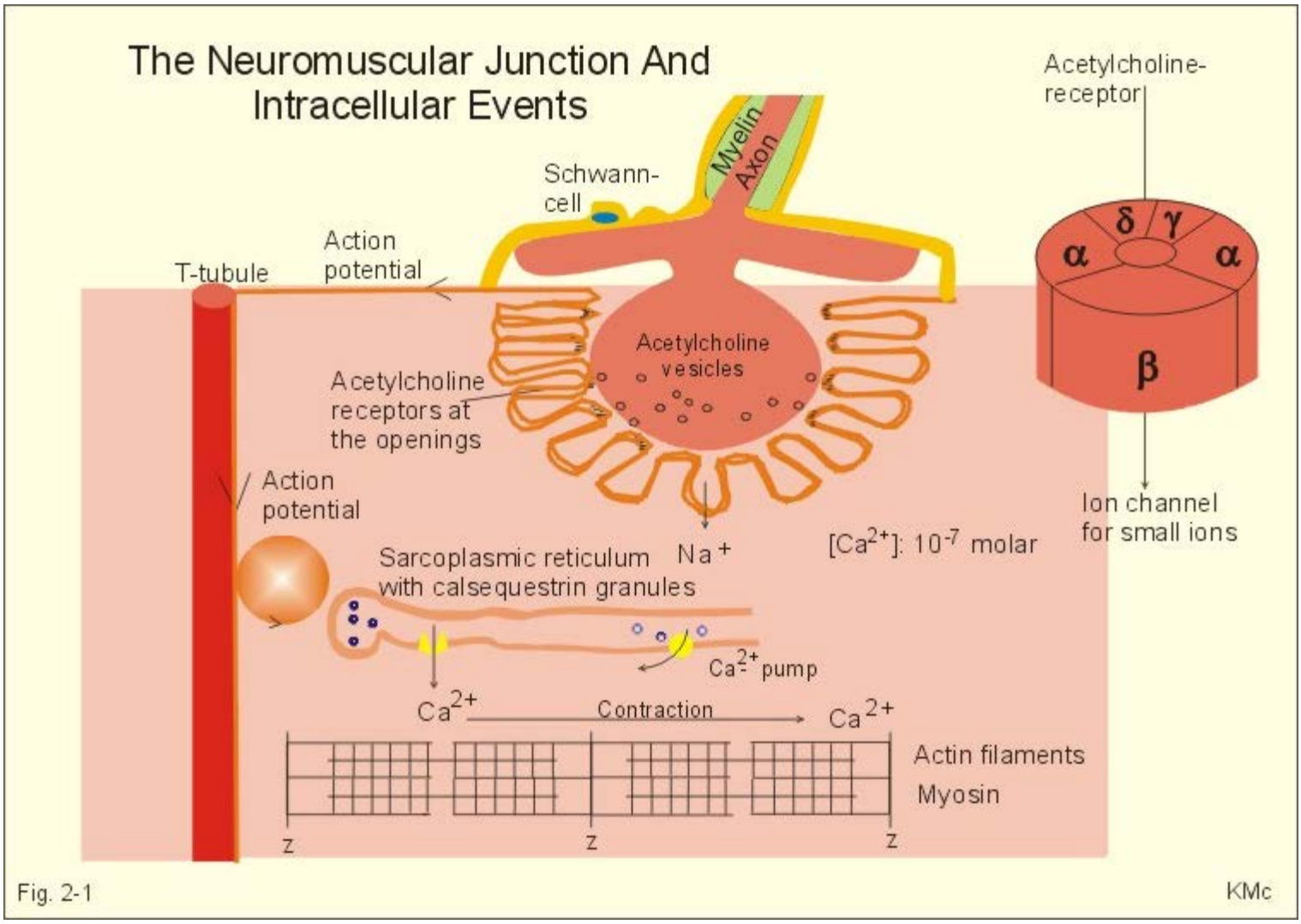

The Neuromuscular Junction And Intracellular Events
Acetylcholine-receptor
Myelin
Axon
Schwann-cell
Action potential
T-tubule
Acetylcholine vesicles
Acetylcholine receptors at the openings
Action potential
Sarcoplasmic reticulum with calsequestrin granules
Na +
[Ca2+]: 10-7 molar
Ca2+ pump
Ca2+
Contraction
Ca2+
Actin filaments
Myosin
z
z
z
δ
γ
α
α
β
Ion channel for small ions

Fig. 2-1

KMc

Chapter 6.
The Autonomic Nervous System And Disorders

Study Objectives

- To *define* receptors, autonomic neurotransmitters and blocking drugs, homeostasis, receptors and related concepts.
- To *describe* the anatomy and the physiology of the sympathetic and the parasympathetic nervous system, the visceral afferent system, the enteric nervous system, transmitter mechanisms in autonomic ganglia and at peripheral receptors, the bladder emptying, the pupillary reflexes.
- To *explain* the central autonomic control, the autonomic control of temperature, appetite, thirst, the subsynaptic autonomic mechanisms, emotional disorders, the Kluver- Bucy-syndrome.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The autonomic nervous system mediates neural control of the internal milieu despite substantial environmental changes.*
- *Cannons law: The peristalsis in the small intestine is polarised, so it always proceeds in the oral-aboral direction.*

Definitions

- **Autonomic neurotransmitters** are adrenergic and cholinergic substances (Box 6-1 and 6-3).
- **Autonomic blocking drugs** (sympatholytics and parasympatholytics) block the normal effect of sympathetic and parasympathetic neurotransmitters.
- **Cholinergic receptors** are *nicotinic* (with a fast EPSP within ms) and *muscarinic* (with a slow EPSP lasting several seconds). Both cholinergic receptors are transmembrane proteins and both open an ion channel in the protein.
- **Homeostasis** refers to all processes helping to keep in internal milieu of the body constant despite environmental alterations.
- **Mydriasis** refers to dilatation of the pupil by sympathetic stimulation of the dilatator muscle.
- **Miosis** refers to contraction of the sphincter muscle (parasympathetic) resulting in a small pupil.
- **Receptors for neurotransmitters** are specific cellular components, who react with a neurotransmitter, a hormone or a drug (agonist) to produce a biological response in the cell.
- **SIF cells** are *small intensity fluorescent cells*, which possess muscarinic receptors and contain vesicles filled with dopamine. Adequate stimulation releases dopamine, which interacts with dopamine receptor (D_2) on the postsynaptic cell body and modulates the effect of acetylcholine. The modulation takes place through a permeability increase for small ions (K^+ out and Cl^- into the cell), hyperpolarizing the cell membrane.
- **The nicotinic receptor** responds to acetylcholine with a rapid influx of Na^+, whereby the membrane is depolarised.
- **The muscarinic receptor**. In the muscarinic M_1 receptor, IP_3 is second messenger and increases cytosolic Ca^{2+}. Activation of the M_2 receptor implies activation of an inhibitory G-protein, which inhibits adenylcyclase. The result is reduced concentration of cAMP, which operates in smooth muscle contraction, with secretion from glands or with a slow EPSP.

Essentials

This paragraph deals with *1. The autonomic system in general, 2. The sympathetic system, and 3. The parasympathetic system.*

1. The Autonomic System In General

The autonomic system directly influences *smooth muscles, glands and the heart* through its two subdivisions, the sympathetic and the parasympathetic system. The two subdivisions function in a dynamic balance aiming at homeostasis.

The *enteric nervous system* is lying within the walls of the gastrointestinal tract and includes neurons in the pancreas, liver and gallbladder, thus being an entity in itself. However, the enteric nervous system is clearly an important part of

the autonomic nervous system that controls gastrointestinal motility, secretion and bloodflow.

The central autonomic system

The central autonomic nervous system outflow arises in the hypothalamus, the brainstem, and the spinal cord. The motor and premotor cortex, the cingulate gyrus and the hypothalamus can modulate the function of the autonomic medullary control neurons in the lateral horn of the grey matter. Circulatory changes during exercise and in various stressful situations are influenced or governed by the cortex and deeper brain nuclei. The central autonomic system also modulates release of certain peptides and catecholamines that affect both blood volume as well as the total peripheral vascular resistance.

The cerebral cortex assimilates all inputs of visual, olfactory, labyrinthine, locomotor origin, as well as from other specialised sensors (stretch receptors, chemo-, baro-, osmo-, and thermo-receptors).

The integration of these inputs into an appropriate response takes place in the hypothalamus and in the ponto-medullary centres. From here the efferent signals pass to the periphery via the sympathetic and the parasympathetic pathways.

The primary afferent projections from the baroreceptors reach the solitary tract nucleus (STN), and from here we have connections to the important dorsal motor nucleus of the vagus (DMNV in Fig. 6-4). A high baroreceptor activity stimulates the DMNV, so that the vagal inhibition of the heart is increased. More importantly, the high baroreceptor activity inhibits the sympathetic drive to the heart and vessels thus reducing blood pressure (Fig. 6-4).

The central autonomic structures co-operate in situations of survival character: Fright, flight or fight- response, feeding and drinking in starvation, reproduction and sexual satisfaction for continuation of life, thermoregulation at extreme temperatures and emotional behaviour in crises.

The Fright -Flight Or Fight Response

Aggression and defence responses are elicited in emergency situations. The sympatho-adrenergic system gives rise to the fright, flight or fight-reactions in acutely stressful situations. The sympathetic reactions dominate over the parasympathetic and the subject is aggressive or anxious. The brain releases corticotrophin-releasing factor to the hypothalamic-pituitary portal system. The hypothalamic-pituitary axis secretes adenocorticotropic hormone, the cardiac rate and contractile force increases, the blood is distributed from viscera to the active skeletal muscles by visceral vasoconstriction and preferential vasodilatation. The subject hyperventilates, the gastrointestinal activity is reduced, and there is increased glycogenolysis and lipolysis. The airways dilatate, and the adrenal medullary (catecholamines) and cortical secretion (cortisol) increases. This response is seen in humans exposed to psychological-emotional stress. Stress in general is comprised of severe emotional and physical burdens (fear, pain, hypoxia, hypothermia, hypoglycaemia, hypotension etc).

Cannons emergency reaction is an immediate sympatho-adrenergic response to life-threatening situations, with both sympatho-adrenergic and parasympathetic overactivity. The last phenomenon includes vagal cardiac arrest with involuntary defecation and urination.

Feeding and drinking

Bilateral destruction of the ventromedial hypothalamic nuclei leads to hyperphagia and failure of body weight control. Such animals become obese, and they have high plasma [insulin].

Bilateral lesions in the lateral hypothalamic regions cause a temporary hypophagia.

The cells of the ventromedial nuclei have a special affinity for glucose, and these cells are responsible for insulin secretion from the pancreatic b-cells. Signals from the dorsal motor nucleus of the vagal nerve increase insulin secretion, and sympathetic stimulation inhibits the release of insulin. The ventromedial nuclei seem to function like a glucostat.

Stimulation and ablations of the limbic system affect food intake. Obviously from clinical practice, psychological factors, emotional disturbances, motivations and conditioned behaviour are all affecting our drive for food intake.

Concerning the control of food intake see also Chapter 20 and 27.

Sexual behaviour

Hypothalamic and other limbic system co-operation are responsible for a wide variety of autonomic and somatic phenomena associated with emotions. Stimulation of the midbrain septum yields pleasurable sensations and sexual drive in patients. The dorsomedial nucleus of the hypothalamus is probably a major sex centre responsible for the sexual act. Stimulation of the ventromedial and preoptic regions also releases sexual activities. See also Chapter 29.

The Thermocontrol

Thermoreceptors can initiate generalised reactions to heat and cold. The signals from both superficial and deep thermoreceptors must act through the hypothalamus to arouse appropriate, generalised reactions.

Cooling or heating the denervated lower extremities of spinal men evoked vasoconstriction and shivering or

vasodilatation and sweating of the innervated upper body shortly after cooled or warmed arterial blood reached the brain. The anterior hypothalamus is responsible for sensing blood temperature variations. The anterior hypothalamus, in particular the preoptic area, has been shown to contain numerous heat-sensitive cells and less cold-sensitive receptors. Such central thermoreceptors are also found at other levels of the CNS. After destruction of the hypothalamus, the midbrain reticular formation takes over the temperature control. Sections eliminating both the hypothalamus and the mesencephalon leave the medulla and spinal cord to control temperature. The posterior hypothalamus does not contain thermoreceptors. Concerning thermocontrol see also Chapter 21.

The brain and the immune defence system

Internal and external stress affects the prefrontal cortex, whereby the limbic system with the hypothalamus is activated. Hypothalamic nuclei release corticotropin-releasing hormone (CRH) to the portal blood. The blood reaches the adenohypophysis, where CRH triggers the release of adenocorticotropic hormone (ACTH), endorphins and met-enkephalin. ACTH works through different pathways in order to protect the body. ACTH stimulates the adrenal cortex to release corticosteroids, which produce immuno-suppression. Immuno-suppression reduces the number of inflammatory effector cells, including helper T cells and killer cells.

On the other hand, cancer therapists assume that relaxed lifestyle and positive reinforcement may have stimulated the immune defence in some patients with malignant diseases, and explain miraculous remissions. Higher brain centres may even affect the reticuloendothelial production of killer cells through the peripheral nerves to the lymph nodes and bone marrow. See also Chapter 32.

Fig. 6-1: The peripheral autonomic nervous system. β -receptors stimulate glycogenolysis in the liver and lipolysis in lipid tissues.

Autonomic nerves are composed of two neurons termed the preganglionic and the Postganglionic neuron based on anatomical location relative to the ganglion. A preganglionic neuron has its cell body in the spinal cord or brainstem and is modulated by higher centres and by spinal reflexes (Fig. 6-1). The preganglionic axon leaves the CNS from the cranial, thoracic, lumbar or sacral regions and synapse in the autonomic ganglia with the cell body of the postganglionic neuron. The postganglionic neurons innervate the effector organs (Viscera).

Viscera function involuntarily and their activity must be modulated by the autonomic nervous system with excitatory or inhibitory signals. All autonomic nerves have ganglia outside the CNS in contrast to the somatic nervous system, where neural connections are located entirely within the CNS. Most somatic nerves that control motor function are myelinated and have a high conduction velocity, whereas most postganglionic neurons are unmyelinated with a low conduction velocity. However, the preganglionic neurons are mostly myelinated with a high conduction velocity (Fig. 6-1).

Receptors for neurotransmitters are specific cellular components, whose interaction with the neurotransmitter, a hormone or a drug produces a biological response in the cell.

Acetylcholine (ACh) is the transmitter between the pre- and the post-ganglionic neurons, not only in the sympathetic nervous system, but also in the parasympathetic system. The cholinergic receptors are nicotinic or muscarinic. The cholinergic receptors of the ganglia and in the somatic motor endplate are *nicotinic*. Nicotine and acetylcholine activate nicotinic cholinergic receptors. When the action potential arrives at the preganglionic fibre, acetylcholine is released from its terminals and diffuses across the synaptic cleft to bind to the specific nicotinic receptors on the membrane of the postganglionic neuron. Nicotinic receptors are linked to cation channels lined with negative charges.

These channels open enough to allow mainly hydrated Na^+ to enter the cell rapidly (for about 1 ms) and depolarise the membrane (Fig. 6-2).

Fig. 6-2: The nicotinic cholinergic receptor

The resulting current elicits an *excitatory postsynaptic potential* (EPSP). Repolarisation is also fast (ms).

Acetylcholine is also the neurotransmitter for the sympathetic innervation of sweat glands, and they are completely blocked by *atropine*. The acetylcholine receptors of the sweat glands are *muscarinic*, since acetylcholine and muscarine (Fig. 6-3) activate them.

Fig. 6-3: The muscarinic cholinergic receptor

These slowly working surface-receptors are linked to a long lasting cascade of events starting with binding of the hormone to the receptor, activation of G-proteins (see Chapter 7), enzyme activation, production of second-messengers, protein kinase activation, and phosphorylation of specific proteins such as channels. All these processes are simplified in Fig. 6-4, and the result is opening of K^+-channels, with efflux of K^+, so the membrane is hyperpolarised. In this example, acetylcholine is an inhibitory transmitter.

2. The Sympathetic Nervous System

The preganglionic sympathetic nerve fibres originate in small multipolar neurons in the lateral horn of the grey matter in the thoracic and lumbar spinal cord. The central sympathetic outflow converges on these preganglionic neurons.

Their axons are thin myelinated fibres that leave the spinal cord through the ventral root. The preganglionic fibres then leave the spinal nerve forming myelinated white rami communicantes, through which they reach the nearest ganglion in the paravertebral ganglia of the paired sympathetic trunk. Typically, each fibre will end here forming synapses with up to 20 postganglionic neurons. A few preganglionic fibres pass the sympathetic trunk without interruption to form the splancnic nerves that reach the three unpaired prevertebral ganglia (coeliac = solar plexus, superior mesenteric and inferior mesenteric) of the lower intestinal and urinary organs. Most sympathetic ganglia are remote from the organ supplied. The postganglionic fibres are all unmyelinated, and they leave the sympathetic trunk through the grey rami communicantes and thus reach the effectors supplied by the sympathetic system. The effectors are the smooth muscles of all organs (blood vessels, viscera, lungs, hairs, pupils), the heart and glands (sweat glands, salivary and other digestive glands). In addition, the sympathetic postganglionic fibres innervate adipocytes, hepatocytes and renal tubular cells.

The sympatho-adrenergic system is a functional and phylogenetic unit of the sympathetic system and the adrenal medulla. The adrenal medulla is a modified sympathetic ganglion. Any increase in sympathetic activity increase the secretion of adrenaline and noradrenaline from the medulla into the circulation. The preganglionic fibres to the adrenal medulla pass all the way to the special postganglionic cells in the adrenal medulla. The synapse is cholinergic (nicotinic) as it is for all preganglionic synapses. The postganglionic cells of the adrenal medulla have developed to cells filled with chromaffine granules, and are called chromaffine cells. These cells do not conduct signals, but synthesise adrenaline (and noradrenaline) which is released into the blood. Sympathetic stimulation triggers the conversion of tyrosine to dihydroxyphenylalanine (DOPA). A non-specific decarboxylase catalyses the conversion of DOPA to dopamine, which is taken up by the chromaffine granules in the cells. The granules contain the crucial enzyme, dopamine b-hydroxylase. This enzyme is activated by sympathetic stimulation, and catalyses the formation of noradrenaline from dopamine.

A few granules store noradrenaline (NA), while the remaining granules liberate NA to the cytosol, where NA is methylated by phenylethanolamine N-methyltransferase to adrenaline. Adrenaline is taken up by chromaffine granules and stored as the predominant adrenal hormone.

Adrenergic Receptors

The sympathetic system exerts either excitatory or inhibitory actions through adrenergic receptors. Adrenergic receptors are *membrane-receptors*. The dual response to adrenergic stimulation was known before Ahlquist in 1948 proposed that adrenergic receptors could be divided into two groups, a- and b- receptors, on the basis of *blocking drugs* (Box 6-1). The basic idea of Ahlquist is that noradrenaline (NA) act predominantly on *vasoconstricting α-receptors,* and isoprenaline (Iso) predominantly on *vasodilatating β-receptors*. Both types of receptors are stimulated by *adrenaline* (Ad).

Box 6-1: Adrenergic receptor subtypes.
The symbol > indicates the rank order of sensitivity.

	Adrenergic receptors			
	α-receptors	**β-receptors**		
Stimulated by:	NA>Ad		Iso>Ad ³NA	
Blocked by	**Phenoxybenzamine**		**Propranolol**	
	α-receptors	**β-receptors**		
a_1-receptors	a_2-receptors	b_1-receptors	b_2-receptors	
Stimulated by	NA>Ad	NA>Ad	Iso>Ad=NA	Iso>Ad>NA Salbutamol
Blocked by	**Prazosin**		**Metoprolol**	**Butoxamine**

The rank order of sensitivity of a series of chemically similar compounds for activating a receptor (agonists) or inhibiting the receptor response (antagonists) is considered diagnostic of the receptor subtype. More and more closely related subtypes are distinguished, so there is already three subtypes for each of the following receptors: a_{1ABC}-receptors, a_{2ABC}-receptors, and b_{123}-receptors.

1. The a-receptors are blocked by Phenoxybenzamine and Phentolamine.

The a_1-receptors are located on the surface of target cells (vascular smooth muscle, sphincter muscles of the gastrointestinal tract and bladder, and radial iris muscles). They are highly sensitive to NA, less sensitive to Ad, and almost insensitive to isoprenaline (Box 6-1).

The a_1-receptors act through phospholipase C and through intracellular [Ca^{2+}] elevation. Ca^{2+} binds to calmodulin in the cytosol. The complex activates protein kinase, which catalyses the phosphorylation of proteins. They become enzymatically active, and trigger vasoconstriction.

In contrast, the presynaptic a_2-receptors are located on the presynaptic membrane (sympathetic end bulbs). NA released into the synaptic cleft diffuses to the a_1-receptors on the target cells, but part of the NA diffuses back to the a_2-receptors on the presynaptic nerve terminals. Here, NA activates membrane adenylcyclase, reducing [cAMP] in the cells, and thus inhibiting release of more NA from the vesicles by negative feedback. Hence, a function of a_2-receptors is auto-inhibitory feedback. These receptors are also found in gastric smooth muscle cells and the b-cells of pancreatic islets. Stimulation decreases gastric motility and attenuates insulin secretion.

2. The b-receptor is blocked by propranolol (Box 6-1).

The b-receptors are located on effector cells that are most sensitive to isoprenaline, but less so to Ad and NA. All b-receptors act through activation of adenylcyclase and cAMP. b_1-receptors are equally sensitive to NA and Ad, whereas b_2-receptors are more sensitive to Ad than to NA (Box 6-1).

b_1-receptors are located in the myocardium - primarily on pacemaker cells. The b_1-receptors of the heart are stimulated by NA which increases cAMP production with increased chronotropic (increased heart rate) and inotropic effect (increased force). Heart patients use Cardioselective b_1-blockers such as Metoprolol, because Metoprolol decreases cardiac arrhythmias and tachycardia.

b_2-receptors are found primarily on bronchiolar smooth muscle cells, vascular smooth muscle, uterine smooth muscle, salivary glands, the intestine and the liver. When NA binds to b_2-receptors, it causes inhibition of the target organ. Therefore, NA causes vasodilatation, bronchodilatation and uterine relaxation. Similarly, sympatomimetics such as b_2-stimulators (salbutamol) increase cAMP production, resulting in bronchodilatation, increased salivary secretion, uterine relaxation and enhanced hepatic glucose output. b_2-stimulators are used to eliminate bronchial asthma attacks. Butoxamine is a selective b_2-blocker.

Box 6-2: Responses elicited in effector organs by sympathetic and parasympathetic activation

Effector organ	**Adrenergic response**	**Cholinergic response**
Heart		
Rate of contraction	Increase, b_1	Decrease, M2
Force of contraction	Increase, b_1	Decrease, M2
Arteries and arterioles		
in myocardium	Vasodilatation, b_2 (a 1constr)	Vasodilatation, M
in skeletal muscles	Vasodilatation, b	

	$_2$	
in lungs	Vasodilatation, b $_2$	
Bronchial muscles	Bronchodilatation, b $_2$	Bronchoconstriction, M
Gastrointestinal		
motility	Decrease, a $_2$ (b $_{2,b\ 3}$)	Increase, M
sphincters	Contraction, a	Relaxation, M
secretion	Decrease, a	Increase, M_1
Exocrine glands		
Salivary	Small secretion, a $_1$	Secretion, M_2
Lacrimal		Secretion, M_2
Digestive	Decreased secretion, a	Secretion, M_2
Airway		Secretion, M
Sweat	Secretion, a 1	Secretion, M
Pancreatic acini	Decreased secretion, a	Secretion, M
Langerhans islets	Decreased secretion, a 2 Increased secretion, b 2	
Lipid cells	Lipolysis, b $_1$ b $_3$	
Liver glycogenolysis	Increase, a $_{1\ b\ 2}$	
Eye		
Ciliary muscle	Relax., b (far vision)	Contraction, M (near vision)
Dilatator muscle of pupil	Contract., a $_1$ (Mydriasis)	
Sphincter muscle of pupil		Contraction, M (Miosis)
Kidney	Renin secretion, b $_2$	
Ureter-motility	Increase, a $_1$	
Urinary bladder		

detrusor	Relaxation, b	Contraction, M
sphincter	Contraction, a_1	Relaxation, M
Genital organs		
male	Ejaculation, a_1	Erection, M
uterus (pregnant)	Contraction, a_1	
Adrenal medulla		Secretion, N

The near-vision response is also called the convergence response. Near vision -even with only one eye - triggers accommodation and pupillary contraction (Box 6-1). The ciliary muscle and the pupillary sphincter muscle are innervated of the parasympathetic oculomotor nerve, and the two muscles (with M-receptors) contract simultaneously for near vision. This leads to increased refractive power or accommodation, and to pupillary contraction (miosis).

When a person closes his eyelids the pupils enlarge, and when he opens the pupils again the pupils become smaller. This is due to the pupillary light reflex, where retinal ganglion cells are stimulated by light, send signals through the optic nerve to the olivary pretectal nucleus neurons. These light-sensitive neurons are connected to the parasympathetic preganglionic neurons in the oculomotor Edinger-Westphal nuclei on both sides. The light reflex contracts the pupillary sphincter muscle. Argyll-Robertson's pupillary syndrome refers to small, light-refractive pupils with maintained convergence response to near vision. The syndrome is seen in neurosyphilis, when the pupillary light reflex is spoiled by interruption of the fibres from brachium to the olivary pretectal nucleus neurons.

Sympathetic preganglionic neurons in the intermediolateral cell column of segment T1-T2, send ascending axons to the superior cervical ganglion. Postganglionic axons follow the ciliary nerve into the eye. The nerve terminals end on a 1-receptors on the dilatator pupillae muscle, and noradrenaline is neurotransmitter. The sympathetic fibres also contain vasoconstrictors to the facial skin and stimulate facial sweat glands (see Horners syndrome).

Catecholamines are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The synthesis is described in Chapter 29.

Catecholamines increase heart rate and cardiac output by stimulation of the adrenergic b_1-receptors in the myocardium.

Catecholamines, released by the adrenal medulla, support the sympathetic system by modifying the circulation during exercise. During exercise the blood is directed to the working muscles from other parts. Noradrenergic nerve fibres innervate blood vessels all over the body. Sympathetic innervation accounts for vascular tone and vasoconstriction.

The most important exercise response in humans is a tremendous vasodilatation in the vascular bed of muscles. The vasodilatation is probably due to a decrease in the a-adrenergic tone of muscular arterioles, and to the action of adrenaline on b_2-receptors.

Catecholamines dilatate the bronchial airways by stimulating adrenergic b_2-receptors. They increase both tidal volume and respiratory frequency. The result is increased ventilation. Catecholamines acting on b_1-receptors cause increased cardiac output. Catecholamines relax the smooth muscles of the digestive tract (b_2-receptors), but contract the sphincters. Catecholamines stimulate metabolism (by activation of the thyroid hormone, T_3) and lipolysis. Adrenaline stimulates hepatic glycogenolysis via b_2-receptors.

Finally, adrenaline stimulates the ascending reticular system (ie, the reticular activating system or RAS) in the brain stem, thus keeping us alert and causing arousal reactions with desynchronisation of the EEG (Chapter 10).

The resistance vessels of the striated muscles in hunting predators (and perhaps in humans) are also innervated by another system. This is the cholinergic, sympathetic vasodilator system. It is capable of a rapid and appropriate bloodflow response during hunting.

Acute stress activates the splancnic nerves and liberates large amounts of adrenaline from the medulla. Diabetics who are developing acute hypoglycaemia, secrete large amounts of catecholamines. Acute muscular activity starts a large catecholamine secretion in exercising persons.

Besides catecholamines, ACTH is also released during stress by increasing hypothalamic signals. ACTH stimulates the glucocorticoid and to some extent the mineralocorticoid secretion through cAMP. Small amounts of glucocorticoids

are permissive for the actions of catecholamines.

Plasma catecholamines are rapidly removed from the blood and have a half-life in plasma of less than 20 s. This is the combined result of rapid uptake by tissues and inactivation in the liver and vascular endothelia (see Chapter 29).

The Autonomic Control Of The Cardiovascular System

The brainstem is the primary site for the autonomic cardiovascular control.

High-pressure baroreceptors are distension-activated stretch receptors located in the walls of the carotid sinus and the aortic arch. Increased arterial blood pressure increases the signal frequency in the sensory baroreceptor neurons that project into the medullary cardiovascular centre (ie, the solitary tract nucleus, nucleus ambiguus and the dorsal motor nucleus of the vagus, DMNV). Impulses generated in the baroreceptor neurons with increasing blood pressure, activate the vagal efferents to the heart and inhibit the sympathetic tone towards the heart. As a consequence the heart rates and force of contraction decreases. Impulses generated in cardiac baroreceptors by cardiac filling, also activate the force of contraction (Fig. 6-4).

The postganglionic fibres from the 3 upper cervical ganglia of the sympathetic trunk pass to the heart as the *cardiac nerves* to the *cardiac plexus*.

Fig. 6-4: The autonomic control of blood pressure and heart rate.

The sympathetic effect is dominant. Increased sympathetic activity constricts the veins, which increases cardiac output by augmenting cardiac filling. Arteriolar constriction reduces cardiac output by increasing the arterial blood pressure (ie, afterload). Other sensory inputs from skeletal muscles, lungs, gastrointestinal viscera, hypothalamus and forebrain help to co-ordinate the autonomic cardiovascular responses related to exercise, respiration, and feeding and temperature control. Hormones, such as angiotensin II, can also modulate the autonomic responses through neurons in the circumventricular organs of the brain, These organs (such as the area postrema) lack the blood-brain barrier.

The sympathetic system innervates the sinus node, the coronary vessels and the myocardial syncytium. Each fibre ends in many terminals, and from the terminals the transmitter noradrenaline is released to the b_2 -receptors of the smooth muscle cells of the coronary vessels and of the myocardium (Fig. 6-4). As a result of increased sympathetic tone, the contractility of the myocardium is increased. Thus, the *end systolic volume* falls from its usual volume- as an example from 70 ml to 40 ml, and the *end diastolic volume* increases due to increased venous return of blood from 140 to 180 ml. Hereby, the *stroke volume* is increased from *70* to *140 ml* of blood in the example. A combination of the doubling of stroke volume with a threefold increase in heart rate, results in a *6-fold* rise in cardiac output.

Sympathetic stimulation depolarizes the sinus node, so that the threshold potential is reached faster than normal. Hereby, the heart rate is increased, and may reach 220 beats/min in young persons. Such a high frequency is due to a *maximal sympathetic activation* of the heart combined with a *reduction* of the vagal tone.

Sympathetic Activation

Activation of noradrenergic fibres leads to *peripheral sympathetic vasoconstriction*, so that blood is shunted to central areas. The heart is stimulated through b_1-receptors so that its frequency and contractility is increased. Other organs are also stimulated to make the person fit for *fight or flight* in any stressful situation.

The postganglionic sympathetic fibres have noradrenaline and ATP containing vacuoles in their nerve terminals. Hence, they release noradrenaline and ATP. The noradrenaline is produced in the chromaffine granules of the neuron.

Fig. 6-5: A sympathetic ganglionic synapse with a small intensity fluorescent cell (SIF cell).

Acetylcholine is released from the preganglionic cell and binds to nicotinic receptors on the postganglionic cell. Acetylcholine also binds to small SIF cells (Fig. 6-5) with muscarinic receptors and vesicles that contain dopamine. Dopamine interacts with dopamine receptors (D2 and D4) on the postganglionic cell and modulates ganglionic transmission by increased permeability to small ions and *hyperpolarisation*.

Liberation of noradrenaline and ATP to the blood does not only lead to constriction of arterioles and arterial vessels, but also constriction of veins and venules. Without venous constriction, the large venous compliance would cause an inordinate amount of blood to be stored in the veins upon sympathetic arteriolar constriction. The consequence would be decreased venous return, which decreases cardiac output and perfusion of vital organs.

Activation of presynaptic purine receptors by adenosine inhibits adrenaline release from the postganglionic terminals innervating the blood vessels. This results in massive vasodilatation.

Exercise and stress demand mobilisation of energy to muscles and heart. Activation of b_2-receptors in the arteriolar wall by circulating catecholamines from the medulla also contributes to vasodilatation in the striated muscles. The total peripheral vascular resistance is reduced during exercise to 20-30% of resting values.

During stress the cutaneous circulation is reduced at first, but then the cutaneous bloodflow rises due to the increased heat production. The brain vessels are only modestly constricted by sympathetic stimulation.

3. The Parasympathetic System

The parasympathetic system has two subdivisions. The cranial division in the brainstem innervates the blood vessels of the head and neck and of many Thoraco-abdominal viscera. The sacral division in the sacral cord innervates the smooth muscles of the walls of the viscera and their glands (the large intestine, liver, kidney, spleen, the bladder and the genitals).

The parasympathetic system only innervates a small percentage of the resistance vessels. Only arteries in the brain and of the penis, the clitoris, and the labia minora receive parasympathetic innervation. Hence, the parasympathetic system has a minimal effect on the arterial blood pressure.

Parasympathetic fibres travelling in the vagus nerve are of utmost importance in affecting the cardiac rate. Vagal fibres innervate the sino-atrial- and the atrio-ventricular-nodes as well as the atrial muscle walls.

The parasympathetic system also innervates the tear and the salivary glands, and the muscles within the eye.

Excitation of the vagus decreases heart rate and atrial contractile force, increases intestinal motility, contracts the gall bladder and bronchi, and relaxes the sphincters of the gastrointestinal tract. The vagal decrease in heart rate is due to the rhythm shift to special P cells, which have a slow rate of depolarisation. Acetylcholine (ACh) is liberated on the cardiac cell membranes, ACh-activated K^+ - channels are opened (via cholinergic receptors and G-regulatory proteins), and K^+ leaks out of the cells, thus opposing the pacemaker current. Vagal stimulation slows down the AV-conduction, causing the co-ordination of atrial and ventricular rhythm to be disrupted. Vagal stimulation can lead to death. Thus external massage of the carotid sinus can cause collar death by greatly increasing vagal stimulation.

The effect of acetylcholine released in the autonomic ganglia can be simulated by nicotine. Conversely, the effect of acetylcholine released by parasympathetic nerve terminals at the target organs can be simulated by muscarine. These observations suggest the presence of two different types of cholinergic receptors. Cholinergic receptors are activated by ACh and by metacholine (MeCH).

Box 6-3: Cholinergic receptor subtypes. ACh stands for acetylcholine, CCh for carbacholine, MeCH for metacholine, DMPP for dimethylphenylpiperazine, and HHSD for hexahydrosiladifenol

Cholinergic receptors				
	Nicotinic		Muscarinic	
Stimulated by:	Nicotine, ACh, MeCH, **DMPP**		Muscarine, ACh, **CCh** MeCH	
Blocked by:	**Hexa**- and **deca**methonium d-tubocurarine		**Atropine**, scopolamine	
Two types of nicotinic receptors			Three muscarinic subtypes	
	Ganglionic	Neuromuscular		$M_1, M_2 M_3$
Stimulated by	Nicotine, ACh **DMPP**	Nicotine, ACh		
Blocked by:	**Hexa**methonium d-tubocurarine	**Deca**methonium Atropine	Pirenzepine Gallamine HHSD	Atropine Dicyclomine

The most important ganglionic blocking drug for blockade of both sympathetic and parasympathetic transmission is *hexamethonium* (Box 6-3).

Cholinergic receptors are located in all autonomic ganglia (nicotinic type), in postganglionic terminals at target organs with parasympathetic innervation (muscarinic type), and in the motor endplate (nicotinic type).

Nicotinic receptors are those activated by acetylcholine, nicotine and nicotinic agonists (ex. dimethylphenylpiperazine, DMPP). Nicotine stimulates all autonomic ganglia simultaneously. Hence, sympathetic vasoconstriction in the limbs and viscera is accompanied by increased gastrointestinal activity and slowing of the heart via the vagus. Nicotinic receptors are blocked completely by d-tubocurarine, and hexa- or decamethonium (Box 6-3). The motor endplate has a different type of nicotinic receptor than the ganglions, since its receptors are not blocked by hexamethonium, but are blocked by d-tubocurarine and decamethonium (Box 6-3).

Acetylcholine, muscarine and muscarinic agonists (pilocarpine and carbacholine, CCh), activate muscarinic receptors. At least 5 different muscarinic receptor molecules have been identified (M_1, M_2, M_3 ..). Activation of the M_1 type is illustrated in Fig. 6-3. Activation of the M_2 type activates an inhibitory G-protein, which inhibits adenylcyclase.

Muscarinic receptor activation is linked to G-protein activation and second-messenger systems.

Muscarinic receptors are blocked completely by atropine, and by antimuscarinic drugs such as homatropine and scopolamine (Table 6-3). These drugs do not block the nicotinic effect of ACh on the postganglionic neurons or on the motor endplate.

1. The sympathetic system consists of short preganglionic and long postganglionic nerve fibres. The parasympathetic system contains long preganglionic and short postganglionic fibres.
2. The chemical transmitter at the target organ is noradrenaline in the sympathetic and acetylcholine in the parasympathetic system.
3. The sympathetic system contains adrenergic receptors (a and b), whereas the parasympathetic system has cholinergic receptors (muscarinic or muscarinergic and nicotinic or nicotinergic).
4. Activation of the cholinergic system serves anabolic functions (ie, stay and play), whereas activation of the noradrenergic system serves catabolic functions (ie, fight, fright or flight).
5. Activation of a_1-receptors increases intracellular [Ca^{2+}], which leads to phosphorylation of protein kinases and thus to a response. Activation of a_2-receptors triggers an inhibition of the membrane adenylcyclase, reducing [cAMP] in the cells. b_1- and b_2-receptors activate adenylcyclase, which increases cAMP production in the cell. Muscarinic receptors are completely blocked by atropine. Activation of M_1-receptors increases intracellular [Ca^{2+}]. Activation of M_2 inhibits adenylcyclase, and through an inhibitory G-protein reduces the formation of cAMP.

Pathophysiology

This paragraph deals with 1. Mushroom poisoning, 2. Carbamate and organo-phosphate poisoning, 3. Xerophtalmia and xerostomia, 4. Tachycardia and bradycardia, 5. Smoking, 6. Phaeochromocytoma, 7. Primary or essential hypertension, 8. Horners syndrome, 9. The Kluver- Bucy-syndrome and emotional disorders. - Altzheimers disease and Parkinson's disease are also related to the autonomic nervous system and described in Chapter 7.

1. Mushroom Poisoning

Among the poisonous mushrooms at least two are related to the autonomic nervous system:

Amanita palterina (false blusher) contains a substance with atropine-effect, which completely blocks the muscarinic cholinergic receptors. Atropine-effects are described below.

Amanita muscaria (red fly agaric) contains atropine-like substances, muscarine and other hallucinogens. Atropine-effects are prominent: Motor unrest, delirium (red fly agaric was used by the Vikings to run berserk), mouth dryness, pupillary dilatation, and tachycardia. Some cases are dominated by muscarine-effects: Glandular secretion (sweat, saliva, tear-flow), miosis, eye pains, bradycardia, cramps, respiratory failure, lung oedema, and coma. Muscarinic symptoms are treated with slow intravenous injection of atropine (1 mg in 1 ml). - Ventilation with the mouth-to mouth-method may become fatal for the rescuer.

2. Carbamate And Organo-Phosphate Poisoning

These substances (carbaryle, dimethoate, melathion, parathion etc) are used as insecticides in agriculture. They are anti-cholinesterases so they accumulate acetylcholine in the tissues. The clinical picture is due to the muscarinic and nicotinic effects of acetylcholine. The muscarine-effects are described above. Other symptoms are nausea, vomiting, muscle weakness, paresthesia, bronchospasm, shock and respiratory arrest. Atropine (1 mg iv.) is given repeatedly to obtain complete atropine blockade beginning with pupillary dilatation. Artificial ventilation is often imperative. Cholinesterase- reactivators are tried in desperate situations.

3. Xerophtalmia and xerostomia

Xerophtalmia and xerostomia is dryness of the conjunctiva and the cornea, and dryness of the mouth, respectively. Most of these disorders are of unknown origin (ie essential or primary), although an autonomic cause may be suspected. Xerophtalmia and xerostomia occurs in connection with anaesthesia due to the use of atropine-like substances in order to reduce secretion.

4. Tachycardia and bradycardia.

A balance between the parasympathetic and the sympathetic system normally determines cardiac rhythm. The parasympathetic system predominates. A relative dominance of the sympathetic tone towards the heart leads to tachycardia (ie, a heart rate above 80 beats per min), and a further relative dominance of the parasympathetic system leads to bradycardia (ie, a heart rate below 50 beats per min). Fluctuations of the autonomic tone leads to phasic changes of the sinus node activity. During inspiration, the parasympathetic dominance falls and the heart rate becomes rapid, whereas during expiration parasympathetic dominance increases and the heart slows down. This phenomenon is found in children, and it can even be found in healthy subjects of high age. The phenomenon is called sinus arrhythmia.

5. Smoking

Cigarette smoking is common among teen-agers, and more girls than boys of 15 years smoke cigarettes. Activation of *nicotinic cholinergic receptors* is fast and does not require G proteins. Nicotinic receptors stimulated by acetylcholine open a Na^+-channel and depolarise the cell membrane. Nicotine stimulates nicotinic receptors on the postganglionic neuron of all autonomic ganglia.

The number of cigarettes smoked and the number of years smoked seem to increase the number of nicotinic cholinergic receptors in brain tissue. Smoker's dependency may depend upon the number of nicotinic receptors. Smokers, deprived of the usual daily dose of nicotine in cigarettes, become depressed in mood, they feel stress and they are un-concentrated. This is called abstinence and the cause is lack of nicotine. The mood of the smoker is immediately improved by smoking. Although smoking is nicotine addiction, nicotine seems to be one of the less dangerous molecules in smoke. Polycyclic aromatic hydrocarbons and nitrosamines are potent carcinogens and mutagens. Such substances release proteinases from granulocytes and macrophages, whereby elastin is destroyed resulting in multi-site lung degeneration or emphysema. The life-threatening dangers are due to cancer and atherosclerosis: Lung cancer, chronic bronchitis and emphysema, cerebral stroke, ischaemic heart disease, peripheral vascular disease, bladder cancer, and memory problems. Male smokers have an increase in the number of abnormal spermatozoa, and pregnant female smokers have an increase in neonatal mortality. Female smokers above 30 years using anti-pregnancy pills have a 40-fold higher risk of cerebral stroke than their non-smoking control group.

6. Phaeochromocytoma

Some patients suffer from attacks of severe hypertension due to adrenaline hypersecretion. The hypertension is caused by release of large amounts of adrenaline from a *medullary tumour* of chromaffine cells. The attacks are sometimes fatal. The diagnosis is important, because the patient can be cured by surgical abolition of the tumour.

7. Primary or essential hypertension

in general begins as a condition with sympathetic overactivity (see Chapter 9).

8. Horner's syndrome

refers to a condition with miosis, facial vasodilatation and loss of facial sweating and enophtalmus due to damage of the sympathetic nerve supply from the T1-T2-segments to the eye and facial skin.

9. The Kluver- Bucy-syndrome

refers to an *emotional disorder* with bilateral temporal lobe lesions. The temporal cortex, hippocampus and the amygdaloid body are damaged. Mental blindness (visual agnosia) is the inability to recognise objects seen. Besides mental blindness, the syndrome consists of loss of short-term memory, and hypersexual behaviour incompatible with normal social adaptation. The hypersexual behaviour is related to the visual agnosia.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Altzheimers disease is a primary cortical brain atrophy with premature ageing of the brain. Lack of the acetylcholine-producing enzyme choline acetyltransferase and of acetylcholine is characteristic.

B. Mushroom poisoning can be dominated by muscarine-effects: Glandular secretion (sweat, saliva, tear-flow), miosis, eye pains, bradycardia, cramps, respiratory failure, lung oedema, and coma.

C. Cannons law: The peristalsis in the small intestine is polarised, so it always proceeds in the aboral-oral direction.

D. Cannons emergency reaction is an immediate sympatho-adrenergic response to life dangerous situations, with

both sympatho-adrenergic and parasympathetic overactivity. The last phenomenon includes vagal cardiac arrest with involuntary defecation and urination.

E. The sympathetic system contains adrenergic receptors (a and b), whereas the parasympathetic system has cholinergic receptors (muscarinic or muscarinergic and nicotinic or nicotinergic).

II. Each of the following five statements have True/False options:

- **A**. Mydriasis refers to contraction of the pupil by sympathetic stimulation.
- **B**. SIF cells are small intensity fluorescent cells, which possess muscarinic receptors and contain vesicles filled with dopamine. Adequate stimulation releases dopamine, which interacts with dopamine receptor (D_2) on the postsynaptic cell body and modulates the effect of acetylcholine.
- **C**. Apraxia refers to a condition with lack of ability to recognise and interpret a sensory stimulus.
- **D**. Phenoxybenzamine and phentolamine block the a-receptor.
- **E**. Adrenaline stimulates the ascending reticular system in the brain stem, thus keeping us alert.

III. Each of the following five statements have True/False options:

A. Argyll-Robertson's pupillary syndrome with small, light-refractive pupils and maintained convergence response to near vision is a typical sign in acute syphilis.

B. Besides catecholamines, ACTH is also released during stress by increasing hypothalamic signals. ACTH stimulates the glucocorticoid and to some extent the mineralocorticoid secretion through cAMP.

C. All autonomic nerves have ganglia outside the CNS in contrast to the somatic nervous system.

D. When stimulated nicotinic receptors work through a slow cascade of events.

E. Catecholamines dilatate the bronchial airways.

6. Case History

A 19-year-old female is in hospital with a cranial lesion caused by a fall from her horse. The following clinical signs are found: 1) speech troubles and hoarseness, 2) swallowing problems and paresis of the soft palate, 3) rapid heart rate, and 4) dilatation of the stomach with vomiting.

Lesion of a certain cranial nerve can explain all symptoms and signs.

- *1. What is the name of the nerve?*
- *2. What is special about this particular lesion?*

Try to solve the problems before looking up the answers .

Highlights

- *The autonomic nervous system mediates neural control of the internal milieu despite substantial environmental changes. The autonomic system directly influences smooth muscles, glands and the heart through its two subdivisions, the sympathetic and the parasympathetic system. The two subdivisions function in a dynamic balance aiming at homeostasis.*
- *The sympathetic system consists of short preganglionic and long postganglionic nerve fibres. The parasympathetic system contains long preganglionic and short postganglionic fibres.*
- *The chemical transmitter at the target organ is noradrenaline in the sympathetic and acetylcholine in the parasympathetic system.*
- *Carbamate and organo-phosphate poisoning. These substances (carbaryle, dimethoate, melathion, parathion etc) are used as insecticides. They are anti-cholinesterases, so they accumulate acetylcholine. The clinical picture of poisoning is due to the muscarinic and nicotinic effects of acetylcholine. Atropine (1 mg iv.) is given repeatedly to obtain complete atropine blockade beginning with pupillary dilatation.*
- *The sympathetic system contains adrenergic receptors (α and β), whereas the parasympathetic system has cholinergic receptors (muscarinic or muscarinergic and nicotinic or nicotinergic).*
- *Activation of the cholinergic system serves anabolic functions (i.e., stay and play), whereas activation of the*

noradrenergic system serves catabolic functions (i.e., fight, fright or flight).

- *Activation of α_1-receptors increases intracellular [Ca^{2+}], which leads to phosphorylation of protein kinases and thus to a response.*
- *Activation of α_2-receptors triggers an inhibition of the membrane adenylcyclase, reducing [cAMP] in the cells.*
- *β_1- and β_2-receptors activate adenylcyclase, which increases cAMP production in the cell.*
- *Muscarinic receptors are completely blocked by atropine. Activation of M_1-receptors increases intracellular [Ca^{2+}]. Activation of M_2 inhibits adenylcyclase, and through an inhibitory G-protein reduces the formation of cAMP.*
- *The intrinsic enteric nervous system consists of two sets of nerve plexi. The submucosal (Meissner) plexus mainly regulates digestive glands, whereas the myenteric (Auerbach) plexus, located between the longitudinal and the circular muscle layers, is primarily connected with gut motility.*
- *Sympathetic activity (excitement or pain) causes a large pupil, and parasympathetic activity (light or near-sight) causes a small pupil.*
- *Parasympathetic activity controls salivation, gastrointestinal functions (with the enteric nervous system), emptying of the bladder and defaecation.*
- *The limbic system including the hypothalamus controls vital autonomic functions and emotional behaviour.*

Further Reading

- Brodal, A (1981) *Neurological anatomy.* Oxford University Press, New York.
- Loewy, A.D. and K.M. Spyer (editors). "Central Regulation of Autonomic Functions." *Oxford University Press*, N-Y., 1990.

Return to content

Return to chapter 6

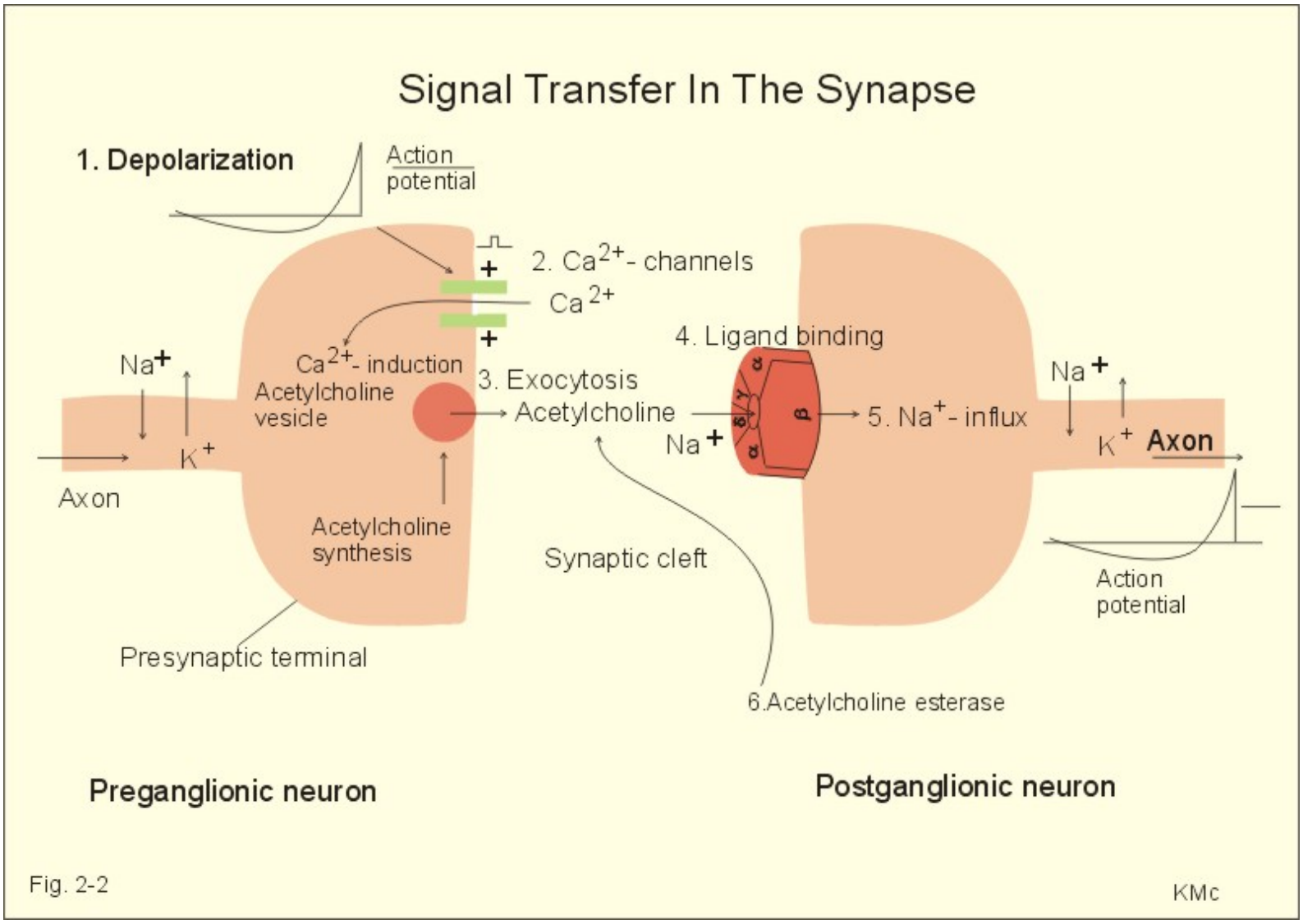

Signal Transfer In The Synapse
1. Depolarization
Action potential
2. Ca^{2+}- channels
Ca^{2+}
4. Ligand binding
Ca^{2+}- induction
Acetylcholine vesicle
3. Exocytosis
Acetylcholine
Na^+
K^+
Axon
5. Na^+- influx
Na^+
K^+
Axon
Acetylcholine synthesis
Synaptic cleft
Action potential
Presynaptic terminal
6.Acetylcholine esterase
Preganglionic neuron
Postganglionic neuron

Fig. 2-2

KMc

Chapter 7.

Neurological And Psychiatric Diseases

Study Objectives

- To *define* anorexia and bulimina nervosa, astrocytes, bradykinesia, delirium, delusions, dementia, depression, dystonia, Huntington‘s disease, mania, manic-depressive psychosis, meningitis, microglia, multiple sclerosis, myoclonus, oligodendroglia, paranoid delusions, Parkinson‘s disease, schizophrenia, sleep apnoea, sleeplessness, and status epilepticus.
- To *describe* or *explain* the pathophysiology of brain injuries, brain inflammations, brain tumours, Parkinson‘s disease, manic-depressive psychosis, schizophrenia, epilepsy, status epilepticus, and common nervous and stress-related psychiatric disorders.

Principles

- *Glutamate is a major neurotransmitter, which act on several classes of excitatory amino acid receptors.*

- *Many neurological disorders are mediated by neuronal injury involving excessive stimulation of glutamate receptors. Glutamate antagonists are used in clinical trials.*

Definitions

- **Akathisia** is an extrapyramidal defect with swaying and twisting body dyskinesia.

- **Anorexia nervosa** is an eating disorder in adolescent females and males resulting in severe malnutrition. The patient has an intense wish to be thin. Biological and psychological factors are involved. In some cases there is a regression into childhood, where the girl tries to escape from the problems of puberty and adolescence.
- **Astrocytes** are specialised neuroglia, which separate nerve pathways, buffer extracellular potassium and repair nerve injuries.
- **Bradykinesia** is a term for slowing of voluntary movements found in Parkinsons disease.
- **Bulimina nervosa** refers to a condition, where the patient is preoccupied with food and periodically eats excessively. The patient sometimes avoids overweight by self-induced vomiting just after binge eating.
- **Chorea** refers to rapid involitional hyperkinesia with jerky movements of the limbs.
- **Cognitive brain functions** are intellectual processes such as calculation, judgement, language, learning ability, memory, orientation and thinking.
- **Computed tomography** (CT) is a technique using X-rays moving across a slice of brain (tomia, to cut). Normal brain tissue absorbs X-rays differently from tumour tissue, infarcted tissue, and coagulated blood and brain oedema. The radiation passing in a certain direction is recorded with scintillation detectors, and a computer processes the signals.
- **Conversion** refers to Freud’s hypothesis that mental energy can be converted into physical symptoms and signs (abdominal pain, blindness, double vision, deafness, muteness, fits with dramatic movements, artistic gait disturbances, hysterical paresis with normal muscle tone and deep reflexes, crude tremor, sensory loss, stigmatisation).
- **Delirium** is an acute impairment of consciousness also called *toxic confusion*.
- **Delusion** is an abnormal belief arising from distorted judgement.
- **Dementia** (*senility* or ageing of the brain) disturbs almost all cognitive brain functions, whereby the personality of the patient is completely changed.
- **Depression** is characterised by early morning waking with unresponsive sadness, guilt feeling, suicidal feelings, and lack of a precipitating factor.
- **Dissociation** refers to an apparent dissociation between different mental activities. An example is a mentally protective cover of enjoyment (euphoria) in terminal cases of painful cancer or AIDS (French: belle indifference).
- **Dystonia** are abnormal involitional muscle contractions that produce abnormal movement patterns and postures.

- **Hallucinations** are sense impressions experienced in the absence of external sense stimuli.
- **Hemiparesis** means weakness of the limbs of one side – frequently occurring in upper motor neuron lesions.
- **Hemiplegia** means total paralysis of the limbs of one side of the body.
- **Huntington's disease** is a chorea-condition with hypotonia, dementia and involuntary movements. This is an autosomal genetic defect on chromosome 4.
- **Magnetic resonance imaging** (MRI) is a scanning technique, where protons in a strong magnetic field are bombarded with radiofrequency waves in order to produce images. MRI scanning can picture brain tumours, multiple sclerosis lesions, and syringomylia among others. MRI scanning can even separate white from grey matter. MRI scanning is replacing myelography, because it can visualise spinal cord compression, spinal cord tumours and malformations.
- **Mania** refers to a psychiatric disorder with periodic elevations of mood with overactivity, restlessness, fast talk, excessive energy, increased sexuality, overwhelming self-confidence, and insomnia.
- **Manic-depressive psychosis** covers severe abnormalities of mood. Mood ranges from severe depressive psychosis over moderate and minor depression, sadness, normal mood, happiness, euphoria, hypomania and severe mania.
- **Meningitis** refers to inflammation of the meninges.
- **Microglia cells** proliferate and move to the site of nerve injury, where they transform to large phagocytes, which remove debris.
- **Multiple or disseminated sclerosis** refers to a common neurological disease caused by inefficient myelin production in the oligodendroglia.
- **Myoclonus** often occurs at night and refers to brief contractions or jerks of one or more muscles. Myoclonus is often related to metabolic or drugs toxicity.
- **Neurosis** refers to a psychiatric disorder in which the personality as a whole is unimpaired and without psychotic symptoms. Neurosis is an amplified, more than normal reaction to mental stress such as anxiety, depression and irritability.
- **Oligodendrocytes** produce myelin sheaths around axons in the CNS just like Schwann cells do in the peripheral nervous system.
- **Paranoid delusions** (*paranoia*) are abnormal beliefs dominated by fear of persecution.
- **Parkinsonism** is a dopamine-deficiency state of the forebrain with bradykinesia, tremor, and rigidity.
- **Psychosis** refers to a psychiatric disorder impairing the whole personality and functioning of the individual (insight, sense of reality, delusions, and hallucinations).
- **Repression** means exclusion of memories, impulses, and emotions from consciousness, because these elements would cause anxiety and stress.
- **Sleep apnoea** is periodic breath holding during sleep. Sleep apnoea often occurs with snoring and airway obstruction in obese patients or in patients with chronic obstructive lung disease.
- **Schizophrenia** means splitting of the mind or disconnection of cognitive and emotional psychic functions. Schizophrenia is a psychosis with hallucinations, dissociation of ideas, intense fear, and paranoid delusions.
- **Status epilepticus** is an emergency condition, where consciousness is not regained between grand mal seizures lasting more than half-an-hour.
- **Tics** refer to focal myoclonus with repeated twitching of facial or neck muscles. Tics may even begin in childhood for unknown reasons and they are extremely resistant to therapy.
- **Tremor** or *shaking* can be caused by hyperthyroidism and by Parkinsonism, but it is also a typical side effect of alcohol, narcotics and drug abuse.

Essentials

This paragraph deals with *1. Nerve cells, 2. Ion channels, 3. Neurotransmitters, and 4. Signal transduction.*

1. Nerve cells

The cells of the Central Nervous System (CNS) consist of neuroglia and of neurons.

Neuroglial cells outnumber all the neurons in the CNS and they constitute half of the brain volume. Glial cells are known to sheath and protect neurons. Glial cell membranes contain receptors and ion channels. They help control the environment of neurons and thus contribute to the function of neurons. We have three types of neuroglial cells. *Microglia cells* are small cells scattered throughout the nervous system. Microglia proliferate after injury and move to

the site of injury. Here they transform to large phagocytes, which remove the debris. *Oligodendrocytes* produce myelin sheaths around axons in the CNS just like Schwann cells do in the peripheral nervous system. *Astrocytes* separate nerve pathways, buffer extracellular [K^{+}] , and repair nerve injuries.

2. Ion-channels

Two classes of proteins span the cell membrane and control ion transfer. The first class is Na^{+}-K^{+}-pumps and other ATP-demanding pumps that actively move ions across the membrane against their electrochemical gradient (Fig. 7-1). The second class is *channels* or pores through which specific ions can pass. Ions traverse such an open channel along the electrochemical gradient. The small ion permeation through the cell membrane at rest is referred to as *leak current* (Fig. 7-1). The typical Na^{+}-channel *opens promptly* in response to *depolarisation* (voltage-gated opening) and also closes rapidly, although the cell is still depolarised. The channels then remain inactivated for a short period. Opening of Na^{+}-channels increases the flux of Na^{+} into the neuron, and depolarizes the membrane, so the effect is excitatory.

Fig. 7-1: Ion channels in a neuronal membrane, where the Na^{+}-channel is closed and the K^{+} -channel is open at rest.

Closure of K^{+}- or Cl^{-} -channels decreases the flux of K^{+} out of the neuron or decreases the flux of Cl^{-} into the cell. These events also depolarise the membrane, and again the effect is excitatory.

Obviously, closure of Na^{+}-channels or opening of K^{+}- or $Cl^{—}$channels have an inhibitory effect by hyperpolarisation.

Voltage-gated Na^{+}-, K^{+}-, and Ca^{2+}-channels comprise subunits with membrane spanning domains (Fig. 7-1). There is amino acid sequence homology in the transmembrane helices of these channels. The *channel protein* includes a charged group, which is sensitive to the electric field across the membrane. During depolarisation the gate opens, which changes the whole channel, rendering it much more conductive to specific ions. Each channel continues to open, close and reopen several times during depolarisation. The *fast* Na^{+}-channels close rapidly and are inactivated during depolarisation due to a *channel polypeptide* located on the cytosolic side (Fig. 7-1).

Opening of Na^{+}-channels requires or results in a rapid change of potential. Partial and slow depolarisation, inactivate a critical fraction of the Na^{+} -channels. This is called *voltage- inactivation*. Voltage inactivation of Na^{+}-channels is involved in the accommodation and in the refractory periods.

3. Neurotransmitters

Neurotransmitters are signal molecules used by neurons to communicate with each other and with target cells. Chemical synapses are specialised. The presynaptic terminal contains mechanisms for production and storing of neurotransmitters that are released in response to depolarisation.

The postsynaptic membrane carries protein receptors that can detect and identify different neurotransmitters and initiate appropriate responses to stimulation. Finally, there are adequate mechanisms for degradation and reuse of transmitters to ensure rapid onset and offset of arriving signals. Chemical synapses are the sites of action for many drugs.

Neurotransmitters can be divided into two groups: Classical rapid acting non-peptide neurotransmitters (Box 7-1) and putative, slowly acting peptides (Box 7-2).

During development some process of differentiation determines the type of neurotransmitter that a given neuron will synthesise, store, and release. Thus, a single neuron releases the same neurotransmitter from all its synapses - an assumption, which has been generally accepted for years as *Dale's law*.

Recent advances indicate that some neurons can release more than one neurotransmitter. Up to 4 neuropeptides have been localised to a single neuron.

Two or more transmitters released together are called *co-transmitters*. One member of each pair of transmitters appears to be a peptide. Perhaps these peptides act by enhancing the message transferred with the rapid neurotransmitters.

Classical neurotransmitters are substances such as acetylcholine, noradrenaline, dopamine, GABA (gamma-aminobutyric acid), glycine etc (Box 7-1). Their diffusion pathway is short, and they have no other function than neurotransmission.

Catecholamines (dopamine, noradrenaline, and adrenaline) are neurotransmitters both in the sympathetic system and in the CNS. Noradrenaline is the transmitter for most postganglionic sympathetic fibres (some of these fibres use acetylcholine).

In the CNS catecholamines are found in several brain nuclei: *Dopaminergic neurons* are found in the substantia nigra, *noradrenergic neurons* in locus coeruleus, and *serotonergic neurons* in the raphe nuclei and in many midbrain structures.

Serotonin (5-hydroxytryptamine) is a transmitter in brainstem nuclei (in particular the Median raphe) concerned with wakefulness and behaviour. Adrenaline, noradrenaline, dopamine, and serotonin serve as fast neurotransmitters in the CNS in the same way as the Enzyme- inactivated acetylcholine. The most important excitatory amino acid (EAA)-receptors are the glutamate receptors, the N-methyl-D-aspartate (NMDA)-receptors, and the alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-receptors. The glutamate receptor is a typical non-NMDA-receptor. NMDA-activated ion channels are only active, when the membrane is depolarised, and they are specific for Ca^{2+}- Na^{+}- and K^{+}-penetration. AMPA-activated ion channels are specific for Na^{+}-and K^{+}-permeability and depolarise the cell membrane.

Box 7-1: Classical, rapidly acting non-peptide transmitters with their cell-surface receptor type and action/location

Substance	Receptor type	Action/ location
Acetylcholine	**Cholinergic**	**Excitatory/Autonomic ganglia etc**
Adrenaline	**Adrenergic**	**Excitatory/Locus coeruleus etc**
Noradrenaline	**Adrenergic**	**Excitatory mainly (awake, mood)**
Dopamine	**Dopaminergic**	**Inhibitory/Substantia nigra**
Histamine	**Histaminergic**	**Excitatory/Hypothalamus**
Serotonin	**Serotonergic**	**Inhibitory/Median raphe of brain stem**
GABA	**GABA-receptor**	**Inhibitory/CNS**
Glycine	**GABA-receptor**	**Inhibitory/Spinal cord**
Aspartate	**NMDA-receptor**	**Excitatory/CNS**
Glutamate	**non-NMDA-receptor**	**Excitatory/Cortex**
Nitric oxide (NO)	**NO diffuses into cells**	**Behaviour and memory/CNS**

		Long-term potentiation

Neuropeptides

Several of the neuropeptides are well-known hormones. They are synthesised in the soma of the neurons and reach the axon terminals by fast axonal transport.

The secretion of both gonadotropins (LH and FSH) from the anterior pituitary is controlled by the hypothalamic luteinizing hormone-releasing hormone (LHRH). The thyrotropin-releasing hormone (TRH) and somatostatin from the hypothalamus control the pulsatory secretion of thyrotropin or thyroid stimulating hormone (TSH) from the anterior pituitary. The secretion of growth hormone (GH) and prolactin from the anterior pituitary is controlled by two hypothalamic hormones: GH-inhibiting hormone (GHIH or somatostatin) and GH-releasing hormone (GHRH).

Many of these peptides are cut off from a big mother molecule: *pro-opio-melanocortin* (POMC). Cleavage of POMC in the anterior pituitary lobe releases adrenocorticotropic hormone (ACTH) and b -lipotropin. Cleavage of ACTH in the intermediate pituitary lobe releases melanocyte-stimulating hormone (a -MSH) and corticotropin-like intermediate lobe peptide (CLIP). Cleavage of b -lipotropin releases b -MSH and b -endorphin. Endorphins and enkephalins bind to opiate receptors and are called endogenous opiates.

Cleavage of a pre-pro-hormone produced in the hypothalamus releases the two octapeptides, oxytocin and vasopressin together with two neurophysin molecules. The two octapeptides are moved to the neurohypophysis by axoplasmatic transport.

Box 7-2: Putative, slowly acting peptide neurotransmitters are water- soluble and binds to cell-surface receptors. Some neuropeptides have non-peptide co-transmitters (-co).	
Substance/co-transmitter	**Action/location**
LHRH	**Gonadotropin release/anterior pituitary**
TRH	**TSH release/ "**
Somatostatin/noradrenaline (-co)	**GH-inhibition/ "**
ACTH	**Stimulates secretion of adrenal cortex hormones/ "**
β -lipotropin	**Lipolysis/Fat cells**
β -endorphin	**Pain release/opiate receptors**
α -MSH	**Melanocyte stimulation/skin**
β -MSH	**" " / "**
Prolactin	**Development of mammary gland/breasts**
Luteotropin (LH)	**Rupture of follicle/ovaries**
Thyrotropin	**Activates adenylcyclase/thyroid follicles**
Growth hormone (GH)	**Regulates growth/the body as a whole**
Oxytocin	**Stimulates myoepithelial cells/milk ducts and uterus**

Vasopressin (ADH)	**Vasoconstrictor. Stimulates renal water reabsorption**
Enkephalins/adrenaline (-co)	**Pain release/opiate receptors**
Substance P/serotonin (-co)	**Smooth muscle contraction/neurons, endocrine cells**
Gastrin	**Gastric acid secretion/neurons, endocrine cells**
CCK/dopamine (-co)	**Bile and pancreatic enzyme secretion/neurons**
VIP/acetylcholine (-co)	**Smooth muscle relaxation, secretion/neurons**
Insulin	**Reduces blood glucose/pancreatic β -cells**
Glucagon	**Increases blood glucose/ pancreatic α -cells**
Angiotensin II	**Aldosterone secretion. Arteriolar constriction**
Bombesin	**Pancreatic enzyme secretion. Synaptic transfer**
Bradykinin	**Vasodilatator. Synaptic transfer**
Calcitonin	**Bone is remodelling. Synaptic CNS transfer**

CCK stands for cholecystokinin and VIP for vasoactive intestinal polypeptide.

The *gut-brain peptides* are described in Chapter 22, and the *hypothalamo-pituitary peptides* in Chapter 26.

4. Signal transduction

Signal transduction is a cascade of processes from the *receptor-hormone* binding to the final cellular response. Many hormones and neurotransmitters raise the concentration of a second messenger in the target cell via guanyl triphosphate (GTP) and act through it. The receptor-hormone complex activates a GTP-binding protein (so-called G-protein) which controls and amplifies the synthesis of the second messenger. Hereby, each hormone molecule can produce many molecules of second messenger such as cAMP or cGMP. Furthermore, each protein kinase unit can phosphorylate many molecules of its substrate, resulting in a great amplification factor.

G-proteins function as molecular switches, regulating many cellular processes, such as activation of intracellular enzymes (protein kinase, phosphorylase), activation of membrane enzymes and channels, and activation of gene transcription.

G- protein-linked receptors form a family, which has evolved from a common ancestor. Most G-proteins are membrane bound heterotrimers (a, b, g) and exist in an activated state, where it has high affinity for GTP, and an inactive state, where the molecule prefers GDP.

Hydrophilic (lipophobic) hormones, such as acetylcholine and many peptides, bind to membrane receptor proteins, and the hormone-receptor binding activates the enzyme phospholipase C via active G-protein.

Multiple receptor subtypes can co-exist on a single cell. The b-adrenergic receptors are both stimulated by noradrenaline and both activate a stimulatory G-protein (Gs in Fig. 7-2). Gs activates adenylcyclase, which increases the production of the second messenger cAMP.

Fig. 7-2: A single cell with both β_1- and β_2 -adrenergic receptors. Both receptors activate adenylcyclase through stimulatory G-proteins (Gs).

The result is an additive cellular response.

A single cell with b_1-adrenergic receptors activating adenylcyclase through a stimulating G-protein, and a_2 -adrenergic receptors, inhibiting adenylcyclase via an inhibitory G- protein, results in opposite signals when stimulated by noradrenaline (Fig. 7-3).

Fig. 7-3: Antagonistic reactions to noradrenaline in a single cell.

Noradrenergic stimulation of another single cell with b_1-adrenergic receptors activating adenylcyclase through a stimulating G-protein, and a_1 -adrenergic receptors which activate phospholipase C leads to production of two phosphorylated derivatives of phosphatidylinositol (PI): PI-phosphate (PIP) and PI-diphosphate (PIP_2). Phospholipase C cleaves (PIP_2) into inositoltriphosphate (IP_3) and diacylglycerol (DAG) (see Fig. 7-4).

Fig. 7-4: Independent reactions to the stimulation of two subtypes of adrenergic receptors on a single cell.

IP_3 is a second messenger that binds to Ca^{2+}-channels in the endoplasmic reticulum (ER), so that Ca^{2+} is released to the cytosol. DAG and Ca^{2+} are second messengers that activate *protein kinase C*, which is involved in the regulation of cellular metabolism, growth and many other processes. Inactive cytosolic protein kinase C is activated by Ca^{2+}, and binds to the inner surface of the membrane, where DAG activates it. Ca^{2+} and proteinkinase C catalyses the transfer of phosphate from ATP to the effector proteins. Independent reactions are generated by the presence of these two subtypes of adrenergic receptors.

Specific receptor-ligand bindings also activate phospholipase A_2 via a G-protein. Phospholipase A_2 cleaves membrane phospholipids, and releases arachidonic acid (AA) in the cells. AA activates a precursor to platelet activating factor (PAF) termed lyso-PAF. AA is also the precursor for the synthesis of endoperoxides, prostacyclin, thromboxanes (mediates platelet aggregation and vasoconstriction) and leucotrienes.

Insulin and related growth factor peptides bind to membrane receptors that are glycoproteins protruding from the membrane. The insulin receptor is typical for this receptor family. Peptide binding to the outer receptor subunit stimulates a protein tyrosine kinase on the inner receptor subunit. This phosphorylates tyrosine residues, both on the receptor itself and on other proteins. The tyrosine kinase activity is essential for signal transduction.

Examples of growth factors are: EGF (epidermal growth factor), FGF (fibroblast growth factor), IGF-II (insulin-like growth factor-II), NGF (neural growth factor) and PDGF (platelet-derived growth factor).

Protein tyrosine kinase activity is abnormally high in certain types of cancer and cellular modification. This can be caused by growth factors or by a mutation of the tyrosine kinase part of the trans-membraneous receptor. Mutations of one gene localised on chromosome 10 can lead to four different syndromes: Multiple endocrine neoplasia, Hirschprung's disease, medullary thyroid carcinoma, and Phaeochromocytoma (se Chapter 28).

The final step is often phosphorylation or dephosphorylation of a particular key or effector protein. Protein kinases and dephosphorylation accomplish phosphorylation by protein phosphatase. Second messengers (cAMP, cGMP, IP_3, DAG, and Ca^{2+}) control the activities of protein kinases such as cAMP-dependent protein kinase A, cGMP-dependent protein kinase, calmodulin-dependent protein kinase, and protein kinase C. Calmodulin binds 4 Ca^{2+}.

The phosphorylation level of an enzyme or an ion channel determines and triggers the physiological response.

Protein phosphatase reverses the effect of protein phosphorylation. The phosphatase dephosphorylates the key proteins, and thus opposes or stops the physiological response.

The free radical gas nitric oxide (NO) is a neuronal messenger in both the central and the peripheral nervous system. The NO gas is membrane permeant, and can bypass normal signal transduction in synapses.

Two types of NO synthase (NOS) have been identified: constitutive Ca^{2+}- calmodulin dependent enzyme, and inducible Ca^{2+} - independent enzyme. Both enzymes are flavoproteins containing bound flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). Both enzymes require the cofactors NADPH and tetrahydrobiopterin (BH_4).

NOS catalyses the conversion of L-arginine to citrulline and NO in two step when activated by the Ca^{2+} - calmodulin complex, muscarinic agonists, or other activators (Fig. 7-5).

Fig. 7-5: The biosynthesis of NO with cell-cell effects on target cells, such as smooth muscle cells etc. Non-adrenergic non-cholinergic (NANC) relaxation of the gastrointestinal tract (GIT) and the genito-urinary (GU) system is shown.

NO diffuses to the target cell, where it activates *guanylcyclase* resulting in the formation of cGMP (Fig. 7-5). NO is labile. Hence, a carrier for NO has been postulated. The biological effect of NO is mediated by an increase in cGMP levels, and the effects on target cells are shown below (Fig. 7-5). *Nitric oxide effects* are further developed in Chapters 2, 3, 4, 9, 12, 22, 25 and 30.

The *NO biosynthetic pathway* can be interfered with at several points. Nitrovasodilatators, such as nitro-glycerine, have been used for over a century to treat cardiac cramps or angina pectoris (see Chapter 9). Nitrovasodilatators act by releasing NO and thereby causing coronary vasodilatation. Nitric oxide synthase is inhibited by *L-arginine analogues* (Fig.7-5).

One of the effects of NOS-inhibitors is an increase in blood pressure. NOS-inhibitors are effective in treating endotoxic shock. This is a condition caused by increased NO synthesis by inducible NO synthase, where the sympathetic vasoconstrictors are often ineffectual. The cofactors, like BH_4, can also be manipulated, e.g., by *anticancer drugs*.

Pathophysiology

The first part I concerns neurological disorders of cerebrovascular and brain parenchymal origin (eg, brain trauma, epilepsy, movement disorders, multiple sclerosis, inflammations), and end-up with the differential diagnosis between dementia and delirium - a situation of life-threatening consequences.

The second part (II) is confined to psychiatric disorders (eg, neuroses and psychoses).

I. Cerebrovascular Disorders

These disorders include parenchymal brain damage, and the main groupings are *1. Stroke and minor stroke*, *2.Brain lesions*, *3.Epilepsy*, *4.Movement disorders*, *5.Multiple sclerosis*, *6. Inflammations* and *7.Intracranial tumours*. *8. Dementia contra delirium*

1. Stroke and minor stroke

Thrombo-embolism of the middle cerebral artery is a common cause of stroke

(suddenly occurring unconsciousness with hemiplegia). The *middle cerebral artery* is the artery most often occluded by thrombo-embolism or by atherosclerotic material (Fig. 7-6). Risk factors for stroke are related conditions such as inactivity, obesity, hypertension, smoking, hypercholesterolaemia, hypertriglyceridaemia and oral contraception (see also Chapter 10).

Thrombo-embolism causing a stroke usually leads to *cerebral infarction*. Occlusion of the internal carotid artery or the middle cerebral artery causes infarction of the internal capsule with *aphasia* (lesion of the dominant hemisphere), contralateral *hemiplegia*, and *areflexic flaccid limbs*. The lesion blocks the corticospinal tract as it traverses the internal capsule. Glutamate is released in the ischaemic tissue.

The *lateral descending system* consists of the corticospinal, the corticobulbar and the rubrospinal tracts. Interruption of the lateral descending system to the brainstem and spinal cord causes contralateral paresis, weakness of the finger

muscles with loss of fine movements, and loss of the abdominal and cremasteric reflexes. Following the initial spinal shock, a series of release signs are found: *Positive* sign of Babinski, *spasticity* (eg, a motor condition dominated by increased tonic and phasic stretch reflexes), *foot clonus*, and abnormal flexion reflexes. This syndrome is termed the *upper motor neuron disease* or the *pyramidal tract syndrome*. The positive sign of Babinski is a slow dorsiflexion of the big toe and fanning of the other toes, when the sole of the foot is stroked laterally from the heel and forward.

Fig. 7-6: A stroke patient with thrombo-embolism of the right middle cerebral artery.

A pure interruption of the *corticospinal tract alone* (the medullary pyramid) causes *decreased muscle tone* and loss of finger movement control, but it does not induce spasticity and flexion reflexes as the lesion of the lateral descending system. Lesions of the medial descending system (ie, the vestibulospinal, reticulospinal, and tectospinal tracts) causes impaired control of the axial muscles, loss of balance during walking, and loss of rightening reflexes. The fine finger movements are normal.

A patient with thrombo-embolism in the right hemisphere suffers from left-sided hemiplegia. Head and eyes (conjugated eye deviation) are typically turned toward the lesion.

A coma patient, who lacks conjugated eye deviation, is likely to suffer from brainstem injury with severe damage of the reticular activating system. Such a condition has a grave prognosis.

Arteriosclerotic brain arteries, micro-aneurysms and larger aneurysms rupture and it bleed into the brain tissue. This primary intracerebral haemorrhage is seen in patients with hypertension. The clinical picture is the same as in *thrombo-embolic stroke*, although cortical tissue damage with unconsciousness is more common here. Coma is the deepest stage of unconsciousness, where the patient is completely without reactions.

Rupture of an arteriosclerotic brain artery or an aneurysm causes bleeding. Bleeding can interrupt the corticospinal tract as it traverses the internal capsule. Such a block of the excitatory pathways to the spinal cord results in severe contralateral paresis, weakness of the finger muscles with loss of fine movements, and loss of superficial reflexes (the abdominal and cremasteric reflexes).

This is a typical result of interruption of the lateral descending system, and often termed the *upper motor neuron disease* or the pyramidal tract syndrome. The capsula interna damage interferes with other cortical efferents to the basal ganglia, the thalamus, and the pons. Therefore, the symptoms and signs are much broader than those after injury of the corticospinal system only are.

The stroke patient can slide into deep unconsciousness termed *coma*. Coma is the deepest stage of unconsciousness. The comatous patient is completely without reactions to even the strongest stimulus. The EEG is dominated by delta waves. When coma proceeds into *brain death*, the EEG trace shows no electrical activity.

Glutamate is released during cerebral ischaemia, such as the ischaemia occurring after a stroke. In animals, N-methyl-D-aspartate-receptor (NMDA) antagonists can prevent ischaemia-induced neurodegeneration.

Microembolism or fall in cerebral perfusion may cause a syndrome called *transient ischaemic attacks*. Small emboli (clots, atherosclerotic material, air or fat) occlude the small arterioles and the brain capillaries.

Hypertension can lead to fibromuscular hyperplasia of the walls of parenchymal brain arteries and arterioles. The proliferation reduces the calibre of the arterioles and leads to microinfarction. Multiple microinfarcts impair cognitive functions and lead to dementia.

2. Head injuries

If the retinal artery is temporarily blocked by a microembolus, the patient experiences a sudden transient loss of vision (amaurosis fugax). Temporary bloodflow reduction in the posterior cerebral artery to the medial surface of the temporal lobe causes transient amnesia (memory loss). Transient aphasia is caused by bloodflow reduction to the language comprehension area (Wernicke’s area) of the dominant hemisphere. Transient hypoperfusion of this area causes sensory aphasia (ie, difficulties in understanding written or spoken language, although single words may be

recognised). When this occurs in the non-dominant hemisphere, the result is apraxia such as dressing apraxia.

The principal causes of head injuries are road traffic accidents and alcohol abuse. Head injury often results in a simple brain concussion, but due to intracranial bleeding the condition sometimes becomes life threatening. The bleeding is located extradurally (epidurally), subdurally or subarachnoid with different degrees of brain parenchymal damage.

Computerised tomography (CT) scanning has revolutionised the diagnostic work with head injuries. CT scanning reveals non-invasively the location of blood, skull fractures and cerebral contusion. The location of blood is epidural, subdural, subarachnoid, intraventricular or intracerebral.

Head injuries are divided into *2a. Simple concussion and brain contusion, 2b. Epidural haematoma, 2c. Subdural haematoma, 2d.Subarachnoid haemorrhage, and 2e. Intracranial mass lesion.*

2a. Simple concussion and brain contusion

Simple concussion is defined as a *transient loss of consciousness* followed by complete recovery. A short period of amnesia is often related to the loss of consciousness. This is a migraine injury, where the duration of the unconsciousness indicates the severity of brain damage.

Brain contusion refers to brain damage with prolonged coma, amnesia and focal signs. Later on such patients often suffer from chronic impairment of higher cerebral functions and hemiparesis. *Post-traumatic* epilepsy is frequently caused by head injury with coma following depressed skull fractures, brain contusion or intracranial haematoma. Actually, depressed skull fracture causes a high incident of post-traumatic epilepsy.

Traffic accident victims with severe brain damage may develop the so-called *punch-drunk syndrome* (dementia with extrapyramidal signs), which typically is found among professional boxers (Chapter 18) and alcoholics.

2b. Epidural haematoma

The *middle meningeal artery* and its branches are located in the temper-parietal region. Skull fractures in this region or in regions traversing a dural sinus often cause bleeding into the epidural space.

Extradural or epidural haematoma is caused by rupture of the middle meningeal artery due to a skull fracture (Fig. 7-7) or by tearing of dural veins to the sigittal sinus. The skull fracture sometimes is accompanied by CSF loss (eg, rhinorrhoea and otorrhoea).

Following the head injury with a period of unconsciousness, the patient may appear in a good condition, but suddenly he looses consciousness or develops hemiplegia. Early surgical drainage is lifesaving.

A blow to the *temporo-parietal region* may lead to fracture with transection of one or more branches of the middle meningeal artery (Fig. 7-7). *Pulsate bleeding* at the high systolic pressure can dissect the dura mater from the calvarium and form an *epidural compartment* filled with blood. This is a gradual process, because the dura adheres firmly to the bones, and there is usually an asymptomatic interval of 5-6 hours. Since the supratentorial volume is fixed, the expanding haematoma displaces an equal volume from the supratentorial compartment, firstly by reducing the CSF volume, secondly by pressing brain tissue through the orifice as a *trans-tentorial* and possibly also as a *subfalcine hernia* (Fig. 7-7).

Fig. 7-7: Development of epidural haematoma with herniation and displacement of the falx cerebri. The unilateral pupillary dilatation is caused by oculomotor nerve palsy.

As the rising intracranial pressure exceeds the pressure in the large venous sinuses, the veins are compressed and the venous stasis secondly impedes the arterial bloodflow to the brain. The result is *cerebral ischaemia* (hypoxia and hypercapnia), that occurs even with a high systolic arterial pressure. The high arterial pressure elicits a decline in heart rate via the arterial baroreceptors. Hereby, the ventricular filling is increased, whereby myocardial contractility is increased. The cortical impairment is recognised clinically as confusion and disorientation.

The brain tissue is displaced by the growing haematoma, and the tissues of the uncus of the hippocampus are pressed through the tentorial orifice as a *trans-tentorial herniation*. Hereby, the oculomotor nerve (III. cranial nerve) is pressed against the edge of the tentorium and the resulting nerve palsy is shown as a fixed, dilatated pupil (Fig. 7-7).

The clinical picture is that of a *simple concussion* with a brief period of unconsciousness followed by recovery for some hours. Simple concussion is defined as transient loss of consciousness. The loss is due to traumatic malfunction of neurones in the *reticular formation* of the brainstem. After 4-8 hours of seemingly recovery, the patient suddenly looses consciousness and develops hemiplegia. The transtentorial herniation is recognised, when the patient is found with an ipsilateral dilatated pupil (Fig. 7-8), but terminally both pupils are fixed and dilated. In the terminal phase tetraplegia develops.

The herniated brain tissue also compresses and displaces the brain stem (midbrain, pons and medulla) resulting in venous stagnation of blood and *ischaemia*. Impaired function of the neurones in the *reticular formation* and the *cardio-respiratory control centres* leads to unconsciousness and cardio-respiratory failure. Lack of oxygen for even a short period results in neuronal damage and necrosis, which is irreversible.

An *epidural haematoma* must be recognised and evacuated as soon as possible. Otherwise the bleeding progress until death ensues.

2c. Subdural haematoma

Subdural haematoma is an *accumulation of blood in the subdural space* caused by venous bleeding. The cause is head injury with *latency* between the event and the symptoms (headache, confusion, stupor, coma, delirium, hemiparesis, epilepsy etc). CT scanning confirms the diagnosis. The latency is sometimes so short, that the clinical picture resembles that of extradural haematoma.

The arachnoid is bound to the cerebral hemispheres, but unattached to the dura mater. Veins from the cerebral hemispheres cross the subarachnoid space, penetrate the arachnoid and the dura, and finally enter the dural sinuses. Injuries applied to the frontal or occipital regions initiate shock waves through the liquid brain tissue, whereby the cortical veins are cleaved just before the blood has reached the saggital sinus.

Fig. 7-8: Development of subdural haematoma with transtentorial and subfalcine herniation.

In this way blood accumulate in the subdural space (Fig. 7-8). The bleeding may be from only one vein, and the development may be slow. Latency between the time for injury and the occurrence of the first symptom can be weeks or months. Headache and confusion are unspecific indications in the elderly. *Cognitive functions* are often impaired by bilateral subdural haematoma. Manifest dementia is sometimes misinterpreted as senility. CT, MRI or arteriography confirms the diagnosis. Surgical drainage is performed.

New bleeding may develop acutely with terminal transtentorial herniation. Some types of subdural haematoma resolve spontaneously.

2d. Subarachnoid haemorrhage

Subarachnoid haemorrhage is a *spontaneous arterial bleeding* into the subarachnoid space. The clinical picture can be that of delirium. Diagnosis is confirmed with CT scanning, and neurosurgical closure of the aneurysm is sometimes possible.

Bleeding into the subarachnoid space is most often spontaneous rather than traumatic. The circle of Willis and adjacent vessels is the most frequent site for *saccular* or *berry* aneurysms (Fig. 7-9).

Fig. 7-9: The circle of Willis with saccular aneurysms (black).

The aneurysms *rupture spontaneously*, often at rest and the patient experience a sudden, devastating headache followed by loss of consciousness. The *neck is stiff* and the back is stiff as well.

Subarachnoid or intraventricular blood is clearly demonstrated by CT scanning, and in such cases lumbar puncture is unnecessary (Fig. 7-10).

Fig. 7-10: Subarachnoid and intraventricular bleeding in the left lateral ventricle (left). The blood-filled lateral ventricle is also projected to the base of the brain (right).

Previously, the diagnosis was confirmed by the presence of blood in the CSF. Today, the lumbar puncture is often avoided, as a spinal tap causes a sudden pressure differential between the supra- and infra-tentorial compartments. This may elicit transtentorial herniation with brainstem compression and death. Angiography is performed on patients fit for neurosurgical closure of the bleeding site.

2e. Intracranial mass lesions

located *supratentorially* (above the tentorium cerebelli) can compress the brain towards the tentorium as to block the upward flow of CSF and thus its absorption. Such mass lesions are brain tumours, encephalitis, meningitis, haemorrhages, aneurysms, brain abscesses, and the effect is similar to the effect of brain contusion just analysed.

Hereby, the CSF-pressure below the tentorium cerebelli increases. A rise in CSF-pressure below the tentorium results in *papiloedema*, because it creates a high pressure inside the optic nerve sheath and thus pushes fluid into the optic disc or papilla. Ophtalmoscopy reveals blurring of the edges of the papilla and dilated retinal veins without the normal pulsation. The papilla looks like the top of a champignon. Lesions of the vessels result in visible, retinal haemorrhages. Continuous intracranial pressure monitoring is important during treatment of comatous patients with severe head trauma.

Intracranial mass lesions are almost always surrounded by *cerebral oedema*. Mass lesions that include the cerebral cortex often lead to *epilepsy*.

Cerebral oedema is caused by increased pressure in the brain capillaries or by lesions of their walls.

A rise in cerebral arterial pressure above the upper limit for *autoregulation* (ie, almost constant bloodflow despite rising driving pressure) results in brain oedema. *Brain oedema* compresses intracranial vessels, whereby brain bloodflow is reduced and *brain ischaemia* develops. This is the start of a vicious cycle, because the hypoxia increases the capillary permeability and dilates also the arterioles. Hereby, the brain oedema develops further. Hypoxia also blocks the Na^+-K^+-pump, whereby the brain cells swell (eg, intracellular overhydration). Intravenous infusion of a concentrated osmotic solution such as mannitol drags oedema fluid from the brain tissue, and benefit the patient. A patient in coma, suspected of increased intracranial pressure, can be treated with 1- 2 g mannitol per kg iv., while further procedures are carried out.

Intubation and hyperventilation should also be instituted to any comatous patient. Reduction of P_{aCO2} to 25 mmHg (3.3 kPa) will rapidly reduce intracranial pressure by decreasing cerebral bloodflow and blood volume.

Brain stem compression occurs when intracranial mass lesions above the tentorium damage the ascending reticular activating system (RAS). The high pressure pushes the basal parts of the temporal lobes through the incissura tentorii, and the cerebellar tonsils through the foramen magnum. Brain tissue incarceration with brain stem compression is a serious cause of coma, which must be diagnosed and treated immediately. Suspicion of increased intracranial pressure is contraindication of lumbar puncture, because it may cause brain stem compression by transtentorial herniation. Accordingly, ophtalmoscopy with the exclusion of *papillary oedema* is necessary before any lumbar puncture.

3. Epilepsy

Epileptic seizures are partial or general. Epilepsy is an *abnormal paroxystic discharge from cerebral neurones* resulting in a condition with clinical consequences.

The normal EEG waves are due to synaptic potentials by groups of neurons including pyramidal cells. An epileptic seizure is characterised by *high voltage-high frequency discharge* (100-200 µV) from large groups of neurons or from the entire cortex.

Partial or focal seizures can be caused by an *epileptic focus* anywhere in the cortex. The causes of focal seizures are acquired lesions such as cysts, tumours, scar tissue, infections, ischaemic lesions. The epileptic discharge causes involuntary muscular contractions on the contralateral side. Foci in the somatosensory cortex produce sensory hallucinations called an epileptic *aura*. These hallucinations precede the epileptic seizure. The aura varies and is particular for a certain patient. Epileptic foci in the visual cortex cause visual auras, while epileptic foci in the vestibular cortex produce an aura-feeling of spinning. *Psychomotor epilepsy* originates in the limbic system and causes emotional hallucinations and muscle contractions. Focal seizures are characterised by high *epileptic spikes* in the EEG. *Motor seizures* originate in the motor cortex of the opposite side, and they follow a specific pattern in each patient. They are called *Jacksonian seizures* and often precede the generalised types.

Generalised epileptic seizures involve most of the brain and imply *loss of consciousness*. *Generalised absence* (petit mal) is a transient loss of consciousness. These short attacks are recognised by *spike-doom waves* in the EEG.

Primary generalised tonic-clinic seizure (grand mal) is characterised by an extreme and widely distributed electrical activity, with tonic-clinic convulsions of the entire body. Presumably, a basic neuronal circuit activates the cortex of both hemispheres in generalised seizures. The hyperactive nerve cells release K^+ and glutamate during a seizure.

Small children with high fever (purexia) often react with generalised epileptic seizures called *febrile convulsions.* Diabetics in hypoglycaemia (ie, a blood glucose concentration below 3 mM) may develop generalised convulsions.

Epileptogenesis. The genesis and spread of epileptic discharges are poorly understood. The increased cortical excitability, with *high voltage-high frequency discharge* over the entire cortex, is not explained. Several mechanisms are probably involved:

1. The GABA-receptor complex contains the receptor site for not only GABA, but also for anti-epileptic drugs, including barbiturates and benzodiazepines, that potentiate or mimic the GABA-effect. GABA is the major inhibitory neurotransmitter in the brain. GABA opens chloride-channels, whereby the neurons are hyperpolarised, which reduces the likeliness of epileptic firing.

2. Epileptic seizure activity is either initiated or propagated through N-methyl-D-aspartate-(NMDA)-receptors binding glutamate or aspartate. The NMDA-receptor bind glutamate and the result are a high-frequency neuronal discharge. The hyperactive neurons release K^+ and excitatory amino acids or EAA's (glutamate and aspartate). NMDA-receptors and their ionic pores often work with a Mg^{2+}-sensitive glycine-receptor as a co-transmitter to glutamate. NMDA-receptors are the only ligand-gated channels that are also voltage-gated and Ca^{2+}-permeable. Drugs that effectively block the NMDA-receptor can reduce the abnormal excitatory spread of transmission and the focal epileptogenesis.

3. Block of Na^+-channels. The Na^+-K^+-pump and other Na^+-channels normally re-establish the ionic distribution after a discharge, and thus allow the cell to depolarise again. Some antiepileptic drugs do not alter the first action potential but reduce the repetitive firing-pattern. Carbamazepine and phenytoin block Na^+-channels, which prolongs the relative refractory period, and reduce repetitive firing of neurons.

During seizures, the extracellular [K^+] increases substantially, so the resting membrane potential is reduced. This makes the neurons more excitable and promotes the spread of the discharge. Fortunately, phenytoin blocks better at high K^+-concentrations around the neurons.

Adenosine inhibits the initiation of seizures in experimental animals. Carbamazepine promotes the adenosine-inhibition and thus blocks or reduces epileptogenesis.

Emergency therapy of a seizure is to keep the airways patent, apply diazepam suppositories to patients with prolonged seizures, and intravenous glucose in case of hypoglycaemia. The patient must be protected from harming himself during the few minutes of generalised cramps.

Long-term therapy of primary generalised tonic-clonic epilepsy (grand mal) and partial epilepsy is frequently made by use of carbamazepine or phenytoin. Generalised absence (petit mal) is frequently treated with sodium valproate.

Status epilepticus is life threatening due to cardio-pulmonary insufficiency and must be treated immediately with cardio-pulmonary support and diazepam intravenously.

4. Movement disorders

Disorders of the neurotransmission in the extrapyramidal system results in movement disorders of two types. Loss of movement with increase in muscular tone is termed akinetic-rigid syndromes, whereas disorders with involuntary movements are called dyskinesias.

4a. Parkinson´s idiopathic disease

or *shaking palsy* is characterised by tremor at rest, rigidity, akinesia or bradykinesia and postural changes.

Parkinson's disease is characterised by well-preserved cholinergic activity, but a reduction of the dopamine content in putamen and substantia nigra, of noradrenaline and 5-hydroxytryptamine in putamen, and of the GABA-synthesising enzyme glutamic acid decarboxylase in substantia nigra and in the cerebral cortex.

The main pathological mechanism is degeneration of dopaminergic neurons in substantia nigra. The consequence is a severe lack of dopamine in the striatum. The lack of dopamine in substantia nigra hyperactivates the GABA pathways to the thalamus, which activates the motor cortex neurons. This increases the discharge of alpha-motor neurons in the spinal cord resulting in plastic rigidity, akinesia and dyskinesia.

L-DOPA is a precursor of dopamine that is capable of crossing the blood-brain barrier. Administration of this drug, which is transformed to dopamine in the brain, relieves much of the rigidity and akinesia by inhibiting the striatum. Transplantation of dopaminergic neurons into the striatum has been explored.

Increased dopamine activity relieves rigidity. Too much dopamine causes chorea (see below).

Increased acetylcholine activity or reduced dopamine activity causes rigidity and bradykinesia in healthy persons. This is why reserpine, which depletes neurons for dopamine, and butypherones, which block the secretion from dopaminergic neurons, all causes so-called drug-induced Parkinsonism. Drug-induced Parkinsonism is a common side effect in patients treated neuroleptic drugs or in patients given metochlopramide. Akathisia refers to a condition with restlessness and uncontrolled repetitive movements in patients with drug- induced Parkinsonism. Drug-induced Parkinsonism is refractory to usual drug therapy. The patients respond immediately upon seponation of the inducing drug.

The tremor is often pill rolling between the fingers and thumbs, and not necessarily bilateral. The stiffness or rigidity is called lead-pipe or plastic rigidity, because the high muscle tone is equal throughout the range of passive movements and the same by flexion and extension. Sometimes the resistance to passive movements is jerky, so-called cog-wheel-movements.

Akinesia (hypokinesia, bradykinesia) means inability to initiate normal movements. The face is mask-like with rare blinking and monotonous dysarthria.

Postural changes include a forward posture with short step gait and no arm swinging. The patient easily loose balances and falls stiffly to the ground.

Fig. 7-11: Neurotransmitters in the basal ganglia. ACh stands for acetylcholine.

Dementia or cognitive disturbances are present in some patients. Dementia is ageing of the brain with a resulting loss of mental powers.

A balance between the effects of dopamine and glutamate is a necessity for the normal functioning striatum.

Patients are treated with an amino acid precursor of dopamine: L-dihydroxyphenylalanine (L-DOPA). This molecule can cross the intestinal-blood barrier and the blood- brain barrier easily. Hereby, L-DOPA reach the inside of neurons (carrier-mediated transport).

L-DOPA is normally synthesised in dopaminergic neurons from dietary L-tyrosine. Exogenous L-DOPA is methylated by catechol-O-methyl transferase (COMT) to 3-O-methyl DOPA, or it is decarboxylated by decarboxylase to dopamine. Finally, dopamine is catabolised to homovanillic acid by COMT and monoamine oxidase (MAO). Most of the exogenous L-DOPA is lost by decarboxylation in patients where carbi-DOPA is not administered concomitantly. Carbi-DOPA inhibits the decarboxylase and reduces the side effects of L-DOPA.

When Parkinson patients are treated with too much of the dopamine precursor L-DOPA, they may develop hallucinations, fear and paranoid delusions (schizophrenia-symptoms), because excess dopamine causes schizophrenia or schizophrenic symptoms and signs. Chloropromazine and haloperidol decrease the dopaminergic effects and are called anti-schizophrenics. These drugs block both the D_1 and the D_2 dopamine receptors, so they reduce the dopaminergic effects, but also cause extrapyramidal side effects on the top of cholinergic and adrenergic blockade.

Amantidin increases the dopamine release from nerve terminals in the striatum, so this drug is effective mainly in the early phases of Parkinson´s disease, before all dopaminergic neurons are degenerated (Fig. 7-11).

Anticholinergic drugs, which block muscarinic, cholinergic receptors, are still in use to treat early Parkinson symptoms such as tremor. The side effects are dry mouth, bladder weakness, constipation, and confusion and memory loss.

Parkinson patients develop DOPA resistance following years of dopaminergic medication. The use of antagonists for the glutamate receptor in Parkinson patients decreases the activity of the glutamate pathway to the cortex, and reverses the akinesia and rigidity.

4b.Wilsons disease (hepatolenticular degeneration)

is an *autosomal recessive anomaly* of the copper metabolism. The abnormal gene is located on chromosome 13. Normally, copper is absorbed from the gastrointestinal tract and in the liver it is incorporated into caeruloplasmin. Normally, copper is excreted into the bile.

Wilson-children have low serum caeruloplasmin and serum copper. They fail to excrete copper. Copper is accumulated in the basal ganglia of the brain, the liver (liver cirrhosis), and the cornea/lens. Wilson-children show an *akinetic-rigid syndrome* with liver cirrhosis, haemolysis, anaemia and visual disturbances. Early diagnosis with long-term treatment (penicillamine) has improved the prognosis.

4c. Dyskinesias (Chorea and Myoclonus)

Chorea is a type of involitional hyperkinesia, with jerky movements of the limbs. The movements look almost like voluntary, purposeful movements.

A special type of chorea is *athetosis* (aethymology: "not fixed"), which refers to slow, twisting involuntary movements of the fingers. Athetosis is seen in children following brain damage with bilateral damage of the nucleus subthalamicus.

St. Vitus dance or *Sydenhams chorea minor* is a complication following rheumatic fever with encephalitis. The movements are usually unilateral. Recovery typically occurs spontaneously.

Hemiballismus or *hemichorea* describes violent, throwing movements of one arm or one side of the body - as if the patient tried to throw a ball. The cause is partial lesion of the contralateral *nucleus subthalamicus* - often due to thrombosis. Ballism means flailing movements of the limbs.

Huntington's chorea

Huntington's chorea is an *autosomal dominant genetic defect on chromosome 4*. Its characteristic disorders are hypotonia, dementia and involuntary hyperkinesia. Huntington's chorea is due to a defect in GABAergic and

acetylcholinergic interneurons of the striatum and the cerebral cortex. The two transmitters are normally synthesised by the enzymes, glutamic acid decarboxylase and acetylcholine transferase, but their concentrations in the interneurons is markedly reduced. The diagnosis is confirmed with genetic testing. The prognosis is bad with rapid mental deterioration in particular in young patients. The dementia of the Huntington's chorea is caused - as for most types of dementia - by cortical degeneration.

GABA receptors are usually inhibitory. When GABA no longer inhibits the globus pallidus from the striatum, this leads to a stronger inhibition of the thalamus, which is probably the cause of the involitional choreiform movements. Chorea is opposed to rigidity. Low doses of a dopamine agonist may reduce the choreiform movements of patients with Huntington's disease.

Myoclonus refers to brief contractions or jerks of one or more muscles. Myoclonus is often called nocturnal myoclonus, because it occurs at night. Generalised myoclonus resembles epileptic seizures, and is related to epilepsy following brain damage by hypoxia. Myoclonus is related to drug toxicity or metabolic toxicity from renal or hepatic insufficiency.

Tics are repeated twitching of facial or neck muscles. Tics are also called mimic or focal myoclonus. The tics may begin in childhood for unknown reasons. Tics are extremely resistant to any therapy.

Tremor can be caused by hyperthyroidism and by Parkinsonism, but it is also a typical side effect of alcohol, narcotics and drug abuse. Some cases of essential tremor can be reduced by beta-blockers.

5. Multiple sclerosis

Multiple or disseminated sclerosis refers to a common neurological disease caused by inefficient myelin production in the oligodendroglia.

The cause is unknown, but the acquired defect in the oligodendroglia cells results in demyelinized areas or plaques in the CNS. The prevalence increases progressively with the distance from the equator, and the patients have an abnormal immune response with large concentrations of antibodies to virus infections.

The demyelinized plaques are mainly localised to the brainstem, cerebellum, periventricular region and optic nerves. Motor neurons of the spinal cord and peripheral nerves are rarely affected by demyelinisation.

Blurring of vision in one eye is usual with disc swelling of the optic nerve at ophtalmoscopy. There is also diplopia, vertigo, nystagmus, and dysphagia, when the brainstem is affected. Later paraparesis and tetraparesis develops. Death ensues by lung infections or uraemia.

Magnetic resonance imaging (MRI) with scanning of the brain and CNS can visualise demyelinized plaques in the periventricular white matter or elsewhere. MRI is an expensive technique, where protons are activated with radiofrequency waves to create images. In the CNS, the white and the grey matter are distinguished.

This is a disabling disorder for which there is no cure. Interferon has been tried with some effect on the lesions visualised with MRI. Palliative treatment necessitates teamwork.

6. Inflammations

Meningitis refers to inflammation of the meninges. Clinically, the meningitis syndrome is characteristic. The meningitis syndrome is a patient with high fever, headache, photophobia and vomiting. The patient - often a child - is placid and inactive, consciousness may be impaired and neck stiffness develops.

Bacteria, viruses, fungi, chemicals or drugs cause meningitis, or unusual organisms in immuno- compromised patients.

Immediate administration of intravenous benzylpenicillin is life saving in cases of acute meningococcal or other bacterial meningitis together with urgent investigations.

Encephalitis is inflammation of the brain tissue caused by the same organisms as meningitis. Herpes simplex encephalitis is treated with acyclovir intravenously. Acyclovir inhibits DNA synthesis and thus the proliferation of the virus. Japanese B encephalitis is avoided by vaccination of travellers to the Far East. Other causes to acute viral encephalitis are Coxsackie virus, Echo virus and mumps virus.

AIDS in the CNS is caused by the HIV itself or the CNS disease is caused by other infectious agents - in particular fungi, TB, or Escherichia coli which damage brain cells. The clinical picture is meningitis, myelitis or encephalitis.

Neural syphilis may occur as *tabes dorsalis*. Tabes dorsalis is caused by demyelinisation of the dorsal roots of the spinal cord. Lancinating pains, ataxia, loss of reflexes, muscle wasting, neuropathic joints, Argyll Robertson's light-stiff pupils and optic atrophy.

Tertiary syphilis can be avoided if the primary syphilis infection is treated correct. Usually, injection of 1 g i.m. Benzylpenicillin for 2 weeks is enough.

Rubella encephalitis caused by rubella virus may progress following some years, because of antibody production against rubella viral antigen.

Creutzfeld-Jacobs Disease (CJD) and *KURU* (among cannibals in Papua, New Guinea) are known from the spongioform encephalopathy seen at autopsy - the brain looks like an Emmenthaler cheese.

The pathology is similar to that of bovine spongiform encephalopathy of cattle and sheep ("SCRAPIE"). CJD is inherited of transmitted to man with a *prion*. The prion is an *abnormal neuronal membrane protein*, which can mutate like a virus. The prion is resistant to usual sterilisation procedures, and the incubation period is not always for years. Prions are transferred when eating neural tissue from sick cows or sheep and in other ways. The first signs in humans are various neurological insults such as sudden blindness, difficulties in gait and balance, memory and concentration disturbances, and slowly progressing dementia until a rapid death. – The hereditary form of *CJD* is caused by mutation of the human gene (PRNP) for the neuronal membrane protein.

Today, *KURU* is history. The high incidence in New Guinea 30 years ago was due to cannibalistic rituals. *Gadjusek* showed that *KURU* was infectious, and received the Nobel Prize in 1976.

7. Intracranial tumours

Most intracranial tumours are *primary* tumours (75%) and approximately one-quarter are *secondary* (metastases).

The primary malignant tumours are *gliomas* (eg. astrocytomas and oligodendrogliomas), and the primary benignant tumours are *meningeomas* and *neurofibromas*.

The metastases originate from primary tumours in the breasts, bronchi, kidneys, prostate, stomach, and thyroid etc.

Magnetic resonance imaging (MRI) is a scanning technique, where protons in a strong magnetic field are bombarded with radiofrequency waves in order to produce images. MRI scanning can picture brain tumours, multiple sclerosis lesions, and syringomylia among others. MRI scanning can even separate white from grey matter. MRI scanning is replacing myelography, because it can visualise spinal cord compression, spinal cord tumours and other malformations.

Gliomas originate in the neuroglia. *Astrocytomas* are gliomas originating from astrocytes. Astrocytomas are usually located in the cerebrum in adults, and in the cerebellum in children

Oligodendrogliomas originate from the oligodendroglia and grow slowly in the cerebral tissue

Meningeomas originate from the arachnoid matter usually along the venous sinuses above the tentorium. They are benign and grow slowly.

Neurofibromas arise from Schwann cells usually around the 8^{th} cranial nerve (acoustic Schwannomas).

Symptoms and signs of brain tumours are treated already in *2e. Intracranial mass lesions.*

8. Dementia contra delirium

Dementia (*senility* or ageing of the brain) disturbs almost all cognitive brain functions, whereby the personality of the patient is completely changed (cognitive functions as calculation, comprehension, judgement, language, learning ability, memory, orientation, and thinking).

Dementia develops *slowly* and has no diurnal variation. Cortical atrophy is found by using CT or MRI scanning of the brain. The clinical differential diagnosis to delirium and depression is sometimes difficult to establish, but it is consequential to the delirious patient, if the diagnosis is misinterpreted as dementia (Box 7-3). Dementia is an exclusion diagnosis.

Box 7-3: Differences between the syndromes dementia and delirium			
Syndrome	**Dementia**	**Delirium**	**Depression**
Attention and cognition	Variable	Globally impaired	Variable
Consciousness	Normal	Impaired	Normal
Diurnal variation	None	Worst at night	Morning worst
Development	Insidious	Sudden	Variable
Hallucinations	None	Often visual	Auditory
Speech	Perseveration	Difficulty finding words	Normal
Delusions	Absent	Fleeting	Systematised
Primary causes	Cortical trophia/	Illness -intoxication	Loss of NA
Therapy	Palliative	Causative treatment	SSRI

Alzheimer's disease

is a possible cholinergic system disease. Altzheimers disease is a primary (ie, unknown aetiology) cortical brain atrophy. Lack of the acetylcholine-producing enzyme choline acetyltransferase, and of acetylcholine has been demonstrated by neurochemical studies. Alzheimer's disease is a form of presenile dementia or premature ageing of the brain (ie. occurring before the age of 70). The disease is rapidly progressing to complete loss of mental powers, in particular loss of memory and normal emotional behaviour. CT scan shows *cortical atrophy* and excludes brain tumours. At autopsy argentophilic plaques filled with *amyloid protein A4* are found in the hippocampus, basal ganglia, thalamus and the cortex. The gene defect causing familial Alzheimer disease is located on *chromosome 21*, close to the *pro-A4 gene.*

The cholesterol transport to the tissues is also affected. Alzheimer's disease is probably caused by neuronal degeneration in the nucleus basalis close to the globus pallidus, and possibly also to lack of somatostatin and substance P in deep brain centres. Normally, cholinergic axons from the nucleus basalis project to the cortex, and their functions relate to memory and to the limbic system functions.

There is no specific treatment of dementia. Anxiety and depression is treated symptomatically (Box 7-3).

Delirium

Delirium is an acute impairment of consciousness also called *toxic confusion*. Sense impressions are misinterpreted,

the mind and memory work incoherently, and the patient is frightened and suspicious because of hallucinations. Relatives often mention senility, but they may also inform about an *acute start*, so delirium is recognisable. Besides being acutely developing, delirium is also worst at night with visual hallucinations and incoherent speech and perseveration.

In contrast, the dementia patient is conscious, cannot find the right words and the development has been slow.

The basis for the delirious pattern is organic brain disease caused by *intoxication* (eg. alcohol, drugs, poisons), brain damage by infections, lesions, subarachnoidal haemorrhage or tumours, systemic infections (malaria, septicaemia, TB), and *metabolic brain damage* (eg, hepatic or renal failure, hypoxia, vitamin B_2, B_6, B_{12} deficiency).

The treatment of delirium concentrates firstly on the *underlying* disease (including electrolyte disorders, ischaemia etc), and secondly on the symptomatic aspect.

II. Psychiatric Disorders

The international Classification of Disease and Related Health Problems (ICD 10, WHO) is used.

The most serious disorders are *psychoses*, and the most common disorders are *neuroses*.

The description concentrates on the two classical psychoses schizophrenia and manic-depressive psychosis. The following personality disorders (neuroses) are described: Phobic anxiety neurosis, obsessive-compulsive disorders, dissociative-conversion disorders (hysteria), and eating disorders (anorexia nervosa and bulimina nervosa). The pathophysiology of affect and stress is also considered.

1. Schizophrenia

means splitting of the mind or disconnection of psychic functions (emotional and cognitive). Schizophrenia is a *psychosis with hallucinations, dissociation of ideas, intense fear, and paranoid delusions* (paranoia). Schizophrenia is possibly caused by hypersecretion of *dopamine* or by blockage of the *glutamate producing neurons* from the cortex to the striatum. The balance between these two neurotransmitters in the striatum is seriously disturbed.

The clinical syndromes covered by this term include - according to WHO - auditory hallucinations (eg, hearing voices), thought withdrawal with abnormal posture, delusional perceptions with paranoia and external control of emotions with persecution from the outside. The patients' feel that their thoughts and emotions are broad casted and they not only hear voices commenting their lives, but also their own thoughts are spoken aloud.

The cause is sometimes clarified as a biochemical brain damage with hypersecretion of dopamine from neurons in the *mesolimbic dopaminergic system* close to substantia nigra or by blockade of the glutamate-producing neurons from the cortex to the striatum. An imbalance between the effects of dopamine and glutamate spoils the normal function of the striatum. A special gene located in chromosome 5 increases the risk of schizophrenia. Dopamine agonists such as amphetamine, and other psychotic drugs (LSD, mescaline, and ecstasy) can cause schizophrenic psychosis.

The genetic involvement is demonstrated by a 50% risk for the monozygotic twin of an affected person. There is a 40% risk for two affected parents for having a schizophrenic child. The gene is probably located on *chromosome 5*. Some schizophrenics have limbic dysfunction of the left hemisphere.

Schizophrenia begins in young adults of both sexes, and may be more than one entity. Schizophrenics are frequently vulnerable to highly expressed emotions.

Chronic schizophrenics are characterised by lack of drive, underactivity, social withdrawal, and emotional emptiness. Catatonia (stupor, stereotypes, and automatic obedience) was previously seen in many institutional patients, and may still be seen among understimulated patients.

Dopamine blockers - blocking D_1 and D_2 dopamine receptors - are the drugs of choice in acute schizophrenia. These drugs belong to the phenothiazine family (chlorpromazine, trifluoroperazine and promazine).

Side effects are unavoidable as the drugs block both D_1 and D_2 receptors, as well as adrenergic and cholinergic receptors. The side effects are *extrapyramidal* (acute dystonia, Parkinsonism, akathisia and tardive dyskinesia), *autonomic* (hypotension and ejaculation failure), and *anticholinergic* symptoms (dry mouth, urine retention, constipation and blurred vision).

2. Manic-depressive psychosis

covers severe abnormalities of mood. Mood ranges from severe depressive psychosis over moderate and minor depression, sadness, normal mood, happiness, euphoria, hypomania, and severe mania.

The diagnosis manic-depressive psychosis describes patients with periodic attacks of mania or depression, separated by periods of normal behaviour.

The diagnosis also includes patients with depressive periods alone, or with only manic periods.

Endogenous depression is characterised by early morning waking with unresponsive sadness, guilt feeling, suicidal feelings, and lack of a precipitating factor. Severe depression disturbs mood, talk and initiative. One type of mental depression is related to reduced formation of noradrenaline in the locus coeruleus, and of serotonin in the midline raphe nuclei of the brainstem, which seriously damage the limbic system. - Other types of depression are called exogenous or reactive depressions, because they are considered to be due to exogenous or environmental factors.

Medical drugs that inhibit the production of noradrenaline and serotonin often cause depression.

Hypomania is mild mania with euphoria, overactivity and disinhibition.

The *genetic aetiology* of manic-depressive psychosis is confirmed by the concordance of two thirds of monozygotic twins, and by the fact that more than 20 % of dizygotic twins are concordant. There is a clear overweight of females. *Winter depression* from autumn to spring is frequent in areas with lack of light in the winter months. Up to 20% of the population north of the polar circle suffers from winter depression. Light therapy several hours daily are so effective that overdoses may release manic phases.

Monoamine-neurotransmitters are depleted in depression but increased in manic phases.

Stressful social life events (marriage, divorce, moving house, loss of job, vacation, etc.) often precipitate depression.

Selective serotonin reuptake inhibitors (SSRI) are often preferred in treatment of depressive states, because of rapid effect and lower rate of serious side-effects including addiction. These substances inhibit serotonin reuptake within the synaptic cleft, and are named "happiness pills" in the media. Happiness pills do not exist.

Depression was previously treated with *monoamine oxidase inhibitors* (MAO-inhibitors). They inhibit the enzyme monoamine oxidase A&B and thus the breakdown of monoamines. Hereby, adrenaline, dopamine and 5-hydroxytryptamine are accumulated in the brain. Tricyclic antidepressants block reuptake of monoamines and are likewise effective in the treatment of depressive patients.

Electroconvulsive therapy (ECT) is a physical treatment with rapid effect, often used for cases with suicidal or other deep depressions.

Therapeutics (such as *lithium compounds*) that inhibits the action of noradrenaline or serotonin is effective prophylactic agents against manic phases. In this model mania is caused by overproduction of monoamines, and depression by reduced formation of monoamines in the brain nuclei mentioned above. Actually, *lithium carbonate* is used in the prophylactics of manic phases. A plasma-lithium concentration of 0.5-1 mM is necessary to obtain an acceptable result.

Psychoses are treated with antipsychotics often supplemented with benzodiazepines. The typical antipsychotics are traditionally divided into high-, medium-, and low-potency drugs that are blocking the D2 and D1 dopamine receptors,

with secondary blocking of the serotonine – histaminergic- adrenergic – and cholinergic receptors. The main drugs in this category are fluopentixole, haloperidole, zuclopenthizole, chlorpromazine and levopromazine.

Side effects are unavoidable as the drugs block a range of receptor types. The side effects are extrapyramidal (dopaminergic – acute dystonia, parkinsonism, akathasia and tardive dyskinesia), serotonine related (weight gain), histaminergic (sedation), autonomic (hypotension, ejaculation failure, salivation), and anticholinergic (dry mouth, urine retention, constipation and blurred vision).

A new class of antipsychotics which do not fit into the high/low potency classification have been introduced and are gaining world wide use because of their low degree of side effects. Drugs without cholinergic activity do not lead to extrapyramidal side effects. Some drugs have less adrenergic activity and they are generally more limbic than pyramidal in their selectivity compared to the classical antipsychotics. The main drugs in this category are: Amisulpride, risperidone, sertindole, closapine and olanzapine.

Nervous and stress-related personality disorders (Neuroses)

Phobic anxiety neurosis

Anxiety neurosis is a chronic condition or it occurs as attacks of panic. *Acute overactivity* of the sympathoadrenergic system results in precardial pain and palpitations (cardiac neurosis = neurocirculatory asthenia), chest constriction, flatulence and frequent defecation and urination, lack of libido, dizziness, headache, and sleep disturbances.

Attacks of panic anxiety occur in young, nervously sweating persons, who feel that they are dying from cardiac disease or from hyperventilation with tetany and carpopedal spasms. Hyperventilation reduces the carbon dioxide tension in the alveolar air (decreased P_{ACO2}) and thus the Ca^{2+} -concentration in the ECV, which opens Na^{+}-channels, reduces the membrane potential and increases the neuromuscular irritability (see *tetany* in Chapter 17).

Symptoms and signs of anxiety (ie, sweating, palpitations, tremor, tachycardia, flatulence, and urination) are caused by increased release of adrenaline and noradrenaline from the adrenal medulla. Drugs containing b -*adrenergic blockers* are of benefit to the anxious patient, because they block the sympathetic nerves and adrenergic synapses in the CNS. Cognitive-behavioural therapy is also applied with effect – sometimes combined with selective serotonin reuptake inhibitors (SSRI). Actually, SSRI substances are the first choice in anxiety conditions.

Obsessive-compulsive disorders

These patients have *obsessional thoughts* and perform *compulsive actions* to the extent that their social lives are seriously impaired. The patient feels an irresistible obsession to perform a given act - such as washing the hands or superstitious check of a closed door - again and again. To the patient the behaviour is often quite meaningless, but still it is necessary to carry on the ritual. A frequent complaint is that dirt and excretions are nasty, and many obsessions concern excretory processes.

In *cognitive-behaviour therapy* the patients learn not to perform the compulsive rituals.

Eating disorders

Anorexia nervosa is an eating disorder in adolescent females resulting in severe malnutrition. The patient has an intense wish to be thin. Biological and psychological factors are involved. In a few cases there is regression into childhood. The girl or boy tries to escape from the problems of puberty and adolescence.

The increased concordance in monozygotic twins indicates a genetic aetiology.

This disorder occurs among amenorrhoic teen-age girls, who express an abnormal fear of being fat. Typically, they have an extremely low body weight. The girl is usually bright and knowledgeable, and her parents are overprotective. The patient sometimes realises the presence of problems in accepting the role as a maturing female.

Low plasma Gonadotropin levels with impaired response to LHRH are frequently found.

Positive reinforcement for even small weight gains is sometimes of help. The basic psychological problems must be treated with cognitive-behavioural or other psychological treatment. Tricyclic antidepressants are beneficial in cases of *depression.*

Bulimia nervosa is diagnosed in persons who are preoccupied with food and periodically eats excessively. They may avoid overweight by self-induced vomiting just after *binge eating.* Bulimia nervosa may be associated with anorexia nervosa. *Behavioural therapy* is sometimes successful – sometimes combined with selective serotonin reuptake inhibitors (SSRI).

Dissociative-conversion disorders (hysteria) is characterised by psychologically mediated psycho-somatic disorders. These disorders have no physical pathology; they are not sympathetic overactivity and are produced without the consciousness of the patient.

Freud believed that mental energy was converted into physical disorders such as abdominal pain, blindness, double vision, deafness, muteness, fits with dramatic movements, artistic gait disturbances, hysterical paresis with normal muscle tone and deep reflexes, crude tremor, sensory loss, stigmatisation, - all with *secondary gain*. The disorders are explained as the result of repression, dissociation and *conversion* of mental energy into physical disorders. *Repression* means exclusion of memories, impulses, and emotions from consciousness, because these elements would cause anxiety and distress. *Dissociation* means an apparent dissociation between different mental activities. An example is a protective mental cover of enjoyment in terminal cases of painful cancer (French: *Belle indifference*).

The classical *hysterical triad* include mydriasis (large pupils), lack of pharyngeal reflexes and lack of plantar reflex.

The mental disorders are *amnesia* for long periods, *sleep walking* or *somnambulism* (see below), *imitation* with multiple personalities, globus hystericus and *pseudo-dementia*.

Psychotherapy is a causal treatment, although sometimes impossible to carry through.

Sleep disturbances

include insomnia, somnambulism and sleep apnoea.

Insomnia is *subjective sleep deficiency*. The patient complains that he sleeps too little, or has the impression that he cannot sleep. Such patients sleep more than they think, when studied in sleep laboratories, and their health is not impaired. There is a natural decline of the sleep duration with age, and the use of drugs should be restrained. Monotonous sounds such as music or the sounds from ocean waves have proven to be an optimal "sleeping drug" for many individuals. The common complain of insomnia among the elderly is often curable by regular motion passes (eg, walking, swimming, jogging etc).

Insomnia as *early morning waking* is a sign of depression, but it is also seen as nocturnal confusion in dementia and in delirium, where the cause may be organic brain damage or drug abuse.

Sleepwalking or *somnambulism* is a form of personality dissociation with unknown aetiology. Some of the patients have hysterical patterns (see above).

Sleep apnoea often occurs with snoring and airway obstruction in obese patients or in patients with chronic obstructive lung disease.

Affect and stress reactions to psychological or physical stress are seen in otherwise healthy individuals. One example is described above as *panic attacks*.

Self-Assessment

Multiple Choice Questions

- **I. Each of the following five statements have True/False options:**

- **A.** Cellular responses, mediated by drug receptors linked to ion-channels, are rapid compared to responses mediated by G-protein systems.

- **B**. Drugs acting via receptors have side effects, because they are bound to several receptors, distributed in several tissues, and the receptors are linked to different secondmessengers, which produce different cellular responses.

- **C**. Akathisia is an extrapyramidal defect with swaying and twisting body dyskinesia.

- **D**. The circumventricular organs (ie, hypothalamus, the pineal gland, and the area postrema have a tight blood-brain-barrier.

- **E..** Dopamine agonists. such as amphetamine and other psychotic drugs (LSD, mescaline, ecstasy), can cause schizophrenic psychosis.

II. Each of the following five statements have True/False options:

A. Dopaminergic neurons are found in the substantia nigra, noradrenergic neurons in locus coeruleus, and serotonin sensitive neurons in the raphe nuclei.
B. Catecholamines are neurotransmitters both in the sympathetic and the parasympathetic nervous system as well as in the motor endplate.
C. A single neuron releases only one neurotransmitter from all its synapses.
D. Insulin and related growth factors bind to membrane receptors that are glycoproteins protruding from the membrane.
E. The free radical gas nitric oxide (NO) is a neuronal messenger in both the central and the peripheral nervous system. The NO gas is membrane permeant and can bypass normal signal transduction in synapses.

III. Each of the following five statements have True/False options:

A. A constant small ion permeation through the cell membrane at rest is referred to as leak current.
B. The typical Na^+-channel opens promptly in response to repolarisation.
C. Opening of Na^+-channels increases the flux of Na^+ into the neuron, and depolarizes the membrane, so the effect is excitatory.
D. Immediate administration of intravenous benzylpenicillin is life saving in cases of acute meningococcal or other bacterial meningitis.
E. Most intracranial tumours (gliomas, meningeomas and neurofibromas) are primary tumours and only 25% are metastases (secondary tumours).

Case History A

A professor in linguistics, 59 years old, consults his doctor because of speech and movement problems. The patient is intellectually well functioning, but his speech has changed from motivating to a slow monotonous sequence of words. His gait is slow with small steps, and the standing position is difficult for this previous long distance runner. His facial expression is motionless, and he seems to have difficulties in initiating normal movements. There is tremor of the hands and fingers of the pill-rolling type. When the doctor examines the patient for rigidity, he finds high tonus (plastic rigidity) and cogwheel-movements.

1. *What is the main pathological mechanism of this disease?*
2. *What is it called?*
3. *Why is the muscle tone so high?*

Case History B

A female, 26 years of age, suffers from an epileptic seizure during her work as a nurse on a neurological department. A colleague saw that the nurse suddenly stopped while walking, her eyes and head turned left, her left hand moved in a

curious way, and she uttered a cry and felled. The whole body became rigid for a minute, during which time she developed cyanosis. Then the muscles started to jerk rhythmically for a few minutes. She was unconscious during the seizure and remained so for an hour after the seizure. An EEG was taken during and after the seizure. A blood sample was taken and the blood glucose was determined to 5 mM.

- *1. Describe the type of epilepsy starting the seizure and the development into a second type of seizure.*
- *2. Describe the most likely EEG findings during and after the seizure.*
- *3. What is the pathophysiological basis for a grand mal seizure?*
- *4. Was hypoglycaemia involved in the seizure?*

Case History C

A professor in economics, 56 years of age, finds it increasingly difficult to concentrate during his work. His wife and two adult children find him totally different from his normal personality; he is with- drawn, depressed and forgetful.

Two months on antidepressants prescribed by his GP does not improve the condition, which is dominated by lack of memory. Psychiatric and neurological examination disclose no evidence for depression or increased intracranial pressure due to focal brain damage. CT scan shows cortical atrophy and excludes brain tumours. The mental powers are rapidly deteriorating. The total cholesterol concentration in blood plasma is increased.

- *1.What is the most probable diagnosis (two must be considered)?*
- *2.Is there a definite criterion for one of these in this patient?*
- *3.What are the prognoses for these two disorders?*
- *4.Is the disorder of this patient inherited?*
- *After a year the patient passes away.*
- *5.What are probable findings at autopsy?*

Try to solve the problems before looking up the answers.

Highlights

- *Opening of Na^+-channels increases the flux of Na^+ into the neuron, and depolarizes the membrane, so the effect is excitatory.*
- *Closure of K^+- or Cl^- -channels decreases the flux of K^+ out of the neuron or decreases the flux of Cl^- into the cell. These events also depolarise the membrane, and again the effect is excitatory.*
- *Obviously, closure of Na^+-channels or opening of K^+- or Cl^-channels have an inhibitory effect by hyperpolarization.*
- *Two types of NO synthase (NOS) have been identified: constitutive Ca^{2+}- calmodulin dependent enzyme, and inducible Ca^{2+} - independent enzyme. Both enzymes are flavoproteins containing bound flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD).*
- *The free radical gas nitric oxide (NO) is a neuronal messenger in both the central and the peripheral nervous system.*
- *Nitrovasodilatators act by releasing NO and thereby causing coronary vasodilatation in patients with angina pectoris. Nitric oxide synthase is inhibited by L-arginine analogues.*
- *Signal transduction is a cascade of processes from the receptor-hormone binding to the final cellular response. Many hormones and neurotransmitters raise the concentration of a second messenger in the target cell via guanyl triphosphate (GTP) and act through it.*
- *The receptor-hormone complex activates a GTP-binding protein (so-called G-protein), which controls and amplifies the synthesis of the second messenger.*
- *G-proteins function as molecular switches, regulating many cellular processes, such as activation of intracellular enzymes (protein kinase, phosphorylase), activation of membrane enzymes and channels, and activation of gene transcription.*
- *G- protein-linked receptors form a family, which has evolved from a common ancestor. Most G-proteins are*

membrane bound heterotrimers (α β γ) and exist in an activated state, with high affinity for GTP, and an inactive state, where the molecule prefers GDP.

- *Hydrophilic (lipophobic) hormones such as acetylcholine and many peptides bind to membrane receptor proteins, and the hormone-receptor binding activates the enzyme phospholipase C via active G-protein.*
- *Protein tyrosine kinase activity is abnormally high in certain types of cancer and cellular modifications. This can be caused by growth factors or by a mutation of the tyrosine kinase part of the transmembraneous receptor. Mutations of one gene localised on chromosome 10 can lead to four different syndromes: Multiple endocrine neoplasia, Hirschprung's disease, medullary thyroid carcinoma, and Phaeochromocytoma.*
- *Stroke is commonly caused by thrombo-embolism of the middle cerebral artery.*
- *Simple concussion is defined as a transient loss of consciousness followed by complete recovery. A short period of amnesia is often related to the loss of consciousness. This is a migraine injury, where the duration of the unconsciousness indicates the severity of brain damage.*
- *Brain contusion refers to brain damage with prolonged coma, amnesia and focal signs. Later on such patients often suffer from chronic impairment of higher cerebral functions and hemiparesis.*
- *Post-traumatic epilepsy is frequently caused by head injury with coma following depressed skull fractures, brain contusion or intracranial haematoma. Actually, depressed skull fracture causes a high incidents of post-traumatic epilepsy.*
- *Epidural haematoma is caused by skull fractures traversing a dural sinus in the temper-parietal region, resulting in bleeding into the epidural space.*
- *Subdural haematoma is an accumulation of blood in the subdural space caused by venous bleeding.*
- *Subarachnoid haemorrhage is a spontaneous arterial bleeding into the subarachnoid space, often with an acute clinical picture of acute delirium. The circle of Willis and adjacent vessels is the most frequent site for saccular or berry aneurysms that rupture.*
- *Intracranial mass lesions located supratentorially can compress the brain towards the tentorium as to block the upward flow of CSF and thus its absorption.*
- *Epilepsy is an abnormal paroxystic discharge from cerebral neurones resulting in a condition with clinical consequences. Epileptic seizures are partial or general.*
- *The normal EEG waves are due to synaptic potentials by groups of neurons including pyramidal cells. An epileptic seizure is characterised by high voltage-high frequency discharge from large groups of neurons or from the entire cortex.*
- *Partial or focal seizures can be caused by an epileptic focus anywhere in the cortex. The causes of focal seizures are acquired lesions such as cysts, tumours, scar tissue, infections, and ischaemic lesions. The epileptic discharge causes involuntary muscular contractions on the contralateral side. Foci in the somatosensory cortex produce sensory hallucinations called an epileptic aura.*

Further Reading

- *The Journal of Neuroscience*. Semi-monthly journal published by the Society for Neuroscience, 11 Dupont Circle, NW, Washington DC 20036, USA.
- Hopkins AP (1993) *Clinical Neurology, a Modern Approach.* Oxford University Press, Oxford.
- Sims ACP and DW Owens (1993) *Psychiatry*. 6th edition. London: Bailliere Tindall.

Return to chapter 7

Return to content

A Gaba-receptor In An Inhibitory Synapse

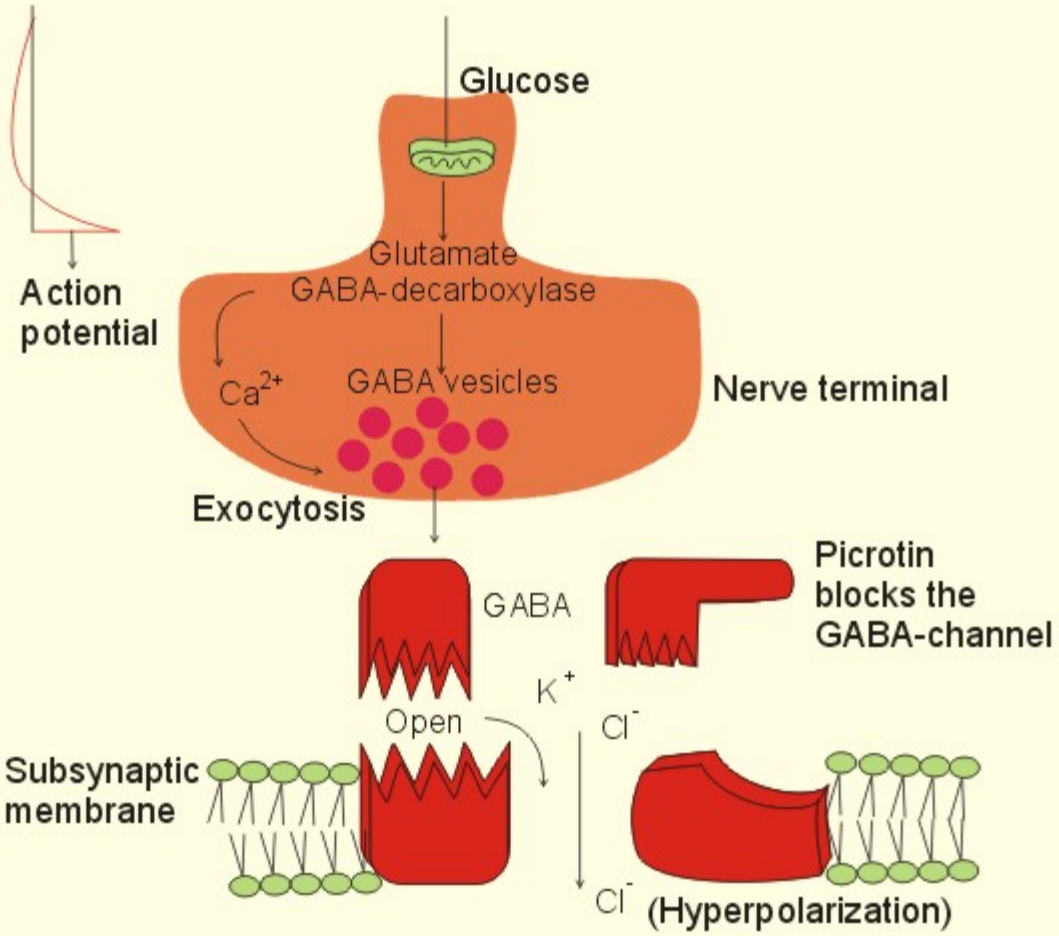
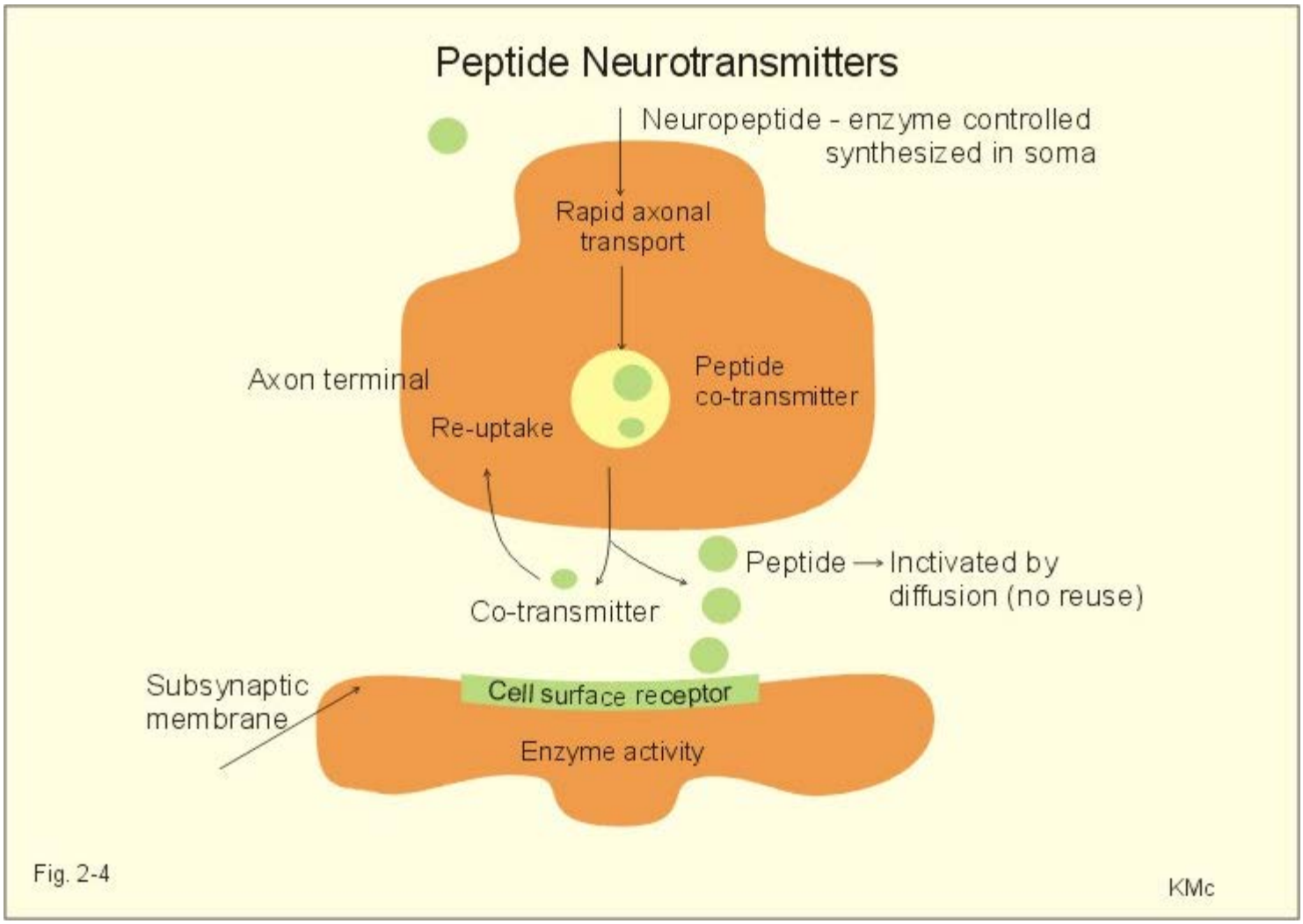


Fig. 2-3

Kmc


Peptide Neurotransmitters
Neuropeptide - enzyme controlled synthesized in soma
Rapid axonal transport
Axon terminal
Peptide co-transmitter
Re-uptake
Peptide → Inctivated by diffusion (no reuse)
Co-transmitter
Subsynaptic membrane
Cell surface receptor
Enzyme activity
Fig. 2-4
KMc

Dopamine Receptors

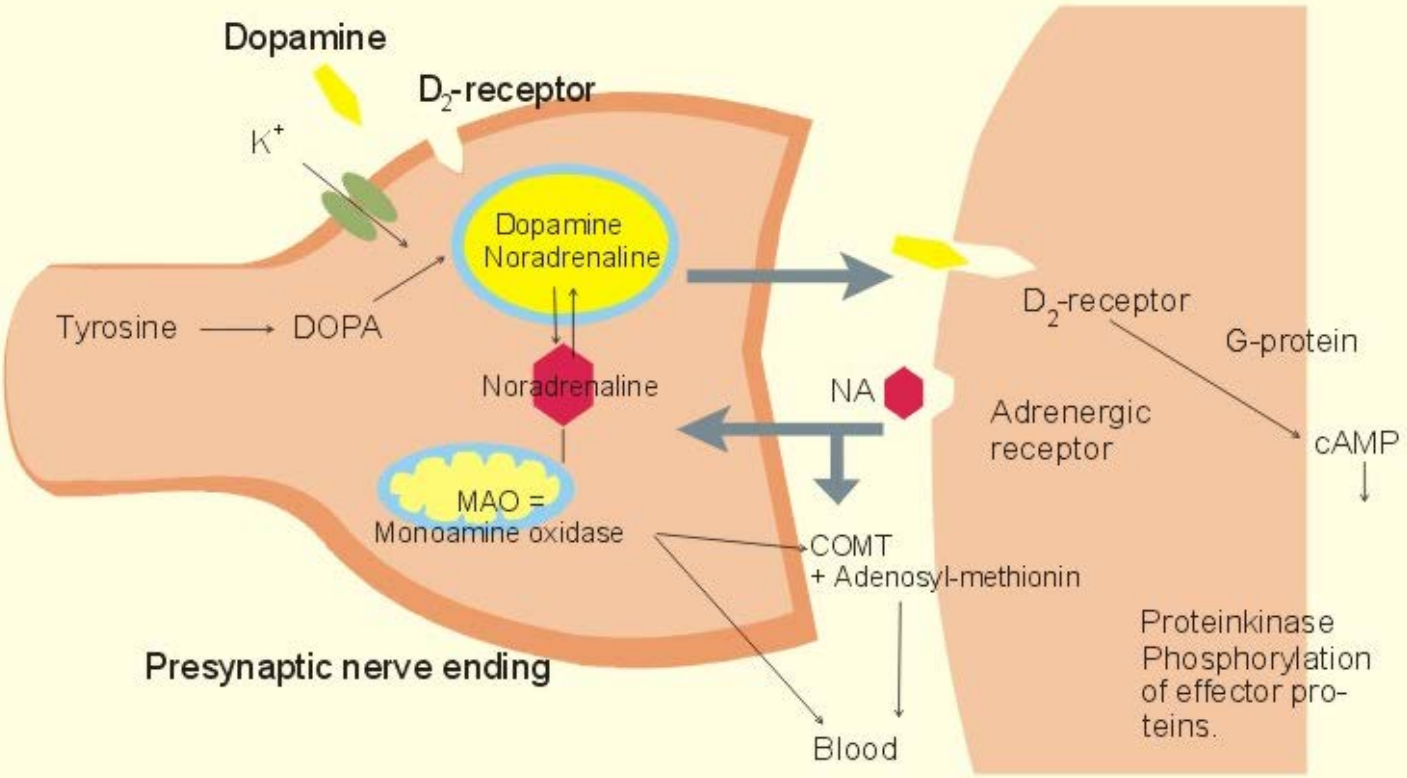


Fig. 2-5

KMc

Chapter 4.

Brain Function, Locomotion And Disorders

Study Objectives

- To *define* akinesia, amnesia, aphasia, arousal, coma, rigidity, a motor unit and three different unit types, habituation and non-associative learning, conditioning, and long-term potentiation.
- To *describe* the primary motor cortex, the corticospinal pathways, damage to the corticospinal pathways, the control of nucleus ruber, the symptoms rigidity and spasticity, nerve conduction velocity, monoaminergic transmission, postural control, neck-and labyrinthine reflexes, the control of voluntary movements, the cerebellar cortex and its pathways, cerebellum and motor learning, damage to the cerebellum, damage to the basal ganglia, cause and therapy of Parkinson´s syndrome, the main functions of the brain lobi and the hippocampus with effects of typical lesions, synaptic plasticity in brain growth and brain damage.
- To *draw* a model of the basal ganglia with pathways, a muscle tendon with efferent and afferent pathways, and a model of recurrent inhibition of motor neurons.
- To *explain* the components in a reflex arch, the muscle tendon, the Golgi tendon organ, the flexor reflex, the crossed extensor reflex, reciprocal innervation and inhibition, alpha-gamma-coactivation, the effect of gamma-efferents on muscle length, the effects of a spinal cross sectional lesion, the orientation reflex, exteroceptive and proprioceptive reflexes, and muscular force. To explain the electromyogram, autonomic movements, damage to the capsula interna, damage to the pyramidal system, the techtospinal pathways, nucleus raphe, locus coeruleus, the EEG during different conditions, sensory and motor aphasia, and hemispheric dominance.

- To *use* the concepts in problem solving and case histories.

Principles

The functional unit of the nervous system is the neurone with its cell body, dendrites and axon, which terminates in a synapse.

Action potentials passing down the axon release chemical neurotransmitters at the synapse.

Definitions

- **Agnosia** refers to lack og ability to recognise and interpret a sensory stimulus. Agnosia is related to a lesion of the sensory cortex.
- **Akinesia** or *hypokinesia* or *bradykinesia* means inability to initiate normal movements. Akinesia is a typically finding in Parkinsons disease.
- **Anterograde amnesia** is a lack of ability to learn anything new. This is a consequence of bilateral removal or damage of the hippocampi.
- **Aphasia** is a condition with disorders of the language function. Lesions of the left hemisphere produce deficits in the language function of most people.
- **Apraxia** refers to lack of ability to perform certain practical actions (unbutton a jacket etc) – often found with parietal-lobe syndromes. The apraxia of gait (failure of skilled walking) is due to frontal-lobe disease.
- **Arousal** is a high level of consciousness also called alertness.
- **Ataxia** refers to uncoordinated movements in particular found as ataxic gait in cerebellar disorders.
- **Athetosis** refers to slow, serpentine, writhing involitional movements of the hands or of most of the body. Athetosis is seen following neonatal insults (cerebral palsy).
- **Coma** is an unresponsive state of unconsciousness from which the patient cannot be awakened even with the most vigorous stimuli.
- **Chorea** refers to rapid involitional hyperkinesia with jerky movements of the limbs. Chorea is found in Huntingtons disease.
- **Dominant hemisphere** is the hemisphere that *controls the expressed language*. Lesions of the left hemisphere produce deficits in language function of most people. These deficits are called aphasia.

- **Electroencephalogram** (EEG) is a recording of a *rhythmic electrical activity* from the surface of the skull. In humans, the EEG is recorded from a grid of standard leads.
- **Flexor withdrawal reflex** is a nociceptive (pain) reflex involving all muscles of a limb in flexor withdrawal. This is an attempt to protect the limb from further damage. The reflex can activate extensor muscles of the opposite limb (the crossed extension reflex).
- **Habituation** refers to a gradual diminution of the response to a repeated stimulus without behavioural consequences.
- **Hemiballism** refers to violent swinging movements of one side of the body almost simulating the throwing of a ball. Hemiballism is caused by lesion of the contralateral nucleus subthalamicus.
- **Learning** is a change of behavior caused by neural mechanisms affected by experience.
- **Long-term memory** (*long term potentiation*) is a prolonged storage and retrieval of new information.
- **Memory** refers to the neural storage mechanisms for experiences.
- **Non-associative learning** means that the learning is unassociated to the stimuli.
- **Orientation reflex** is a fundamental change of behaviour, where the eyes, head and body are turned toward an alarming external stimulus.
- **Prosophenosia** refers to the inability to recognize faces following extensive damage of both occipital and temporal lobes.
- **RAS** is an abbreviation for a large region of the reticular formation of the brainstem termed the *reticular activating system* (RAS). Stimulation of this system causes arousal and an *arousal reaction* in the electroencephalogram.
- **Retrograde amnesia** refers to a condition, where the patient cannot recall information from the *memory*.
- **Rigidity** is a clinical condition with muscle stiffness caused by a high tonus level in the alpha-motor units of limb muscles. The muscle resistance is increased towards slow, passive movements of the limb and it is equal in opposing groups of muscles. This condition is called *lead-pipe rigidity* and it is found in Parkinson´s disease.
- **Sensitisation** is the opposite of habituation. The increased response upon repetition of a stimulus has important behavioural consequences in order to avoid the threat.
- **Sensory aphasia** refers to damage of the Wernicke area with difficulties in understanding written or spoken language, although single words can be heard.
- **Sopor** or *clouding of consciousness* is a term for reduced wakefulness.
- **Spasticity** is a clinical phenomenon following lesion of one pyramidal tract. Loss of the inhibitory effect of the corticospinal pathway increases the spinal reflex activity of the gamma-loop. The muscle tone is increased towards rapid, passive movements of the limbs resulting in a sudden *clasp-knife effect*. Stroke, spinal cord lesions, neonatal insults (cerebral palsy) or multiple sclerosis causes spasticity.
- **Stupor** is a sleepy state from which the patient can be aroused by vigorous stimuli.
- **Tone.** Skeletal muscle tone is a low level contractile activity in some motor units driven by reflex arcs from muscle receptors. Normally, the muscles feels relaxed and flaccid during passive movements of the limbs. – Increased muscle tone is called hypertonia. Hypertonia is found as spasticity in cerebral palsy and as rigidity in Parkinsons disease. Low muscle tone is called hypotonia and found in cerebellar disorders.
- **Tremor.** Rest tremor with pill-rolling movements of the fingers is found in Parkinsons disease. Intention tremor or action tremor is characteristic of cerebellar lesions.

Essentials

This paragraph deals with *1. The cortex and the reticular activating system, 2. Higher brain functions, 3. The limbic system and the hippocampus, 4. Spinal organization of the motor control, 5. Descending motor pathways, 6. Motor control by the brain*.

1. The Cortex and reticular activating system

Six cell layers are recognized in typical regions of the cerebral *neocortex* and they are numbered layers I-VI (Fig. 4-1). The neocortex first appears with the mammals, and its structure is phylogenetically younger than the *allocortex*.

All six layers contain glial cells and more or less of the three typical neurons: Stellate cells, pyramidal cells and fusiform cells. The superficial layers receive and process information, whereas the deep layers are the sites of origin of

most cortical efferents.

The six neocortical layers are as follows:

I. The molecular layer contains numerous dendrites, axons and axon terminals almost without cell bodies.
II. The external pyramidal layer contains mainly densely packed stellate cells, which are GABAergic (inhibitory) interneurons.
III. The external pyramidal layer contains small pyramidal cells. Pyramidal cells use excitatory amino acids (aspartate, glutamate) as transmitters. Layers I, II, III connect adjacent cortical regions and integrate cortical functions.
IV. The internal granular layer resembles layer II with many stellate cells. Most of the sensory signals project to layer IV.
V. The internal pyramidal layer resembles layer III. The pyramidal cell bodies increase in size inwards.
VI. The multiform layer consists of long spindle-shaped or fusiform cells arranged perpendicular to the cortical surface.
VII. The perpendicular collections of neurons, axons and dendrites in the cortical areas form the so-called cortical columns.

Fig. 4-1: The cerebral neocortex with pyramidal, stellate and fusiform cells.

The pyramidal and fusiform cells of layers V and VI provide the output from the cortex. The pyramidal cells have long axons passing to other cortical regions to the brain stem and to the spinal cord.

Thalamocortical afferents mainly project to layers I, IV, and VI, whereas corticothalamic projections have their origin in pyramidal cells in layers V and VI. These connections form a reverberating thalamocortical system, which excite the cortex and contribute to the patterns of the electroencephalogram (EEG).

A large region of the reticular formation of the brainstem is termed the ascending reticular activating system (RAS), which determines our state of consciousness, by its connection with the thalamocortical system (Fig. 4-2). The RAS transmits facilitatory signals to the thalamus. The thalamus excites specific regions of the cortex, and the cortex then excites the thalamus in a reverberating circuit with fast acetylcholine and long-lasting neuropeptides as transmitters. Such a positive feedback loop is what wakes us up in the morning. External stimuli and internal factors (inhibitory interneurons) serve to create a balance of different activity levels during the day. RAS maintain the ascending thalamic activity, but also a certain descending activity level in our antigravity muscles and reflexes. An inhibitory region in the medulla can inhibit RAS and thus both its ascending and descending activity.

A high level of consciousness is called arousal or alertness, which is recognized in the EEG as a high frequency-low voltage shift (see below). An orientation reflex, a fundamental change of behaviour following an external stimulus, often accompanies arousal. The eyes, head and body are turned toward the external stimulus.

Impaired consciousness is caused by malfunction of the neurons in the RAS, and the impairment has at least three levels. Sopor or clouding of consciousness is a term for reduced wakefulness, stupor is a sleepy state from which the patient can be aroused by vigorous stimuli, and coma is a unresponsive state of unconsciousness from which the patient cannot be awakened even with the most vigorous stimuli. Brain stem compression at the mesencephalon leads to coma and death.

Fig. 4-2: The RAS and the thalamocortical system.

The electroencephalogram (EEG) is a rhythmic electrical activity recorded from the surface of the skull. In humans, the EEG is recorded from a grid of standard leads (Fig. 4-2). During neurosurgery the electrical activity is recorded from the surface of the cortex as an electrocorticogram. In normal adult persons, the dominating frequencies are 8-13 Hz (a-rhythm) over the parietal and occipital lobes, as long as the subject is awake and relaxed with his eyes closed. With open eyes, the EEG becomes desynchronized with low amplitude (10 mV) and the dominant frequency increases to 50 Hz. The theta- (3-7 Hz) and delta rhythms (0.5-2 Hz) are observed during light and deep sleep, respectively.

A thalamocortical rhythm produces coordinated extracellular currents, when the brain is not exposed to external stimuli (Fig. 4-2). The EEG recording is due to large synaptic potentials by whole groups of mainly pyramidal cells. The EEG pattern is desynchronized by sensory inputs through the thalamus. The level of alertness (in RAS) also modifies the EEG pattern.

Each pyramidal cell - as with each Purkinje cell in the cerebellum - posses an extraordinarily large number of synapses (10^6). The potentials recorded on the surface of the skull are 50-100 mV. The large pyramidal cells form a dipole with one pole directed toward the surface of the cortex, and the other toward the white matter.

When an external stimulus evokes an EEG change, the change is termed a cortical evoked potential. The large numbers of synaptic potentials in the cortical region are summated to form an evoked potential recorded on the skull by an electrode placed over the associated cortical area. However, evoked potentials are small, so measurement requires repeated stimulation and signal averaging. The evoked potentials over the auditory, visual, and somatosensory cortex (areas I and II) are used clinically to assess the integrity of the respective sensory pathway.

Circadian periodicities are changes in biological variables that occur daily. The circadian controller is the so-called biological clock, probably located in the suprachiasmatic nucleus of the hypothalamus. The biological clock receives many projections from sense organs including projections from the retina signaling light and darkness. These signals are transmitted further to the pineal gland according to one hypothesis. Darkness probably stimulates melatonin secretion by the pineal gland, which inhibits the secretion of gonadotropic hormones from the anterior pituitary, and thus reduces sexual drive. Melatonin secretion decreases with age.

Destruction of the biological clock disrupts many biological rhythms, such as oscillations in body temperature, other vegetative functions and the sleep-wake cycle.

The astronomic 24-hour cycle is shorter than the biological sleep-wake cycle (normally 25 hours). When flying east the astronomic cycle is shortened further acerbating the discrepancy between the two cycles. This increases the problems of adjusting the circadian systems, which often require a week to regain their normal phase relation to the biological clock. Problems caused by changes of biological rhythm are summarized in the term jet lag. Melatonin is used clinically to reduce the jet lag.

The endogenous circadian periodicity of the sleep-wake cycle is normally 25 hours - see above. Sleep is divided into four stages based on EEG. The relaxed individual with eyes closed has 8-13 Hz a-rhythm. As he falls asleep, he passes through the four stages of sleep. During these stages the muscles are relaxed, all vital functions are decreased, and the gastrointestinal motility is increased.

Stage 1 is light sleep, where a-rhythm is interspersed with theta rhythm. Stage 2 is somewhat deeper sleep dominated by slow waves and by sleep spindles (periodic spindle-shaped bursts of a-rhythm) and by large, irregular K-complexes. Stage 3 is characterized by delta waves and by occasional sleep spindles. Stage 4 is recognized by the very slow delta waves with frequencies around 0.5-1 Hz. The subject is difficult to wake up.

A different form of sleep with complete loss of muscle tone occurs periodically every 90-min during stage 1 sleep. This is termed rapid eye movement sleep or REM sleep. Eye movement artifacts and a desynchronized EEG (low voltage, fast activity as in the arousal reaction when awake), is characteristic for REM sleep. The subject is difficult to wake up, so the condition is therefore also termed paradoxical sleep.

Fig. 4-3. Differences in sleeping pattern between three age groups.

Spontaneous erection occurs during REM sleep, and an irregular heart rate and respiration are often observed. Dreams occurring during REM sleep are often recalled by the person when awake.

Children and young adults have all 4 stages of sleep and several periods of REM sleep (Fig. 4-3). The depth of the non-REM sleep diminishes through the night and the REM periods increase in duration (Fig. 4-3).

Stage 4 sleep disappears with age, and stage 3 sleep decreases in duration (Fig. 4-3). The REM sleep is also reduced, and wake periods occur in increasing number. This is why elderly people believe that they do not sleep sufficiently.

The passive theory of sleep claim sleep to be caused by reduced activity in RAS. However, transecting the brainstem in the midpontile region produces an animal that never goes to sleep. Stimulation at the nucleus of the solitary tract can induce sleep, suggesting that sleep be an active process related to centres below the midpontile level.

The question is difficult to address. An educated guess is that sleep is an active, energy saving condition, preferable to most animals. The metabolic rate during sleep falls to 75% of the basal metabolic rate.

2. Higher brain functions

Each hemisphere consists of the following four lobes: the frontal - occipital - parietal and temporal lobes.

The frontal lobe, located in front of sulcus centralis (central fissure), is involved in motor behaviour. The frontal lobe contains the primary motor (area 4), the premotor (area 6), and the supplementary motor areas (frontal eye areas 8 and 9 of Fig. 4-4). These cortex areas are responsible for planning and execution of voluntary movements.

The motor speech areas (44 and 45 or Broca's area) are located close to the motor cortex, on the inferior frontal gyrus of the dominant hemisphere in humans (the left hemisphere is controlling the expressed language in most people). Lesions here cause motor aphasia (difficulties with speech and writing). Patients with lesion of Broca's area (in the dominant hemisphere) frequently suffer from paralysis of the opposite side (right) of the body.

Fig. 4-4. The human cerebral cortex of the left hemisphere controls the expressed language.

The frontal cortex is also involved in personality and emotional behaviour - including attention, intellectual and social behaviour.

The occipital lobe is located behind the parietal and temporal lobe, and involved in visual processing and visual perception. Adjustments for near vision are controlled by the primary visual cortex in area 17 and in the cortex around the calcarine fissure occipital lobe. The conscious visual perception takes place in the primary visual cortex. The secondary visual cortex is in area 18 and 19, where visual impressions are compared, interpreted and stored (Fig. 4-4).

The important primary somatosensory area I is located on the postcentral gyrus (area 1,2 and 3 in Fig. 4-4). There is a distinct spatial representation of the different areas of the body in the postcentral gyrus (the sensory homunculus). The secondary somatosensory area II is located in the rostral part of area 40, close to the postcentral gyrus (Fig. 4-2). The somatic association or interpretation areas (areas 5 and 7) are located in the parietal cortex just behind the somatosensory area I (Fig. 4-2).

Each side of the cortex receives information exclusively from the opposite side of the body.

Widespread damage to the somatosensory area I causes loss of sensory judgement including the shapes of objects (astereognosis).

Auditory and vestibular signals are processed and perceived by the superior temporal gyrus (area 41 in Fig. 4-2). Area 42 is the secondary auditory centre, where auditive signals are interpreted and stored.

The medial temporal gyrus helps control emotional behavior in the limbic system and all the functions of the autonomic nervous system.

Signals from the auditive (area 42), visual (areas 18 and 19) and somatic (areas 7 and 40) interpretative areas are integrated in the posterior part of the superior temporal gyrus. This large gnostic area is specially developed in the dominant hemisphere, where it is called the general interpretative or language comprehension area (Wernicke's area). Damage in Wernicke's area causes sensory aphasia (i.e., difficulties in understanding written or spoken language, although single words can be heard).

Learning processes

Learning is a change of behavior caused by neural mechanisms affected by experience. Memory refers to neural storage mechanisms for experiences. The hippocampus is involved in learning and memory.

1. Non-associative learning means that the learning is unassociated to the stimuli.

Habituation refers to a gradual diminution of a response by repetition of a stimulus, because experience show that the stimulus is unimportant. Sensitization is the opposite of habituation. Firstly, a strong threatening stimulus triggers a certain response, but repetitions of the stimulus increase the size of the response in order to avoid the threat. This evaluation is called the reward and punishment hypothesis. The neural processes are probably related to the function of the hippocampus.

In the snail aplysia a facilitating interneuron releases serotonin onto the presynaptic terminal of a neuron. This stimulates adenylcyclase and the formation of intracellular cAMP in the presynaptic terminal. The resulting protein kinase activation causes phosphorylation and blockage of K^+-outflux. The K^+-outflux is necessary for recovery from the action potential. Lack of K^+-outflux prolongs the presynaptic action potential considerably. This causes a prolonged Ca^{2+}-influx into the presynaptic terminal with increased release of neurotransmitter and facilitated synaptic transmission.

2. Associative learning is the process of learning by associations between stimuli. The free radical nitric oxide (NO) modulates learning.

Conditioning refers to a neural process of associative learning, where there is a temporal association (optimum 0.5 s) between a neutral stimulus (eg, a sound before food) and an unconditioned stimulus (food) that elicits a response (gastro-intestinal secretion). Repetition of the sound-food manoeuvre develops into a conditioned reflex, where the sound alone elicits salivary secretion.

In operant conditioning the response is associated with reinforcement, which changes the probability of the response. Positive and negative reinforcement increases the probability of the response, whereas punishment reduces its probability. Learning is highly improved by happiness. Light stress is an advantage in learning something new. However, substantial stress is not helpfull in the recall process, and stress can completely block the memory.

Strategic behavior is the basis for our social life. Strategic or motivated behavior is related to homeostasis in general (defence, reproduction, temperature and appetite control). Previously, strategic behavior was explained by negative feedback with the purpose as a fixpoint, and with the human brain playing a minor role. Today it is generally accepted that the cerebral drive is a dominant determinant for strategic or motivated behavior. The drive that arouses individuals from inactivity originates in the limbic system (including the hypothalamus), that is acting in close relation to the thalamus and the cerebral cortex. The limbic system is connected to the autonomic control functions of the brainstem reticular formation by the medial forebrain bundle. These vital functions are thermocontrol, appetite control and sexual behavior (Chapter 6).

The dominant hemisphere is the hemisphere that controls the expressed language. Lesions of the left hemisphere produce deficits in language function of most people. These deficits are called aphasia. The left planum temporale in the floor of the lateral fissure of Sylvii is larger than that of the right hemisphere in most people - not only right-handed. The right hemisphere is dominant for functions related to language (intonation, body language), and to mathematically related functions. Each hemisphere controls the contralateral side of the body.

Information between the two hemispheres is transferred through the anterior commissure and the corpus callosum. The language centres on the left hemisphere cannot influence the right hemisphere unless the corpus callosum is intact. The two hemispheres can operate relatively independently with language. One hemisphere can express itself through spoken language. The other communicates non-verbally.

If an animal with intact corpus callosum and optic chiasm learns a visual discrimination task with one eye closed, the

task can still be performed with the untrained eye alone, even when the optic chiasm is transected before the animal is trained. Therefore, visual information is transferred as long as the corpus callosum is intact.

Surgical transection of the corpus callosum has been performed to prevent epilepsy from spreading. When such a patient fixate his vision on a point on a screen, it is possible to stimulate only one hemisphere by showing an object to one side of the visual field. Similar objects (key, ring, nail, fork etc) can be manipulated (but not seen) through an opening below the screen. Healthy persons can locate the correct object with either hand. Split-brain patients, with the picture of the object transferred to the right hemisphere, can locate the correct object with the left hand (ie, right hemisphere), not with their otherwise preferred right hand.

Jigsaw puzzles are solved with such manipulo-spatial capabilities. Right-handed patients with split brain can solve three-dimensional puzzles, if the visual signals can reach the motor cortex for the hand to explore. The visual and motor cortex are connected to each other only in the same hemisphere, when the corpus callosum is cut.

Memory research has characterized three temporal stages in human memory processes.

- 1. An immediate memory holds sensory information for a few hundred milliseconds to seconds for analysis and further processing. The immediate memory is erased by new incoming signals, so we can only remember a few new telephone numbers at a time. Accumulation of Ca^{2+} in the presynaptic terminals with each signal possibly causes prolonged release of neurotransmitter at the synapse (synaptic potentiation).

- 2. The short-term memory is covering seconds to a few minutes, and the short-term memory receives selected information from the immediate memory. Information is erased as new items displace old data. If a person sees a rapid succession of slides, it is the last slide that remains in the short-term memory. We store recent events in the short-term memory, by a neural activity with improved synaptic efficacy that lasts for seconds to minutes. The improved synaptic efficacy is possibly due to synaptic potentiation, presynaptic facilitation, or impulses circulating in neuronal circuits for a restricted period.

- 3. The long-term memory is a large and permanent memory. The long-term memory receives information from the immediate and the short-term memory. Recycling of information through the short-term memory is termed rehearsal. The likelihood of a successful storage in the long-term memory increases with the number of cycles. When the long-term memory is searched for a certain information, it may take minutes to recall the memory. The long-term memory is subdivided into the intermediate long-term memory, which lasts for days or weeks and can be disrupted, and the long lasting long-term memory, which lasts for years.

- 4 The long lasting long-term memory is the storage in the brain of highly overlearned information as one's own name and address. This memory is difficult to disrupt, and it is seldomly affected in retrograde amnesia (see below).

- The long-term memory and consolidation of memory relate to effector protein synthesis at the synapses. Electron microscopy suggests an increased number of vesicular release sites in the presynaptic terminals.

Retrograde amnesia is a term used for a condition where the patient cannot recall information from the immediate and short-term memory. The mild form of retrograde amnesia is typical following head lesion with loss of consciousness (cerebral commotion). The short-term memories have only been rehearsed a few times and probably stored only discretely.

The long term-memories are widespread in the cortex as structurally maintained modifications of the synapses after many rehearsals. Only in severe cases is the long-term memory involved.

3. The limbic system, the hippocampus and emotions

The limbic system is the neuronal network that controls emotional and motivational behavior. Motivational behavior include control of vegetative functions such as body temperature, respiration, circulation, osmolality of body fluids, sexual behavior, smell, thirst, appetite and body weight.

Hypothalamus constitutes the major part of the limbic system, and is located in the middle of the other limbic elements.

Fig. 4-5. The limbic system. The corpus callosum is transected, and we are looking at the medial aspects of the right hemisphere.

The limbic cortex begins in the frontal lobe as the orbitofrontal cortex, extends upward as the subcallosal gyrus, over the corpus callosum and into the cingulate gyrus (Fig. 4-5). The limbic cortex finally passes caudal to the corpus callosum down towards the hippocampus, para-hippocampal gyrus and uncus at the medial surface of the temporal lobe (Fig. 4-5). The fornix connects the hippocampus to the mamillary body. The mamillothalamic tract connects the mamillary body to the anterior nucleus of the thalamus. Thalamus connects to the cingulate gyrus, and its cortex is associated with the hippocampus. Stria terminalis connects the amygdaloid body to the midbrain septum and to the mamillary body (Fig. 4-5).

The limbic paleocortex links the subcortical limbic structures to the neocortex. Hereby, the limbic system relates behavior and emotions to the intellectual cortex functions.

Another important pathway is the medial forbrain bundle, which connects the limbic system to the autonomic control functions of the brainstem reticular formation.

The hippocampus connects with the cerebral cortex, the midbrain septum, the hypothalamus, the amygdaloid and the mamillary bodies and acts both as a store and a recall centre (Fig. 4-5). The hippocampus is the decision-maker, determining the importance of incoming signals. Hippocampus becomes habituated to indifferent signals, but learns from signals that cause either reward (pleasure) or punishment. Hippocampus is the "brain librarian" (helps the cortex to store new signals into the long lasting long-term memory). The signal molecule, nitric oxide (NO), modulates aspartate responses related to hippocampal long-term potentiation.

Bilateral removal of the hippocampi in epileptic patients permanently disrupts the ability to learn anything new (anterograde amnesia). Other lesions of the hippocampi reduce previously learned memory material (retrograde amnesia - see above). Long-term alterations imply a rise in the number of synapses. Cholinergic synapses in the midbrain septum are essential to our memory, and these neurons are dependent upon the nervous growth factor. Repeated activation of a sensory pathway increases the reaction of pyramidal cells. Such a reaction may last for weeks in the hippocampus and be involved in storage and retrieval of new information in the long-term memory.

Our memory (cortex and hippocampus) works as a filter. Perhaps only 1 per mille of all received signals contain useful or emotional information and are catched in the memory. Unfortunately, we are unreliable witnesses, because we invent emotional "information" concerning a factual experience. The easiest facts to remember are those that make sense. All facts, concepts and acquired skills are stored in a ready-to-use fashion. Feelings play a large role in memory, and strong impressions that are charged with emotion etch themselves into our memory.

A recollection is split up into numerous subunits in different regions of the brain. Later, all subunits are brought together by the hippocampus into a complete memory (eg, a certain smell act as a strong clue to a clear memory from way back). One individuals recollection of a particular incident can trigger off anothers, whereby new associations can be created.

4. Spinal organization of motor control

Motor activity can be voluntary or involuntary. Voluntary movements are planned and started by feedforward control, and when maintained for a while they are regulated by feedback loops. Involuntary movements comprise reflexes, such as the stretch reflexes, and autonomic functions, such as the respiratory muscle movements. We have motor centres in the cerebral cortex, the brainstem, the spinal cord, the cerebellum, and the basal ganglia. Motor centres all receive sensory information in an organized neural structure termed a somatotopic map (see the motor homunculus).

We have 200 different skeletal muscles, which are controlled by more than 300 000 motor units.

A motor unit is comprised of a a-motor neuron, all its axon terminals, and the skeletal muscle fibres it innervates. The number of muscle fibres in a motor unit varies from 2 in highly regulated eye muscles (entirely red fibres) to 2000 in the quadriceps femoris muscle. The motor unit is the final common pathway, because all muscle fibres of the unit contract when a motor unit is activated. Adjacent motor units interdigitate, so they can support each other. The muscle power is increased by recruitment of more motor units and by increased frequency of discharge in each unit.

We have three types of motor units (a-motor neurons) in a mixed muscle such as the gastrocnemius. The three types of motor units are characterized in Chapter 2, Box 2-2.

The myotatic stretch reflex

A spinal reflex is a stereotyped motor reaction to an input signal. The myotatic stretch reflex is the most crucial monosynaptic reflex for the maintenance of the erect body posture in humans.

Fig. 4-6: The phasic myotatic stretch reflex and reciprocal innervation (F-).

The reflex has two components. Firstly, the primary annulospiral endings (group Ia) of the muscle spindles trigger the phasic stretch reflex. Secondly, both primary and secondary endings elicit the tonic stretch reflex.

- 1. The phasic stretch reflex is elicited in the clinic by a light tap on a muscle tendon. When the patellar tendon from the quadriceps muscle is stretched quickly by the tap, a discharge is elicited in the afferent fibres (Ia) from the primary endings of the muscle spindle (Fig. 4-6). This is the phasic myotatic stretch reflex or the so-called patellar reflex. These Ia fibres synapse directly (monosynaptically) on a-motor neurons that supply the extensor muscles of the knee (E+ in Fig. 4-6). The response elicited is a brief contraction of the latter. Of all the presynaptic terminals arriving to the motor neuron up to 90% are located on the surface of the dendrites. The remaining 10% synapse on the soma of the motor neuron.

- The Ia afferent fibres also synapse with small group Ia inhibitory interneurons in the grey matter of the spinal cord, as the one synapsing with the upper a-motor neuron in Fig. 4-6. This neuron innervates the semitendinosus muscle, which flexes the knee joint (F- in Fig. 4-6). The reflex inhibition of antagonist muscles when synergistic muscles are contracted is called reciprocal innervation. In pathologic conditions, the phasic stretch reflexes may be depressed or hyperirritable.

- 2. Passive bending of a joint triggers the tonic stretch reflex. This elicits a discharge in both groups Ia and II afferents from the muscle spindle. The tonic stretch reflex contributes to the erect body posture and helps maintain posture by increasing the tone of the physiologic extensor muscles (ie, antigravity muscles).

Renshaw inhibition and presynaptic inhibition

Renshaw inhibition. Cajal found that the a-motor axons give off thin recurrent (antidromal) collaterals in the grey matter of the spinal cord (Fig. 4-6). These collaterals synapse with Renshaw interneurons in the ventral horn (Fig. 4-6). The Renshaw cells synapse with a-motor neurons of synergistic muscles, and thus inhibit monosynaptic reflexes (postsynaptic inhibition). Stimulation of each a-motor unit inhibits adjacent motor units (ie, recurrent inhibition). This is also called the principle of lateral inhibition, whereby the motor response is confined to selected units only.

Descending signals from the brain can either amplify the postsynaptic inhibition or reduce its effect. Renshaw cells make it possible for the higher brain centres to influence spinal reflexes by central inhibition or facilitation.

Presynaptic inhibition. Presynaptic terminals contain a large number of voltage-gated Ca^{2+}-channels. Ca^{2+} must enter the presynaptic terminal from the extracellular space before the vesicles can release their neurotransmitter at the synapse. Presynaptic inhibition takes place at presynaptic contact sites on the presynaptic terminals. Activation of these sites closes many Ca^{2+}-channels, and thus inhibits transmitter release.

The Golgi tendon organ

The Golgi tendon organs are the serially located terminals of group Ib fibres wrapped around bundles of collagen fibres in the tendons. Golgi tendon organs monitor the force in the tendon; they are activated either by stretch or by contraction of the muscle. The adequate stimulus is the force developed in the tendon.

The inverse stretch reflex or the Golgi tendon reflex completes the stretch reflex by a force-controlling feedback. The Golgi tendon organs monitor force in the tendons. Golgi tendon organs are in series with the muscle fibres - not parallel as the muscle spindles. If the extensor muscles of the thigh are fatigued, as during standing, the force in their tendons begins to decrease. This reduces the discharge of the Golgi tendon organs. This acts as a compensating feedback, which excites the a-motor neurons and increases the force of contraction. The inverse stretch reflex helps maintain the force of muscular contraction and posture during standing. During the rapid contraction of the myotatic stretch reflex, the inverse stretch reflex reduces the force of contraction. The stretch reflexes regulate the length of the muscle, and provide a length-force feedback to the CNS.

The muscle spindle

The muscle spindle monitors muscle length and rate of change of length (velocity); they are particularly abundant in muscles that are capable of fine movements and in large muscles that are dominated by slow twitch fibres. The organ is shaped like a spindle, which lies in parallel to the large, regular, extrafusal muscle fibres. Each organ contains two main types of intrafusal muscle fibres: Nuclear bag fibres which swell in the equatorial region due to all the nuclei located here, and thin nuclear chain fibres which have central nuclei arranged in line (Fig. 4-7). The primary afferent fibres (Ia) twine around the equatorial regions of both the bag and chain fibres like a corkscrew or annulospiral; the annulospiral nerve endings signal length and velocity. The secondary afferent fibres originate mainly from the nuclear chain fibres and with a few branches originating from the nuclear bag fibres (Fig. 4-7). They monitor only the length of the muscle.

Two types of g-motor neurons innervate the muscle spindle. The dynamic g-motor axons form plate endings (P_2) on the nuclear bag fibres, while static g-motor axons form creeping trail endings on nuclear chain fibres (Fig. 4-7). The intrafusal fibres receive a Ab-motor fibre, which terminates with P_1 plate endings on both extra- and intrafusal muscle fibres (Fig. 4-7). The Ab-motor fibres may be involved in a-g-coactivation.

When the extrafusal fibres contract, the muscle spindles shorten, whereby the discharge rate of their afferents decreases.

Fig. 4-7: The structure of a muscle spindle with a bag and a chain fibre.

Activity of the g-motor neurons causes the polar spindle regions to contract on either end. This elongates the equatorial regions so that muscle spindles can adjust to stretch (Fig. 4-7).

Descending commands from the brain often cause contraction of both extrafusal and intrafusal fibres simultaneously so that the muscle spindle is sensitive to stretch at all muscle lengths. When the muscle is stretched, the muscle spindles are simultaneously stretched with it, and the discharge rate of the afferents is increased.

The flexion reflexes

The flexion reflexes are triggered by various flexion reflex afferents including nociceptors. The flexion reflexes have a long latency, because it involves polysynaptic interneurons. The afferent discharge causes excitatory interneurons to activate a-motor neurons that innervate ipsilateral flexor muscles. The afferent discharge also causes inhibitory interneurons to inhibit a-motor neurons, supplying the ipsilateral extensor antagonists.

The flexor withdrawal reflex is crucial. This reflex is also called a nociceptive reflex or a pain reflex, and involves all the muscles of a limb in flexor withdrawal in order to protect from further damage. In addition, the reflex can activate the extensor muscles of the opposite limb. This contralateral activity is termed the crossed extension reflex by

reciprocal innervation.

The locomotor pattern generator controls flexion reflexes involved in locomotion.

Severe visceral disease can trigger contraction of the chest and abdominal muscles, which reduces pain by limiting movement of the body. When examining the abdomen of such a patient it will be observed that the muscles are tense. This sign is called defence musculaire, which is a viscero-somatic protective reflex.

Coordination of limb movements

We possess pattern generators or neural circuits in the spinal cord, for every limb and for respiration, chewing etc. The midbrain locomotor centre, via the reticular formation and through the reticulospinal tracts, organizes the commands. Such spinal pattern generators also account for other movement patterns like scratching, dancing etc.

5. Descending motor pathways

Clinical dichotomy traditionally subdivides the descending fibres into the pyramidal and the extrapyramidal pathways; this is based on the fact that the corticospinal tract passes through the medullary pyramids. Therefore, interruption of the corticospinal or pyramidal tract was supposed to cause pyramidal tract disease (see later). The problem, however, is that the loss of the corticospinal tract does not explain all the classical signs of pyramidal tract disease.

The concept of extrapyramidal pathways raises other problems. The concept of extrapyramidal tract diseases is generally used to designate one or more disorders of the basal ganglia. While, extrapyramidal pathways do play a role in basal ganglia diseases (as in cerebellar disease), the main motor pathway involved in basal ganglia diseases is the corticospinal tract!

The descending motor pathways can also be dichotomized based on their endpoint in the spinal cord, and hence which muscles they control and how. Pathways ending in the lateral horn of the spinal cord (on motor neurons or interneurons) are called the lateral descending motor system (the rubrospinal tract and the lateral corticospinal tract). Pathways ending on the medial ventral horn interneurons are termed the medial descending motor system (containing reticulo-, tecto-, and ventriculo-spinal tracts).

The lateral corticospinal, the corticobulbar (to the facial motor and hypoglossal nucleus) and the rubrospinal tracts control the manipulative movements of the limbs and the lower face and tongue muscles. The corticospinal and corticobulbar tracts originate from areas 4, 6, 8, 9, and somatosensory area I (areas 1, 2, 3 in Fig. 4-4). The large and small pyramidal cells and the giant pyramidal cells of Betz are the cells of origin of these tracts. The corticospinal tract descends through the internal capsule and brainstem. At the medullary pyramid 80% of the fibres cross to the opposite side and descend in the dorsal lateral funiculus as the lateral corticospinal tract. The fibres of this tract end on motor neurons and interneurons in the lateral horn of the spinal cord. These motor neurons innervate distal muscle groups. Interruption of the lateral corticospinal tract implies loss of the fine control of the digits. Interruption of the corticobulbar tract to the facial motor and hypoglossal nucleus implies loss of voluntary movements of the lower face and tongue. Interruption of the rubrospinal tract from the red nucleus combined with corticospinal lesions give rise to difficulty in separating finger, hand and arm movements. The red nucleus is closely linked to the deep cerebellar nuclei.

The lateral or dorsolateral descending system allows the primary motor cortex to modify the reflexes and pattern movements at the level of the spinal cord.

The medial or ventromedial descending system involves the ventral corticospinal tract and much of the corticobulbar tract ending in the medial group of brainstem and spinal cord interneurons. The ventral corticospinal tract continues caudally in the ventral funiculus on the same side and ends bilaterally on the medial interneurons. They control the axial muscles and bilateral activity including chewing and wrinkling of the eyebrows.

Other medial system pathways originate in the brainstem:

- 1. The lateral vestibulospinal tract excites motor neurons that innervate proximal postural muscles. It receives input from all compartments of the vestibular apparatus and from cerebellum to the lateral vestibular nucleus.

- 2. The medial vestibular tract receives signals from the semicircular ducts and from cerebellum, and excites motor neurons in cervical and thoracic segments. Thus, it controls the head position in response to angular accelerations of the head.

- 3. The pontine reticulospinal tract excites motor neurons to the proximal extensor muscles to support posture.

- 4. The medullary reticulospinal tracts have mainly inhibitory effects on many spinal reflexes.

- 5. The tectospinal tract from the superior colliculus causes contralateral movements of the head in response to touch and auditory stimuli. This tract allows the integration of hearing and vision with motor performance.

- 6. Pathways from the solitary nucleus and the interstitial nucleus of Cajal are involved in the pharyngeal stage of swallowing. The solitary nucleus receives all sensory signals from the mouth including taste, and is involved in cardiovascular and respiratory control.

The ventromedial system is important for the normal muscle tone and body posture.

Monoaminergic descending pathways

- 1. The neurons of the pontine locus coeruleus and nucleus subcoeruleus contain nor-adrenaline (NA). These nuclei project to and inhibit interneurons and motor neurons of the spinal cord through the lateral funiculi.

- 2. The neurons of the raphe nuclei in the medulla, which are connected to the limbic system also, contain serotonin. The serotonergic nuclei project to and inhibit dorsal horn interneurons reducing pain transmission, and they also project to and excite ventral horn motor neurons of the spinal cord, thereby enhancing motor activity.

- 3. There is also a descending dopamine pathway.

The three monoaminergic pathways function as motor system amplifiers.

6. Motor control by the brain

The primary motor cortex (area 4) on the precentral gyrus controls distal muscles of the extremities. Area 4 is organized parallel to the somatosensory cortex. The face is represented laterally near the Sylvian lateral fissure, and the legs on the medial part of the hemisphere. The cortical representation is somatotopic and disharmonic, as indicated by the motor homunculus.

The premotor cortex helps control proximal and axial muscles.

The supplementary motor cortex is involved in motor planning and in coordination of movements. The frontal eye fields initiate saccadic eye movements.

Corticospinal neurons discharge before voluntary muscle contraction, and the size of the discharge is related to the size of the contractile force. The somatosensory cortex and the posterior parietal association cortex receive feedback from the sensory neurons system, which helps correct motor feed-forward commands.

The role of the cerebellum

The little brain, also termed the motor autopilot, helps regulate movements and posture, influences muscle tone, eye movements and balance.

Cerebellum is particularly concerned about the timing of rapid muscular activities including the interplay between agonist and antagonist muscle groups. Motor learning is programmed in the cerebellum. Cerebellum compares the

proprioceptive input from the actual movements, with the movements intended by the motor control areas of the brain. Cerebellum controls the sequence of movements, and makes corrective adjustments just like an autopilot.

The cerebellar cortex is characteristically folded and consists of three phylogenetically different structures related to three afferent pathways (inputs). The large neocerebellum in higher mammals is also called the pontocerebellum and consists of the hemispheres and vermis caudal to the primary fissure. The paleocerebellum or spinocerebellum consists of vermis of the anterior lobe, pyramis, uvula and paraflocculus. The small archicerebellum or vestibulocerebellum is simply the flocculonodular lobe.

Three important outputs from the cerebellum also divide it into three functional units. The vermis of the cerebellar cortex projects to the fastigial nucleus, the pars intermedia to the globose and emboliform nuclei, and finally the hemisphere, which projects to the large dentate nucleus (Fig. 4-8).

Fig. 4-8: Neuronal connections between the cerebellar cortex and the deep cerebellar nuclei.

The cerebellar cortex is build up of three layers. The superficial molecular layer with axons, dendrites and many synapses, the Purkinje-cell layer and the granular layer (Fig. 4-8). The small granule cells send their axons into the molecular layer, where they divide and send so-called *parallel fibres* in each direction along the folium. These fibres excite the dendrites of the Purkinje and the Golgi cells. The Golgi cells inhibit the granule cells by feedback inhibition. Stellate and basket cells are interneurons that inhibit dendrites and cell bodies of the Purkinje-cells, respectively. Each Purkinje cell is stimulated from a climbing fibre, which projects from the inferior olive. All neurons with cell bodies in the cerebellar cortex are inhibitory except for the granule cells. The cerebellar cortex modulates the activity of the deep cerebellar nuclei.

The incoming pathways to the cerebellum end as *mossy fibres* on the granule cells. Each mossy fibre reach many granule cells. The input signals through the mossy fibres evoke *simple spikes* (single action potentials) in Purkinje-cells. The climbing fibres produce repetitive or complex discharges in Purkinje cells. *Complex spikes* of long duration and low frequency are involved in the cerebellar programming of motor learning. The Purkinje-cell axons terminate in the deep cerebellar nuclei or in the lateral vestibular nucleus.

This is the basis for *cerebellar coordination* and fine, rapid adjustments of complex movements. The cerebellar hemisphere affects movements on the same side of the body, because of its crossed connection to the motor system. The motor system projects contralaterally.

Discrete electrical stimulation of cerebellum does not cause movements or sense impressions, so it is also termed the *silent brain*.

The vestibulocerebellum projects to the vestibulospinal and reticulospinal tracts, which coordinate balance and eye movements. The vestibulo-ocular reflex produces conjugate eye movements in the direction opposite to that of the head movement. The vestibulo-collic reflex increases the neck muscle tone damping the induced movement.

The spinocerebellum receives proprioceptive input from the spinal cord (the spinocerebellar tracts). The spinocerebellum controls the axial muscles through the medial descending motor system, and the proximal limb muscles through the rubrospinal tract of the dorsolateral system.

The pontocerebellum receives decision signals and motor control signals from the cerebral cortex by way of pontine nuclei. The pontocerebellum is involved in motor planning, and controls the distal limb muscles through the lateral corticospinal tract.

The basal ganglia

The main function of the basal ganglia is to initiate and stop movements. The basal ganglia inhibit the thalamus, and thus reduce the thalamic stimulation of the motor *Cortex*.

Fig. 4-9: The basal ganglia and their interplay. Transmitter stimulation is marked by +, and inhibition by -. The

affected cell bodies or axons at disease states are marked with a bar.

The basal ganglia also contribute to cognitive (i.e., intelligence, knowledge, and motor learning) and affective (i.e., emotional) functions.

The basal ganglia include the globus pallidus and striatum. Striatum consists of the nucleus caudatus and the putamen. These deep brain nuclei function in collaboration with several thalamic nuclei, substantia nigra and the subthalamic nucleus (Fig. 4-9).

The striatum receives afferent fibres from the cortex (Glutamate + = glutaminergic excitatory fibres), and dopaminergic (inhibitory) fibres from substantia nigra (Dopamine -). Striatum projects to the globus pallidus and to the substantia nigra. These connections are GABAergic and inhibitory (GABA - in Fig. 4-9). Globus pallidus receives afferent GABAergic fibres from striatum, and projects to the thalamus with GABAergic efferents. In the striatum, there are excitatory cholinergic pathways.

Pathophysiology

This paragraph deals with 1.Pure lesion of the medullary pyramid, 2.Abnormal muscle tone, 3. Spinal transection syndrome, and 4. Cerebellar disease. -

Capsular stroke, Parkinson´s disease, dyskinesias and epilepsy are all dealt with in Chapter 7, which is a systematic description of neurological and psychiatric disorders. Read chapter 7 before trying to solve the case histories.

1. Pure lesion of the medullary pyramid

The control of fractionated finger movements is absent. There is a positive sign of Babinski. Flexion reflexes are not found, and neither is spasticity. On the contrary, muscle tone is decreased. In summary, a pure interruption of the corticospinal tract alone does not show the same signs as capsular stroke.

The main deficits caused by medial lesions are reduced muscle tone in the physiologic extensors, loss of balance during walking and standing, and loss of rightening reflexes (they tend to restore head and body position). However, fine finger movements are quite normal.

2. Abnormal muscle tone

Spasticity is used in clinical neurology to describe muscles resisting fast, passive movements of the limbs, especially in extreme articular positions. When the limbs are moved in extreme articular positions, the increased muscle resistance suddenly disappears. Spasticity includes hyperactive stress reflexes and foot clonus. The resistance dominates in the physiological extensors (antigravity muscles). Spasticity is typical for stroke, where the capsula interna is damaged. The resulting disruption of the lateral descending system is extended by damage of other cortical efferents to the basal ganglia, the thalamus and pons (see Chapter 7).

Rigidity is muscle stiffness caused by prolonged activity in the motor units. The muscle resistance is increased towards passive movements of the limbs in any direction (lead pipe rigidity). This condition is found in Parkinson's disease (see Chapter 7).

3. Spinal transection syndrome

The spinal shock is immediately recognized by several characteristic symptoms: flaccid paralysis with loss of stretch reflexes, areflexia, loss of autonomic functions, and of all sensation below the level of transection. After a few weeks the spinal shock fades away and the reflexes return and become hyperactive (foot clonus), including mass reflexes and flexion reflexes. A spastic paralysis or paresis replaces the flaccid paralysis.

4. Cerebellar disease

Cerebellum can suffer from damages at two locations: 1. Damage to the flocculonodular lobe causes nystagmus and difficulties in gait and balance (i.e., resembling lesion of the vestibular apparatus). 2. Damage to the vermis or the intermediate region and hemisphere, results in motor disturbances of the trunk and limbs, respectively.

Cerebellar disorders include cerebellar incoordination, dysequilibrium, and loss of muscle tone.

Cerebellar incoordination comprises ataxic gait, as seen in alcohol intoxication and in disseminated sclerosis. Another type of ataxia is dysmetria, where there is an inability to move the limbs to the desired position. Many patients manifest their ataxia as dysdiadochokinesis, which is a disturbance of the normal ability to make repeated supinations and pronations of the lower arms. Complicated muscle function is stepwise - not smooth. Intention tremor is seen when the patient is asked to touch a target. Speech is slow and slurred, a defect termed dysarthria or scanning speech.

Dysequilibrium results in balance problems, and the patient falls to the affected side. Gyratoric vertigo is a genuine rotational or merry-go-round vertigo with the associated loss of equilibrium. This cerebellar vertigo is similar to that following lesion of the vestibular apparatus.

Loss of muscle tone is called hypotonia. The hypotonic lack of damping causes the leg to swing back and forth, when the patellar reflex is triggered - so-called pendular knee jerk.

Cerebellar nystagmus is involuntary movements of the eyeballs around their natural position - often accompanied by rotational vertigo, when the flocculonodular lobe is damaged.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. The frontal cortex is involved in motor and emotional behavior.

B. The somatic association or interpretation areas (area 5 and 7) are located in the temporal cortex.

C. Recycling of information through the primary memory is termed rehearsal.

D. Retrograde amnesia following brain commotion is a loss of the short-term memory.

E. The limbic system relates behavior and emotions to the intellectual cortex functions.

II. Each of the following five statements have False/True options:

A. The EEG arousal reaction is a low frequency-high voltage shift.

B. Circadian periodicities are changes in biological variables occurring once a day.

C. N-methyl-D-aspartate-(NMDA)-receptors bind aspartate, dopamine and glutamate.

D. Dreams occur during REM sleep, and the person always reproduces them when awake.

E. Dominating EEG frequencies of 8-25 Hz are characteristic of light sleep.

III. Each of the following five statements have True/False options:

A. Fast fatigable motor units consist of type IIB twitch fibres with few mitochondria and small amounts of myoglobin.

B. The Renshaw cells synapse with a-motor neurons of antagonistic muscles, and thus inhibit

monosynaptic reflexes.

C. The cerebellar hemisphere affects movements on the opposite side of the body.

D. The Golgi tendon organs are the serially located terminals of group Ib fibres wrapped around bundles of collagen fibres in the tendons.

E. The ventromedial descending system involves the ventral corticospinal tract and much of the corticobulbar tract ending in the medial group of brainstem and spinal cord interneurons.

4. Case History A

An outstanding Russian composer, 63 years of age, recovered from a cerebral insult. However, he could no longer understand spoken or written language, although his speech was fluent. The composer also maintained his ability to compose excellent music.

- *1. What is the name of this deficit in language function?*
- *2. Where in the brain is the lesion* ***localized and in what side of the brain?***

4. Case History B

A male of 65 years suddenly falls and is found in deep coma by the doctor. There is a left-sided hemiplegia with short arm-long leg as a flexion reflex. The paralysis and areflexia turns into spastic hemiparesis with a positive sign of Babinski. The deep stretch reflexes (patellar- and Achilles-tendon reflexes) are enhanced. There is loss of superficial reflexes (the abdominal and cremasteric reflexes). When the Achilles-tendon reflex is triggered it releases foot clonus. When the patient is awake from coma his facial nerve paresis is examined. He can knit his brows and turn his eyes upwards.

- *1. What is the pathophysiologic basis for this condition?*
- *2. What are spasticity and foot clonus?*
- *3. Is the facial nerve paresis central or peripheral?*

Try to solve the problems before looking up the answers .

Highlights

- *The reticular activating system (RAS) transmits facilitatory signals to the thalamus. The thalamus excites the cortex, and the cortex then excites the thalamus in a reverberating circuit. Such a positive feedback loop is what wakes us up in the morning. During the day external stimuli and internal factors including inhibitory interneurons balance the different activity levels.*
- *Impaired consciousness is caused by malfunction of the neurons in the RAS, and the impairment has at least three levels. Sopor or clouding of consciousness is a term for reduced wakefulness, stupor is a sleepy state from which the patient can be aroused by vigorous stimuli, and coma is a unresponsive state of unconsciousness from which the patient cannot be awakened even with the most vigorous stimuli.*
- *Circadian periodicities are changes in biological variables that occur daily. The circadian controller is the so-called biological clock, probably located in the suprachiasmatic nucleus of the hypothalamus. The biological clock receives many projections from sense organs including projections from the retina signaling light and darkness.*
- *Children and young adults have all 4 stages of sleep and several periods of REM sleep each night. The depth of the non-REM sleep diminishes through the night and the REM periods increase in duration.*
- *The motor speech areas (44 and 45 or Broca's area) are located close to the motor cortex, on the inferior frontal gyrus of the dominant hemisphere in humans (the left hemisphere is controlling the expressed language in most*

people). Lesions here cause motor aphasia (difficulties with speech and writing). Patients with lesion of Broca's area (in the dominant hemisphere) frequently suffer from paralysis of the opposite side (right) of the body.

- *The medial temporal gyrus helps control emotional behaviour in the limbic system and all the functions of the autonomic nervous system.*
- *The hippocampus is involved in learning and long lasting long-term memory. This is what makes hippocampus the decision-maker.*
- *Motor centres all receive sensory information in an organized neural structure termed a somatotopic map (motor homunculus).*
- *The motor unit is the final common pathway, because all muscle fibres of the unit contract, when a motor unit is activated. Adjacent motor units interdigitate, so they can support each other. The muscle power is increased by recruitment of more motor units and by increased signal frequency in each unit.*
- *Renshaw cells make it possible for the higher brain centres to inhibit or facilitate spinal reflexes.*
- *Cerebellum or little brain is also termed the motor autopilot, because it helps regulate movements and posture, and influences muscle tone, eye movements and balance.*
- *Cerebellar disorders include cerebellar incoordination, dysequilibrium, and loss of muscle tone.*
- *The main function of the basal ganglia is to initiate and stop movements. Disorders of the basal ganglia, such as lack of dopamine in substantia nigra, result in a clinical syndrome with rigidity, hand tremor, and akinesia (Parkinson´s disease).*

Further Reading

Schuman, E.M., and D.V. Madison. "Nitric oxide and synaptic function." *Annu. Rev. Neurosci.* 17: 153-183, 1994.

Thomson, R.F. "The brain. A Neuroscience Primer." *Second Edition*. Freeman, 1993.

McIntosh, A.R., C.L. Grady, L.G. Ungerleider, J.V. Haxby, Rapoport, S.I., and B. Horwitz. "Network analysis of cortical visual pathways mapped with position emission tomography." *J. Neurosci.* 14 (2): 655-666, 1994.

Return to top

Return to content

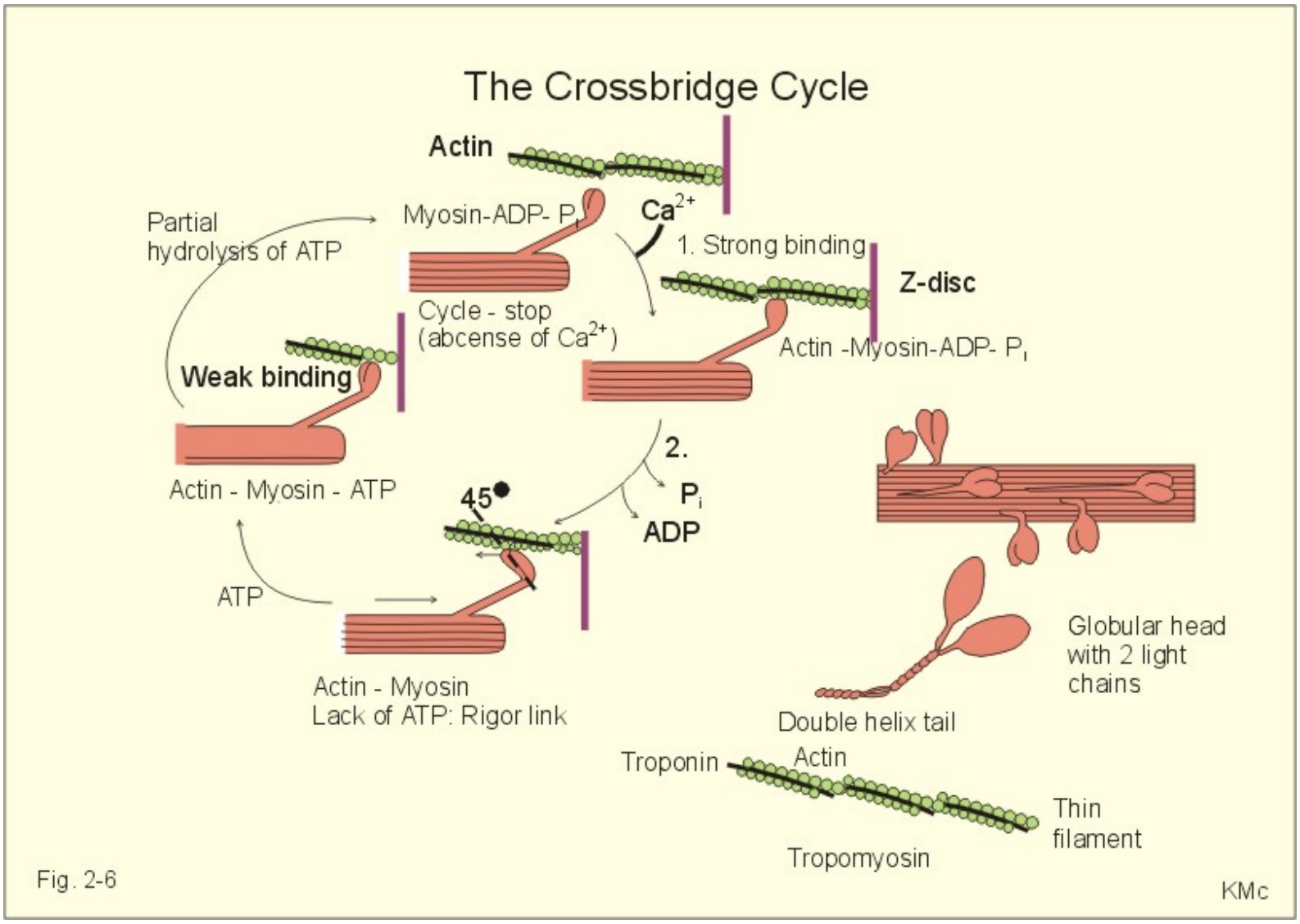

The Crossbridge Cycle
Actin
Myosin-ADP- P_i
Ca^{2+}
1. Strong binding
Z-disc
Partial
hydrolysis of ATP
Cycle - stop
(abcense of Ca^{2+})
Actin -Myosin-ADP- P_i
Weak binding
Actin - Myosin - ATP
2.
P_i
ADP
45
ATP
Actin - Myosin
Lack of ATP: Rigor link
Globular head
with 2 light
chains
Double helix tail
Troponin
Actin
Thin
filament
Tropomyosin
KMc

Fig. 2-6

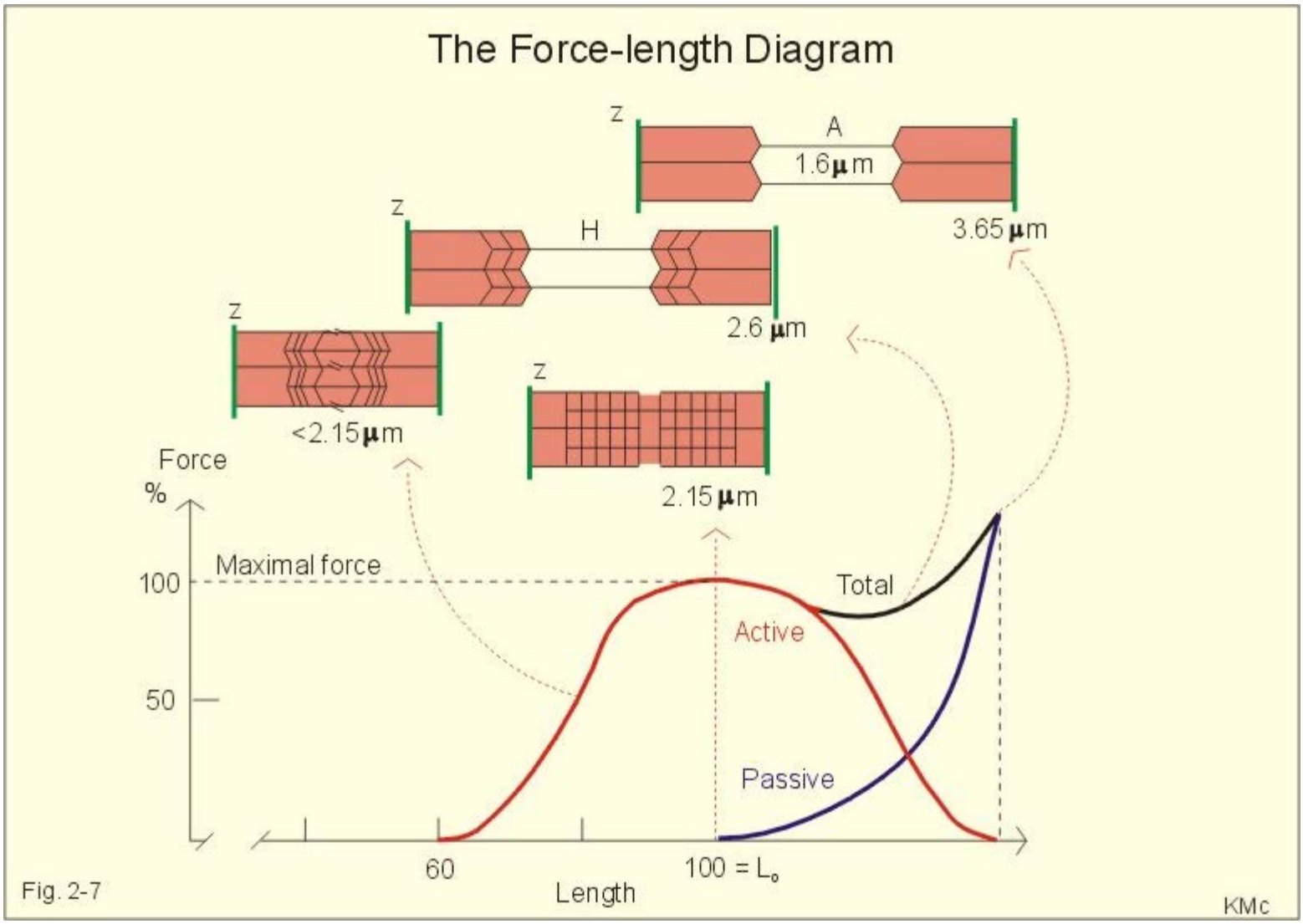
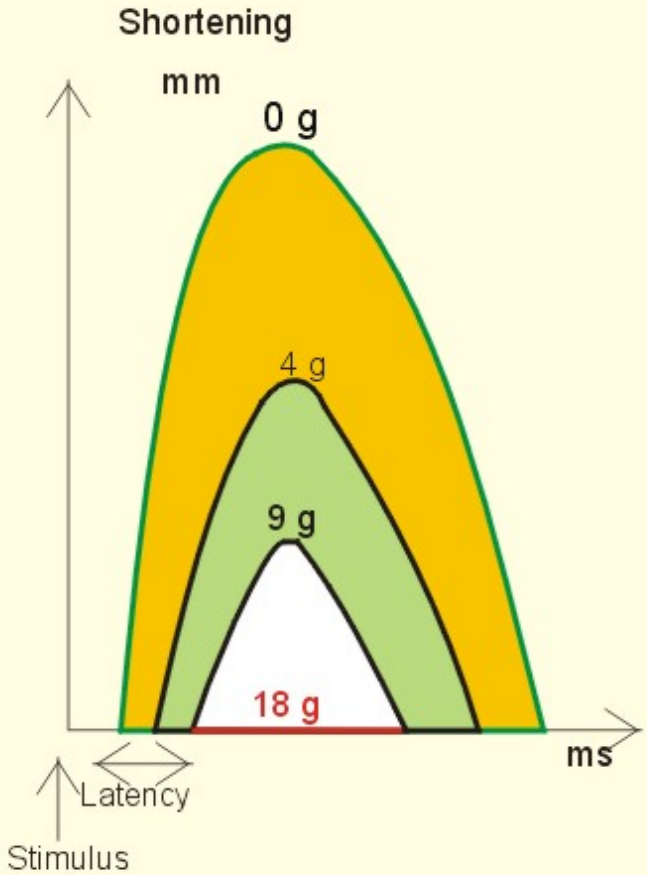

The Force-length Diagram
Z
A
1.6 µm
3.65 µm
H
2.6 µm
<2.15 µm
2.15 µm
Force
%
Maximal force
100
50
Total
Active
Passive
60
100 = L₀
Length
KMc

Fig. 2-7

Shortening Curves And The Force-velocity Diagram



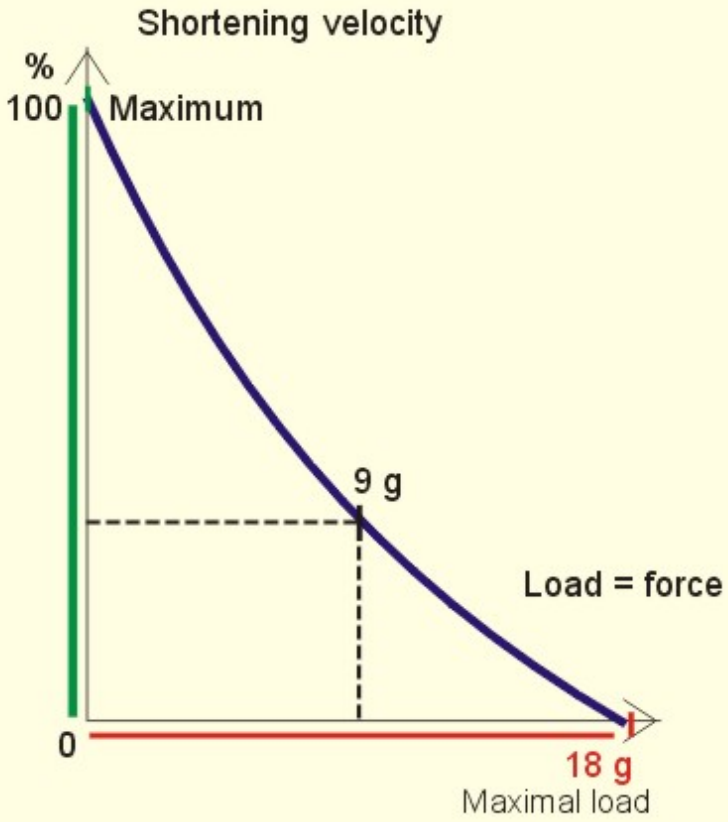


Fig. 2-8

KMc

Chapter 18.

Exercise, Sports And Doping

Study Objectives

- To *define* factors of importance to oxygen uptake, cardiac output, and ventilation during exercise.
- To *describe* the rise in ventilation, oxygen uptake and cardiac output during increasing exercise intensity, the concepts anaerobic threshold, oxygen deficiency and oxygen debt.
- To *calculate* the relationship between the major variables.
- To *explain* the metabolism and limits of exercise, typical sport injuries, doping, the effects of training including health consequences, and hypotheses of the cardiopulmonary regulation.
- To *use* the concepts in problem solving and case histories.

Principles

- *The human body has a redundancy of overlapping cardio-pulmonary control systems during exercise.*
- *The redundancy-hypothesis, with neural factors dominating at the start of work and peripheral feedback control during steady state, is a possible explanation of the hyperpnoea of exercise and the related increase in cardiovascular activity .*

Definitions

- **Anaerobic threshold.** This is the exercise level above which the energy requirements can be satisfied only by the combined aerobic metabolism and anaerobic glycolysis. Lactic acid is produced and stimulates the peripheral chemoreceptors. Hereby, ventilation starts to increase out of proportion to the rise in oxygen uptake.
- **Blood doping.** Blood boosting is an artificial improvement of performance through an increase in the haemoglobin binding capacity. Blood doping (one litre) definitely improves the oxygen transport with the blood and also the maximal oxygen uptake, which is beneficial to distance runners.
- **Doping**: Athletes who use drugs or other means with the intention to improve performance artificially are doped by definition.
- **Endurance capacity** or *fitness number* is given as the *maximal oxygen uptake* in ml of oxygen STPD min^{-1} kg^{-1}.
- **Energy equivalent** of oxygen on a mixed diet is defined as the heat energy liberated in the body per litre of oxygen used (20 kJ of energy per litre at an RQ of 0.8).
- **Flow Units** (FU) measure relative bloodflow as the number of ml of blood passing an organ per 100 g of tissue and per min.
- **Mean Arterial Pressure** (MAP) is the arterial blood pressure measured as the sum of the diastolic pressure plus 1/3 of the pulse pressure (see below).
- **Mechanical efficiency** is the ratio between external work and the total energy used during work.

- **Oxygen debt** is defined as the extra volume of oxygen that is needed to restore all the energetic systems to their normal state after exercise.

- **Oxygen deficiency** is defined as the difference in oxygen volume between an ideal, hypothetical oxygen uptake and the actual uptake in real life. The missing oxygen volume at the initiation of exercise is the oxygen deficit.

- **Pseudo-doping.** Many drugs reputedly increase athletic performance, but the fact remains that such effects rarely show up in double-blind controlled trials. - On the contrary, serious side effects occur with a biologically high and statistically significant frequency.

Essentials

This paragraph deals with *1. Athletes and training, 2. Fitness testing, 3. Limits of exercise performance, 4. The anaerobic threshold, 5. Ventilation and oxygen uptake, 6. Cardiopulmonary control, and 7. Oxygen debt and deficiency.*

1. Athletes and training

At the start of exercise, signals from the brain and from the working muscles bombard the cardiopulmonary control centres in the brainstem. Both cardiac output and ventilation increase, the a-adrenergic vasoconstrictor tone of the muscular arterioles falls abruptly, whereas the vascular resistance increases in inactive tissues. The systolic blood pressure increases, whereas the MAP only rises minimally during dynamic exercise. The total peripheral vascular resistance (*TPVR*) falls during moderate exercise to 0.25-0.3 of the level at rest, because of the massive vasodilatation in the muscular arterioles of almost 35 kg muscle mass. This is why the major portion of cardiac output passes through the skeletal muscles (Fig. 18-1) and why the diastolic pressure often decreases during exercise. The coronary bloodflow increases, and at some intensities of exercise we see increases in the skin bloodflow (Fig. 18-1).

Fig. 18-1: Distribution of cardiac output during exercise. HBF means hepatic bloodflow, and CBF is cerebral bloodflow.

A top athlete increases his cardiac output from 5 to 30-40 l of blood per min, when going from rest to maximal dynamic exercise (Fig. 18-1). However, the muscle bloodflow can rise 25 fold in the total muscle mass. Accordingly, the total muscular oxygen uptake rises 85 fold from rest to maximal exercise (see Box 8-1 with calculations).

Training improves the capacity for oxygen transport to the muscular mitochondria, and improves their ability to use oxygen. After long-term endurance training the athlete typically has a lower resting heart rate, a greater stroke volume, and a lower *TPVR* than before. The maximum oxygen uptake progressively increases with long-term training, and the extraction of oxygen from the blood is increased. The lung diffusion capacity for oxygen probably increases by endurance training. The capillary density of skeletal muscles, the number of mitochondria, the activity of their oxidative enzymes, ATPase activity, lipase activity and myoglobin content all increase with endurance training. Endurance training also produces a rise in ventricular diastolic volume. Strength training (weight lifting) produces a rise in left ventricular wall thickness without any important increase in volume

During dynamic exercise the stroke volume increases as does heart rate, and the residual ventricular volume decreases (Fig. 18-2).

Fig. 18-2: The pressure-volume loop of the left ventricle in a healthy male at rest (red curve) and during dynamic exercise (blue curve).

Although the peak ventricular pressure during systole rises considerably and thus the arterial peak pressure, the diastolic pressure falls because of the massive fall in total peripheral vascular resistance. The contractility of the heart is depicted as the slope of the pressure-volume curve. The contractility increases considerably from rest to exercise (Fig. 18-2).

2. Fitness testing

A simple objective method of estimating the *endurance capacity* or *fitness number* (maximal oxygen uptake, $V°_{O2}$max) in a person, is to measure the heart rate (HR) at a standardised work on a cycle ergometer. The test rests on the assumption, that there is a linear increase in HR with increasing oxygen uptake or work rate (Fig. 18-3). The net mechanical efficiency is relatively constant in each individual (approximately 20%). On a mixed diet the energy equivalent for oxygen is 20 kJ per l (STPD), so it is easy to calculate the volume of oxygen corresponding to any maximal work rate extrapolated from Fig. 18-3.

The test subject wears light clothes, and is not allowed to smoke, eat or work at least three hours before the measurement, which takes place in a comfortable not too warm room.

Fig. 18-3: The relationship between work intensity and steady state heart rate at work. The fitness number is in ml STPD oxygen per min and per kg of body weight.

The work intensity on the ergometer is chosen to produce a heart rate between 130-150 beats per min, and must be continued for at least 5 min in order to secure respiratory steady state. Respiratory steady state means that the pulmonary oxygen uptake is equal to the oxygen uptake of the tissues. This implies that ventilation and heart rate at work is also stable.

As a standard work rate of 100 W is chosen for females, and 150 W is standard for untrained males. An optimally performed fitness test results in a heart frequency of 130-160 beats per min. The submaximal test is designed by P.-O. Åstrand (and is sometimes performed at 2 work rates - Fig. 18-3). Since the rise in HR is linearly correlated to the work rate, a line is drawn through the points (Fig. 18-3). The line is extended until it reaches the horizontal line (maximal HR). Here, the *maximal heart rate* is the mean of the maximal heart rate of persons of the same age and sex.

The rise in mean arterial pressure (MAP) during dynamic exercise is often minimal, because of the arteriolar dilatation with a large, rhythmic bloodflow through the working muscles. The *TPVR* is typically reduced to 1/4 of the value at rest.

In contrast, static exercise often results in a doubling of the MAP, because a large muscle mass is contracting and the contraction is maintained. Static work is typically accomplished with a low cardiac output, so the *TPVR* is relatively high. This is dangerous to elderly people with known or unknown degrees of atherosclerosis.

3. Limits of exercise performance

The limitation of performance (measured as oxygen uptake) depends upon the type of work and upon the person. Several factors are involved. The mass balance principle provides an overview (see Eq. 18-1).

Both the maximum cardiac output and the maximum arterio-venous O_2 content difference are limiting factors. Healthy persons have redundant ventilation and diffusion capacity in their lungs imposing no limitation.

3.1. *Pulmonary ventilation*. *Increasing work rate* (with 15 W more each min) leads to a marked increase in pulmonary ventilation without any ceiling being reached even at maximal oxygen uptake ($V°_{O2}$max). The steeper rise in ventilation is shown by its deviation relative to the thin line towards the right (Fig. 18-4). Light exercise often increases ventilation by an increased tidal volume (V_T). With increasing work rate also the respiratory frequency must rise from 10 towards 50 respiratory cycles per min. The tidal volume can increase to half the value of the vital capacity (6 l), which corresponds to an exercise ventilation of (3*50=) 150 l per min. *At exhaustion* the ventilation is much greater than at the point, where maximal oxygen uptake is already reached. At this maximum many individuals can increase ventilation further voluntarily. The alveolar gas tensions, P_{AO2} and P_{ACO2}, are essentially maintained during most work rates. At maximal work rate the P_{AO2} increases and P_{ACO2} decreases 5-10%. This fact illustrates effective gas exchange or adequate ventilation during non-exhausting exercise. Thus ventilation is not the limiting factor in these healthy persons.

3.2. The oxygen utilisation in the tissues is not a likely limitation in healthy people. A group of skiers increased their

maximal oxygen uptake further, when they started arm work $_{during}$ continued running. Obviously, the *maximal oxygen uptake* measured during running is not always maximal.

3.3 Pulmonary diffusion capacity for oxygen with failure of the lungs to fully oxygenate blood. This is certainly not a limiting factor in healthy persons with a redundant lung $_{diffusion}$ capacity. The arterial blood is fully saturated with oxygen even during the most strenuous exercise at sea level. The lung diffusion capacity (D_{L02}) increases, because the number of open lung capillaries is increased, the surface area increases and the barrier-thickness is reduced. In addition, O_2 transport is boosted further by the rise in cardiac output from 5 to 30 l of blood per min

Oedema or interstitial pulmonary fibrosis leads to thickening of the alveolar-capillary barrier, which will impede O_2 exchange. The reason is that the pulmonary vascular volume is reduced (reduced capillary transit time), and thus the diffusion equilibrium point is moved towards the end of the capillary. If patients with lung diseases try to exercise, this problem is further aggravated by the still more reduced capillary transit time. Thus exercise would impose a significant diffusion limitation on O_2 transfer.

3.4. Cardiac output. Limited transport capacity for oxygen caused by limited peripheral bloodflow is the only logical explanation. Limitations in reducing *TPVR* or in the pumping capacity of the heart could cause the limited muscle bloodflow. When work is maintained at peak cardiac output and maximal oxygen uptake, the blood pressure falls as more vasodilatation occurs and there are no signs of even a slight relative increase in the low *TPVR*. - The major limitation to exercise in well-trained athletes is the heart's pumping capacity in delivering oxygen to the working muscles.

4. The Anaerobic threshold

The *anaerobic threshold* (AT) is the exercise level at which the energy requirements can be satisfied only by the combined aerobic metabolism and anaerobic glycolysis. The lactic acid formed in the muscle cells diffuse into the blood and causes a metabolic acidosis, which stimulates the peripheral chemoreceptors. Hereby, ventilation starts to increase out of proportion to the rise in oxygen uptake (Fig. 18-4).

Just after the AT is passed, the ventilation increases proportional to the increase in carbon dioxide output (ie, so-called *normo-capnic* buffering). Accordingly, ventilation increases linearly with carbon dioxide output but out of proportion with the oxygen uptake. The carbon dioxide output and ventilation will increase faster than oxygen uptake, because bicarbonate react with the lactic acid produced, so CO_2 is liberated, added to the metabolic CO_2 production and eliminated by hyperventilation, causing P_{aCO2} to fall (ie, hyperventilation)

The rise in blood [lactate] is gradual, and Fig. 18-4 does not show any sign of a lactic acid threshold at the anaerobic threshold. Note the total rise in plasma [lactate] of 10 mM, which is equal to the fall in plasma [bicarbonate] from 24 to 14 mM (Fig. 18-4).

Exercise levels above the maximal aerobic capacity is called *supra-maximal work*. Here, the anoxia leading to a metabolic or lactic acidosis contribute with a large ventilatory drive, as shown in the steep component of $\dot{V}_E$ (Fig. 18-4 and 18-5).

Fig. 18-4: Ventilation and arterial blood concentrations (pH, lactate and bicarbonate) at rest and during an incremental work test on a cycle ergometer up to 100%.

Lactate is produced even at light exercise, but only minimal amounts are liberated to the blood (Fig. 18-4). Untrained subjects at any oxygen uptake, have higher ventilation and heart rate than the trained. The AT in untrained persons is often about 50% of maximal oxygen uptake, whereas the AT of athletes approaches 80%. Patients with heart disease increase their blood [lactate] at a minimal activity.

The oxyhaemoglobin dissociation curve is moved progressively to the right as exercise intensity increases due to the

rise in 2,3-diphosphoglycerate (DPG) concentration (Fig. 8-3) and to the rise in temperature.

Above the AT, when oxidative metabolism is high, extra mechanical output is financed by anaerobic energy generation. The end product is lactic acid (Fig. 18-4). The *lactic acidosis* causes a further shift to the right of the oxyhaemoglobin dissociation curve easing oxygen delivery to the mitochondria. Lactate, nitric oxide and adenosine also dilatate muscle vessels and increase the number of open capillaries, thus improving the diffusion of oxygen from capillary blood to the mitochondria.

5. Ventilation and oxygen uptake

Results from an untrained person with a maximal oxygen uptake of 2.7 l STPD min^{-1} (AT: 1.3 l STPD min^{-1}), and from a top athlete with 6 l STPD min^{-1} (AT: 3.6 l STPD min^{-1}) are shown in Fig. 18-5. Several studies have shown oxygen uptake to remain at maximal level despite increasing work rates, and with carbon dioxide output increasing too. These curves also illustrate that ventilation - in these persons -is not the limiting factor for maximal oxygen uptake.

If the athlete is suddenly breathing oxygen instead of atmospheric air, while working at a high level (5-6 l STPD min^{-1}), a drastic fall in ventilation will occur within 30 s. This is not a chemoreceptor response, since there is no stimulus. The oxygen breathing reduces the blood [lactate], but not within 30 s. Oxygen breathing abruptly increases the diffusion gradient and thus the rate of diffusion from haemoglobin to the muscle mitochondria. Exhaustive exercise with a severe metabolic acidosis may cause the steep rise in ventilation without a further rise in oxygen uptake (Fig. 18-5). In this case, oxygen seems to diffuse at a reduced rate from haemoglobin to the muscle mitochondria.

Fig. 18-5: Ventilation and oxygen uptake in an untrained person with a maximum oxygen uptake of 2.7 l per min. Results from a top athlete, with a $V\dot{}_{O2}max$ of 6 l min^{-1} breathing air (•) or oxygen (o) is shown for comparison.

Strenuous exercise is also associated with a rise in plasma concentration of catecholamines, dehydration and a rise in core temperature approaching 41° C. The sensitivity of most receptors is increased in an overheated body. Increased activity of the *arterial chemoreceptors* causes hyperventilation in exercise situations where plasma-K^+ is high and P_{aO2} is dangerously low. The athlete approaches exhaustion and collapse.

6. Cardiopulmonary control

The proportional increase in ventilation and cardiac output with increasing oxygen uptake suggests a common control system. The integrator consists of sensory and motor cortical areas, and the brain stem neighbour-centres for respiratory and cardiovascular control. The link between the respiratory and the circulatory control system is probably established in the neural network of the brain stem centres.

The nucleus of the tractus solitarius is the site of central projection of both chemoreceptors and baroreceptors. The respiratory and the cardiovascular systems are connected during most forms of dynamic exercise (Fig. 18-6), but they can also operate differently. There is a sharp rise in ventilation within the first breath at the on-set of exercise, and cardiac output also increases abruptly (Fig. 18-6). Both variables increase progressively over minutes until a steady state is reached. At the offset of exercise, ventilation and cardiac output falls instantly (Fig. 18-6).

The cardiopulmonary adjustments to exercise comprise an integration of I. neural and II. humoral factors.

I. The neural factors consist of: 1) Signals from the brain, 2) Reflexes originating in the contracting muscles, and 3) the central & peripheral chemoreceptors.

1. Signals from the brain to the active muscles passes the reticular activating system (RAS) in the reticular formation of the medulla, which includes the respiratory (RC) and cardiovascular centres. This signal transfer is called irradiation from the motor cortex to the RC, and proposed as an explanation of the exercise hyperpnoea. The mesencephalon and

hypothalamus are also involved in the Krogh irradiation hypothesis now called central command. Cortical activation of the sympathetic nervous system accelerates the heart, increases myocardial contractility, dilatate the muscular arterioles and contract other vascular beds such as the splancnic region. Speculative mechanisms as irradiation or central command are so-called feedforward hypotheses.

Fig. 18-6: The exercise hyperpnoea and the rise in cardiac output follow the same pattern.

2 Afferent signals from proprioceptors in the active muscles through thin myelinated and unmyelinated fibres in the spinal nerves (type III and small unmyelinated type IV) to RC are the best-documented feedback hypothesis.

3. Central and peripheral chemoreceptors are sensitive to the final product of metabolism, carbon dioxide. The carbon dioxide molecule is most likely the controlled variable, perhaps as P_{aCO2}. The pH, P_{aO2}, and P_{aCO2} are normal during moderate steady state exercise, where the central chemoreceptors dominate. However, during transitions from rest to exercise and during severe exercise the peripheral chemoreceptors are stimulated. Stimulation of peripheral chemoreceptors increases the rate and depth of respiration and causes vasoconstriction.

II. The humoral factors that influence skeletal muscle bloodflow, cardiac output and ventilation are metabolic vasodilatators and hormones. Neural and chemical control mechanisms oppose each other. During muscular activity the local vasodilatators supervene. The local vasodilatators have not been identified. Ischaemic mitochondria in fast oxidative muscle fibres release many vasodilatators such as adenosine, AMP, and ADP. However, it is possible to block many of the neural and humoral factors without disturbing the proportional exercise hyperpnoea and the rise in cardiac output**.** These experiences suggest that the human body have a redundancy of overlapping control systems. The *redundancy-hypothesis*, with neural factors dominating at the start of work and peripheral feedback control during steady state, is a logical compromise.

7. Oxygen debt and deficiency

The *O_2 deficit* is defined as the difference in O_2 volume between an ideal, hypothetical O_2 uptake and an actual uptake as it occurs in real life (see Fig. 18-7). The missing O_2 volume is the oxygen deficit.

The energy demand increases instantaneously at the start of a working period, but the actual O_2 uptake via the lung lags behind for 2 min. The oxygen demand deficit is provided for by the O_2 stores (oxymyoglobin) and by anaerobic energy.

Fig. 18-7: The oxygen deficit and the oxygen debt at exercise.

The *oxygen debt* is defined as the extra volume of O_2 that is needed to restore all the energetic systems to their normal state after exercise (Fig. 18-7). The non-lactic O_2 debt following moderate work is characterised by maintained blood lactate concentration around the normal resting value of 1 mM. The non-lactic debt is maximally 3 l, used for regeneration of the Phosphocreatine and for refilling the O_2 stores. The lactacide O_2 debt following supramaximal work (100-400 m dash) can amount to 20 l and the blood [lactate] to as high as 20-30 mM. This O_2 debt is used for oxidation of 75% of the lactate produced, and for the formation of 25% of the lactate to glycogen in the liver. Restoration of Phosphocreatine etc following activity, is a process referred to as repayment of the O_2 debt. However, it is very uneconomical, since the debt is often twice as high as the O_2 deficit.

Pathophysiology

The pathophysiology of sports is related to the ultimate limits of human performance. Severe exercise for prolonged periods, such as a 20-fold rise in metabolic rate in a marathon runner, sometimes result in life-threatening conditions: *Histotoxic hypoxia* with blockage of ATP production, dehydration, hyperthermia and metabolic acidosis with a pH_a below 6.9.

Following a short paragraph on *1. Muscle fatigue*, two consequences of aggressive attitudes in competitions are dealt with here: *2. Sport injuries* and *3. Doping*. The final point is *4. Fit for life*.

1. Muscle fatigue

Muscular contraction releases a great ionic leak (Na^+ -influx and a K^+-outflux) through the skeletal muscle membrane, which elicits the action potential (Fig. 18-8). Thus the muscle cell loses K^+ and gains Na^+ during intensive exercise. Contraction stimulates the Na^+-K^+-pump acutely, and training increases its activity. Still, at high intensity exercise the ionic leaks can exceed the capacity of the Na^+-K^+-pump for intracellular restoration.

During intensive exercise the Osmolarity of the contracting muscle cells increases together with the capillary hydrostatic pressure. As a consequence, the ECV and plasma volume can fall by 20% within a few min. The plasma [K^+] can rise to 8 mM due to efflux from the contracting muscle cells and from red blood cells into a reduced plasma volume. Training reduces exercise-induced hyperkalaemia.

Muscle fatigue following prolonged muscle activation increases proportional to the performance and to the loss of muscle glycogen. The insufficient and uncoordinated muscle contractions are due to the lack of glycogen and to failing neuromuscular transmission. *Exhaustion* of the stores of neurotransmitters in presynaptic terminals can occur within seconds to minutes of repetitive stimulation. Weight lifting, football dash and 100 m dash use up the phosphagen system within seconds.

Exhaustion often causes a serious drawback in the systematic practice of an athlete. The body stores are totally depleted, and deleterious consequences may occur.

Fig. 18-8: Skeletal muscle cell maintaining homeostasis by the activity of Na^+ -K^+ -pumps.

During exercise the striated muscle cells loose K^+ to the ECV and the blood. The Na^+ -K^+ -pump contains Na^+ -K^+ -ATPases, which are temporarily inefficient in maintaining homeostasis during exercise (Fig. 18-8). The rise in extracellular K^+ is probably related to muscular fatigue and dependent upon the maximal work capacity. Following exercise there is an extremely rapid homeostatic control in healthy well-trained persons. The activity of the Na^+ -K^+ -ATPases seems optimised in well-trained persons - not necessarily the concentration of Na^+ -K^+ -ATPases in skeletal muscle biopsies.

Even minor diseases, such as a common cold, may reduce cardiac output in an endurance athlete, thus causing muscle ischaemia during the usual practise and extreme muscle fatigue. Isolated muscular fatigue is thus due to depletion of ATP stores, whereby the actin-myosin filaments form a fixed binding and develop rigor or cramps. Neuromuscular fatigue is probably caused by progressive depletion of acetylcholine stores during prolonged, high frequency muscular activity.

Fatigue can never be fully explained by a simple rise in plasma-[K^+] only. Many other signals are integrated in the CNS before a person feels fatigued.

Endurance athletics in a hot and humid environment can increase the temperature of the body core to more than 41 ^oC. Such a level is dangerous to the brain and CNS symptoms and signs develop severe fatigue, headache, dizziness, nausea, confusion, staggering gait, unconsciousness, and profuse sweating. When the victim suddenly faints, this is termed heat stroke, which can be fatal.

2. Sport injuries

Five typical categories of sport injuries are considered here.

1. ***Runners*** are almost always damaged when working at a *too high velocity* or high velocity combined with turning or jumping. The force applied to the feet of a 75 kg person while walking is around (Gravity acceleration * body weight)

= (9.807 m s^{-2} * 75 kg) = 750 kg m s^{-2} or 750 Newton. The force applied to the feet while running is 3-4 fold larger

Four typical injuries of runners are shown in Fig. 18-9.

Fig. 18-9:Two athletes showing four frequent leg and foot injuries attended by running.

The typical injuries are 1) muscle fibre lesions (myopathy with tender muscles), 2) tendosynovitis (shin splint) of the tibial posterior muscle, 3) tendinitis or rupture of the Achilles tendon, and 4) subluxation of the peroneus muscle tendon. 5) Dome fractures are osteochondral fractures from the talus with pain during running. This often occurs as a complication after a foot distortion, which does not heal. 6) Stress fractures are consequences of walking long distances but are also found after distance running and basketball.

These injuries occur during activities (athletes, ball players) with acceleration and deceleration by running or jumping in different directions. Quite often, the athlete is damaged following a break in the training. Even a few days of absence are enough. The athlete starts out too rapidly in order to compensate for the break in the training schedule.

2. ***Brain injuries*** *(boxing)* are known from serious accidents during many types of sport - in particular boxing. Even the elegant boxing legend, Muhammad Ali, was seriously injured during a long - although rather successful - carrier.

Acute brain damage or brain contusion includes deeper brain structures with neuronal damage, increased intracranial pressure and brain ischaemia (Fig. 18-10). Head injury during boxing can result in epidural haematoma (cranial fracture with rupture of the middle meningeal artery). The boxer hits the floor, is unconscious, wakes up and appears in good condition. Suddenly, he collapses again, and develops hemiplegia or die. The development of subdural haematoma is insidious venous bleeding sometimes with a latency of weeks between the head injury and the clinical phenomena (Chapter 7). CT scanning confirms the diagnosis. In chronic subdural haematoma there is a slow development of headache, drowsiness, confusion, sensory losses, hemiparesis, stupor and coma.

Fig. 18-10: Professional boxer with typical damages from the carrier.

Incomplete recovery from brain damage impairs higher cerebral function, with damages of locomotion (hemiplegia), and of psychological functions (Fig 18-10). The end result for the so-called punch-drunk boxer is chronic traumatic encephalopathy with dementia, post-traumatic epilepsy and other neurological disorders (Chapter 4).

*3.****Ball play damages***. Cruciate ligament lesions are common from ball play (ie, handball, football, baseball, basket and volleyball).

Basketball players often land on the toe tip from height and eventually develop exostoses. The exostosis hallucis is called basketball toe. The nail is tender and the exostosis has to be surgically removed.

The tibial anterior muscle originates on the tibia and passes to the navicular bone. Tendinitis in the tendon of this muscle leads to oedema, pain and crepitation.

Fig. 18-11: Soccer, baseball and basketball players are shown with typical injuries from the sport.

Baseball finger or mallet finger is an avulsion of an extensor tendon of the finger usually including a small flake of bone (Fig. 18-11).

Foot distortion (distorsio pedis) frequently includes rupture of the talofibular- calcanofibular- and bifurcate ligament or even fracture (Fig. 18-11).

Orthopaedic specialists must handle Malleole and other complicated fractures.

Turf toe is overextension of the basal joint of the large toe - frequently during ball play. In this case the large toe is protected with spica plast.

4. Skiing injuries range from trivial to fatal. The incidence of knee sprains is high, because improvements of binding design seem to be unsuccessful. The ski acting as a moment arm (Fig. 18-12) magnifies external rotation of the knee. Slalom skiing is the type of skiing with most fractures. The medial collateral ligament of the knee often ruptures.

Fig. 18-12: Typical skiing and tennis injuries are shown in a male and a female.

Another common ski injury is the *skiers thumb*. During a fall the ski pole and the wrist strap tend to concentrate forces to extend the thumb at the mid phalangeal joint until the ligaments burst.

5.Tennis injuries are haematoma subungualis (tennis toe) with bleeding under the nail of the big toe. This is a painful condition - not reserved for tennis players only. The haematoma pressure is relieved by puncture through the nail. The so-called tennis fracture is a fracture of the base of the 5.th metatarsal bone (Fig. 18-12)

The tarsal tunnel syndrome is also frequent in tennis players with pains along the medial side of the foot and toes. This involves the tibial posterior nerve in the channel behind the inner Malleole.

Tennis elbow is a painful disease of the aponeurotic fibres through which the common extensor origin is attached to the lateral humerus epicondyle. Tennis players from the strain use the name tennis elbow (Fig. 18-12); only few of the sufferers actually play tennis.

Conclusion:

*The demand of **fast progress** is linked to competitive sports. A better strategy is to practice at a **relaxed level**, until stamina is developed and hard training is tolerated. Relaxed training is often so comfortable that it becomes a lifestyle. Tender muscles are avoided by prewarming, and a careful **muscle stretch program** following exercise.*

3. Doping

Doping derives from the word dope, which means a stimulating drug. Athletes, who use drugs or other means with the intention to improve performance artificially, are doped by definition

The list of forbidden drugs counts more than 3500, and it is still growing.

Pseudo-doping

Many drugs reputedly increase athletic performance, but the fact remains that such effects rarely show up in double-blind controlled trials. On the contrary, serious side effects occur with a biologically high and statistically significant frequency.

Pseudo-doping with Ginseng and a multitude of other extracts and substances is often quite harmless, and - just as many potent drugs - without proven beneficial effect on athletic performance.

Anabolic steroids

Anabolic steroids are used to increase muscle strength in females and in male athletes with a poor natural testosterone production - possibly a pure placebo effect. Compared to placebo in double-blind studies there is no detectable steroid-effect on the maximal oxygen uptake, size of the muscles or erythropoiesis. However, both steroids and placebo improves the mood and motivation, so both groups trained more and were eating more than before.

As an example, the muscles of body builders are extremely large, but not necessarily equally strong (Fig. 18-13). Some side-effects of dope are lesions of muscle fibres, hypogonadism, liver disorders, and psychosocial deroute (see illustration for further information).

The reversible side effects and irreversible sequel are indisputable. Doping addicts have a high risk of cardiovascular diseases (arterial hypertension, atherosclerosis, heart attacks and strokes), muscular disorders, liver disease, and - in males - testicular failure. Both the sperm formation and the testosterone production are suffering, often irreversibly.

Body building is considered to be the most doping related discipline - in particular by the use of anabolic steroids - and the results are often monstrous (Fig. 18-13).

Fig. 18-13: A body builder, a Sumo wrestler and an obese super-heavy weight champion with a world record (235 kg). All have serious health problems.

In wrestling, discos and super-heavy weight lifting the use of anabolic steroids is frequently disclosed.

A previous world record holder in super-heavy weight lifting developed extreme adiposity when increasing his natural body weight from 80 to 183 kg. The use of steroids resulted in muscular lesions and severe psycho-social crises. The adiposity developed into restrictive lung disease and arthrosis in the knees and other articulations. The athlete was actually a patient with a normal thoracic skeleton, but the lungs were compressed by fat accumulation. During his career he developed the Pickwick syndrome (ie, a fat patient with reduced ventilation, somnolence, sleep apnoea, secondary polycythaemia and cyanosis).

The Japanese Sumo wrestlers have the same problems created by the required extreme adiposity, and many excellent wrestlers have obvious difficulties in walking.

Blood doping

Blood boosting is an artificial improvement of performance through an increase in the haemoglobin binding capacity. Blood doping (one litre of the athletes own blood) definitely improves the oxygen transport with the blood and also the maximal oxygen uptake, which is beneficial to distance diciplines.

Approximately 6 weeks before the competition (Olympic Games or World Championship) the athlete deposits 1000 ml of his own blood as separated red blood cells. The haemoglobin binding capacity is regained by maintained training, and a few hours before the competition, he receives a blood transfusion with his own erythrocytes. Of course, a sudden improvement of the maximum oxygen capacity of more than 10 %, is unfair in endurance disciplines (long distance running, cycling, skiing etc), but it may cause viscosity problems and thrombus formation (see below).

High altitude training and erythropoietin

High altitude training is a physiological method to obtain the same increase in haemoglobin as in blood doping. The idea is to obtain an advantage not present for most of the other competitors, and thus it is unethical, but impossible to disclose. Training at high altitude implies a larger degree of hypoxia than the same sea level training, so two hypoxic metabolites are produced: *Erythropoietin* and *2,3-DPG*. Erythropoietin increases erythropoiesis and thus the haemoglobin concentration, whereas 2,3-DPG form haemoglobin in the deoxy-conformation and increases the P_{O2} gradient when delivering oxygen to the muscle cells.

A serious development occurred following the introduction of industrially produced human erythropoietin (EPO). Natural production is increased, if there is hypoxia in the kidneys.

Erythropoietin is clearly beneficial to endurance athletes, but most types of doping have deleterious effects. The synthesised erythropoietin, when administered to athletes, definitely stimulates the red bone marrow to increase the production of erythrocytes. The effect on the maximum oxygen capacity is indisputable, but the price is often death, because of fluid loss, Haemo-concentration, drastically increased blood viscosity and thrombus formation all over the circulatory system. The death of a whole group of young racing bicyclists, within a short period of time, was probably caused by erythropoietin.

Stimulants

Ephedrine, amphetamine and other psychomotor CNS stimulants are still used by athletes in the hope of increased velocity (so-called speed). Amphetamine or speed pills have improved results in running, bicycling, swimming, weight-throwing and other disciplines compared to placebo. The same stimulants have increased blood pressure and heart rate in athletes exercising heavily in hot climates, until they died from cerebral bleeding or ventricular fibrillation. This has taken place several times in the history of Tour de France.

Cocaine and coffeine seem to suppress natural fatigue, and is also on the doping list. Suppression of natural fatigue leads to exhaustion and circulatory collapse sometimes with cerebral bleeding and ventricular fibrillation.

β- **Adrenergic blocker>**

b-Adrenergic blockers are drugs that reduce heart rate (negative chronotropic effect) and the force of contraction (negative inotropic effect). Both mechanisms reduce the myocardial oxygen demand. In precision sports, where relaxation without tremor is essential, these drugs have a proven beneficial effect in double-blind controlled clinical trials, and they are therefore on the doping list. Precision sports include archery, standard pistol, skeet shooting, rifle shooting, ski jumping, billiards, etc.

Participation is ski-shooting competition is hardly advantageous on b-blockers, because the abuser gets too tired to accomplish endurance performance.

Diuretics

The athletes in disciplines with specific weight classes - such as boxing, wrestling, weight lifting etc - reputedly use diuretics in order to cause a rapid weight loss, with the advantage of competing against smaller persons. Uncontrolled use disturbs the normal distribution of ions in the cells and body fluids, and reduces the blood volume and increases viscosity. In extreme cases there is circulatory collapse and death.

Peptidergic hormones

Gonadotropins - in particular the luteotropic hormone (LH) - stimulate release of testosterone from the Leydig interstitial cells of the testes. Human chorion Gonadotropin (hCG) also binds to the Leydig cells and releases testosterone in males.

Corticotropin (ACTH) from corticotropic cells of the adenohypophysis stimulates production and secretion of adrenal cortical hormones (mainly glucocorticoids).

Somatotropin (human growth hormone, HGH) from somatotropic cells of the adenohypophysis increases and regulates growth, partly directly and partly through evoking the release of somatomedins from the liver. HGH increases protein synthesis, lipolysis and blood glucose. HGH induces gigantism in growing individuals and acromegaly in adults. Uncontrolled use may lead to cardiomyopathy, diabetes, adiposity, articular pain, hypertension and early death.

Monstrous growth of the shoulders and bodies of female swimmers is disclosed by vision alone, and the sight is clearly different from a naturally top- trained female.

Some of the female track runners have written history by winning WM and the Olympics for females year after year, although they looked like a male.

Pregnancy/abortion as doping

Pregnancy seems to increase muscle strength in female athletes. Female top athletes - just following the period, where they gave birth to their first child - have set several world records. Of course, this is acceptable as a natural and unintended event.

However, in some countries female athletes have become pregnant for 2-3 months, in order to improve their performance just following an abortion.

Genetic do*ping* In countries where the political will, is not balanced by ethics, recombinant DNA technique may be used in the future to clone groups of individuals with remarkable talents for special athletic performances.

Smokin*g* has acute and deleterious effects on both the cardiovascular and the respiratory system, but athletes have used it. The substances involved are not at the doping list. The CO blocks off part of the haemoglobin, and limits the transport capacity for oxygen to all mitochondria, which is especially inhibitory to the heart and the skeletal muscles. Nicotine constricts terminal bronchioli and arterioles in many vascular beds. Nicotine also paralyses the cilia of the epithelial cells of the respiratory tract. Chronic smoking leads to life-long chronic bronchitis and emphysema or to lung cancer.

4. Fit for life

A high endurance capacity or fitness is healthy. The mortality increases with low endurance capacity in males (Fig.

18-14). A similar pattern is recorded for females. An endurance capacity (fitness number, $V°_{O2}$max) of 34 ml O_2 min^{-1} kg^{-1} or more seem compatible with a reasonable health status and mortality risk.

Fig. 18-14: The endurance capacity ($V°_{O2}$max) in relation to mortality. The total mortality is given as Number of deaths per year per 10 000 males.

Physical inactivity with an endurance capacity (fitness number) below 34 ml min^{-1} kg^{-1} is a risk factor for the development of atherosclerosis, other risk factors and sudden death in males - and probably also in postmenopausal females.

Equations

- The principle of mass balance states that cardiac output is equal to the oxygen uptake ($V°_{O2}$) divided by the arteriovenous oxygen content difference:

Eq. 18-1: $Q° = V°_{O2} / (C_{aO2} - C_{v\text{-}O2})$ - or $V°_{O2}\,max = Q°\,max * (C_{aO2} - C_{v\text{-}O2})$.

- The following is valid for exercising healthy males (up to 70% of their maximal oxygen capacity, $V°_{O2}max$):

Eq. 18-2: $Q°$ litre $min^{-1} = 3.07 + 6.01 \times V°_{O2}$.

- This calculation of $Q°$ allows for estimation of the rarely available mixed venous carbon dioxide concentration ($C_{v\text{-}CO2}$) from Fick's principle:

Eq. 18-3: $C\bar{v}CO2 - C_{aCO2} = V°_{CO2} / Q°$; or $C_{v\text{-}CO2} = C_{aCO2} + V°_{CO2} / Q°$.

- The diffusion-limited oxygen uptake ($V°_{O2}$) equals the product of lung diffusion capacity (D_{L02}) and the mean alveolar oxygen tension gradient (DP_{O2}):

Eq. 18-4: $V°_{O2} = (D_{L02} \times DP_{O2})$.

- Poiseuilles law relates bloodflow ($Q°$) to total peripheral vascular resistance (*TPVR*) and the mean arterial driving pressure (MAP):

Eq. 18-5: $Q°$ = MAP/*TPVR*.

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A. Fatigue can never be fully explained by a simple rise in plasma-[K^+] only.

B. Fitness tests rest on the assumption, that there is an exponential increase in HR with increasing oxygen uptake or work rate.

C. Erythropoietin clearly improves the endurance capacity of athletes.

D. Erythropoietin is produced and secreted from the kidneys.

E. At maximal works the lung diffusion capacity for oxygen rises to 9 ml STPD s^{-1} kPa^{-1} .

Case History A

A 32-year old marathon runner (body weight 60 kg) is examined on a treadmill. He is running at a velocity of 16 km per hour. The following variables are measured:

$Q\circ$= 25 l per min, C_{aO2} = 200 and $C_{v\text{-}O2}$ = 40 ml (STPD) per l of blood. The concentration of lactic acid is measured in the blood every second min for 18 minutes. The level increases from 1.1 to a steady state value of 15 mM.

The total work rate of the heart (pressure-volume work and kinetic work) is 16 Watts or 16 J per second, and the mechanical efficiency of the heart work is assumed to be 20%.

The energy equivalent for oxygen on a mixed diet is 20 kJ per l (STPD) of oxygen.

1. *Calculate the oxygen uptake of the heart during this work.*
2. *Assume that the arterio-venous oxygen content difference for the heart is equal to that of the whole body. Calculate the coronary bloodflow during running.*
3. *Has the athlete accumulated an important oxygen debt during the 18 minutes of running?*

Case History B

A male world record holder in 100-m dash is examined on a treadmill with a velocity capacity up to 35 km per hour. His body weight is 78 kg. While standing relaxed on the treadmill before exercise, his oxygen uptake is measured to 300 ml STPD min^{-1}. The ventilatory exchange quotient (R) in respiratory steady state, is measured to 1.0. At a given signal the athlete jumps on the running treadmill and performs a 20-second dash similar to a 200 m dash on the track. The inspired, atmospheric air has an oxygen fraction (F_{IO2}) of 0.2093 and a carbon dioxide fraction (F_{ICO2}) of 0.0003. As soon as he jumps off the treadmill, he is connected to a system of rubber bags, where his expired air is collected over the next hour until his resting oxygen uptake is re-established. All the expired air in the rubber bags is mixed and analysed. The mixed expired air fractions are F_{EO2} = 0.1746 and F_{ICO2} = 0.035. The volume of mixed expired air is measured at ATPS, and by calculation corrected to a STPD volume of 1090 litres.

1. *Calculate the oxygen debt repaid over the 1-hour post-exercise period.*
2. *What assumptions is maid in order to perform this calculation?*
3. *Calculate the ventilatory exchange quotient in the post-exercise period and compare the result to the pre-exercise R-value.*
4. *What is the basis of oxygen debt?*

Case History C

A well-trained male long distance skier, weight 74 kg, has a $V\circ_{O2}max$ of 6 l STPD min^{-1} .The maximal arterio-venous oxygen content difference is 150 ml STPD l^{-1} of blood.

During maximal work his lung diffusion capacity for oxygen (D_{LO2}) rises to 9 ml STPD s^{-1} kPa^{-1}.

1. *Calculate maximum cardiac output and describe the consequences for the lung perfusion.*

2. *Calculate the fitness number and explain what it means.*

3. *Define the mean oxygen tension gradient for lung diffusion and calculate its size. Explain how the gradient can increase to this extent.*

Case History D

A female 20 years of age, with a body weight of 62 kg, is exercising on a bicycle ergometer during steady state. Her cardiac output ($\dot{Q}$) is *measured to 25 l* min^{-1} *by the mass balance principle with carbon dioxide as indicator, and her arteriovenous* O_2 *content difference is measured to 170 ml STPD* l^{-1}.

1. *Calculate her oxygen-uptake per min (*$\dot{V}_{O2}$*).*

2. *What assumption must be made in order to calculate her fitness?*

Case History E

An adult male has a lung diffusion capacity for oxygen of 22 ml STPD per min and per mmHg, and a mean alveolar O_2 *tension gradient of 12 mmHg. His oxygen concentration in the arterial blood (*C_{aO2}*) is 200 ml per l and the renal bloodflow (RBF) is 1200 ml per min. The renal* O_2 *consumption is 15 ml per min.*

1. *Calculate his oxygen uptake in ml STPD per min.*

2. *How is it possible for this person to increase the oxygen uptake to 4900 ml STPD per min?*

3. *Calculate the arteriovenous oxygen content difference in the kidneys.*

4. *Is the oxygen delivery to the kidneys redundant?*

Try to solve the problems before looking up the answers.

Highlights

- *The total muscular oxygen uptake can rise by a factor of 80 from rest to maximal exercise.*
- *At the start of exercise, signals from the brain and from the working muscles bombard the cardiopulmonary control centres in the brainstem. Both cardiac output and ventilation increase, the* α*-adrenergic tone of the muscular arterioles falls abruptly, whereas the vascular resistance increases in inactive tissues.*
- *The circumventricular organs of the brain contain many fenestrations, and they are located close to the control centres of the hypothalamus and the brainstem.*
- *Training improves the capacity for oxygen transport to the muscular mitochondria, and improves their ability to use oxygen. After long-term endurance training the athlete typically has a lower resting heart rate, a greater stroke volume, and a lower peripheral resistance than before.*
- *The* $\dot{V}_{O2}$*max progressively increases by endurance training, and also the extraction of oxygen from the blood increases.*
- *At maximal work the lung diffusion capacity for oxygen (*D_{LO2}*) rises to 9 ml STPD* s^{-1} kPa^{-1} *(from 3.6 at rest). The lung diffusion capacity for oxygen increases by endurance training.*
- *According to the redundancy hypothesis, the rise in cardiac output and in ventilation during exercise is caused by*

an integration of neural and humoral factors.

- *The demand of fast progress is linked to competitive sports. A better strategy is to practice at a relaxed level, until stamina is developed and hard training is tolerated.*
- *Relaxed training is often so comfortable that it becomes a lifestyle. Tender muscles are avoided by prewarming, and a careful muscle stretch program following exercise.*
- *Amphetamine or speed pills and other CNS stimulants have improved results in running, bicycling, swimming, weight-throwing and other disciplines compared to placebo. The same stimulants have increased blood pressure and heart rate in athletes exercising heavily in hot climates, until they died from cerebral bleeding or ventricular fibrillation.*
- *Pregnancy/Abortion as doping. Pregnancy seems to increase muscle strength in female athletes. Female top athletes have set world records, just following the period, where they gave birth to their first child. - In some countries female athletes have become pregnant for 2-3 months, in order to improve their performance just following an abortion.*
- *Blood doping definitely improves the oxygen transport with the blood and also the maximal oxygen uptake.*
- *Doping with erythropoietin stimulates the red bone marrow to increase the production of erythrocytes. The beneficial effect on the maximum oxygen uptake is indisputable, but the prize has been death because of thrombus formation.*
- *Doping addicts have a high risk of cardiovascular diseases (arterial hypertension, atherosclerosis, heart attacks and strokes), muscular disorders, liver disease, and - in males - testicular insufficiency. Both the sperm formation and the testosterone production are suffering, often irreversibly.*
- *Lack of fitness is a risk factor for the development of atherosclerosis and for sudden death.*

Further Reading

Medicine and Science in Sports and Exercise. Monthly journal published by the Am. College of Sports Medicine. Williams & Wilkins Co, 428 East Preston Street, Baltimore MD 21202-3993, USA.

Apps DK, Cohen BB and CM Steel. Biochemistry. Bailliere Tindall, London, 1994.

Wood, S C, and R C Boach: Sports and Exercise Medicine. Marcel Dekker Inc, N.Y. 1994.

Katzung BG. Basic & Clinical Pharmacology. Appleton & Lange, Stanford, Connecticut, 1998.

Return to chapter 18

Return to content

Chapter 20

Metabolism & Nutritional Disorders

Study Objectives

- To *define* heat energy, basal metabolic rate, Gibbs energy for ATP-formation, mechanical efficiency, metabolic rate, and the energy equivalent for oxygen.
- To *describe* direct and indirect calorimetry, factors influencing metabolic rate and basal metabolic rate, and conditions with unsteady respiratory state.
- To *draw* a curve for the combustion rate of alcohol.
- To *calculate* a metabolic variable from relevant variables given.
- To *explain* the alcohol metabolism and toxicity. To explain the control of appetite, dietary thermogenesis, energy balance, net combustion and RQ relations. To explain the first law of thermodynamics applied to humans.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The first law of thermodynamics. The internal energy of a system can change for any transition between two equilibrium states and is equal to the heat exchanged by the system and the work done by or on the system. – As a consequence, the metabolic heat energy transfer equals the heat loss plus the stored heat energy.*
- *The surface law: The basal metabolic rate (BMR) per body surface area is much more uniform than the BMR per kg of body weight in individuals of the same species but of different form and size. The best expression for comparison is the BMR per kg of lean body mass. The lean body mass is the fat free mass.*
- *Van´t Hoof´s rule: The rate of energy conversion in chemical reactions increases in proportion to the rise in temperature.*

Definitions

- **Basal metabolic rate** *(BMR)* is defined as the metabolic rate measured with the subject awake in the morning, fasting, at neutral ambient temperature and resting horizontally in the respiratory steady state.
- **Body mass index** (BMI) is the weight of the person in kg divided by the height (in m) squared. The normal range is 19-25 kg per square metre.
- **Brocas index** is the predicted body weight in kg, which equals the height of the person in cm minus 100 for males and 110 for females.
- **Dietary thermogenesis** is the increase in metabolic rate following food intake.
- **Energy balance** is a condition, where the energy input equals the energy output, so the energy stores of the body are unchanged.
- **Gibbs energy** is the free chemical energy in food, which is available for life.
- **Heat energy** is energy transfer caused by a temperature gradient.
- **Ideal weight** refers to the weight associated with the highest statistical life expectancy. The ideal weight is determined with the Brocas index or with prediction tables.

- **Inactivity** is defined as a low *endurance capacity* (ie, a maximal oxygen uptake below 34 ml * min^{-1}*kg^{-1}). Inactivity is probably involved in development of *life-style risk factors.*
- **Lean body mass** is the fat free body mass.
- **Marasmus** is the result of universal starvation in a child, who has low body weight, muscle wasting, and the look of an old person.
- **Mechanical efficiency** is the ratio between external work and the total energy used during work.
- **Metabolism** is defined as the sum of all chemical processes in which energy is made available and consumed in the body.
- **Metabolic rate** *(MR)* is defined as the decrease in internal energy (enthalpy) of the body in a given time period. The metabolic rate refers to the measurement in energy units with direct or indirect calorimetry.
- **Net metabolism** is the stochiometric sum of the net reactions in the body.
- **Nitrogen balance** is a condition, where the nitrogen input from absorbed amino acids equals the nitrogen output in the urine.
- **Obesity** implies the excess storage of fat, and is defined as an actual body weight exceeding the ideal weight by more than 20%, or by a body mass index above 30 kg per square metre (WHO).
- **Protein deficiency** (kwashiorkor) is starvation in children, which subsists on a protein-poor diet rich in carbohydrates.
- **Respiratory Quotient (RQ)** and **ventilatory exchange ratio (*R*)** is defined in Chapter 14.
- **Teratogens** refer to all chemical, physical and biological agents that cause developmental abnormalities (teraton means monster).
- **Vitamins** are *essential organic catalysts* in the diet, necessary for normal metabolic functions in humans, but not synthesized in the human body.
- **Respiratory steady state** is a condition where RQ equals *R*.

Essentials

This paragraph deals with *1. Energy exchange, 2. Metabolism, 3. Alcohol, 4. The respiratory quotient and R, 5. Net mechanical efficiency, 6. Energy sources, 7. Direct calorimetry, 8. Native diets, and 9. Control of energy balance.*

1. Energy exchange

It is generally believed that nutrients are necessary in order to *produce* energy in the human body. However, this is *impossible*. The *first law of thermodynamics* states that energy can neither be created nor destroyed but is *transferred* from one form to another or from one place to another.

Life is thermodynamically the maintenance of an infinite row of non-equilibrium reactions in such a way that appear to be in a stationary condition, a *steady state.* Real life is *chaos,* a steady state only maintained as long as we derive chemical energy from food. Only part of the dietary energy is available for ATP formation in humans. Cellulose, for example, passes the digestive tract without being absorbed. The absorbable chemical energy passes through the intestinal mucosa, and is in the body transformed to energy rich phosphate bindings in ATP (*Gibbs-energy*, D*G*).

ATP is broken down to ADP during muscular contractions. Muscular contractions stimulate the oxidation of fatty acids and carbohydrates in the muscle cells which liberate more energy for rephosphorylation of ADP to ATP. The energy is used for the maintenance of chemical syntheses, electrochemical potentials and for the net-transport of substances across membranes.

The *Gibbs energy* is the free chemical energy available in food. However, 75% is lost as heat energy, and the *mechanical efficiency* of exercise is therefore only 25%. The ratio between *external work* (*W*') and the total energy used during work (-DU) is called the *mechanical efficiency*. In this case DU equals D*G*. The *mechanical efficiency* is always less than one and often only 0.25 as stated above. The energy, which is not transferred to external work, is released as heat energy (-*Q*) or is accumulated in the body as heat. At the onset of exercise 50% of the total energy from hydrolysis of ATP is converted into mechanical energy in the myofibrils. The remaining 50% are lost as *initial heat*. As shown above the mechanical efficiency is only 25%, however. This is because energy recapturing recovery processes (oxidative regeneration of ATP etc) occur outside the myofibrils. Hereby, half of the energy is dissipated as so-called *recovery heat*.

Heat energy is *low prize energy*. In contrast to ATP energy, it is not available for work in the body. The sum of heat energy generated and work performed is constant and equal to the Gibbs energy.

When no work is performed *W*' is zero, and all body reactions are reflected by the liberated heat energy (-*Q*), which is equal to the decrease in Gibbs energy (-D*G*).

When the pressure-volume work is zero, we have a special energy concept: the *heat content* or *enthalpy*, *H*, which sums up all energy. The sum of liberated heat energy (*Q*) and liberated work (-*W*') is thus equal to the fall in enthalpy (Eq. 20-1).

The decrease in enthalpy of the human body (-D*H*) is equal to the fall in potential, chemical energy stored in the body.

The *decrease* in Gibbs energy covers almost the total energy, except for the pressure-volume work. Since oxygen consumption is almost equal to the carbon dioxide output, the pulmonary volume change is negligible and this work is negligible.

2. Metabolism

The *metabolism* of a person is defined as the sum of all chemical reactions in which energy is made available and consumed in the body. The bindings between hydrogen and carbon in nutrients are a source of energy for animals. Such substances are changed into *metabolic end products* (eliminated as bilirubin, urobilin, urea, uric acid, creatinine etc.) and to *metabolic intermediary products* (ie, products that participate in other chemical reactions). The *net metabolism* is the sum stochiometry of the single net reactions in the body.

The decrease in enthalpy in a given time period (-D*H*/min) is the *metabolic rate (MR).*

The *oxidation* of fuel (carbohydrates, glycerol, fatty acids) to CO_2 and water is the primary pathway for generation of energy and subsequent heat energy liberation. Protein can also serve as an important energy source during prolonged exercise, but it must first be broken down to amino acids, who are then partially oxidised (to CO_2, water, NH_4^+ etc). The daily production of *metabolic water* is 350 g and of urea 30 g.

Diabetes mellitus and hunger (*hunger diabetes*) are conditions where fatty acids can produce ketone bodies.

During forceful exercise, energy is obtained primarily from *non-oxidative sources* (glycolysis). There is, therefore, a net formation of lactic acid from glycogen. Following anaerobic exercise the *lactate elimination* accounts for an extra O_2 consumption called *oxygen debt* (Chapter 18).

Oxidation of alcohol can contribute to metabolism. The energetic value of alcohol is 30 kJ/g. An adult person of *70 kg* body weight can combust *7 g* of alcohol per hour (see calculation below). The chemical energy liberated is $(7 \times 30) = 210$ *kJ* per hour or 70% of his resting MR (300 kJ per hour or 83 Watts).

Most of the chemical reactions in our body are *degradative* or *catabolic* - they break a molecule down to smaller units. These reactions are often also *exothermic* (heat releasing) and *exergonic* (the content of Gibbs energy decreases during these reactions). The *synthetic* or *anabolic* reactions (the formation of protein from amino acids) are obviously coupled to these degradative reactions. Synthetic reactions are most often also *endothermic* and *endergonic*.

3. Alcohol

Alcohol diffuses easily in the human body. 20% of the intake by drinking is already absorbed in the stomach. The absorption is *fast* and is stimulated by CO_2 (champagne).

Alcohol distributes in the *total water* of the body within one hour. The distribution volume depends upon the fat mass, because fat tissue only contains 10% of water. The Swedish scientist Widmark called the fraction of the body weight, which is distribution volume for alcohol, **r**. The values for **r** varies considerably, but the *mean*-**r** for females is 0.55 and for males it is 0.68 kg per kg of body weight.

The blood alcohol concentration is measured in permille (ie, one g of alcohol per kg of distribution volume). The most important elimination of alcohol is by *oxidation*. The rate of alcohol oxidation is constant (b = 0.0025 permille per min) and is independent of the blood alcohol concentration. The absolute amount of alcohol *eliminated* per minute is: (b × r × body weight) - see Eq. 20-4.

The constant rate is due to the primary, partial oxidation to acetate via acetaldehyde in the liver by *alcohol dehydrogenase*: $C_2H_5OH + 0_2$ « $CH_3COOH + H_20$. Acetate is broken down in nearly all tissues. The total oxidation of alcohol: $C_2H_5OH + 3\ 0_2$ Þ $2\ CO_2 + 3\ H_20$ implies an RQ of 2/3. A healthy person with a metabolic rate (*MR*) of one mol O_2 per hour can partially oxidise almost 1/6 mol of alcohol per hour, by using almost 1/6 of his *MR* in the liver (one mol = 46 g alcohol; 46/6 or about *7 g* alcohol per hour).

Fig. 20-1: Absorption and oxidation of alcohol.

If this standard person receives an alcohol infusion of 7 g per hour and has a normal hepatic bloodflow of 90 l per hour (1.5 × 60 min), his maximal alcohol elimination rate corresponds to a blood [alcohol] of (7/90) = 0.08 g per l. This is a blood [alcohol] threshold below, which the oxidation rate decreases with time (Fig. 20-1).

The excretion of alcohol molecules takes place through *expiratory air, urine and sweat*. This excretion is generally considered to be negligible compared to the oxidation. This is actually true at rest (Box 20-1). A resting athlete with a blood [alcohol] of one permille or 1 g per kg has a small alcohol partial fraction in pulmonary blood and alveolar air (1/2000- 1/2100). With an alveolar ventilation of 5 l BTPS per min at rest, this person excretes *0.15 g* each hour via expiratory air (Box 20-1). The concentration of alcohol in plasma water, sweat and urine is 20% (1.2 fold) higher than the blood [alcohol]. A resting athlete with a diuresis of one ml/min or 0.060 l per hour excretes alcohol by renal ultrafiltration at a rate of only 0.072-g per hour. If the person also has a sweat loss of 0.1 l each hour, he further excretes *0.12 g* per hour. The total excretion at rest is only *0.342 g* each hour (Box 20-1).

Box 20-1: Oxidation of alcohol is generally agreed to be the single essential elimination method. Look here for excretion, which is considered to be negligible. The person is a male athlete with a blood alcohol of one permille.

Excretion of alcohol by Route	Resting condition g each hour		Exercise g each hour	
1. Expiratory air	(1×5×1/2000*60)	= 0.15	(1*80*60*1/2000)	=2.4
2. Urine	(1.2*0.060)	= 0.072	-	
3. Sweat	(1.2*0.1)	= 0.12	(1.2*4)	= 4.8
Total	**0.342**		**7.2**	

However, during *one hour of exercise* in a warm climate, when ventilation is 80 l per min, and when water loss is 4 l per hour (sweat and evaporation), the total alcohol excretion of the athlete is *7.2 g* per hour. This total *excretion* is actually larger than the amount broken down by *maximal oxidation: 7 g* each hour (calculated above).

The rate (b) increases with *increasing* temperature, with *increased* metabolism (thyroid hormones, dinitrophenol), and decreases under the influence of enzyme inhibitors.

Hepatic alcohol dehydrogenase metabolises alcohol to acetaldehyde, which is oxidised to acetate by aldehyde

dehydrogenase in the mitochondria. Acetate is then oxidised to carbon dioxide and water, primarily in the peripheral tissues. Fructose increases b. Both enzymes are dependent on nicotinamide adenine dinucleotide (NAD^+). One mole of alcohol oxidised to acetate produces 2 moles of reduced nicotinamide adenine dinucleotide (NADH).

The sum [NAD^+ + NADH] is constant. Hepatic alcohol oxidation causes [NADH] to rise, so that NADH inhibition becomes the rate-limiting factor for oxidation.

There are two other enzymes, apart from *hepatic alcohol dehydrogenase*, that can oxidise alcohol. These are catalase and MEOS (Microsomal Ethanol Oxidation System). A small amount of alcohol dehydrogenase is found in the gastric mucosa.

4. The respiratory quotient and R

RQ is the *hypothetical, metabolic ratio* between carbon dioxide output and oxygen consumption of all the cells of the body. RQ is an indicator of the type of foodstuff metabolised.

R is the *ventilatory ratio* between CO_2 output and O_2 uptake for the person quantified by gas exchange equipment.

Respiratory *steady state* is a condition where *R* equals RQ, and the gas stores of the body are unchanged. See Chapter 16.

Compared to the oxidation of carbohydrates (RQ = 1), fat oxidation has a distinctly low RQ (0.7), and protein is oxidised with a RQ of 0.8.

Carbohydrates are rich in oxygen compared to the minimum in fats. Overfeeding with carbohydrates results in a partial conversion to fat. The corresponding release of oxygen from carbohydrates diminishes the oxygen uptake, and *R* becomes larger than one.

The diminished glucose metabolism during fasting and in diabetics lowers the *R* towards 0.7, because of the increased conversion rate of fat.

Hyperventilation decreases the amount of exchangeable CO_2 in the large body stores, without altering oxygen uptake. The tissues and blood cannot store additional oxygen. As a consequence the $V°_{CO2}$ / $V°_A$ -ratio (= F_{ACO_2}) is reduced. This implies a fall in P_{ACO2} and in P_{aCO2}. *R* is distinctly increased during hyperventilation - often up to 2-3.

Hypoventilation reduces *R* towards zero at apnoea.

Metabolic acidosis is characterized by low pH and by negative *base excess* in the extracellular fluid (Chapter 17). A high pH and positive base excess characterise *metabolic alkalosis*. Metabolic acidosis is compensated by hyperventilation implying a rise in *R*, and metabolic alkalosis is compensated by hypoventilation with a fall in *R*.

R does not change when a person on a mixed diet (RQ = 0.83), or when a person on a high fat diet (RQ = 0.7) exercises moderately, because the fat combustion dominates.

R will fall, however, when a person on *carbohydrate rich* diet (RQ = 0.96-1) works for hours.

Strenuously heavy exercise implies a substantial, initial rise in *R* (*R*>3), because the lactate liberated will release CO_2, which is then eliminated in the lungs in much larger volumes than oxygen is taken up.

Glycogen: $(C_6H_{10}0_5)n + 6n\ 0_2 = 6n\ CO_2 + 5n\ H_20$, that is RQ = 1.

Glucose: $C_6H_{12}0_6 + 6\ 0_2 = 6\ CO_2 + 6\ H_20$, that is RQ = 1.

The enthalpy released per mol of glucose is 2826 kJ. One mol of glucose has a mass of 180 g, and 6 mols of oxygen have a volume of $(6 \times 22.4) = 134.3$ l STPD. The *enthalpy* per g of glucose is thus 2826/180 = 15.7 kJ/g, and the *energy equivalent*, which expresses the energy with respect to the oxygen consumed, is 2826/134.3 = 21 kJ per l STPD.

The dietary protein-nitrogen is equal to the nitrogen excretion in the urine when the person is in nitrogen balance. Protein-retention during growth, training, protein-rich diet, pregnancy and reconvalescens are called *positive* nitrogen balance (not

urea accumulation in uraemia). Protein-loss during inactivity, bed rest, fever, blood loss, burns and lesions is called *negative* nitrogen balance.

5. Net mechanical efficiency

The *net mechanical efficiency (E_{net})* is the ratio of *external work rate* (N × m/s = J/s) to net chemical *energy expenditure* (J/s or Watts) during work. E_{net} is 20-25% in isolated muscles and also in humans during aerobic cycling. Its size increases with the amount of training, because the untrained individual does not use the muscles effectively. Legwork has the largest E_{net}, since arm work necessitates fixation of the shoulder belt. The work rate is measurable with a cycle-ergometer (Eq. 20-5).

6. Energy sources

The predominant source of energy is *oxidation* of fuel in the mitochondrion. Hereby, high-energy compounds such as creatine phosphate and ATP are formed. Glucose is oxidised by nicotinamide-adenine-dinucleotide (NAD^+), so by glycolysis two pyruvate molecules are formed in the cytosol, transported to the mitochondrion, and transformed to a *Co-enzyme-A derivative (acetyl-CoA)*, which then is involved in the*Tri-Carboxylic Acid (TCA)* cycle (Fig. 20-2).

Provided a certain oxygen flux from the lungs to the mitochondria is present, the electron transport chain (the glycero-phosphate shuttle) will reoxidize ($NADH+H^+$) and $FADH_2$ to NAD^+ and FAD (Fig. 20-2).

In the glycolysis, one glucose molecule is converted to 2 molecules of pyruvate, with the other products being 2 ATP and ($2\ NADH + H^+$).

Through the oxidation of pyruvate in the TCA-cycle, three ($NADH+H^+$), one $FADH_2$, and one GTP are formed. If *complete oxidation* occurs in the glycerophosphate shuttle of the mitochondrion, one NADH equals 3 ATP, and one $FADH_2$ equals 2 ATP. Since the NAD^+ reduced in the glycolysis is cytosolic, it usually equals 2 ATP only, depending on the shuttle used.

When pyruvate is transformed to acetyl-CoA, one molecule of ($NADH^+ + H^+$) is formed.

The total production by use of the *glycero-phosphate shuttle* in oxidative phosphorylation is *36* ATP molecules per glucose molecule (6 from the glycolysis, 6 from the transformation and 24 from the TCA cycle). - If the *malate-aspartate shuttle* is used, a total of *38* ATP molecules are formed per molecule of glucose oxidised.

Oxidation of one glucose molecule typically implies the use of six oxygen molecules. Accordingly, the P: O_2 ratio is 36/6 = 6, which is equal to a P:O ratio of 36/12 = *3*. The free fatty acids (FFA) from the cytosol (intramuscular or extramuscular origin) are transformed to acetyl-CoA (Fig. 20-2).

The pyruvate production rises with increasing glycolysis rate, and pyruvate is the substrate for alanine production. Alanine is liberated to the blood and its concentration increases linearly with [pyruvate] during rest and exercise.

During anaerobic conditions - an insufficient oxygen supply - ($NADH + H^+$) is reoxidized by the pyruvate- lactate reaction, and the glycolysis continues. The anaerobic ATP production does not block the aerobic ATP production, but functions as an *emergency supply*.

Fig. 20-2: Biochemical pathways for ATP production.

The largest rise in blood [lactate] takes place at work intensities above 50% of the maximum oxygen capacity. Lactic acid is a fixed acid - in contrast to the volatile H_2CO_3 - produced during exercise, and in a muscle cell with a pH of 7 such an acid is essentially totally dissociated (pK = 3.9). Since the proton associated with lactate production reacts immediately with bicarbonate within the cell, its CO_2 production must increase by one mol CO_2 for each mol of *bicarbonate* buffering *lactic acid*.

Lactate accumulates in the muscles and blood, if the glycolysis proceeds at a rate faster than pyruvate can be utilised by the mitochondria, or if ($NADH + H^+$) is not reoxidized rapidly enough.

We posses *100 mmol* of glucose (stored as glycogen) per kg of wet muscle weight, or 3.5 mol in the muscle tissue. Muscle tissue does not contain glucose-6-phosphatase. Our normal 5-l of circulating blood only contains 5 mM, or as a total 25 mmol (5 g) of glucose. During exercise the muscle uptake of glucose increase considerably, but the blood [glucose] does not fall. The blood [glucose] is kept normal by an increased flux of glucose from the liver (Fig. 20-3).

Fig. 20-3: A schematic overview of carbohydrate metabolism.

1. With increasing intensity and duration of exercise, the sympatho-adrenergic activity and the blood [catecholamines] increase. This is a strong stimulus to the *hepatic glucose production*. The liver contains 50-100 g of *dynamic glycogen*. This liver glycogen is easily broken down into glucose by *glycogenolysis* and released to the blood. Any fall in blood [glucose] during exercise will increase the blood [glucagon] and decrease [insulin] toward zero. Glucagon is bound to hepatocyte receptors, and via cAMP a *glycogenolytic cascade* is started (Fig. 20-3). Hereby the hepatocytes produce large amounts of glucose, sparing muscle glycogen and delaying the onset of fatigue. The lack of insulin inhibits the glucose transport across the cell membranes.

2. Glucose is also produced by *gluconeogenesis* in the liver from glycerol, lactate, pyruvate, and glucogenic amino acids. The gluconeogenesis is stimulated by pituitary ACTH and by cortisol from the adrenal cortex.

   With prolonged exercise the blood [glucose] will fall at the end, when hepatic and muscle glycogen stores are depleted, and the *compensating gluconeogenesis* is also running out of energy sources.

3. Complete exhaustion is delayed considerably in trained athletes, because they utilise lipids, so the glycogen stores are spared by *oxidation* of free fatty acids (FFA).

Skeletal muscles contain *lipid stores* (20-g triglycerides/kg wet weight or 700 g in a person with 35-kg muscles). A standard 70-kg man also contains *extramuscular fat stores* of triglycerides (15 kg).

Sympathetic activity and catecholamines *increase* lipolysis (i.e., hydrolysis of the stored adipose tissue to FFA and glycerol) via activation of *adenylcyclase,* increase in cAMP, phosphorylation and activation of the *hormone sensitive lipase*. Increased blood [lactate] and *glucose intake* reduces lipolysis during exercise.

The fat stores are the *ideal energy stores* of the body, because a large quantity of ATP is available per g; this is due to the relatively low oxygen content of lipids - the point being that the necessary oxygen is inhaled at request.

At rest we have a slow turnover of muscle protein, but during exercise alanine is released in appreciable amounts by transamination of pyruvate in the muscle cells, and the blood [alanine] is doubled - without any important change in other amino acids. Alanine is produced via the *pyruvate-alanine cycle*, and the amino groups are from valine, leucine and isoleucine. The blood to the liver transports the muscle alanine, where its carbon skeleton is used in the *gluconeogenesis*. The blood [alanine] also stimulates the pancreatic islet-cells to *increased* glucagon secretion. Glucagon activates the *glycogenolytic cascade* (see above) in the liver cells, further stimulating glucose output from the liver. These are the two factors in the *alanine-liver cycle* of exercise.

The ventilatory, the cardiovascular and the metabolic systems are coupled, and determined by the following factors:

The primary factor is the *size* of P_{aO_2}, but the *blood oxygen store* is of similar importance in keeping P_{aO_2} as high as possible. The blood oxygen store depends upon the haemoglobin concentration, the haemoglobin-oxygen affinity incl. 2,3-DPG, temperature, and P_{aCO_2}.

The *total oxygen flux* to a certain population of mitochondria also depends upon the bloodflow (ie, cardiac output, muscle bloodflow, lung perfusion etc.).

Indirect measures of enthalpy (*MR* in kJ/min) are easily applicable both at rest and in an exercise setting. Expired air is collected in a Douglas bag (volumetric principle) for subsequent air analysis, and the *volume of oxygen* consumed per min is calculated. It is convenient also to determine the carbon dioxide production in the same period, because their ratio is the respiratory quotient (RQ).

A person on a mixed diet has a RQ of 0.83 and a heat energy yield of 20 kJ per l or 0.45 kJ/mmol of O_2. The metabolic

rate (in kJ/min) is calculated by multiplying the *estimated* volume (l/min) of O_2 consumed with *20* kJ per l. The heat energy *yield* varies with RQ and is found in a table (see Symbols). A metabolic ratemeter - a spirometer (Fig. 13-1) with CO_2 absorber - is practical for determination of oxygen uptake.

A more detailed calculation of the metabolic rate is performed as shown with Eq. 20-6 and ->.

Disadvantages of indirect calorimetry are that it ignores the O_2 debt, and that the method depends upon maintained nitrogen balance and gas stores.

7. Direct calorimetry

The total output of heat energy from the body is most precisely measured in a whole-body calorimeter. The Atwater-Rosa-Benedict's *human calorimeter* has been used to verify the *first law of thermodynamics* in humans. The heat energy delivered from the chamber is only equal to the metabolic rate (*MR)*, provided the external work is zero, and neither equipment nor the human body alters temperature.

Fig. 20-4: The human calorimeter combined with a metabolic ratemeter.

The major single factor is *muscular activity*, which can increase *MR* with a factor of 20 even for hours in marathon running. Inactive persons can have a daily *MR* of 9600 kJ, whereas heavy occupational labour requires 20 000 kJ (20 MJ).

Dietary intake can increase *MR* by 20-30% (see Specific Dynamic Activity, below).

Increased *energy demand* in heart and lung diseases, rapid growing cancer will increase MR importantly. Energy is also lost in other disease states such as proteinuria, glucosuria, ketonuria, diarrhoea, and exudate loss (of plasma) through lesions in the skin or in the mucosa. An extra physiologic *energy loss* takes place during pregnancy and during nursing.

Deposition of heat energy in the body (as in fever and hyperthermia) can increase *MR*.

No work is done under basal conditions, so that all energy is ultimately liberated in the body as heat energy. The liver and the resting skeletal muscles account for half of the basal metabolic rate.

Measurement of the *basal metabolic rate (BMR)* requires the subject to be awake in the morning, fasting and resting horizontally. The ambient temperature must be *neutral*, which is the temperature at which compensatory activities are minimal. Prediction tables for *BMR* in different races are available, and the variables are age, sex, height, and weight and thyroid function.

BMR is rarely used for diagnosis of thyroid disease, because radioimmunoassays (Chapter 26) for thyroid hormone analysis are specific and uncomplicated in use.

The surface law states that the *BMR per body surface area* is much more uniform than the *BMR* per kg of body weight in individuals of the same species but of different form and size. The best expression is the BMR per kg of lean body mass. The lean body mass is the fat free mass. Among different animal species the large animals (elephants) have the smallest relative surface area (ie, surface area per kg), so elephants must have small BMR per surface area compared to mice. This is because the surface-volume ratio decreases with increasing body weight. Besides, small animals also have a thin body shell. The body surface area is estimated with Eq. 20-8.

BMR decreases with age in both sexes (Fig. 20-5).

Fig. 20-5: The basal metabolic rate in females and males decreasing with age.

The *female* surface-related *BMR* values are approximately *10% below* the male values throughout life. Let us compare a female and a male both 21 years of age. The female has a Height of 1.68 m, weight 58 kg and a surface area of 1.66 m^2, whereas the male values are:1.8 m, 76 kg and 1.95 m^2. Calculations from the values read at Fig. 20-5 result in *BMR*s of 70 and 90 Watts, respectively. Now, let the couple live for 50 healthy years maintaining height and weight. At the age of 71, their *BMR*s are reduced to 60 and 75 Watts, respectively.

Intake of meals as such *increases* metabolic rate. This is the *specific dynamic activity of the diet (SDA)* or *dietary*

thermogenesis. SDA is less than 10% of the intake energy for carbohydrates and for fat, but 30% for proteins (Fig. 20-6).

Fig. 20-6: Dietary thermogenesis or so-called specific dynamic activity of foods.

Glucose loaded person forms glycogen and fatty acids out of glucose within an hour, even before the glucose can be oxidised. Accordingly, the SDA caused by glucose can be due to an obligate formation of glycogen and fatty acids. The thermogenic response to carbohydrate seems to include a *muscular* component activated by adrenaline via b_2-receptors and a *non-myogenic* component activated by noradrenaline (NA) via b_1-receptors.

Proteins have no SDA in hepatectomized animals, so hepatic intermediary processes must cause the SDA of proteins. These intermediary processes include formation of *urea* from NH_4^+, breakdown of *amino acids* etc.

In general, SDA can also be related to *mass action* due to increases supply of nutrients, and to temperature increase by the activity (increases the rate of all enzymatic processes).

8. Native Diets

Native diets in Africa and the Orient are rich in fibre, which are plant substances (ie cellulose, hemicellulose A & B, and lignins) resistant to digestion. Dietary fibre has been used in an attempt to cure obesity. Constipation with or without diverticulosis/ diverticulitis of the colon also responds to dietary fibre.

The most widespread *dietary fibre* is *cellulose*, which is a major component of plant cell walls. Cellulose is a linear glucose polymer, but human intestinal enzymes cannot hydrolyse its b-1,4-linkages.

Hemicellulose A is a heteropolymer with linkages between glucose, galactose, mannose, xylose and arabinose (ie, gums or mucilages). *Mucilages* delay gastric emptying and decrease the rate of intestinal absorption.

Hemicellulose B or *pectin* binds water in the gastrointestinal tract, but in addition salts minerals and heavy metals. Hemicellulose A and B seem to lower LDL concentrations, while maintaining HDL concentrations.

Lignins in natural fibres are cross-linked polymers of oxygenated phenylpropane entities. Lignin provides bulk for the faeces because they are difficult to degrade.

Dietary fibres reduce postprandial blood glucose and insulin concentration.

Delay in gastric emptying caused by some dietary fibres reduces symptoms of the dumping syndrome. This is an unwanted consequence of large *gastrectomies*. Following removal of the major part of the stomach, the food pass quickly down the small intestine and elicit distension by nutrients and osmosis, causing a massive sympathetic activity with discomfort.

Dietary fibres seem to prevent hiatus hernia by softening the food bolus and decrease of the swallowing effort. Softening of the faecal bulk with decreased defaecation strain seems to reduce the frequency of haemorrhoids.

Overconsumption of dietary fibre can produce adverse effects with increased flatulence, diarrhoea and intestinal discomfort.

Fig. 20-7: Continuous fasting leads to numerous serious complications or death.

Fasting is a *total stop* of food intake. After 12 hours of fasting, conditions are optimal to measure *BMR* or to analyse the chemical composition of blood (fasting blood values are predictable and easy to interpret). The 12 hours are the methodological criterion for the correct minimum *BMR*, but continued fasting for day's results in a much lower value (65% of *BMR*). Following the first 2 weeks of hunger is the normal body weight reduced to 85%, whereas the resting *MR* is stable at 65% of *BMR*, which is constant to the end of the fasting period (either voluntary or by death).

Glycogen stores are broken down in a few days, since only small stores prevail in the liver and muscles. Then urine nitrogen *increases* as a sign of renewed protein combustion (*gluconeogenesis*). In general the *fat* combustion dominates, until the fat stores are used. Healthy people contain 5-15 kg of fat, but monstrous amounts have been recorded in a 540-kg male from *Guinness Book of Records*.

Oxidation of fat stores - including the partial hepatic oxidation to ketonic bodies - implies development of ketoacidosis and a diabetic glucose tolerance test (Chapter 27). Such a *hunger diabetes* with ketonaemia and ketonuria as in diabetes, have been found in healthy individuals even after only 24 hours of fasting or after extremely fatty meals.

Serious illnesses develop after a few weeks of fasting, because the cell structure proteins are broken down (Fig. 20-7). The proteins of the cell nuclei produce uric acid, which accumulate in the heart (*cardiac disease*) and in the articulations (*uric acid arthritis* or *podagra*).

9. Control of Energy Balance

Energy balance is a condition, where the energy input equals energy output, so the energy stores of the body are unchanged. A person with a body weight of 70 kg contains 550 MJ of combustible energy (entalphy), and if allowed to eat naturally, at least 10 MJ is consumed every day. If the person is fasting for some days he will lose body weight and his metabolic rate will fall to 6.6 MJ daily, so a certain *input control* is hereby documented. The loss in body weight is rapidly compensated when feeding is resumed. If enough food is available the person automatically eat more and more (towards a doubling) with increasing workload (MJ/day), so also a certain *output control* is documented. The internal feedback signals operating in this output and input control are uncertain.

Signals from gastrointestinal centres inhibit the *feeding* or *hunger* centre in the lateral hypothalamic area through afferent nerves (Fig. 20-8). Chyme in the duodenum containing HCl and fatty acids liberate enterogastrones to the blood (ie, intestinal hormones that inhibit gastric activity and emptying). The enterogastrone family consists of secretin, somatostatin, cholecystokinin (CCK) and gastric inhibitory peptide (GIP). Enterogastrones reduce gastric activity, stimulate the *satiety centre* (Fig. 20-8), and increase the production of bicarbonate-rich bile and pancreatic juice. A glucose-rich chyme in the duodenum liberates members of the incretin family to the blood. The incretin family consists of gut glucagon, glucagon-like peptide 1 & 2, and GIP. incretin produce a rapid rise in insulin secretion, which causes the energy stores to increase. The hunger and satiety centres operate reciprocally.

The lipostatic theory explains the constant body weight by liberation of a lipostatic, satiety peptide called *leptin* (ie, thin) from fat tissue. The plasma concentration of leptin is recorded by hypothalamic satiety centres, and seems to reflect the size of the body fat stores or the body fat percentage. Obese patients, often with excessive high plasma leptin concentrations, reduce their leptin concentrations by Banting. Some patients may lack the normal sensitivity to leptin. The leptin molecule is large (16 kDa) and it probably must pass the large fenestrae of the circumventricular organs in order to reach the hypothalamic control centres (Fig. 20-8). The plasma leptin concentration is highest at night.

Also thermoregulatory signals from cold and heat receptors and the plasma concentrations of nutrients may stimulate the satiety centre.

In workers, a minimal work activity threshold must be passed in order to trigger the hypothalamic weight control, but above this threshold eating increases proportional with the workload, and the body weight is constant. If the work rate is extremely high as in marathon training, the hypothalamic control is broken, and the dietary intake and the body weight cannot cope with the high combustion. The body weight falls drastically, which is a certain sign of overtraining.

The cybernetics of appetite control is not only feedback factors. As in all human behaviours, cerebral feedforward factors can dominate. Cerebral feedforward factors are exercise habits, eating habits, *social inheritage*, and they can be of extreme importance to the individual.

Fig. 20-8: A schematic overview of the regulation of food intake. The hypothalamic-feeding centre is located in the lateral region, whereas the satiety centre is medially located in the hypothalamus.

The hypothalamus controls food intake and metabolism, mainly by *autonomic* effects on the islets of Langerhans (secreting insulin, glucagon, pancreatic polypeptide, and gastrin), *hepatocytes* and *adipocytes*. Neuroendocrine-behavioural disturbances seem to be involved in abnormal eating patterns such as *anorexia nervosa*, *bulimia nervosa* and *obesity*. We seem to regulate our appetite by a *combination* of negative feedback and essential feedforward factors.

Vitamins are *essential organic catalysts* in the diet, necessary for normal metabolic functions in humans, but not synthesized in the human body. *Essential catalysts* refer to the fact that lack of the compound in the diet results in a clearly demonstrable *disorder* in humans.

Vitamins A, D, and K are lipid soluble, so they follow the lipid absorption to the liver, where they are stored. Accordingly, any type of lipid malabsorption results in vitamin deficiency of these vitamins.

The vitamin B complex (B_1, B_2, B_6, B_7, B_{12}, folate) and vitamin C are water-soluble. Since they are only stored in minimal amounts, vitamin deficiency develops rapidly. Exceptional is the *enormous* vitamin B_{12} store in the human liver, so pernicious anaemia takes years to develop.

Pathophysiology

This paragraph deals with *1. Starvation with marasmus, 2. Vitamin deficiencies, 3. Alcohol intoxication, 4. Obesity, and 5. Hyperuricaemia and gout*.

1. Starvation with marasmus

Lack of all elements in the diet of a child - or *universal starvation* - leads to *marasmus*. Marasmus is often complicated by deficiencies in vitamins and essential minerals.

Marasmus is common throughout the third world, because when breast-feeding stops, the child must try to survive on an insufficient diet. The body weight decreases, the fat stores disappear, muscle wasting leads to thin limbs that do not grow in length, and infants and children look like aged persons (Fig. 20-9).

Fig. 20-9: The child has lived through periods of universal starvation alternating with periods on a diet mainly consisting of cassava. The result is a combination of marasmus and kwashiorkor.

The abdomen is tremendously distended, because of *hepatomegaly*, flaccid abdominal muscles and possibly oedemal fluid in the abdominal cavity (*ascites*). The short half-life of the intestinal mucosal cells makes them especially sensitive to lack of nutrients, so villous atrophy develops with malabsorption and diarrhoea.

Marasmus is unexpected in the rich part of the world. However, this is not so. Malignant tumours and severe cardiopulmonary disease imply an enormous loss of energy, and terminal weight loss is unavoidable even where the diet is supposed to be sufficient (hospitals and other institutions).

The basal metabolic rate is low and the core temperature is also controlled on a lower than normal level. The heart rate and arterial blood pressure is also low. Haemopoiesis is deficient and anaemia prevails. The immune defence system is impaired and the patient suffers from numerous infections.

Children suffer from growth failure, and brain development is probably affected.

Children fed with a diet deficient in protein alone (essential amino acids) develop *protein deficiency* or *kwashiorkor*. Following breast-feeding, these children subsist on a protein-poor diet rich in carbohydrates (eg, *cassava*). Thin limbs, hypoproteinaemia and ascites (Fig. 20-9) characterise kwashiorkor. The large liver is fatty, because there is carbohydrates enough to provide the hepatocytes with lipids, but the lack of protein makes the production of lipid transporting proteins (*apoproteins*) inadequate.

In spite of the effort from all international institutions concerned, the fraction of the global population falling below the minimum food intake defined by WHO, is increasing, and has done so for years.

2. Vitamin deficiencies

Vitamin A (retinol) and its analogues are termed *retinoids*. Vitamin A occurs naturally as retinoids or as a precursor, b-carotene, in vegetables. Infants fed with cooked milk in developing countries, and adults suffering from chronic disorders with fat malabsorption may develop vitamin A deficiency. Retinoids have the following three effects:

1. Retinoids are an important constituent of the photosensitive pigment of the retina and enhance night vision. The 11-cis-retinal is the aldehyde of vitamin A_1. Retinal combines with the glycoprotein *opsin* to form *rhodopsin* (visual purple) in the retina during darkness.

   Vitamin A deficiency implies a massive fall in the number of rhodopsin molecules in the outer segment of the rods. This impedes dark adaptation and *night blindness* occurs.

2. Retinoids stimulate cellular growth and differentiation. Retinoids convert keratin-producing cells into mucus-producing cells, transcribe new mRNA and encode for new cell proteins, so a more differentiated cell type develops.

3. Vitamin A promotes growth of the skeleton.

Lack of vitamin A causes diminished vision in dim light, followed by *night blindness*, and eventually *blindness*. Lack of vitamin A leads to *squamous metaplasia* of the conjunctiva and glandular epithelium. The tear ducts are occluded by the metaplasia, causing eye dryness (ie, *xerophthalmia*), and occluded sebaceous glands cause follicular hyperkeratosis. Vitamin A deficiency causes *growth retardation*.

Retinoids are used in cystic acne and in psoriasis. Vitamin A in normal dosage has been proposed as an anticarcinogen. Excess intake may be *teratogenic*.

Thiamine - as the majority of the B complex - is found in green vegetables, milk and liver.

Thiamine is the co-factor for many enzymes in the glycolytic pathway. Thus lack of thiamine leads to inadequate glucose metabolism with accumulation of vasodilatating lactate and pyruvate. The peripheral vasodilatation leads to oedema. The increased work of the heart eventually develops into *cardiac failure,* which increases the venous stasis and worsens the oedematous state.

Thiamine deficiency (beri-beri) is found in persons consuming polished rice (classical beri-beri), in chronic alcoholics, and in marasmus (see above).

Dry beri-beri is symmetrical polyneuropathy (ie, paresthesia, weakness, heaviness, and paresis of the legs). CNS involvement with ischaemic damage results in the **Wernicke-***Korsakoff syndrome* (ie, ataxia, confusion and ophtalmoplegia).

Wet beri-beri describes thiamine deficiency with oedemas of the legs, pleural effusions and ascites. - These disorders respond immediately to thiamine treatment.

Vitamin B_2 (riboflavin) deficiency

Riboflavin is widely distributed in animal and vegetable foods. Riboflavin is destroyed by ultraviolet light but thermostable and not destroyed by cooking. In the human body riboflavin is converted into flavin mono- and di-nucleotides. These compounds are of crucial importance in the electron transport chain.

Riboflavin deficiency is frequently only part of a *combined* vitamin B deficiency, but the classical manifestations are lesions around the natural openings: *1. interstitial keratitis* of the cornea with vascularisation, *2. seborrhoic dermatitis* (face, vulva, and scrotum), *3. angular stomatitis* (ie, *cheilosis* or fissures at the angles of the mouth), and *4. glossitis.*

These lesions respond to riboflavin usually given parenterally as a *vitamin B complex*.

Vitamin B_6 deficiency

Vitamin B_6 activity is found in three compounds found in both vegetable and animal foods: pyrodoxine, pyridoxal, and pyridoxamine. Pyridoxal phosphate is a co-enzyme for transaminases, carboxylases (formation of the neurotransmitter GABA) and other enzymes. Drugs like the antituberculosis drug, isoniazid, and the copper-chelating agent, penicillamine, are B_6-antagonists.

Certain types of polyneuropathy including the CNS and anaemia with saturated iron-stores respond to vitamin B_6 therapy.

Vitamin B_7 deficiency

Niacin deficiency or *pellagra* (ie, roughs skin) is recognized by the combination of the 3 diagnostic Ds: Dermatitis, diarrhoea, and dementia.

Light exposed areas exhibit dermatitis with rough scales. The diarrhoea is colonic and watery. Cortical atrophia and

degeneration of myelinated tracts in the spinal cord cause the dementia.

Niacin is involved in the formation of nicotinamide adenine dinucleotide (NAD) and its phosphate (NADP). These molecules are important in many oxidation/reduction reactions of the intermediary metabolism.

We consume niacin found in different types of grains (poor content in maize), and our endogenous synthesis is from tryptophan found in meat, eggs and milk.

Pellagra is seen in malnourished alcoholics, food faddists, and in patients with the *carcinoid syndrome,* where most of the tryptophan is used for serotonin synthesis.

Vitamin B_{12} deficiency

Vitamin B_{12} is almost ubiquitary in animal foods (meat, fish, eggs and milk) but not in vegetables, so dietary deficiency is only found in extremely rare cases of vegetarianism, starvation or anorexia nervosa. Malabsorption disorders (ie, pancreatitis, coeliac disease) seldomly result in biologically consequential B_{12} deficiency.

Vitamin B_{12} deficiency causes *pernicious anaemia* (Chapter 8). Below is described the absorption of the vitamin in the terminal ileum (Fig. 20-10).

Fig. 20-10: The mechanism of normal vitamin B_{12} (cobalamin) absorption in the terminal ileum and storage in the liver.

The *intrinsic factor-cobalamin complex* is resistant to pancreatic proteases, and is normally carried along the gastrointestinal tract to specific *receptor proteins* on the mucosal surface of the terminal ileum. The complex is recognized and bound to the receptor. The free vitamin B_{12} enters the enterocyte, and the intrinsic factor remains in the lumen. Vitamin B_{12} exits from the enterocyte by facilitated or active transport, and appears in the portal blood bound to the glycoprotein, *transcobalamin II* (Fig. 20-10). The hepatocytes clear the portal blood for vitamin B_{12} by receptor-mediated endocytosis. The hepatic vitamin B_{12} store is enormous in healthy individuals. An average value of 5 mg stored vitamin B_{12} must be compared to a daily requirement of 1 mg.

Transcobalamin II is the main carrier in delivering vitamin B_{12} to the red bone marrow, although most of the vitamin B_{12} is bound to transcobalamin I and III.

Folic acid deficiency

Foliates are present in *leafy green vegetables* such as spinach and broccoli, and in organs such as kidney and liver. Excessive cooking destroys much of the food folate. Pregnancy increases the requirement for folate up to tenfold. Folate is absorbed in the small intestine, and transported to the cells via the blood plasma.

Folate deficiency with poor diet for a few months' results in megaloblastic anaemia and glossitis because the stores of folate are small compared to the enormous liver storage of vitamin B_{12} (Chapter 8).

Vitamin C deficiency (scurvy)

Vitamin C or ascorbic acid is a reducing substance found in fresh fruit and vegetables. Humans cannot synthesise ascorbic acid from glucose as several animals. Ascorbic acid contributes in controlling the redox potential of the cells.

Ascorbic acid is necessary the hydroxylation of *proline* to *hydroxyproline*. This is the single process necessary for the production of *collagen* in all tissues including the vessel walls.

Vitamin C deficiency (scurvy or scorbutus) is found among food faddists and in developing countries, where infants are fed with excessively boiled milk.

The patient with scurvy can only produce *abnormal collagen* without sufficient tensile strength. The capillaries become fragile and bleedings are frequent. They are recognized as bruises of the skin, as haemarthron, as subperiosteal bleedings

and eventually bleeding anaemia develops (Chapter 8). Infections are prolonged and the healing of wounds is poor. Infections of the gingiva (gingivitis) leads to loose teeth, and the lack of normal collagen in growing bones results in arrested bone growth.

Bottle fed infants must receive daily fruit juice, and for poorly fed adults fresh fruits and vegetables are the best preventive means of avoiding scurvy.

There is no advantage in the daily intake of *large doses* of vitamin C to prevent or improve common cold or cancer. In one controlled clinical trial there was an accumulation of cases with kidney stones. *Rebound scurvy* may occur following a sudden stop of the intake of large doses of vitamin C.

Vitamin D deficiency is described in Chapter 30.

Hypervitaminosis D is caused by excess consumption of vitamin preparations. This leads to hypercalcaemia, nephrolithiasis, nephrocalcinosis and *ectopic calcification* of other organs including premature arteriosclerosis.

Vitamin K deficiency

Vitamin K occurs in two forms in nature. Vitamin K_1 is produced in plants, and intestinal bacteria in animals synthesise vitamin K_2.

Insufficient dietary intake of vitamin K is infrequent, and occurs occasionally in the chronically ill patient such as cases of *anorexia nervosa*.

Fat malabsorption is accompanied by *vitamin K deficiency*, because vitamin K is fat-soluble. Newborn babies sometimes suffer from vitamin K deficiency, because the molecule only crosses the placental barrier with difficulty, and because the sterile gut of the baby cannot produce vitamin K_2.

Destruction of the intestinal bacteria by long term antibiotic treatment may also lead to vitamin K deficiency.

Vitamin K deficiency can lead to *terminal bleeding*. This is because vitamin K normally activates four clotting factors: prothrombin, factor VII, factor IX, and factor X. These four proteins probably receive Ca^{2+} binding properties from vitamin K (see Chapter 8).

Therapy with intramuscularly administered vitamin K is rapid and effective.

Vitamin E deficiency

Vitamin E (a-tochoferol) is found in fish, fish oil and vegetable oil from Soya beans and corn. Vitamin E is an *antioxidant*. Vitamin E protects the phospholipids of the plasma membrane against peroxidation by free radicals produced by the cell metabolism.

Prolonged vitamin E deficiency is rare, but leads to CNS lesions, haemolytic anaemia, and muscle disorders. Patients with fat malabsorption or patients receiving parenteral nutrition may develop vitamin E deficiency.

3. Alcohol intoxication

The sequence of events in *acute alcohol intoxication* proceeds with an increasing sense of warmth, flushing of the face, dilated pupils, dizziness and euphoria. There is a general sense of well being with unjustified optimism and the feeling of increased strength and energy. The subject shows a boisterous behaviour with increased psychomotor activity, which is clumsy, and social inhibitions are dissolved.

Negative consequences of alcohol abuse are arrests, automobile accidents, and deleterious effects upon job performance and chronic health problems.

Alcohol interferes with the arrangement of molecules (ion channels, receptors, the GABA-benzodiazepine-channel etc.) in the lipid bilayers of the cell membrane. With increasing intoxication the symptoms and signs of CNS depression become apparent. The subject becomes drowsy, argumentative, angry or weepy, and eventually he is vomiting and complaining of diplopia. Later an examination reveals areflexia, loss of muscular tension, loss of sphincter control, rapid heart rate and

respiratory frequency, decreasing arterial pressure and mean arterial pressure leading to shock. The subject develops hypothermia and increasing stupor, anaesthesia, coma or death.

The intoxication depresses the myocardium and dilates the peripheral vessels. This is why the MAP is falling together with cardiac performance.

Some alcoholics benefit from treatment with disulfiram (Antabuse) or similar drugs. Antabuse inhibits aldehyde dehydrogenase, which results in poisoning from accumulated acetaldehyde.

4. Overeating with obesity

Overeating is related to social patterns and constitutional family traits.

Obesity or *adiposity* implies the excess storage of fat, and is defined by WHO as an *actual body weight* exceeding the *ideal weight* by more than *20%* (if not explained by an above-average muscle and bone mass). The diagnosis is frequently set by inspection of the undressed patient (Fig. 20-11). The fatty stores of the patient in Fig. 20-11 are clinically acceptable. The *ideal weight* is the weight associated with the highest statistical life expectancy. The *Broca index* is a popular and easy method of determining the recommended weight. Brocas index is the predicted body weight in kg, which equals the height of the person in cm minus 100 for males and 110 for females.

Fig. 20-11: The ideal weight and clinically acceptable fatty stores.

Obesity is also established in another way by the help of the *body mass index (BMI).* BMI is the weight of the person in kg divided by the height (in m) squared. The normal range is 19-25 kg per square metre (Fig. 20-11). Marginal overweight is defined as a BMI between 25 and 30 $kg* m^{-2}$. Obesity is defined as BMI above 30 $kg* m^{-2}$, which corresponds to body weights 20% above ideal weight.

Obesity results from a long-term excess of nutritional intake relative to the energy liberation. There are at least three types of obesity: genetic, overeating and inactivity induced.

1. *Genetic obesity*. Genes account for quite some cases of obesity. Genes seem to be causative in 2/3 of all cases of obesity in a lifetime. Genetic movement oeconomists may explain many cases of obesity including familiar obesity, but weight gain does not occur in all pairs of mono- or di-zygotic twins. - Hyperplastic fatness (too many adipose cells) is often found in babies from the rich part of the world, whereas many adults have hypertrophic obesity, which is caused by too large adipose cells.

2. *Overeating*. Intake of poor food, dominated by sweet-fat combinations, explains other cases of obesity. Sugar and fat eaters need not eat very much in order to develop obesity, if they live with a marginal motility pattern. In any type of obesity the low physical activity or inactive life style is typical.

3. *Inactivity*. The major factor in obesity is *physical inactivity*. Persons who exercise can increase their metabolic rate by a factor of 10-20 several hours a day. The second choice of obese persons is to reduce the dietary intake of nutrients. A reduction to half the usual amount of food would be a short, heroic and probably futile project, as well as inefficient, when compared to a metabolic factor of 10-20 during exercise. *Inactivity*, defined as a *low fitness number* (ie, a maximal oxygen uptake below 34 $ml * min^{-1}*kg^{-1}$), is probably involved in western life-style obesity. Obesity is the fate of people dominated by their parasympathetic activities and minimising the use of the sympathetic nervous system. The mortality of a male population increases dramatically with falling fitness (Fig. 18-13).

Obese people have a small dietary thermogenesis, because they avoid physical activity with a high metabolic rate in skeletal muscle and in adipocytes. Decreased sensitivity to leptin has been described in obese patients.

Rare cases of obesity are caused by hormonal, metabolic diseases (insulinoms, hypercorticism, diseases of the thyroid gland, hypothalamic lesions etc). For persons with insulinoms the *high food intake* (hyperphagia) can be a question of life and death.

Obesity and interrelated risk factors

Obesity is clearly a *risk factor*. A risk factor is an epidemiological term for conditions statistically correlated with shortened life expectancy. Obesity paves the way for maturity-onset (type II) diabetes, atherosclerosis, hypertension, myocardial infarction, and stroke.

Type II diabetes is linked to obesity, perhaps because the increased weight or the rise in blood glucose concentration stimulates insulin secretion. A period with high concentration of insulin in the blood plasma decreases the number of insulin receptors on the membranes of muscle and adipose cells (insulin resistance or glucose intolerance). More insulin must be produced from the beta cells of the pancreatic islets, and finally the b-cells are exhausted and a diabetic condition without insulin production developed. Body weight reduction ameliorates the glucose intolerance.

Atherosclerosis, acute myocardial infarction, hypertension, hypercholesterol-aemia, gall-stones, low concentrations of HDL, hyperuricaemia, gout, osteoarthritis of the hip and knee joint, and restrictive lung disease are all related to obesity (Fig. 20-12).

Fig. 20-12: Adipose patient with numerous complications.

There is also an increased incidence of depression, psychological and social problems, intertriginous dermatitis, hernias, impotence, and thrombophlebitis in obese patients.

Amenorrhoea and oligomenorrhoea, reduced fertility is common among premenopausal obese females.

Therapy of established obesity is a frustraneous and highly resistant task. The current therapy of adiposity, including drugs, diet and behavorial modification with exercise, is ineffective - often due to the lack of motivation.

1. Anti-obesity drugs are either centrally or peripherally active.

The *centrally* active drugs either act on catecholamine neurotransmitters (amphetamines), or they act on serotoninergic neurons in the CNS. Initially, all these drugs reduce food intake, and some of them also increase the metabolic rate.

Amphetamines, “holiday pills”, were the first centrally active anti-obesity drugs developed, but there abuse potential is a definite contraindication.

The *peripherally* active anti-obesity drugs, such as *acarbiose*, have only modest effect in controlled clinical trials. Acarbiose is an amylase inhibitor, which reduce the digestion of sucrose. The lipase inhibitor, *tetra-hydro-lipostatin*, blocks the intestinal digestion of lipids, but has only a marginal effect on obesity. From a physiological point of view there is little perspective in new development of anti-obese drugs, because any reduction in nutritional input - even a 10% reduction - is futile.

2. Diet. Most popular, weekly journals publish up to three miraculous diets for obese persons in each issue. This is a contradiction with only marginal possibilities of success. Even a 5-10% reduction in dietary input is experienced as self-torture and tolerated only for a short time. There is no alternative to a healthy mixed diet with a sufficient amount of dietary fibre (see above).

3. Behavioural therapy with exercise. The single factor that can cure obesity is a *balanced degree* of exercise. Exercise can increase metabolic rate from typically 70 Watts at rest to 700 (eg, walking, dancing, sexual intercourse) or 1400 Watts (long distance running). Thus, the most important determinator (factor 10-20) is the increase in metabolic rate from any type of self-induced locomotion.

Running is an alternative for younger and middle-aged persons. Older people must walk, if they have the capacity for it. They may prefer walking in a hilly environment in a relaxed way, so the heart is stimulated. Callisthenics, dancing, skiing, swimming, tennis, cycling, golf are alternatives for persons motivated. The important point is to chose the type of exercise, which is enjoyable - in a relaxed way - for the individual concerned.

Increased awareness of nutritional and fitness intervention in improving the health status of large population groups is essential.

5. Hyperuricaemia and gout

Excessive *production* or inefficient *excretion* of uric acid causes hyperuricaemia.

Hyperuricaemia is a condition with an abnormally high concentration of uric acid in the blood plasma and ECF (above 0.42 mM). Normal values of serum-urate are 0.2-0.42 mM. The saturation threshold over which urate crystals precipitate is around 0.42 mM for tissues with an acid pH.

Hyperuricaemia is asymptomatic for varying periods. When the hyperuricaemia becomes clinically important through recurrent attacks of painful acute arthritis, the condition is called *uric arthritis* or *gout* (Fig. 20-13).

Fig. 20-13: Gout and its complications are shown to the right. Blockage of the uric acid synthesis by allopurinol is shown to the left.

Gout can be primary or secondary.

1. Primary gout. Either increased metabolic urate production, inefficient renal excretion of urate or the two in combination causes genetic or idiopathic gout. Primary metabolic gout relates to two inherited, X-linked enzyme disorders: Hypoxanthine-Guanine-Phospho-Ribosyl-Transferase (HGPRT) deficiency with increased purine synthesis or an abnormally high activity of Phospho-Ribosyl-Pyro-Phosphate -Synthetase (PRPPS). The inherited disorder is exacerbated by diets high in purine or nucleic acids. During purine degradation large quantities of NH_4^+ are liberated. The acidosis leads to crystallisation of urate.

2. Secondary or acquired gout can also be both metabolic and renal. Intercurrent disease with lysis of cell nuclei and release of nucleic acids increases urate production (eg cancer, psoriasis, and excessive weight loss). Impaired renal excretion leads to secondary renal gout.

Most forms of metabolic gout are a result of overproduction of uric acid caused by accelerated purine synthesis from amino acids, formate and CO_2, whereas dietary purines play a minor role. *Xanthine oxidase* oxidises hypoxanthine to xanthine and xanthine to uric acid.

Supersaturated body fluids precipitate thin urate crystals in acid environments. The result is an inflammatory reaction, where leucocytes migrate to the crystals and surround them for phagocytosis.

The acute attack of gout typically occurs in a male with severe pain in the big toe. The pain attack is also called *podagra.* The pain responds to therapy and the patient is asymptomatic for a variable period. Following a series of acute attacks of gout, the pain is persistent, because the urate crystals are permanently present in the joints and other tissues. This is called *chronic gout.*

Toes, ankles and knees are frequently affected. Symptoms and signs of gout include hyperuricaemia, tophi and painful arthritis. Symptoms and signs of gout include hyperuricaemia, tophi and painful arthritis, with extremely tender and swollen joints.

Complications to gout are increased risk of atherosclerosis, hypertension and renal disease including renal calcification and uric acid stones in the ureter.

Persons with hyperuricaemia are at risk, and should be treated with allupurinol (100-300 mg daily) until the plasma-urate is brought down to normal levels (see effect below).

Allupurinol is a *xanthine oxidase inhibitor.* Allupurinol is an analogue to hypoxanthine, but xanthine oxidase (XO) prefers allupurinol as a substrate, so allupurinol is oxidised to oxypurinol (Fig. 20-13). Oxypurinol (alloxanthine) blocks XO, because it binds to the active site on the enzyme. Thus, urate production is inhibited and xanthine/hypoxanthine is accumulated in the blood and ECF. This is fortunate, because these substances are water-soluble and easily excreted in the urine - just as alloxanthine. This is in sharp contrast to the less soluble urate. Uric acid is filtered in the renal glomeruli. Urate is reabsorbed in the proximal tubules by a Na^+ -substrate cotransport with a capacity, which is normally far greater than the amount of urate in the glomerular filtrate. Accordingly, the normal urate secretion takes place by active secretion of urate ions in the distal tubules. The *organic acid-base secretory system* transfer urate ions from the blood to the tubular fluid, but the system has a low capacity for urate.

Patients with *renal gout* suffer from abnormally low distal tubular secretion of uric acid. These patients are treated with

+

uricosuric agents such as *probenecid* and sulfinpyrazone. These molecules compete for the proximal Na -substrate cotransport, so less urate is reabsorbed.

Colchicine is a drug that binds to *tubulin*, a protein in the microtubules of the leucocytes. Hereby, the microtubules disintegrate, which inactivates the leucocytes. Thus, the colchicine prevents the focal infiltration of leucocytes to the damaged tissue and blocks their usual liberation of lactic acid which would further precipitate urate crystals. This effect is slow and not purely beneficial.

Weight loss is often indicated during treatment of gout, but a rapid weight loss is risky in hyperuraemic patients, because the cellular destruction liberates nucleic acids and may elicit an acute attack of gout.

Equations

- **The first law of thermodynamics** states that the sum of liberated heat energy (-*Q*) and liberated work (-*W*') of a system is equal to the fall in internal energy (enthalpy) or heat content (*H*). The decrease in enthalpy of the human body (-D*H*) is equal to the fall in potential, chemical energy stored in the body:

$$\text{Eq. 20-1:}\quad (-D\mathit{H}) = (-\mathit{Q}) + (-\mathit{W}').$$

- **Entropy** is the tendency of atoms, molecules and their energies to spread in a maximum space. The *Gibbs energy* (*G*) is the difference between enthalpy (*H*) and entropy (*S*) when multiplied with the absolute temperature (T):

$$\text{Eq. 20-2:}\quad \mathit{G} = \mathit{H} - T \times \mathit{S}.$$

G determines if a certain reaction occurs, since *G* is minimal at equilibrium. According to the formula, entropy is important at high temperatures, and energy is most important at low temperatures.

- **The Fick cardiac output equation:**

$$\text{Eq. 20-3:}\quad Q^{\circ} = V^{\circ}_{O2/(}\, C_{aO2} - Cv^{-}_{O2)}$$

- **Elimination of alcohol by oxidation**. The rate of oxidation is constant (b = 0.0025 permille per min) and is independent of the blood [alcohol]. The absolute amount of alcohol eliminated per minute is thus:

$$\text{Eq. 20-4: Alcohol oxidation (g/min)} = (b \times r \times \text{body weight})$$

  The fraction of the body weight which is distribution pool for alcohol is called **r** (mean-**r** for females is 0.55 and for males 0.68 kg per kg body weight).

- **Calculation of work rate** on a bicycle ergometer: A measurable blocking force is applied to a wheel with a given radius (r) and with a given rotation-frequency (RPM). The *work rate* or power (force × velocity) is now determined, because the force is known (N) and the distance per s is ($2 \times p \times r$ RPM)/60. The work rate is thus measured in J/s or Watts. Work rate = Force * Distance, or

$$\text{Eq. 20-5: Work rate (Watts)} = N^{*} (2^{*}p^{*}r\ ^{*}RPM)/60.$$

- **Calculation of metabolism** by *indirect calorimetry*: The oxygen uptake and carbon dioxide output is masured volumetrically or gravimetrically together with determination of the nitrogen content in 24 hours urine from the person examined. The urine nitrogen expresses the protein combustion, since protein contains 16% nitrogen. Subtraction of the gas volumes for protein combustion (see data in Symbols) from the total, result in residual volumes of oxygen and carbon dioxide only related to the fat (F g/min) and carbohydrate (C g/min) combustion. Thus F and C can be calculated by solution of two equations with these two unknowns (Eq. 20-6 and -7). By multiplication with the nutritive equivalents for O_2 and for CO_2 (mmol gas per g in Symbols) the mass balance states:

$$\text{Eq. 20-6: F and C related } O_2 \text{ uptake} = (37\ \text{mmol/g} \times \text{C g/min}) + (91\ \text{mmol/g} \times \text{F g/min})$$

Eq. 20-7: F and C related CO_2 liberation = (37 mmol/g × C g/min) + (64 mmol/g × F g/min).

Now the mass of protein, fat and carbohydrate combusted per min is found, and a very precise indirect measure of MR in kJ/min is obtained by multiplication with their energy equivalents (Se Symbols).

- **Body surface area (*BSA*)** is estimated with the approximation formula of the DuBois family:

Eq. 20-8: $BSA\ (cm^2) = WEIGHT^{0.425}\ (kg) \times HEIGHT^{0.725}\ (cm) \times 71.84$

The *BMR* is normally 45 Watts/m^2, so an adult with a body surface area of 1.8 m^2 has a *BMR* of 80 Watts. This corresponds to a daily *BMR* of (60×1440 min $\times 80 \times 10^{-3}$) =6912 kJ.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Vitamins are essential organic catalysts synthesized in the human body.

B. Antibiotics in meats or milk may induce allergy or antibiotic resistance in humans.

C. Obesity is the consequence of inactivity alone, and genes are not involved.

D. Either increased metabolic urate production, inefficient renal excretion of urate or the combination causes primary or idiopathic gout.

E. During therapy with large doses of ascorbic acid, withdrawal may cause symptoms of scurvy.

II. Each of the following five statements have True/False options:

A. Prolonged use of some anti-obesity drugs may imply serious abuse.

B. Osteomalacia does not involve the organic bone matrix.

C. Vitamin A deficiency implies a massive fall in the number of rhodopsin molecules in the outer segments of the rods. This impedes dark adaptation and night blindness occurs.

D. The epiphyseal plate of the growing skeleton is sufficiently mineralised in rachitis.

E. Vitamin K can easily cross the placental barrier.

Case History A

At 7 p.m. a male alcoholic drinks 150 ml of whisky (40 w/v%). His body weight is 58 kg (due to hepatic failure and malnutrition). The fraction of his body weight, which is distribution pool for alcohol, is 0.60. The rate of oxidation of alcohol is reduced to 80% of normal. The rate of oxidation in a healthy person is 0.0025 o/oo per min. At 8 p.m. the patient is involved in a traffic accident and at 9 p.m. his blood [alcohol] is measured to 1.36 g kg^{-1} (o/oo). The patient states to the police that he has been drinking only the whisky at 7 pm.

1. *Calculate the blood [alcohol] at the time of the traffic accident.*

2. *Is the statement concerning alcohol intake correct?*

Case History B

Following an earthquake, three adult females are confined under the ruins of their house in an airtight space of 8 m^3. The initial pressure is 752 mmHg, the temperature of the water-saturated air is 20 ^oC (water vapour tension 20 mmHg) and

the composition of the atmospheric air is normal. The average oxygen uptake is 190 ml STPD min^{-1}, and the average RQ is 0.83. Assume that P_{IO2} = 50 mmHg is the survival threshold.

1 *Calculate the time period, where they have oxygen enough for survival, provided carbon dioxide could disappear?*

2. *Calculate the carbon dioxide output per min.*

3 *Calculate the theoretical amount of CO_2, which should accumulate in the time period for survival from 1. Calculate the theoretical CO_2 fraction in the airtight space.*

4. *Explain the consequences of CO_2 accumulation.*

Case History C

A male, 23 years of age, has a daily metabolic rate of 12 600 kJ (12.6 MJ) and he is eating a mixed diet resulting in a RQ of 0.83. The enthalpy equivalent for oxygen is 0.46 kJ/mmol. His arterial pH is 7.30 and pK for ammonia is 9.3.

1. *Calculate the ratio between ammonia and NH_4^+ in his blood.*

2. *Calculate his daily carbon dioxide output in mol.*

3. *Calculate the amount of carbon dioxide eliminated per day in combination with ammonia, assuming one mol/day of ammonia to be involved in the urea production.*

Case History D

A 70 kg male, 22 years of age, is in a room, where the temperature is 20^o C and P_B are 101.3 kPa. The P_{AO_2} is 14.1 kPa, and P_{AN_2} is 78 kPa. The carbon dioxide output is 660 ml STPD per min. The urinary nitrogen excretion is 10 mg per min. The pressure of water vapour in the alveoli is 6.2 kPa, and F_{IO_2} is 0.2093.

1. *Calculate the alveolar ventilation and F_{ACO_2}. Estimate P_{aCO_2}, and provide reasoning for a possible hyperventilation in this condition.*

2. *Calculate the oxygen uptake for this person.*

3. *Is the pulmonary exchange quotient different from the respiratory quotient (RQ)?*

Try to solve the problems before looking up the answers.

Highlights

- *The first law of thermodynamics states that energy can neither be created nor destroyed but is only transferred from one form to another or from one place to another.*
- *Heat energy is low prize energy. In contrast to ATP energy, it is not available for work in the body. The sum of heat energy generated and work performed is constant and equal to the Gibbs energy.*
- *The oxidation of fuel (carbohydrates, glycerol, fatty acids) to CO_2 and water is the primary pathway for generation of energy and subsequent heat energy liberation. Protein can also serve as an important energy source during prolonged exercise.*
- *Alcohol distributes in the total water phase of the body (60% of body weight) within one hour.*
- *Hepatic alcohol dehydrogenase, catalase and MEOS (Microsomal Ethanol Oxidation System) can oxidise alcohol.*
- *The most important elimination of alcohol is by oxidation. The rate of oxidation is constant (β = 0.0025 permille per*

min) and is independent of the blood alcohol concentration.

- *The blood alcohol concentration is measured in the unit g per kg (permille) of distribution volume. The absolute amount of alcohol eliminated per minute is thus: ($\beta \times r \times$ body weight).*
- *An adult person can oxidise 7 g of alcohol per hour, and at rest only negligible amounts are excreted in sweat, urine and expired air.*
- *During severe exercise by an athlete working in a hot climate, the total excretion of alcohol can be larger than the maximal oxidation (7 g each hour).*
- *The net mechanical efficiency (E_{net}) is the ratio of external work rate ($N \times m/s = J/s$) to net chemical energy expenditure (J/s or Watts) during work. E_{net} is 20-25% in isolated muscles and also in humans during aerobic cycling.*
- *The total production by use of the glycero-phosphate shuttle in oxidative phosphorylation is 36 ATP per glucose molecule (6 from the glycolysis, 6 from the transformation and 24 from the TCA cycle).*
- *The total output of heat energy from the body is most precisely measured in a whole-body calorimeter. The Atwater-Rosa-Benedict's human calorimeter has been used to verify the first law of thermodynamics in humans.*
- *The metabolic rate can increase by a factor of 10-20 during steady state exercise.*
- *Hepatic intermediary processes cause the specific dynamic activity (dietary thermogenesis) of proteins: Breakdown of amino acids, formation of urea etc. The dietary thermogenesis in general is related to mass action and temperature increase when eating.*
- *Serious illnesses develop after a few weeks of fasting, because the cell structure proteins are broken down. The proteins of the cell nuclei produce uric acid, which accumulate in the heart (cardiac disease) and in the articulations (uric acid arthritis or podagra).*
- *The lipostatic theory explains the constant body weight by liberation of a lipostatic, satiety peptide called leptin (ie, thin) from fat tissue. The plasma concentration of leptin is recorded by hypothalamic control centres, and seems to reflect the body fat percentage. Obese patients reduce their plasma leptin concentration by Banting and may lack the normal sensitivity to leptin.*
- *The help of the body mass index (BMI) establishes obesity. BMI is the weight of the person in kg divided by the height (in m) squared. The normal range is 19-25 kg* m^{-2}. *Marginal overweight is defined as a BMI between 25 and 30 kg* m^{-2}. *Obesity is defined as BMI above 30 kg* m^{-2}, *which corresponds to body weights 20% above ideal weight.*
- *Obesity results from a long-term excess of nutritional intake relative to the energy output. There are at least three types of obesity: genetic, over-eating and inactivity induced obesity.*
- *Atherosclerosis, acute myocardial infarction, diabetes, hypertension, hypercholesterol-aemia, gall-stones, low concentrations of HDL, hyperuricaemia, gout, osteoarthritis of the hip and knee joint, and restrictive lung disease are all related to obesity.*

Further Reading

The Journal of Nutrition. Monthly journal published by the Am. Institute of Nutrition, 9650 Rockville Pike, Bethesda, MD 20814-3990, USA.

Silverstone, T (1992) *Appetite suppressants: A review.* DRUGS 43: 820-836.

Maffei, M et al. “Leptin levels in humans and rodents: Measurement of plasma leptin and of RNA in obese and weight-reduced subjects.” *Nature Med* 11: 1155-61, 1995.

Katzung BG. *Basic & Clinical Pharmacology*. Appleton & Lange, Stanford, Connecticut, 1998.

Return to chapter 20

Return to content

Contraction And Relaxation In Smooth Muscle Cell

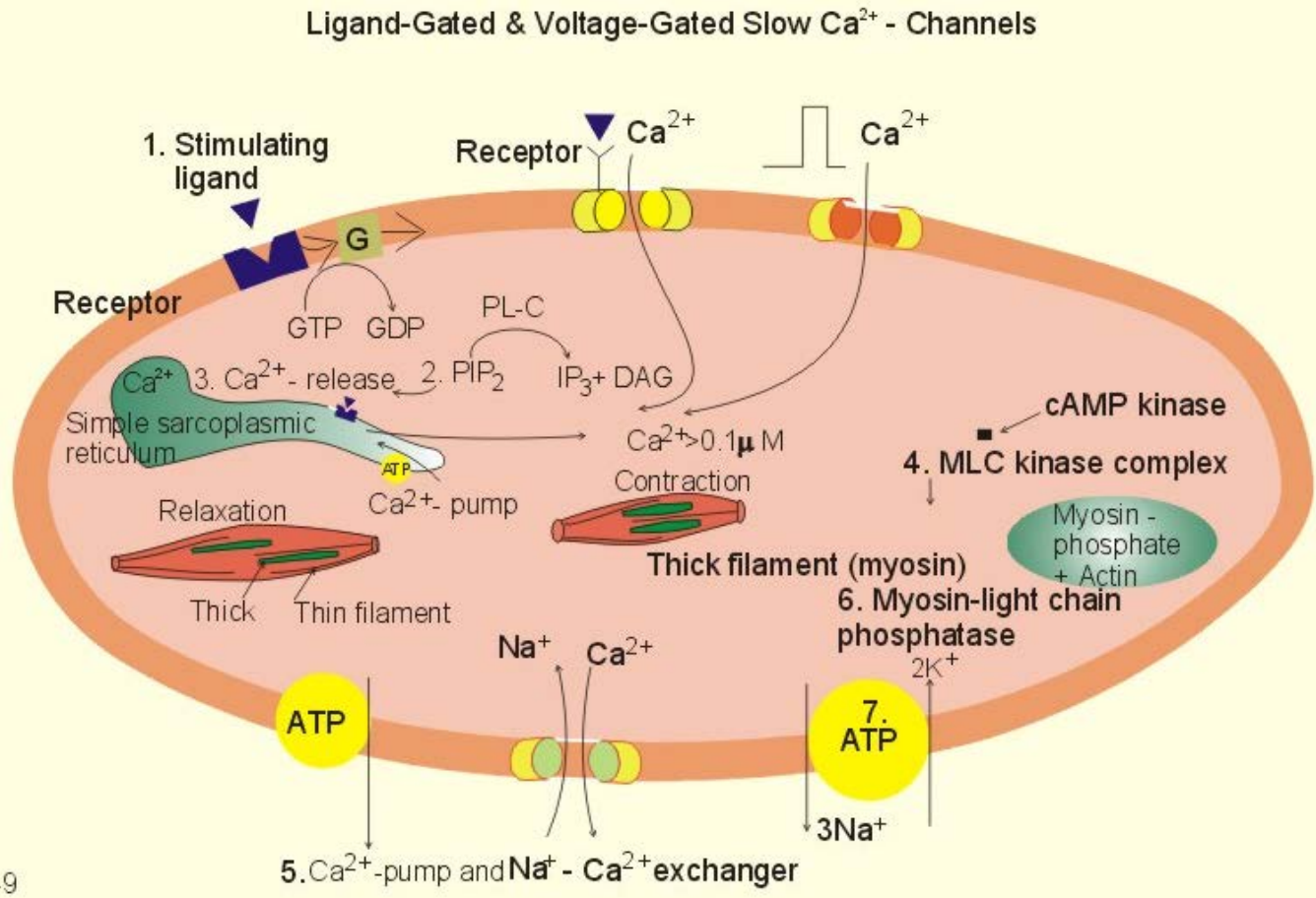


Fig. 2-9

KMc

Chapter 9

Systemic Resistance And Hypertension

Study Objectives

- To *define* arterioles, autoregulation, mean arterial pressure, metabolic vascular control, preferential channels, systemic hypertension, systemic resistance, and inflammatory hyperaemia.
- To *describe* arteriolar calibre, vascular resistance, the measurement of arterial blood pressure and the pressures of the pulmonary system, the role of myoglobin, respiratory arrhythmia, alterations of mean arterial pressure and pressure amplitude.
- To *calculate* one variable when relevant variables are given.
- To *explain* the control of the arterial pressure, hypertension, and reactive hyperaemia.
- To *use* these concepts in problem solving and in case histories.

Principles

- *The haemodynamic features of the cardiovascular system are determined by Newtonian and non-Newtonian relations between the driving blood pressure, the bloodflow and the vascular resistance.*
- *These features are related to the ability of each tissue to control its own bloodflow in accordance with its needs. The control of local vascular resistance is a combination of neural and metabolic factors affecting a basal smooth muscle tone.*

Definitions

- **Arterioles** are vessels that range from 150 to 10 mm in diameter. They control the distribution of blood to different tissues.
- **Arteriolar calibre** is the internal diameter of the arteriole, and the size is determined by the contractile activity of its smooth muscle cells and by the transmural arteriolar pressure.
- **Autoregulation** is an automatic control phenomenon that aims at maintaining a constant bloodflow when the driving pressure is changed.
- **Capillary intermittence**: Krogh presumed that a tissue capillary shifts between a closed and an open state. The capillary diameter varies with the oxygen tension. In well-perfused tissue, high O_2 tension causes vasoconstriction and thus tends to reduce its perfusion.
- **Inflammatory hyperaemia** refers to increased bloodflow with accumulation of leucocytes. This reaction is mainly caused by leukotrienes released by the leucocytes.
- **Malignant hypertension** (*accelerated*) refers to a rapid and serious rise of the arterial blood pressure. The condition can start as paralysis, unconsciousness, or blindness.
- **Metabolic control** is the sum of all metabolic factors that match the oxygen supply to the energy requirement.
- **Myoglobin** is a red, iron-containing, oxygen-binding globin similar to haemoglobin.

- **Reactive hyperaemia** is the increase in bloodflow following temporal vascular interruption of surgical or experimental character.
- **Standard affinity of myoglobin** towards O_2 is the reaction rate at 50% binding. This standard affinity is much higher than that of haemoglobin towards oxygen.
- **Secondary hyperaldosteronism** is recognised by high serum concentrations of renin and aldosterone. This occurs in malignant hypertension or following prolonged use of diuretics. The patients develop cerebral oedema and haemorrhage, cardiac failure and hypertensive nephropathy with proteinuria and microscopic haematuria.
- **Systemic hypertension** - according to WHO - is defined as an arterial blood pressure exceeding 160/95 mmHg (21.3/12.6 kPa) for several months. The pressure increase is either systolic, diastolic or a combination.
- **Systemic resistance** is the total peripheral vascular resistance (*TPVR*), mainly consisting of the arteriolar resistance (in particular that of the essential arterioles in the large striated muscles).
- **Thoroughfare channels** or *preferential channels* shunt the blood directly into the venules bypassing the true capillary bed.
- **Vasomotion** is the rhythmic changes in the arteriolar diameter that causes bloodflow to fluctuate. Vasomotion is brought about by active changes in the tension of vascular smooth muscles. The arteriole can relax completely and then close completely.
- **Vasopressin** is another name for *anti-diuretic hormone (ADH)* from the hypophyseal posterior lobe. ADH controls renal water retention and acts as a moderate vasoconstrictor.
- **VIP** is *Vasoactive Intestinal Polypeptide* from the intestine, the salivary glands and the penile cavernous bodies. VIP is a neurotransmitter and a potent vasodilatator, which is used in the treatment of impotence.

Essentials

This paragraph deals with *1. Autoregulation, 2. Autonomic nervous control, 3. The baroreceptors and other regulators, 4. Oxygen release to the mitochondria, 5. Measurement of blood pressure, 6. Age and MAP*.

1. Autoregulation

The resistance vessels of the coronary system tend to *diminish* any change in the bloodflow in the coronary vessels that are triggered by changes in the driving pressure within a certain range. Increases or reductions in the driving pressure are immediately followed by similar alterations of coronary bloodflow. However, the resistance of the vessels is then changed - metabolically and mechanically - so that the *final* coronary bloodflow is maintained at control levels at all times (changes along the arrows in Fig. 9-1).

Autoregulation has been explained by at least two theories:

1) ***The myogenic theory*** considers autoregulation as a *myogenic* response - an intrinsic property of vascular smooth muscle. Increased stretch of the smooth muscle elicits contraction, whereas diminished stretch elicits vasodilatation. This is illustrated in Fig. 9-1, where an abrupt rise in perfusion pressure from 115 mmHg passively stretches the wall (increases the transmural pressure) and produce an initial increase in bloodflow. Then the vascular smooth muscles contract and the bloodflow falls along the arrows, so that the coronary bloodflow is maintained at 200-250 ml min^{-1}. Similarly, an abrupt fall in perfusion pressure from 115 mmHg has the opposite effect, so the normal bloodflow is re-established.

2) *The* ***metabolic control theory.***

Fig. 9-1: Autoregulation: Changes in bloodflow triggered by changes of the driving pressure has a tendency to be diminished. The example here is the coronary bloodflow, which is described further in relation to Fig. 10-7.

Metabolic control is the sum of all metabolic factors that match the oxygen supply to the energy requirement.

There is a remarkable proportionality between changes of myocardial oxygen consumption and coronary bloodflow. If the oxygen supply is insufficient compared to the myocardial demand, a vasodilatator is releases from the myocytes to the interstitial fluid, so the coronary resistance vessels dilatate.

Adenosine is continuously produced by breakdown of ATP. *Adenosine* is a likely candidate for the role of *metabolic mediator*, because it is such a potent vasodilatator and because it diffuses readily across the cell membranes. Adenosine may work via *presynaptic inhibition* of sympathetic nerve fibres to the smooth muscles of the coronary resistance vessels. Falling perfusion pressure leads to diminished rate of adenosine washout and thus to local vasodilatation. Adenosine dilatates the vessels and causes increased coronary bloodflow. Increased perfusion pressure washes out adenosine, which leads to vasoconstriction and local decrease in bloodflow until it is re-established. - The myogenic and the metabolic control frequently co-operate during autoregulation.

Reactive hyperaemia (ie, increased limb bloodflow following experimental vascular interruption) is probably explained by the metabolic vascular control theory.

Autoregulation protects not only the coronary bloodflow, but also the cerebral, intestinal and renal bloodflow to mention the most important organs.

2. Autonomic nervous control

The sympathetic and the parasympathetic division of the autonomic nervous system control the tone of the resistance vessels by opposing actions. Almost all blood vessels receive efferent nerve fibres from the *sympathetic* nerve system to their smooth muscles.

True capillaries do not contain smooth muscles and do not receive autonomic nerve supply. Metarterioles and capillary sphincters do not receive nerve fibres at all.

The sympathetic vasoconstrictor fibres and circulating catecholamines control both arteriolar, venous and venule tone. The vessels are innervated by postganglionic neurons from the paravertebral sympathetic trunk. The noradrenergic control releases noradrenaline and ATP. The transmitter transport is axonal. Noradrenaline binds to a-adrenergic constricting receptors. Adrenaline binds to both a-adrenergic *constricting* receptors and to b-adrenergic *dilatating* receptors. Consequently, adrenaline elicits vasoconstriction in arterioles where a-receptors predominate, and vasodilatation where b-adrenergic receptors predominate. Adenosine dilatates vessels, because it inhibits release of noradrenaline possibly via presynaptic purine receptors. In the synapse, the neurotransmitter is eliminated by re-uptake, by enzymatic breakdown and by diffusion. The arterioles of the skeletal muscles, the skin, the kidneys and the splancnic region are densely innervated.

Hunting predators are claimed to have sympathetic vasodilatator fibres to the skeletal muscle vessels, which is consequential during hunting, but such fibres have not been found in humans (Uvnaes).

The cholinergic system is almost exclusively *parasympathetic*. The vessels of the head, neck and thoraco-abdominal organs receive parasympathetic nerve fibres (the 3rd, 7th, 9th and 10th cranial nerves). The large intestine, bladder and genital organs receive parasympathetic fibres from the sacral segments 3-5. The nerve fibres to the external genitals are

active during sexual excitation. Acetylcholine is the vasodilatating transmitter for muscarinic and nicotinic cholinergic receptors. Purinergic receptors use vasodilatating transmitters as ATP, AMP and the potent adenosine.

Cholinergic sympathetic fibres innervate sweat glands and release acetylcholine as stimulus.

3. The baroreceptors

Rapid regulators of the arterial blood pressure are the *arterial baroreceptors* originating from the carotid sinuses and the aortic arch. These classical arterial pressor-receptors are well established and work within seconds following dynamic changes in blood pressure. The arterial baroreceptors probably do not regulate chronic blood pressure changes

with constant tone.

The baroreceptor reflex is triggered by stretch of the wall, and the receptors are also called stretch receptors or pressor-receptors. The baroreceptors are mainly located in the walls of the internal carotid arteries (known as the carotid sinuses) and in the aortic arch. Signals are transferred from each carotid sinus via afferent nerve fibres forming the sinus nerve to the glossopharyngeal nerve, and conducted to the nucleus of the solitary tract of the brain stem.

The impulse frequency in the nerve afferents increases with the arterial pressure maintained over a period (Fig. 9-2). The curve is S-shaped with a steep rise in the normal range of arterial pressures, indicating an optimal sensitivity in this area. There is no activity below 60 mmHg.

An increasing rate of pressure change (dP/dt; a sudden rise in pulse pressure amplitude) also increases the firing rate in a single nerve fibre (Fig. 9-2, right). Thus, baroreceptors act as *differential-sensors*. The frequency during the rising systolic pressure is distinctly greater than that in the diastole.

Fig. 9-2: Activity in the carotid sinus nerve at maintained arterial pressure (left) and during a single cardiac cycle with low, normal and high blood pressure.

Baroreceptors convey information about mean arterial pressure (MAP), pulse pressure, and the rate of pressure change (dP/dt). Arterial baroreceptor nerve fibres are *buffer nerves* concerned with *short-term buffering* of the blood pressure.

The afferent signals are conducted to the nucleus of the solitary tract in the medulla. This nucleus is the site confluence for both baroreceptor and chemoreceptor signals. Stimulation here *inhibits* sympathetic structures and *enhances* parasympathetic structures. Thus, a rise in arterial pressure causes vasodilatation and a fall in heart rate, both of which contribute to a lowering of blood pressure. A primary fall in arterial pressure elicits vasoconstriction and a rise in heart rate, both of which contribute to a rising blood pressure.

Change of body posture from lying to erect reduces the arterial pressure in the carotid sinuses, which elicits an immediate reaction with strong sympathetic tone and diminished vagal tone. This minimises the fall in brain blood pressure, and prevents loss of consciousness. – Hypotensive drugs, exposure to weightlessness, and immobilisation interfere with the baroreceptor reflex, which normally protects us during standing. Such individuals may develop *orthostatic hypotension*, when they stand up and they may faint.

Behavioural and emotional control of blood pressure and heart rate is exhibited by the *hypothalamus*. This autonomic control centre also includes a temperature centre from where contraction of skin vessels is instituted in cold environments.

In *hypertension* the baroreceptor system adapts to the rising pressure within days by moving up the set point. Patients with hypertension have stiff arterial walls as a result of the high arterial pressure, so their baroreceptors are less sensitive than in healthy persons. The increased arterial stiffness is not the main phenomenon in hypertension. Most hypertensive patients are dominated by increases peripheral vascular resistance, which mainly affects the diastolic arterial pressure.

Patients with *hypersensitive* baroreceptors in the carotid sinuses to external pressures are in danger of hypotension with fainting and death from external pressure over the neck at the site of the carotid sinus (so-called *carotid collar syncope or collar death*). Tight collars or other types of external pressures elicit fainting due to marked vasodilatation and hypotension. - Another cause is *emotional fainting* (vasovagal syncope) with a strong emotional activation of the vagus tone via hypothalamus.

Three types of regulators are involved in the adjustment of blood pressure. They are classified as short-term, intermediate-term and long-term regulators.

1. *The arterial baroreceptor reflexes* described above operate rapidly.

2. *Transcapillary volume shifts* in response to changes in capillary blood pressure, begin their function within

minutes. When veins are stressed by increased pressure, they slowly expand so that the blood pressure decreases. Conversely, when the intravascular volume decreases, the opposite occurs.

3. *Renal regulation of the body fluid volume*.

When arterial pressure rises, more urine is excreted. Hereby, the plasma and interstitial volume is reduced. The *diminished* plasma volume decreases venous return to the heart, reducing cardiac output, so that elevated arterial blood pressure is brought back towards normal (Fig. 9-6).

A *decrease* in arterial pressure elicits the opposite reaction: The *renin-angiotensin-aldosterone-cascade* is triggered (Chapter 24). Aldosterone from the adrenal cortex promotes Na^+-reabsorption and K^+-secretion from the renal tubules. The reabsorbed Na^+ augments water retention (Fig. 9-6), as does also increased vasopressin (ADH) secretion from the posterior pituitary. A falling arterial pressure also diminishes the release of atrial natriuretic peptide (ANP), and its Na^+ - and water- excreting actions are reduced (Fig. 9-6).

4. Oxygen release to the mitochondria

The factors that ease O_2-diffusion and delivery are:

1. *Myoglobin* in muscle cells releases O_2 during muscular contraction, when the blood supply is blocked. Myoglobin is important as a dynamic O_2 store in muscle cells, although myoglobin is not totally saturated with O_2. During muscular contraction the bloodflow is blocked, and the O_2 tissue tension falls drastically. Myoglobin then gives off O_2 to the cell. The P_{50} for oxymyoglobin is only 5 mmHg (compare to 27 mmHg for oxyhaemoglobin). Bloodflow is re-established during muscular relaxation. Thus, myoglobin is rapidly *reloaded,* even when there is only a small rise in O_2 tension.

2. *Heat energy* releases O_2 during work, since increasing heat energy equals increasing movement of O_2 molecules.

3. *Carbon dioxide:* With rising P_{CO2}, oxygen binding to haemoglobin decreases (Bohr effect, Fig. 8-3).

4. *Binding of 2,3 - DPG* (diphosphoglycerate) to haemoglobin eases the release of O_2 at low tensions (see Chapter 8, paragraph 3).

5. *Mitochondria* located close to capillaries have reduced diffusion pathway.

6. *Short distance* capillary networks, as following capillary recruitment, improve the oxygen delivery.

Oxygen is lipophilic. Since almost the entire capillary surface is identical to the lipid containing plasma membrane of the endothelial cells, oxygen is able to use the total capillary surface for diffusion. The transport of lipophilic molecules is *perfusion limited.*

Oxygen diffuses so easily over the capillary endothelium, that there is *tension equilibrium* between blood and tissues already at the ***arterial*** capillary end.

With rising perfusion the tension equilibrium point is shifted towards the venous part.

Due to the oxyhaemoglobin, the O_2 tension can be maintained through the entire capillary. The oxygen tension varies in the tissues. There is a longitudinal *tension drop* towards the venous end of the capillary, and radial tensions drop in the tissue itself. In brain tissue, the O_2 tension can vary from an arterial level in certain small areas (P_{aO2} of 13.3 kPa or 100 mmHg) towards zero, when bloodflow is insufficient.

Brain and heart tissues are extremely sensitive to a fall in P_{O2}.

Brain tissue is found in the nerve cells of the retina. These nerve cells are deprived of oxygen in 4.5 s (occurrence of *black out*). This can be verified by pressure on the upper eyelid. Consciousness is lost (*grey out*) a few seconds after cardiac arrest. After 90 s, the brain interstitial fluid [K^+] increases drastically from 3 to 60-70 mM, and both action potentials and synapse transmissions are eliminated. There is ion equilibrium over the cell membranes. Intracellular [Na^+] also increases drastically and intracellular brain oedema develops. A high extracellular [K^+] is life threatening.

The EEG of an anoxic brain is recognisable as a *straight* EEG trace (no electrical activity) indicating brain death. Because [Ca^{2+}] rises in the nerve cell, this increases the K^+ conductance, so that more K^+ leaks out into the interstitial fluid.

The kidneys only use 15 ml O_2 each min but they receive 25% of the cardiac output at rest (1200 ml per min containing 200 ml O_2 per l). The kidneys have the lowest arteriovenous O_2 content difference of all the larger organs in our body. The large *safety margin* is important for this vital organ during bleeding or when the renal bloodflow is reduced (more in Chapter 25).

5. Measurement of blood pressure

The arterial blood pressure is measured indirectly in the brachial artery with Korotkoff´s *auscultatory* method. WHO has proposed standardisation of this method. Continuous intra-arterial recordings can obtain exact arterial blood pressure measurements. Comparison with intra-arterial recordings have shown that Korotkoff´s method estimates the systolic pressure too low (about 10 mmHg), and the diastolic pressure differs a few mmHg.

The blood pressure increases in some patients due to the presence of a doctor (ie, *white coat hypertension*). This is revealed by repeated measurements – preferably performed before, during and following exercise.

Ejection of blood from the left ventricle triggers a pulse wave in the wall of the arterial tree, and the volume-pressure variations here distribute with a large velocity along the arterial tree. In young persons the velocity is 5-10 m per s; with age, atherosclerosis and hypertension the arterial tree becomes stiffer and the velocity increases (see Ch. 8 about compliance and also Fig. 8-8).

Fig. 9-3: Changes in pressure in the arterial tree of a supine healthy person.

The systolic pressure increases progressively along the arterial tree, whereas the diastolic and the MAP decrease (Fig. 9-3). The pulse amplitude, which is the difference between systolic and diastolic pressure therefore, increases clearly (Fig. 9-3). The end of systole is marked by a brief sharp fall in pressure (dicrotic notch), caused by the relaxation of the ventricle with backflow of blood as the aortic valves close. This backflow pressure moves with the blood all along the arterial tree (Fig. 9-3).

The blood pressure has to be measured repeatedly, with the patient sitting comfortably in a relaxed environment, and measured at more than three consultations in order to avoid false alarm with white coat hypertension. A diastolic pressure above 95 mmHg (12.6 kPa) expresses an increased MAP and the age of the patient influences the strategy of the treatment.

Fig. 9-4: Normal pressures in the circulation of a supine healthy person.

Essential *diurnal* variations are present, but repeated blood pressure measurements over three consultations seem to define a reasonable diurnal mean level. Continuous recording of the arterial pressure is sometimes necessary.

Patients below 40 years of age, with a diastolic pressure above 100 mmHg must be followed and examined further. Patients above 40 years of age, with a diastolic pressure above 120 mmHg, must be examined further.

Normal values for blood pressures measured in different locations of the circulation are given in Fig. 9-4.

6. Age and MAP

Populations living under *natural conditions* - including Indian troops in Brazil and healthy living persons in the Western Hemisphere - maintain their mean arterial pressure (MAP) throughout life. Their distribution curve for MAP is close to the normal distribution.

The MAP and the systolic pressure measured as an average for the total population, increases with increasing age in the rich part of the World.

As an order of thumb, the systolic blood pressure in mmHg is equal to 100 plus age in years, because these values are close to typical statistical mean values from examination of large population groups. This is because general diseases, with consequences for the systolic blood pressure and MAP, are accumulated with age in the Western Hemisphere. Quite a few of the accumulated disorders (such as atherosclerosis – see Chapter 10) probably occur as a consequence of our life style - operating in a heterogeneous genetic pool.

Previously, systemic hypertension was therefore characterised by a MAP larger than normal for the age. Practically difficult comparisons had to be made with a statistical, so-called normal material. Today, most doctors use the *WHO definition* (see below).

The MAP is a good estimate of the driving pressure, and the cardiac output is the stroke volume multiplied by the cardiac frequency. MAP and cardiac output are easy to determine, so the *TPVR* can be calculated.

With pressure expressed in mmHg and cardiac output expressed in ml per s, the unit for *TPVR* is 1 $mmHg*s*ml^{-1}$. This unit is complicated in writing and the abbreviation is 1 PRU (Pressure Resistance Unit). The normal value for *TPVR* in the systemic circulation at rest is one PRU, and during exercise it is only 0.3 PRU.

Pathophysiology

This paragraph deals with *1. Natural history of hypertension, 2. Symptoms and signs of hypertension, 3. Risk factors (Western lifestyle), 4. Types of systemic hypertension, 5. Therapeutic principles, 6. Future strategy.*

1. Natural history of hypertension

Primary hypertension always has a diastolic element reflecting involvement of the resistance vessels (eg, muscular arterioles etc). *Secondary hypertension*, caused by atherosclerosis or other types of stiff arterial walls, is often purely systolic.

In the early stages of hypertension, the arterial blood pressure is oscillating between hypertensive episodes and normal periods. The hypertensive episodes are typically dominated by sympathetic overactivity with increased cardiac output and almost unchanged *total peripheral vascular resistance (TPVR).* Eventually, the pressure changes the distensibility of the arteriolar walls and thus leads to sustain structural changes of the resistance system. As the hypertension develops the *TPVR* is increased.

Any rise in blood pressure is a strong stimulus to the high-pressure baroreceptors, but these essential sensors do not always work appropriate in hypertension. The expected bradycardia from the high arterial pressure acting normally on the arterial baroreceptors is not seen in hypertensive patients.

The initial sympathetic tone is also depicted in the high resting heart rate, in contrast to the bradycardia found normally, when the blood pressure rises. The abnormal baroreceptor reflex is probably an adaptive consequence of the variable but lasting initial hypertension.

Permanent structural changes of the resistance vessels, with strongly reduced specific compliance (reduced distensibility) and reduced lumen of arterioles and small muscular arteries, eventually leads to permanent hypertension.

The rising *TPVR* implies a rising workload for the left ventricle and thus creates left ventricular hypertrophy.

2. Symptoms and signs of hypertension

The typical patient with hypertension is *asymptomatic*. This is what makes the development of this disorder dangerous.

The first sign of systemic hypertension is sometimes acute myocardial infarction with sudden death. Of all acute cases of myocardial infarction up to 25% only experience a sudden pain, there is cardiac arrest, and the cases are recorded as sudden death from myocardial infarction at section.

Hypertonic patients with coronary artery disease experience angina at exhaustion or from myocardial hypertrophy.

Malignant or accelerated hypertension refers to a rapid and serious rise of the arterial blood pressure. The condition can start as paralysis, unconsciousness, or blindness.

Secondary hyperaldosteronism is recognised by high serum concentrations of renin and aldosterone. This occurs in *malignant hypertension* or following prolonged use of diuretics. The patients develop cerebral oedema and haemorrhage, cardiac failure and hypertensive nephropathy (with proteinuria and microscopic haematuria). Patients with malignant hypertension develop *dissecting aortic aneurysms* and *retinal damage with papiloedema,* so they die rapidly without specific therapy.

Ophtalmoscopy for hypertonic changes of the retina also provides the diagnosis hypertension. These changes include haemorrhages in the retinal nerve fibre layer, exudates as yellow-white spots called *cotton wool spots,* irregular arteriolar diameter, microaneurysms, and papillary stasis (Fig. 9-5).

Fig. 9-5: Hypertonic changes of the retina seen by ophtalmoscopy. The patient has malignant hypertension. - A normal retinal fundus is found in Fig. 6-5.

A necrotic arteriolitis is often found by ophtalmoscopy in malignant hypertension.

3. Risk factors (Western lifestyle)

Causative or *risk factors* for essential hypertension include *genes,* because there is a clear racial and familial accumulation of hypertension. A risk factor is a factor showing statistical covariance with the disease - see also Chapter 10.

Africans have higher arterial blood pressure than Caucasians, and some families accumulate cases of hypertension. Specific genes have not been identified.

The environmental factors are numerous, but *Western Hemispher lifestyle* is the key word, since the occurrence of increasing systemic blood pressure with increasing age is obviously related to accumulation of disease. However, accumulation of hypertension with age is not a law of nature.

Western lifestyle is sedentary, with psychological stress in career and family life. Existential procedures have to be performed rapidly including buying and eating fast food. Persons with a stressful everyday life, with smoking, alcohol and large meals following long work hours, practice little exercise if any, and become obese with hyperlipidaemia, hyperglycaemia and hyperuricaemia.

The hunting human has become a stressed user of automatic tools (cars, mobile telephones, household utilities, TV, PC etc). This lifestyle pattern frequently implies a serious *sympathetic overactivity* with a typical rise in resting cardiac rate and thus in cardiac output.

One essential and measurable variable in the life style pattern is the lack of exercise (eg, physical inactivity). A low maximum oxygen capacity or *fitness number* is measurable with the submaximal exercise test of Åstrand (Fig. 18-3), and reproducible in each individual. The fitness number is expressed as the *maximum oxygen uptake* in $ml*min^{-1}*kg^{-1}$.

A maximum oxygen uptake *below 34* ($ml*min^{-1}*kg^{-1}$) is related to *risk factor accumulation* and early death from hypertensive or other related complications (Fig. 18-14). Such a low maximum uptake is a clear indication of physical inactivity, where dilatation of muscular arterioles is seldom or almost never achieved.

The unknown cause of essential hypertension in the Western Hemisphere may well prove to be *physical inactivity* and

the related life style patterns described above.

In some cases of hypertension there is a clear relation to the renin-angiotensin-aldosterone cascade (Chapter 24)**.** The series of events starts with a rise in *TPVR* due to increased vascular tone. Over months and years, the walls of arteries and arterioles thicken and atherosclerosis is spread in the arterial tree. Such changes reduce the driving pressure in the renal arteries, which leads to a fall in glomerular filtration rate (GFR) and increased NaCl/water retention (Chapter 25). The falling pressure in the renal artery triggers b-receptors on the JG-cells of the juxtaglomerular apparatus (Fig. 25-17). Renin is released from these cells located in the afferent glomerular arteriole. Renin separates the decapeptide, angiotensin I, from the liver globulin, angiotensinogen. When angiotensin I passes the lungs or the kidneys, a dipeptide is cut off from the decapeptide by an angiotensin-converting enzyme (ACE). Hereby angiotensin II (octapeptide) is produced. Angiotensin II stimulates the aldosterone secretion from the adrenal cortex, and thus stimulates the Na^+ reabsorption and the K^+ secretion in the distal, renal tubules. The renin-angiotensin-aldosterone cascade further contributes to the salt and water retention. Angiotensin II is also a circulating vasoconstrictor just as adrenaline and vasopressin found in high plasma concentrations in many hypertonics. ACE inhibitors (see later) are rational choices for hypertonics with high angiotensin II, but also for other categories for reasons unknown (diabetics etc). The *cascade* is further described in Chapter 24 – paragraph 6.

4. Types of Systemic Hypertension

There are two forms of hypertension, I) primary or essential, and II) secondary hypertension.

I) Primary hypertension is a multifactorial syndrome without known cause. Approximately 90% of all cases are classified as primary or essential hypertension, because the causative factors are not clarified in detail. Increased peripheral resistance is responsible for most cases of primary hypertension.

Fig. 9-6: Factors contributing to systemic hypertension. Abbreviations: ECV = Extracellular fluid Volume; TPVR = Total Peripheral Vascular Resistance.

II) Secondary hypertension

In about 10% of all cases the cause of the hypertension is clarified, and these patients are classified as secondary hypertension. This condition must always be suspected in young hypertonics.

Renal, endocrine or cardiovascular diseases cause secondary hypertension or it relates to pregnancy or to drugs. Endocrine disorders are treated systematically in Chapters 26>30.

1. *Renal disorders* (Chapter 25) account for more than 80% of all cases of secondary hypertension. The disorders are chronic cases of glomerulonephritis, pyelonephritis and other permanent damage of the kidneys, where salt and water retention dominates. Hyperparathyroidism and Ca^{2+} overload can lead to *renal failure* and severe hypertension. A renal artery stenosis sufficient to reduce the glomerular pressure leads to renin release from the juxtaglomerular apparatus, aldosterone release and thus to *increased salt-water retention* (see the *renin-angiotensin-aldosterone cascade*, Chapter 24, paragraph 6). Renal artery stenosis (atherosclerosis or fibromuscular hyperplasia), chronic renal inflammation (glomerulonephritis or pyelonephritis), and congenital polycystic kidneys can lead to secondary, systemic hypertension. Renal function is examined with endogenous creatinine clearance and the renal vessels by scanning or arteriography. The plasma renin concentration is measured.

2. *Hyperaldosteronism* has a primary and a secondary form. Conn´s syndrome is *primary* hyperaldosteronism. This condition is characterised by an isolated rise in serum aldosterone, since the cause is hyperfunction of the zona glomerulosa of the adrenal cortex - not the renin release. Secondary hyperaldosteronism is a condition with abnormally high stimulation of the adrenal zona glomerulosa. The serum concentrations of the whole renin-angiotensin-aldosterone cascade are increased.

3. *Cushing's syndrome* describes clinical conditions with increased glucocorticoid concentration in the blood plasma. The classical *Cushings disease* is caused by *excess* liberation of ACTH from the adenohypophysis, but ACTH excess is also known to originate from ectopic ACTH producing tumours or from excess administration of ACTH.

- Non-ACTH related adrenal adenomas or carcinomas, glucocorticoid excess administration, and alcohol abuse (so-called *Pseudo-Cushing*) cause *Cushing's syndrome*. - The *dexamethasone suppression test* is described in Chapter 30.

4. A pituitary tumour producing an excess of growth hormone (Ch.28) causes *acromegaly*. The patient sometimes has a *diabetic* glucose tolerance test (Ch.27). These patients die from heart failure, IHD or hypertension.

5. *Phaeochromocytoma.* This is a tumour of the sympathetic nervous system (Ch. 28) releasing both noradrenaline and adrenaline. The signs are intermittent or constant systemic hypertension, tachycardia with other arrhythmias, orthostatic hypertension and flushing.

6. In the last three months of pregnancy some females develop hypertension, oedema and proteinuria (*pre-eclampsia* or *toxaemia of pregnancy*). If this condition develops into severe hypertension with fits and lung oedema, it is called *eclampsia*. This is a life threatening condition, which must be treated immediately with intravenous hydralazine or minoxidil, and if necessary termination of pregnancy. Hydralazine is orally active vasodilatators, which work by direct relaxation of smooth muscles.

7. *Drugs* such as steroids or oral contraceptives with high oestrogen, sympatomimetics, aldosterone, and vasopressin all cause severe systemic hypertension. Monoamineoxidase-inhibitors combined with tyramine (cheese) or wine sometimes cause hypertension. A careful medical history is helpful.

8. *Cardiovascular disorder* - as coarctation of the aorta - is the cause of hypertension in a few young patients. The coarctation produces a late systolic murmur. These hypertonics have a low pressure distal to the coarctation.

9. *Atherosclerosis* (see Chapter 10) is characterised by a special systolic hypertension frequently found in the elderly without any diastolic hypertension. These patients do not have any arteriolar disease.

5. Therapeutic principles

Systemic hypertension is a health threat to the person as a whole, since the untreated disease shortens life expectancy with approximately 20 years. Target organs for damage are the heart, aorta, brain, eyes and the kidneys.

The positive effect on life expectancy of a moderate reduction of an abnormally high systemic arterial blood pressure is well documented.

The simple *resistance model* presented in Eq. 9-1 is applied for the therapy of systemic hypertension. The *driving pressure* in the systemic circulation is equal to the *cardiac output* multiplied with the *Total Peripheral Vascular Resistance (TPVR).*

The cardiac output is equal to the *cardiac frequency* multiplied with the *stroke volume*, and the stroke volume depends of the *total blood volume*. *TPVR* depends of the degree of contraction of the resistance vessels and of the distensibility (eg, specific compliance) of the arterial system.

Principally, systemic hypertension is therefore treatable through one or more of the following strategies:

1. Reduction of the *total blood volume* (and thus the stroke volume) with diuretics results in reduction of the driving pressure,

2. Reduction of the *cardiac frequency* reduces cardiac output and thus the driving pressure,

3. Reduction of *TPVR* with vasodilatators reduces the driving pressure.

Two strategies of therapy and their combination are available: *Change of life style* with or without drug therapy. *Drug therapy* must usually be continued for the lifetime of the patient.

Life style modifications (relaxed duration exercise and healthy habits):

In healthy individuals, the opening of resistance vessels during exercise typically reduces the *TPVR* to 30% of the value at rest. This vasodilatation expresses an enormous capacity, which is only present in the resistance vessels of the striated muscular system at large. The only natural way to break the vicious circle described above is to maintain the dilatation capacity throughout life by frequent use of the locomotor system. The exercise must include large muscle groups for some time. The exercise must be relaxed and comfortable in order to become a life style. Other beneficial effects of *relaxed duration exercise* (such as walking, golf, jogging, swimming, badminton, tennis etc) is *improved glucose tolerance, weight loss, improved heart function, improved lipid profile, normal gastrointestinal functions and psychological benefits such as improved mood and a healthy sleep pattern.* Healthy food and drinking habits are important, and smoking has to be given up.

Hypotensive drugs can be divided into 5 categories:

5.1. Diuretics

Hypertensive patients seem to handle Na^+ just as healthy persons (see Chapter 25).

Initial administration of diuretics produce a pronounced renal salt and water excretion, which lead to a reduction in ECV, and a fall in systemic blood pressure. The urinary salt and water excretion returns to normal after several days, but the blood pressure remains at the reduced level. This is difficult to explain. Perhaps some diuretics have a direct relaxing effect on vascular smooth muscle in the arterioles or other vessels.

The different groups of diuretics are treated in Chapter 25.

5.2. β-adrenergic receptor blockers

β-blockers antagonise competitively the effects of adrenaline and nor-adrenaline on b-adrenergic *vasodilatating* receptors. The typical non-selective b-adrenergic receptor blocker is propranolol, which is a potent reversible antagonist at both b_1-and b_2-adrenergic receptors. Propranolol acts on the heart and reduces the chronotropic (reduced heart rate) and inotropic effect (reduced force and cardiac output); the reduced cardiac function is most pronounced during high sympatho-adrenergic activity, such as during exercise or stress, so the drug can release acute cardiac failure. The anti-arrhythmic effect of propranolol is probably due to its local anaesthetic action on cardiac cells including pacemaker cells. The effect of propranolol on hypertension is not clarified, since it seems to increase peripheral vascular resistance slightly. Simultaneously, propranolol reduces the release of renin from the juxtamedullary apparatus. This inhibits aldosterone secretion, and thus reduces the potassium secretion of the distal tubular system. The result is potassium retention, which is further aggravated by *β-blockade* of receptors on cell membranes, whereby the adrenaline-stimulated Na^+-K^+ pump is inhibited. Following meals containing carbohydrate and potassium, there is a release of insulin, which stimulates the Na^+ - K^+ pump, and thus the K^+ uptake in cells. Adrenaline also stimulates the Na^+- K^+pump through activation of b_2 - receptors, whereby the plasma-[K^+] is reduced. The normal effect of insulin is hypoglycaemia, which is compensated by lipolysis and glycogenolysis (with FFA and glucose liberation), by increased sympathoadrenergic activity. Propranolol inhibits lipolysis from adipocytes and glycogenolysis from hepatocytes, myocardial and skeletal muscle cells. This is a problem with diabetics or for patients with reduced glucose tolerance. b-blockade may lead to life threatening hypoglycaemia or a serious rise in blood pressure, if adrenaline release dominates. Propranolol is thus contraindicated in persons with diabetes, sinus bradycardia, partial heart block and congestive heart failure. Propranolol increases airway resistance, which is a hazard to patients with COLD or asthma, because of bronchoconstriction.

Many b-blockers act selectively, but all compounds have effects as described below:

Selective b_1-blockers acts on the cardiac b_1-receptors and reduces the force of cardiac contraction and thus lowers the blood pressure.

Blockade of b_1-adrenergic receptors located on the renin-secreting juxtaglomerular cells reduces the renin release and the blood pressure in persons with renin-dependent hypertension (eg, patients with a high renin level in the plasma

from renovascular disease).

Many b-blockers reach the brain tissue through the blood-brain barrier, and others reach the brain cells through the large fenestrae of the circumventricular organs. The CNS-effect is an inhibition of the sympatho-adrenergic output, and beneficial effects on paroxysms of panic and anxiety. The hypotonic CNS-effect is probably dominating, and explains the maintained lowering of blood pressure, although the initial reduction in cardiac output is often only temporary.

5.3. α_1-adrenergic antagonists

inhibit the effect mediated through noradrenaline released from sympathetic presynaptic fibres to the postsynaptic a_1-receptors and produce vasodilatation. Also a central effect of these compounds (doxazosin, prazosin) may be involved. The hypotensive efficiency of these drugs give rise to the main complication, which is a serious fall in blood pressure following the first dose.

5.4. Angiotensin Converting Enzyme (ACE) Inhibitors

Angiotensin converting enzyme is found to have the highest activity in the endothelium of the long pulmonary capillaries. Converting enzyme is a ***kininase II***, which convert the decapeptide, angiotensin I, to the vasoconstrictive octapeptide, angiotensin II. ACE inhibitors (captopril, enalapril, and lisinopril) reversibly inhibit converting enzyme and thus act as a vasodilatator of both resistance and capacitance vessels. Angiotensin II is a potent vasoconstrictor, in particular when its concentration in plasma is high. Patients with 100 pg l^{-1} or more of angiotensin II react beneficial on ACE inhibitors. Also other hypertonics such as diabetic patients reduce their risk of vascular insults by the use of ACE inhibitors for reasons unknown.

5.5. Calcium-channel blocking agents

Ca^{2+}-antagonists (amlodipine, nifedipine, diltiazem, and verapamil) acts as effective vasodilatators, because they relax the smooth muscles of the arterioles. They also inhibit the cardiac contractile force. Ca^{2+} -antagonists inhibits the Ca^{2+} -entry into the cells, because they bind to the proteins of Ca^{2+} -channels in the membrane. The overall effect is beneficial incongestive heart failure, because the vasodilation dimishes *TPVR* and thus reduces afterload. Hereby, the cardiac output is improved despite cardiac contractile depression.

6. Future strategy

- *Systemic hypertension is the most frequently diagnosed and treated risk factor for the development of atherosclerosis (including ischaemic heart disease).*
- *A risk factor is a factor showing covariance with atherosclerosis. The remaining risk factors for atherosclerosis are physical inactivity, hypercholesterolaemia, hypertriglyceridaemia, increased LDL concentration, smoking, diabetes, and familiar factors (genes, social inheritance or life style patterns).*
- *A rational strategy is to control the risk factors for the patients. A successful lowering of arterial blood pressure with a hypotensive drug must not be accompanied by an unrecognised consequential rise in other risk factors.*
- *Relaxed exercise is an alternative therapeutic strategy to antihypertensive drugs in many cases of essential hypertension.*
- *Mild and relaxed exercise has other beneficial effects, namely a consequential reduction of most of the known risk factors for atherosclerosis.*
- *Healthy food, exercise and drinking habits are important to hypertonics, and smoking has to be given up.*

Equations

The driving pressure (D*P*) in the systemic circulation is equal to the cardiac output (Q ∘) multiplied with the *TPVR* according to Poiseuille´s law:

Eq. 9-1: D*P* = *Q*∘ * *TPVR*.

This is a simple resistance model for circulating fluid, and the model is applied for therapeutic strategies.

Fick's first law of diffusion: The flux (*J*) of O_2 is equal to the diffusion coefficient of oxygen (D is 10^{-9} m^2 s^{-1}) multiplied with the concentration gradient (dC) per distance unit (dx) through a given area (A). Fick's first law is written:

Eq. 9-2: $J = (- D \times dC) \times A/dx$ (mol per time unit) with a diffusion gradient (dC) through the area A. Notice that D/dx is a permeability coefficient (m per s).

The first law can also be written: $J = (D \times DP \times A)/dx$.

Self-Assessment

Multiple Choice Questions

I. The following five statements have True/False options:

A. Nitrovasodilatators has side effects such as hypotension, reflex tachycardia and headache.

B. The blood-brain barrier is impermeable to all b-blockers.

C. The ACE-inhibitor, captopril, dilatates both arterioles and capacitance vessels.

D. Depolarisation of the vascular smooth muscle cell membrane opens voltage-gated Ca^{2+} -channels, whereby Ca^{2+}-ions enter the cell, combine with calmodulin and activate myosin light-chain kinase.

E. Thiazides have serious side effects such as hyperglycaemia (glucose intolerance), hypercholesterolaemia, hypokalaemia and hyperuricaemia.

II. The following five statements have True/False options:

A. MAP and Q∘ are easy to determine, so the *TPVR* can be calculated.

B. Angiotensin converting enzyme is a kininase II, which convert the decapeptide, angiotensin II, to the vasoconstrictive octapeptide, angiotensin I.

C. The driving pressure (D*P*) in the systemic circulation is equal to the cardiac output *(Q*∘) *divided* by the *TPVR* according to Poiseuille´s law.

D. Noradrenaline binds to a-adrenergic constricting receptors. Adrenaline binds to both a-adrenergic constricting receptors and to b-adrenergic dilatating receptors. Consequently, adrenaline elicits vasoconstriction in arterioles where a-receptors predominate and vasodilatation where b-adrenergic receptors predominate.

E. Ca^{2+}-antagonists (amlodipine, nifedipine, diltiazem, and verapamil) act as effective vasodilatators, because they relax the smooth muscles of the arterioles.

Case History A

A male, age 50 years, visits an ophthalmologist in order to have measured new lenses for myopia and astigmatism. Ophtalmoscopy reveals irregular vessel diameter, bleeding, yellow-white spots, and papillary stasis. The patient is advised to see his general practitioner, which finds a constant arterial blood pressure of 200/110 mmHg (26.66/14.66 kPa). The heart frequency is 85 beats per min and the cardiac output at rest is normal.

The patient is an office clerk, and also has a sedentary off- duty life. The patient is a heavy smoker using 40 cigarettes per day. His father had high blood pressure and died from cerebral infarction at the age of 62 years.

An X-ray of thorax reveals clear lung fields and left ventricular hypertrophy.

1. *What is the diagnosis?*
2. *What is the treatment of choice?*
3. *What is the main risk for this patient?*
4. *What happens in the lungs and the left ventricle of this patient?*
5. *Compare the left ventricular pressure-volume work rate of this patient to that of a healthy individual. Assume that they are both at rest with a cardiac output of 5 l per min. Assume that the healthy person has a mean arterial pressure of 90 mmHg (12 kPa).*
6. *Convert the work rate units used into watts, and explain the development of ventricular hypertrophy.*

Conversion factors are found in Symbols or here:

1 litre = 10^{-3} m^3. 1 mmHg = 133.3 Pa (N/m^2). 1 watt = 1 Nm/s = 1 J/s.

Case History B

A 59-year old office worker is known to have systemic hypertension. From the initial arterial pressure of 195/115 mmHg, he was brought down to a stable level of 160/95 mmHg by antihypertensive drugs. During work the patient suddenly collapses, and he is brought to hospital in an unconscious state with an arterial blood pressure of 75/45 mmHg. There are no signs of hemiplegia. Assume that the brain is hypoxic, and that the brain is producing lactic acid out of 30% of all glucose molecules combusted here. Among other values the blood glucose concentration is determined to 5 mM, and the arteriovenous glucose concentration difference increases to 300% of normal (0.5 mM). The cerebral bloodflow (CBF) is reduced to 50% of the normal value (650 ml min^{-1}). The total production by oxidative phosphorylation is 36 ATP per glucose molecule, and by anaerobic metabolism 2 ATP per glucose molecule.

1. *What is the most likely diagnosis?*
2. *Calculate the anaerobic and aerobic contribution to brain metabolism.*
3. *Calculate the net glucose flux and the ATP production in a normal brain, and compare the results to those of the patient.*

Case History C

A female, 66 years of age, complains of frontal headache. She has been treated for migraine for the last 40 years. The new headache is different from migraine. The doctor measures her arterial blood pressure to 195/115 mmHg (25.9/15.3 kPa). By ultrasound screening the length of her left kidney is measured to be half the length of the right. Renal arteriography reveals a stenosis of the left renal artery. The stenosis is relieved by balloon dilatation, where a catheter with a balloon at its tip is inflated at the right site. The success of the treatment is confirmed over the following weeks, where her blood pressure reach a level of 145/95 mmHg (19.3/12.6 kPa).

1. *What is the cause of her hypertension?*
2. *Explain the pathophysiological mechanism.*
3. *What is the most likely cause of her renal artery stenosis?*

Case History D

A female, age 22 years, is sitting on a bicycle ergometer with her calf muscles 0.9 m below heart level. She is at rest and the venous pressure is 10 mmHg (1.3 kPa) at the level of the heart. The oxygen uptake(V_{O2}) is 0.247 l STPD per min and the muscle bloodflow is 3 ml per min per 100 g of tissue (3 Flow Units, FU). The total weight of all her skeletal muscles is 30 kg.

Following 5 min of rest, she starts cycling, hereby increasing her oxygen uptake to 4.5 l STPD per min, and her muscular arterioles dilatate to reach a three-fold increase in inner radius. During exercise the arterio-venous O_2 content difference is 170 ml STPD per l, and the oxygen uptake in the skeletal muscles increases from 1 to 100 ml STPD per min per kg.

1. *Calculate the venous pressure in the calf muscles at rest.*

2. *Calculate the relative alteration of the muscular vascular resistance during exercise.*

3. *Calculate the driving blood pressure over the working muscles during exercise, where the arterial blood pressure is 170/70 mmHg (22.7/9.3 kPa) and the venous pressure in the calf muscles is reduced to 20 mmHg (2.6 kPa).*

4. *Calculate the rise in muscle bloodflow during exercise.*

Try to solve the problems before looking up the answers.

Highlights

- *The cardiac output(Q) is the stroke volume multiplied by the cardiac frequency. The heart must pump harder to provide a given Q° with increasing age, because the arteries become increasingly stiff with age.*
- *The distensibility or compliance of the arterial system diminishes with age due to atherosclerosis.*
- *When the pressure wave travels through the arterial tree, the arterial compliance is always less in the distal part of the system.*
- *The mean arterial pressure (MAP) is a good estimate of the driving pressure (ΔP), whereas the pulse pressure varies almost directly with the stroke volume.*
- *MAP and Q are easy to determine, allowing a calculation of total peripheral vascular resistance (TPVR). The systemic TPVR is one PRU at rest and 0.3 PRU during exercise.*
- *The velocity of the systemic arterial pressure wave varies inversely with the arterial compliance, whereby the velocity increases with age and with increasing degrees of hypertension.*
- *The high frequency components of the systemic arterial pressure wave are damped in the periphery, and the systolic peak components are elevated.*
- *The MAP varies directly with the Q and the TPVR.. The resistance model for circulating blood: $\Delta P = Q° * TPVR$ is applicable for therapeutic strategies in hypertension.*
- *The EEG of an anoxic brain is recognisable as a straight EEG trace (no electrical activity) indicating brain death. Because $[Ca^{2+}]$ rises in the nerve cell, this increases the K^+ conductance, so that more K^+ leaks out into the ISF.*
- *The arterial blood pressure is measured indirectly in the brachial artery with Korotkoff´s auscultatory method (standardised by WHO). The systolic pressure is recorded by the occurrence of a tapping sound, and the diastolic pressure is manifested by the disappearance of the sound.*

- *Continuous intra-arterial recordings can obtain exact arterial blood pressure measurements. Comparison with intra-arterial recordings have shown that Korotkoff´s method estimates the systolic pressure too low (about 10 mmHg), and the diastolic pressure differs a few mmHg.*

- *A risk factor is a factor showing covariance with atherosclerosis. The remaining risk factors for atherosclerosis are physical inactivity, hypercholesterolaemia, hypertriglyceridaemia, increased LDL concentration, smoking, diabetes, and familiar factors (eg, genes, social inheritance or unhealthy life style).*

- *Populations living under natural conditions - including Indian troops in Brazil and healthy living persons in the Western Hemisphere - maintain their mean arterial pressure (MAP) throughout life.*

- *In the rich part of the World, the MAP and the systolic pressure, measured as an average for the total population, increases with increasing age.*

- *Systemic hypertension - according to WHO - is defined as an arterial blood pressure exceeding 160/95 mmHg (21.3/12.6 kPa) for several months. The pressure increase is either systolic, diastolic or a combination.*

- *Systemic hypertension is the most frequently diagnosed and treated risk factor for the development of atherosclerosis including ischaemic heart disease.*

- *The cause of essential hypertension in the western world may well prove to be physical inactivity and related life style patterns.*

- *Relaxed exercise is an alternative therapeutic strategy to antihypertensive drugs. Mild and relaxed motion (such as walking, bicycling, golf, jogging, swimming, badminton, tennis etc) is utilised, whenever possible, in the treatment of essential hypertension.*

- *Mild and relaxed exercise has other beneficial effects, namely a consequential reduction of other known risk factors for atherosclerosis: Improved glucose tolerance, weight loss, improved heart function, and improved lipid profile, normal gastrointestinal functions and psychological benefits such as improved mood and a healthy sleeping pattern.*

- *Healthy food and drinking habits are important to hypertonics, and smoking has to be given up.*

- *A rational strategy for the future is to control the risk factors for the patients. A successful lowering of arterial blood pressure must be accompanied by improvement of other risk factors.*

Further Reading

Hypertension. Monthly journal published by the *Am. Heart Association*, 7272 Greenville Av., Dallas TX 75231-4596, USA.

Julian DG, Camm AJ, Fox KS, Hall RJC & Poole-Wilson PA (1995) *Diseases of the Heart*, 2nd Edn. London: Bailliere Tindall.

Katzung BG. *Basic & Clinical Pharmacology*. Appleton & Lange, Stanford, Connecticut, 1998.

Return to content

Return to chapter 9

Chapter 29.

Sexual Satisfaction, Reproduction And Disorders

Study Objectives

- To *define* amenorrhoea-oligomenorrhoea, dyspareunia, gynaecomastia, hypogonadism, impotence, infertility, menarche, menopause, menstruation and phases of the menstrual cycle, oligospermia-azoospermia, sterility, and virilization.
- To *describe* anticonception, anovulatory cycles, bleeding disturbances, castration, cryptorchism, postmenopausal hormonal alterations, puberty, anabolic steroids and doping, genetic and psychosocial sexual disorders.
- To *explain* the effect of anabolic steroids, the normal menstrual cycle, conception, implantation, pregnancy, pregnancy tests, birth and suckling. To explain the normal ovarian and testicular function, gametogenesis, erection, ejaculation and sexual satisfaction (orgasm). To explain the effect of androgen-binding protein, inhibin, aromatase, and the biosynthesis of steroids.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The gonads are concerned with the well being and preservation of the human race.*
- *The sperm decides the genetic sex (genotype). The Y chromosome is a constant determinant of maleness.*
- *Foetal differentiation of the genital ducts and of the external genitalia requires foetal gonadal hormones. The foetal genital tract will always develop into female genitals, if unexposed to embryonic testicular secretion.*

Definitions

- **Amenorrhoea-oligomenorrhoea** are terms used for absence - irregular, infrequent menstrual periods. These signs suggest *female hypogonadism*, when pregnancy is excluded.
- **Azoospermia** describes absence lack of sperm in the ejaculate
- **Dyspareunia** refers to female pain or discomfort during intercourse.
- **Gametogenesis** is the formation of ova and sperm. The primitive germ cells are divided by *meiosis*, so the number of chromosomes is halved (22 autosomes and one sex chromosome).
- **Genetic sex** is determined by the presence or absence of the Y chromosome. The Y chromosome determines the development of testes and maleness. The Y chromosome contains a *sex determining region* (the SRY gene), which encode the *testis determining factor* (TDF).
- **Genital sex** is the *phenotypic* sex (apparent female or apparent male).
- **Gonadal sex** is determined by the presence of normal *ovaries* or *testes*.
- **Gynaecomastia** refers to the occurrence of female breasts in males. The causes are HCG-producing tumours, oestrogens or oestrogenic drugs.
- **Hypogonadism** (male) refers to a condition with small, soft testes producing little sperm and testosterone. The

condition is usually found with subfertility.

- **Impotence** is inability of the male to produce an adequate erection for satisfactory sexual intercourse.
- **Infertility** (subfertility) is a diagnosis used on a couple, which has been unable to conceive during one year of unprotected intercourse. The causes are oligospermia, tubule blockage, ovulatory disorders, or combined problems with both persons in the couple.
- **Menarche** refers to the age at the first menstrual period.
- **Menopause** refers to cessation of periods, which usually occurs around the age of 50 years.
- **Menstruation** is the onset of spontaneous regular uterine bleeding.
- **Oligospermia** refers to reduced numbers of sperm in the ejaculate. The causes are primary testicular disease or blockage of the vas deferens.
- **Puberty** is the transition period from a non-reproductive to a reproductive state.
- **Sterility** refers to individual infertility. Chemotherapy and other drugs may cause sterility. Surgical blockage of the tuba or the vas deferens results in sterility.
- **Virilization** is the occurrence of male secondary sex characteristics in the female.
- *Definitions of other genetic concepts are given in* Chapter 31.

Essentials

This paragraph deals with *1. The sexual drive* , *2. Sex before birth*, *3. The menstrual cycle*, *4. Ovulation/Female orgasm*, *5. Conception*, *6. Breast development*, *7. Labour*, *8. Efferent activity during coitus*, *9. Sex hormones*, *and 10. Male puberty.*

1. The sexual drive

We feel the *sexual drive* or desire for sex *(libido),* when *sex-related areas* in the higher brain centres are stimulated. These centres include the limbic system, Stria terminalis and the preoptic region of the hypothalamus. The desire for sex is increased by androgens in both sexes.

The sex desire of females is variable - for some it increases near the time of ovulation, when *oestradiol* secretion is increasing, while others experience a peak drive near menstruation. The CNS cells involved (see above) must contain *sex hormone receptors*. Sex hormones are steroids. They are lipid soluble and pass the cell membrane easily. After binding to *cytoplasmic receptors* (the steroid-thyroid family), the receptor-hormone complex translocates to the cell nucleus. Here the information is transcribed and translated. The result is release of new proteins with the same information into the cytosol, where the physiologic response is triggered. *Castration* is assumed to reduce female libido minimally, but male libido is most often lost. Removal of one testis need not change the male libido. These clinical observations reflect psychosocial differences, and not necessarily a different libido mechanism in the two sexes. Hypothyroid persons lose their sex drive. The sex desire (libido) is stimulated by a multitude of *sense impressions* (visual, auditive, olfactory, and psychological). *Potency* refers to the ability to engage in intercourse.

The *brain* is an important sex organ. Obviously, any natural *body contact* can be considered part of a healthy sex life - including the penetration of the penis in the vagina.

Sexual satisfaction is synonymous with *orgasm* in Western cultures. *Orgasm* is the psychological climax or the culmination of *total commitment* in a sexual act that is accompanied by a series of physiolo
orgasm involves spinal cord reflexes similar to those involved in male ejaculation (see later). One very important reaction is ovulation, which is an automatic consequence of copulation among many animal species and periodically in

humans.

Sexual enjoyment covers several phenomena. For example the *fetishist satisfaction* of wearing the clothes of the opposite sex. This is the important part of a transves
sense of the term, since they do not need partners. Some individuals prefer *masturbation (onany)* as a substitute for partnership. Many individuals prefer heterosexual contacts; others prefer homosexual activities, while bisexuals may prefer either sex - depending on the circumstances. Sexual activities can vary. Besides, homosexual activity, oral sex, anal sex and many other variants are not uncommon.

2. Sex before birth

Normal sexual development in the embryo involves several processes. The sperm, which can be an X or an Y chromosome sperm, decides the *genetic sex* or sex genotype (Fig. 29-1). The genetic sex is independent of the ovum.

If the ovum is fertilised by an X spermatozoa (22 + X-chromosomes) the offspring is XX, a female. If the ovum fertilises by an Y spermatozoa (22 + Y-chromosomes) the offspring is XY, a male (Fig. 29-1).

Fig. 29-1: The sperm decides the genetic sex. The presence of the Y chromosome is the determinant of maleness.

Sex differentiation in the embryo usually harmonises with the *sex genotype*, but hormonal disturbances can lead to abnormalities. Proliferation of non-germinal and germinal cells in the *genital ridge* creates the gonadal *primordia*, which develops into a cortex surrounding the medulla. Until the 7th week of gestation, each sex has a bipotential system (the sexual indifferent stage) with both Wolffian and Müllerian ducts. The *urogenital sinus* develops into the external genitals in both females and males.

Around the 7th week, the medulla of the primitive gonad begins to differentiate into a *testis*, if an *Y chromosome* is present. This is because the Y chromosome contains the so-called *SRY gene* (the sex determining region of Y), which encodes the *testis-determining factor*.

As the testes grow and their Leydig cells start to produce testosterone, the *Wolffian ducts* develop into the male reproductive tract (epididymis, vas deferens, seminal vesicles and the ejaculatory ducts), whereas the *Müllerian ducts* regress. Testosterone stimulates the growth and differentiation of the Wolffian ducts in the male. The regression of the Müllerian ducts is caused by the *antimüllerian hormone* from the Sertoli cells.

Conversely, in the female, the cortex of the indifferent gonads differentiate into *ovaries*, if only two X chromosomes are present and no Y. In the female foetus, where there is a developing ovary and no antimüllerian hormone, the Müllerian ducts develop into the female reproductive tract (the uterine tubes, uterus and the upper vagina), and the Wolffian ducts degenerate because the ovary does not secrete testosterone.- When a normal female foetus is exposed to androgens during the period of differentiation of the external genitalia, an *apparent male* can result.

Visible differentiation of the gross anatomy does not appear until late in the second month of embryonic life. Testosterone causes the *differentiation* of the foetus to a male. The foetal genital tract will always develop into female genitals, if unexposed to embryonic testicular secretion. The *genital sex* is a phenotypic female. If testosterone is present, male external sex organs develop and the *genital tubercle* elongates to form the male phallos. If testosterone is absent, female organs develop instead. It is the action of *testosterone* and *5-α-dihydrotestosterone* on the urogenital sinus that is behind the normal development of the male external genitalia. In the last months of gestation the growth of the external genitalia depends upon foetal pituitary LH.

One population of cells in the indifferent gonade develops into the granulosa cells of the ovarian follicle and the Sertoli cells of the testicular seminiferous tubules. These cells support and mature the germ cells. – Another population of so-called *interstitial* cells develop into the theca cells of the ovary and the Leydig interstitial cells in the testis. The Leydig interstitial cells secrete testosterone, in response to *human chorionic gonadotropin* (hCG) from the placenta.

The presence of normal ovaries or testes determines the *gonadal sex*. Without normal ovaries or testes any *genetic sex* will develop into an *apparent female*.

Foetal plasma growth hormone (GH) concentrations are high, but GH-receptors are deficient and foetal GH is not essential for linear growth. Prolactin and placental GH act as growth factors and induce the presence of IGF-1 and IGF-2. A small transfer of maternal thyroid hormone is important for early foetal development. At birth, the babys own

thyroid hormone is important for CNS development and somatic growth. Foetal PTH stimulates the Ca -transfer across the placenta and controls plasma-Ca^{2+}. Foetal ACTH is important late during gestation in particular at birth, and the cortisol concentration is high in umbilical cord plasma. Foetal pancreatic a-and b-cells are functional by 14 weeks of gestation, but their release of glucagon and insulin is low.

In 1949 Barr et al. found a densely coloured body in the periphery of the nucleus (the *Barr body* or *sex chromatin*) of the buccal mucosa of females. The Barr body is also present in other individuals with two or more X-chromosomes in each cell. Individuals with one sex chromatin (Barr body) also have a *drumstick* attached to a small fraction of their leukocytes (Fig. 29-6). We find sex chromatin and drum sticks in cells, whether they divide or not. Chromosomes are only visible in dividing cells. The maximum number of sex chromatin and drumsticks is always one less than the number of X-chromosomes (Fig. 29-6).

3. The menstrual cycle

The *menarche* is the age at the *first menstrual bleeding.* It often occurs between the 12th and the 14th year.

LH and FSH are coordinators of gonadal function. The secretion of these pituitary gonadotropins is regulated through negative feedback by the plasma concentration of gonadal steroids. LH stimulates the interstitial cells of the ovaries (and testes), but LH also acts on female granulosa cells. LH binds to a LH-receptor, which spans the cell membrane several times. The LH receptor acts via adenylcyclase and with cAMP as a second messenger. Prostaglandins may increase the cAMP effects. Maintained stimulation by LH down-regulates the number of LH-receptors on the surface of gonadal cells.

FSH acts on ovarian granulosa cells (and testicular Sertoli cells) by binding to FSH-receptors, partially homologue with the LH-receptors. The increase in cAMP following FSH-receptor binding *transcribes* the *aromatase gene* and stimulates oestrogen synthesis. FSH stimulates synthesis of inhibin and peptide/protein products from granulosa and Sertoli cells. FSH amplifies the sensitivity to LH by increasing the number of LH-receptors on granulosa cells.

LH and FSH increase glucose oxidation, lactic acid production and protein synthesis.

The menstrual cycle starts at the *first day of bleeding* (menstruation). The bleeding is due to decrease of oestrogen and progesterone secretion. The FSH and LH secretion start to rise and stimulate the growth of several follicles - in particular following the bleeding. One of these – the dominant follicle – select itself by outstripping the others and grow so fast that the follicle can protrude more than 10 mm from the surface of the ovary. The dominant follicle has an increased oestrogen synthesis due to increased aromatase activity. Oestrogen from the granulosa cells of the dominant follicle binds to specific, *cytoplasmic receptors* (of the steroid-thyroid-family) in the endometrial and other uterine cells. Oestradiol activates and stimulates formation of *oestrogen* and *progesterone receptors*.

Fig. 29-2: The menstrual cycle in a female.

Oestrogen increases the *thickness of the endometrium*, the *size of the myometrial cells* and the *number of gap junctions* thus allowing the myometrium to work as a unit. The oestrogen phase is also called the *proliferative phase*. The concentration of sex hormones in plasma is shown in Fig. 29-2. Oestrogens work synergistically with progesterone to release gonadotropins by *positive feedback* just before ovulation.

Following the rupture of the follicle (*ovulation*), the *corpus luteum* produces increasing amounts of progesterone in addition to oestradiol also from a new developing follicle (Fig. 29-2).

Due to the *priming effect* of oestrogen on progesterone receptors, both hormones stimulate the growth of the endometrial glands, so that they curl like a helix. The progesterone effect in particular provides the endometrial/myometrial tissues with their *high secretion* and *bloodflow*, so the uterus is prepared to receive the fertilised ovum. During sexual stimulation the vaginal fluid secretion increases, as does the bloodflow of the organs involved.

If fertilisation does not occur, the level of oestradiol and progesterone switches off both gonadotropins. The corpus luteum fades out and degenerates with no LH to support it (Fig. 29-2).

The ovarian hormones almost cease to flow, and the uterus is deprived of their stimulating action. Therefore the *uterus shrinks* and sheds its swollen lining.

On the first day of the menstrual bleeding, the low progesterone and *high prostaglandin* level probably releases enough Ca^{2+} to start *spontaneous contractions* of the myometrial cells. Ca^{2+} -ions enter myometrial cells and stimulate their activity in the secretory (progesterone) phase.

The gap junctions *synchronise* these contractions, so that they include the whole myometrium. This can make excretion of blood and necrotic cells (containing prostaglandins) extremely painful. Prostaglandins dominate in menstrual fluid and stimulate the spontaneous activity of the human myometrial cells. A *normal bleeding* corresponds to a *loss of up to 50 ml of whole blood*. The mixture of vaginal fluid and menstrual blood produces a pH close to that of normal blood. The average cycle length is 28 days.

ADH (vasopressin) secretion from the neurohypophysis can cause *pre-menstrual tension* and an unpleasant increase in body fluid volume.

4. Ovulation/Female orgasm

Ovulation

A sudden increase in the plasma level of oestradiol maintained for more than 24 hours can increase FSH output by *positive feedback*. This is called the *positive feedback release ovulation*. The pulsatile release of GnRH from the hypothalamus is possibly stimulated by the high oestradiol concentrations in mid-cycle and oestradiol increase the number of GnRH receptors on the gonadotropic cells of the anterior pituitary. A neural hypothalamic pulse generator has been proposed to be involved in ovulation, and in some cases female orgasm triggers ovulation.

At lower plasma levels oestradiol is a potent inhibitor of GnRH secretion and thus of FSH and LH output (*negative feedback*). The negative feedback forms the basis for the ovulation-inhibition by contraceptives.

LH binds to a membrane LH-receptor and acts via a G-protein, adenylyl cyclase and cAMP. LH mobilises cholesterol and its conversion to progesterone.

FSH acts on ovarian granulosa cells and testicular Sertoli cells by binding to a membrane receptor homologous with the LH-receptor. The binding increases the transcription of the aromatase gene, the oestrogen and the inhibin synthesis.

The primary inhibitor of FSH secretion is the peptide, *inhibin*, that is secreted by the ovary (and testis), and blocks the effect of GnRH.

The oestradiol release from the *dominant follicle* increases sharply in the last part of the follicular phase. This triggers the *preovulatory surge* of gonadotropins (LH and FSH).

The LH surge induces an enzyme that increases the synthesis of leukotrienes, prostaglandins and thromboxanes. These molecules create an inflammation that causes rupture of the follicle. LH continues to act on the follicular granulosa cells, turning them into a yellow endocrine organ, the corpus luteum.

Orgasm

The *time for preplay* including *clitoral and multifocal stimulation* is important for most females. A clitoral orgasm in the preplay often triggers more female orgasms later during the intercourse. Female orgasm is released from the spinal cord reflexes via sympathetic signals in the pudendal nerves.

Two persons with a simultaneous sexual drive must have the necessary time for the sexual act. If they are also in love, it is natural to explore and use all means to satisfy each other.

Years ago, when the Kinsey report was made, the average duration of sexual intercourse was measured in seconds in the US. American males able to ejaculate even faster were assumed to be particularly virile. Today, such a short performance is considered a male disease called *premature ejaculation*.

5. Conception and pregnancy

Conception

Approximately 100-200 million sperms are produced each day of the fertile lifespan. The female foetus may contain 6 million oocytes, but the number decreases throughout her life (less than half a million at puberty and she may have

500 ovulations before the menopause).

The *autonomic moving* spermatozoa passes through the uterus while prostaglandins inhibit their spontaneous activity. The spermatozoa can keep their *vitality* for more than *2 days*, if they reach the fallopian tube. They lose their protective cover in the fallopian tube. The head of the spermatozoa *swell* and liberates *proteolytic enzymes*. These enzymes can dissolve the zona pellucida around the egg (oocyte). All these events in the spermatozoa takes days before it meets with the oocyte. The *oocyte* can only live *12-24 hours* without conception.

Pregnancy

Many sperms bind to the zona pellucida, but only one penetrates the wall – and blocks the entry of other sperms. Fusion of the two sex cell membranes forms the zygote, and the mitosis is complete within 24 hours.

The zygote passes into the uterine tube within a few days, protected against other spermatozoa by an increased permeability for K^+, so that the zygote membrane hyperpolarises. *Peristaltic movements* of the tube and ciliary motion conduct the zygote to the uterine cavity while undergoing *cleavage division*. Each cell is capable of developing into a complete human being up to the *eight-cell stage.*

At the *morula stage*, the cells start to develop into the *inner* cell mass or *blastocyst*, and the trophoectoderm or *trophoblast*. Seven days after conception, the blastocyst loses the zona pellucida and *implants* in the wall of the uterus (*nidation*). Nidation depends on prior conditioning of the endometrial stromal cells by progesterone bringing it into the proliferative phase. The stromal cells accumulate nutrients and swell or *decidualize* around the blastocyst. Endometrial *laminin* and *fibronectin* facilitate adhesion. *Histamine* and *prostaglandins* increase the permeability of the vessels around the nidation site. More than 2/3 of all conceptions result in miscarriage, because of insufficient attachment or other anomalies.

The *foetal trophoblast*, which give rise to the *extra-embryonic tissues* differentiates into two cell types. An inner layer of *cytotrophoblasts*, and an outer layer of *syncytiotrophoblasts*. The cytotrophoblasts synthesise stimulatory hormones such as CRH, GnRH, TRH and steroids.

The syncytiotrophoblasts synthesise first of all *human chorionic gonadotropin* (hCG). The *β-group* of hCG is specific and detected in maternal plasma 6 days following conception by specific antibody methods. The hCG is detectable in the urine within 9 days after conception.

The placenta is a fantastic hormone factory, which produces large amounts of hCG, relaxin. oestradiol, progesterone and human chorionic somatomammotropin (hCS or human placental lactogen, hPL). The hPL is synthesised from the 4. week of gestation. The hPL stimulates maternal lipolysis and inhibits insulin effects, causing hyperglycaemia.

The hCG is chemically related to TSH, FSH and LH. The hCG acts like LH and binds to the LH-receptors. The secretion of hCG is stimulated by GnRH produced by cytotrophoblasts. This is what keeps *corpus luteum* in being, and the pregnancy continues.

During pregnancy, *hCG* thus conserves the *corpus luteum*, taking over the role of LH.

The secretion of hCG stimulates ovarian release of progesterone and oestrogens just like LH. The hCG stimulates production of relaxin, inhibits the maternal secretion of LH and stimulates the maternal thyroid gland causing struma or hyperthyroidism in some pregnant females. Inhibin A from foetal trophoblasts peaks within the first week and suppresses maternal FSH secretion. Inhibin B concentrations remain low throughout gestation. LH and FSH concentrations in foetal plasma peak in mid gestation.

Fig. 29-3: Variations in plasma hormone concentrations during a normal pregnancy (42 weeks).

The plasma [hCG] reaches a peak value after *10 weeks* of pregnancy, when the syncytiotrophoblast count is maximum (Fig. 29-3). - Shortly after delivery hCG disappears.

The first peak on the plasma progesterone curve is progesterone produced by corpus luteum. The placenta takes over the progesterone production during the remaining pregnancy period ending with a peak concentration before birth. Progesterone protects the foetus in the uterine cavity by stimulation of endometrial glands that nourish the zygote and by maintenance of the decidual cells. Progesterone inhibits uterine contractions (inhibits prostaglandin synthesis and oxytocin sensitivity).

The *foetus and the placenta* form a foetoplacental unit. It produces all the hormones necessary for a successful pregnancy. Steroid precursors are delivered from both the foetus and the mother. Oestrogen (oestradiol, oestrone, and oestriol) concentrations rise steadily throughout pregnancy (Fig. 29-3). Oestrogens stimulate the growth of the myometrium and of the ductal system of the breast. Oestriol production independs on the foetal adrenal cortex, so maternal plasma oestriol provides an estimate of the foetal condition.

The placental progesterone blocks the menstrual cycle of the mother. Pregnant females therefore develop *amenorrhoea*.

6. Breast development

During puberty FSH, LH, growth hormone, and insulin are important for the breast development. The thyroid hormones (T_3/T_4) are *permissive*. Before puberty, plasma LH and FSH concentrations are low. There is no reaction to the low concentrations of gonadal steroids and inhibin.

Oestrogens are growth factors for the myometrium and for the ductal system og the breast during pregnancy. At the end of pregnancy there are other hormonal events. Progesterone secretion reaches a peak and then falls. This fall in progesterone allows the pituitary to release *prolactin* (LTH).

Prolactin from the maternal pituitary rises throughout pregnancy. Prolactin acts on the enlarged mammary glands turning them into *milk producers.* Prolactin develops the milk producing acini in the breasts during pregnancy. *Prolactin Inhibiting Factor* (PIF or *dopamine)* from the brain inhibits the prolactin secretion.

The baby's suckling stimulate the secretion of prolactin and oxytocin, but oestradiol and sexual stimulation is also involved. The mechanical stimulation of the breast increases the secretion of prolactin from the pituitary, but the response is strikingly reduced by alcohol.

Prolactin is important for the development of the *mammary gland tissue*, oxytocin, however, governs the *ejection of milk* during lactation. *Oxytocin* causes contraction of the myoepithelial cells in the milk ducts (just as it does in the myometrial cells).

Mother-milk contains long chain fatty acids that are essential for brain development. Suckling babies are protected against juvenile diabetes in comparison to non-suckling babies. Cow's milk contains much more protein and less lactose than human milk.

7. Labour

When the foetus has reached a *critical size*, the myometrial fibres are stretched, which increase their contractility.

At the end of pregnancy the *uterus* is sensitised by *oestrogen*. After a high peak in progesterone secretion the progesterone output *falls*. This fall in progesterone allows the uterus to respond to *oxytocin*, whose release is the *final trigger for parturition*.

The foetal pituitary-adrenal axis signals to the placenta a decrease in the *progesterone-oestrogen ratio* acting on the myometrium. This increases myometrial contractions that are mediated by prostaglandins (PGE_2 and PGE_2-a). A local increase in prostaglandin concentration increases myometrial cell Ca^{2+} and triggers uterine contractions. The density of oxytocin receptors in the myometrium increases throughout gestation and particular at term.

The role of the stable plasma concentration of *maternal oxytocin* at parturition is an enigma. Oxytocin is released according to a pulsatile pattern. The frequency of oxytocin pulsations increases at labour. This fact does not exclude an important role of oxytocin in normal human parturition.

Therapeutic doses of oxytocin initiate labour in most cases at the end of gestation,

The foetal cortisol production prepares the foetus to adapt to extrauterine life by stimulating lung maturation, by increasing the hepatic glycogen stores, and by promoting closure of the *ductus arteriosus* (Fig. 12-7).

Relaxin is an insulin-like polypeptide produced by the corpus luteum and placenta. The hormone relaxes pelvic articulations, suppresses myometrial contractions and softens the uterine cervix in order to facilitate passage of the foetus.

Several other factors are involved in human labour, but the exact trigger mechanism remains unclear.

8. Efferent activity during coitus

The activity in males is described as an example. The typical sequence of efferent events in the male includes erection, emission of semen and ejaculation.

Erection means penile rigidity and elongation due to parasympathetic vasodilatation. Psychological factors trigger penile rigidity, and sexual thoughts can cause erection, emission and ejaculation. The penis contains erectile tissue located in two dorsal *corpora cavernosa* and in a single ventral *corpus spongiosum*. All the cavernous spaces of the three penile corpora receive blood from thick-walled arteries ending centrally in each corpus. The blood leaves the cavernous spaces through thin-walled veins starting peripherally. Tactile stimuli, especially from the very sensitive *glans penis* activate sensory, somatic fibres in the pudendal nerve, whereby impulses reach the *sacral plexus*. Parasympathetic impulses (S_2-S_4) from the sacral plexus elicit dilatation of the arteries and constriction of the veins in penis. The cavernous spaces are hereby filled with blood under high (arterial) pressure within seconds, causing the penis to become hard and elongated for penetration. - Erection occurs quite normally during the REM phases of sleep.

Emission is caused by sympathetic contraction of smooth muscles (in epididymis, ducts and glands), which drive the fluids into the posterior urethra. Oxytocin ejects sperm into semen. Two exocrine glands near the neck of the bladder (the seminal vesicles and the prostate) secrete fluids that nourish the sperm and transport it through the urethra during the sexual act. The prostate gland supply an alkaline secretion containing Ca^{2+}, Zn, and phosphatase to the ejaculate. The seminal vesicles supply fructose and prostaglandins. These two secretions neutralise the acid semen and help propel the spermatozoa towards the ovum. Seminal fluid also contains gonadotropins, sex hormones, inhibins, endorphins, relaxin, proteases and plasminogen activator. Epididymis supply sperm-coating proteins.

Ejaculation: Ejaculation is a sympathetic response. Contractions of skeletal muscles expel the semen from the urethra in a rhythmic pattern. Signals from glans penis reach the lumbar region of the spinal cord through afferent fibres in the internal pudendal nerves. Filling the posterior urethra with semen triggers sensory impulses that travel through the pudendal nerves to the spinal cord. The spinal cord transmits rhythmic signals to the *skeletal ejaculation muscles* (the ischio-and bulbo-cavernous muscles and those of the pelvis). These rhythmic signals stimulate rhythmic contractions that expel the semen from the urethral meatus into the female genitals. – A typical ejaculate contains 300 million spermatozoa in 3 ml.

Inside the female genitalia the sperm is subject to the process of *capacitation,* which takes place within 6 hours. The sperm head is coated with substances from the ejaculate, Ca^{2+} enters the sperm, sperm motility increases, and the ability to penetrate the ovum is enhanced. The acrosomal membrane fuse with the outer sperm membrane, so that pores are formed and proteolytic enzymes can reach the surface of the sperm head.

9. Sex hormones

Sex hormones are oestrogens, progesterone, androgens and eichosanoids. Steroid synthesis in the gonads begins with cholesterol from acetyl Coenzyme A, and is almost identical to that of the adrenal cortex.

Oestrogens stimulate the female genitals and act to produce female secondary sex characteristics when a female enters puberty. Oestrogens and progesterone all enters the cell cytosol easily and bind to cytoplasmic receptors of the steroid-thyroid family. Oestradiol increases the synthesis of *oestrogen*- and *progesterone*-receptors.

These sex characteristics include the progressive growth of fallopian tubes, uterus, vagina, and external genitalia; also the fat deposition in breasts, buttocks, and thighs (Fig. 29-4). The ductal and stromal growth of the breasts is initiated just as the general growth at puberty with increased RNA and protein synthesis in the body cells. Oestrogens stimulate secretion of prolactin from the pituitary lactotrophic cells, increase the thickness of the endometrium and the size of the myometrial cell and their number of *gap junctions*. Oestrogens stimulate the hepatic production of essential proteins (eg, TBG, blood clotting factors, plasminogen, and HDL), but they inhibit formation of antithrombin III and LDL. Retention of salt and water can cause oedema.

Oestrogens consist of *oestradiol,* the principal ovarian oestrogen, *oestriol*, the major placental oestrogen, and *oestrone,* an important ovarian and postmenopausal hormone.

At a certain level oestradiol increases GnRH secretion and FSH output by positive feedback. There is also an increased

LH sensitivity to GnRH. This feedback is already called *positive feedback release ovulation,* where the leading follicle ruptures. At lower oestrogen levels in the blood, it is a potent inhibitor of gonadotropin releasing hormone (GnRH) secretion and thus of FSH output. This is the reason for the *ovulation-inhibition* by many oral contraceptives. In the blood oestradiol is bound to *sex steroid-binding globulin.*

Fig. 29-4: Feedback loops and targets organs in the hypothalamic-pituitary-ovarian axis.

The hypothalamic GnRH secretion shows a cyclic variation in adult females of approximately 28 days, probably a genetic code imposed by the CNS.

1. Peaks of *GnRH release* reach the adenohypophysis through the portal system, and release both FSH and LH to reach the ovary via the systemic circulation (Fig. 29-4).

2. FSH stimulates follicular growth, *inhibin-release* from stromal cells, and aromatase activity in the ovary. *Aromatase* converts ovarian and other androgens to oestrogens.

3. LH stimulates the ovarian *androgen* production.

4. Inhibin is the primary inhibitor of FSH release by blocking the effects of GnRH on the adenohypophysis.

Oestrogens are responsible for the female secondary sex characteristics, the maintenance of libido, anabolic effects, and the negative feedback on the GnRH secretion and the Gonadotropin secretion of the adenohypophysis (Fig. 29-4).

Progesterone secretion rises sharply in the luteal phase of the menstrual cycle, and modulates the effect on oestrogens on the endometrium and the myometrial cells. Since the oestrogens have primed all the progesterone receptors, both hormones stimulate the growth of endometrial glands so they curl. Progesterone stimulates the secretion and high bloodflow of the uterus, so it is prepared to receive the fertilised ovum. Progesterone increases the basal core temperature by 0.5 ^{o}C, which is used as an indicator of ovulation.

In the absence of pregnancy, progesterone secretion falls and switches off the release of GnRH and both Gonadotropins (Fig. 29-4). The corpus luteum degenerates, resulting is sloughing of the endometrium (ie, menstruation).

Pregnancy is maintained by progesterone, and uterine contractions are inhibited. Progesterone has a certain aldosterone effect by competition for the same receptors. Progesterone has a negative effect on the lipid profile by increasing the LDL and reducing the HDL fractions in the blood plasma.

Androgens, such as testosterone, are anabolic, maintains spermatogenesis and libido, and act to produce male secondary sex characteristics. These characteristics are the deepening of the voice at puberty, beard, body hair, sebaceous glands in the skin, as well as the growth of the skeleton, the striated muscle system, the external genitalia and male behaviour-attitude. The primary sex structures are the testes with seminiferous tubules, epididymis, prostate, and seminal vesicles. Testosterone is responsible for the growth, maturation, and maintenance of the primary sex structures. Androgens stimulate the growth and polyamine synthesis in the prostate and the seminal vesicles. Hereby RNA synthesis is stimulated and the result is often hypertrophia and hyperplasia. Testosterone increases LDL and decreases HDL concentrations in plasma.

Fig. 29-5: The feedback system of the hypothalamic-pituitary- testicular axis.

Testosterone is reduced to two other potent androgens (dihydrotestosterone, 5a-androstendiol) in many tissues. Testosterone is thus a prohormone for these potent androgens. Most of the testosterone in plasma binds to *sex steroid-binding globulin*, a small fraction binds to albumin and only 1% is free testosterone. Thyroid hormone and oestrogens increase the concentration of sex steroid-binding globulin and thus reduces the free fraction. Androgens have the opposite effect. Testosterone diffuses easily into the cell cytoplasm and binds to a cytoplasmic receptor belonging to the steroid-thyroid receptor superfamily.

The male sexual system is controlled in the following way:

1. GnRH (= LHRH) is released from hypothalamic cells in a pulsate pattern, and stimulates release of LH and FSH

from gonadotropic cells of the adenohypophysis (Fig. 29-5).

2. LH stimulates the Leydig cells of the testes to produce testosterone. These cell also produce small amounts of oestrogens, oxytocin and subunits of pro-opio-melanocortin.

3. FSH and testosterone stimulate the Sertoli cells of the testicular seminiferous tubules to produce spermatocytes and inhibin.

4. *Inhibin* is a glycoprotein that reduces the pituitary FSH secretion (blocks the effects of GnRH) by negative feedback (Fig. 29-5). *Activins* are synthesized by subunits of inhibin. The *Sertoli cells* produce *inhibin,* as do the granulosa cells in females. Inhibin inhibits FSH but not LH secretion by the pituitary gland. Activin stimulates FSH secretion just as GnRH. Follistatin binds and neutralises activin, so follistatin inhibits FSH-secretion.

5. Testosterone is responsible for the male secondary sex characteristics, the maintenance of libido, anabolic effects, and the negative feedback on the GnRH secretion and the Gonadotropin secretion of the adenohypophysis.

Acne during puberty is due to *testosterone,* but in the female *adrenocortical androgens* are involved. Testosterone promotes protein synthesis (anabolic effect). Anabolic steroids have been synthesized, which have a powerful anabolic action but only a modest androgenic action. These artificial hormones are still used to produce short-term *super-athletes*. Such a misuse of medicine for doping purposes often results in addiction, which has serious psychological, social and physical effects.

The *human hypophysis* produces four sex-related hormones FSH, LH, prolactin and oxytocin. LH is also called *Interstitial Cell Stimulating Hormone* (ICSH) in the male, because it stimulates the Leydig interstitial cells that produce testosterone, which in turn specifically inhibits LH secretion. Removal of the male pituitary causes complete loss of all testicular functions; administration of FSH and ICSH then restores these functions completely.

Eicosanoids are oxygenated, unsaturated 20-carbon fatty acids that originate primarily from arachidonic acid by activation of *phospholipase* A_2. Eicosanoids exert important effects on most human cells. Arachidonic acid is a major component of the phospholipids of membranes. *Arachidonic acid* is converted to prostaglandins and thromboxanes by *cycloocygenases,* to leucotrienes by three types of *lipogenases,* and to epoxides by cytochrome P-450-dependent *mono-oxygenases.*

Prostaglandins and thromboxane (TxA_2) are synthesized in response to stimuli, and they mainly act *locally* as autocrine or paracrine hormones. Prostaglandins (PG) are abbreviated PGD, PGE, PGF, PGG, PGH, and PGI_2 (prostacyclin). Leucotrienes (LT) are abbreviated LTA, LTB, LTC, LTD, LTE, and LTF. The leucotrienes LTC_4 and LTD_4 are vasoconstrictors.

TxA_2 is not only an activator of platelet aggregation, but also an effective bronchoconstrictor, and TxA_2 constricts both the cerebral and the coronary arteries.

Prostaglandin E_2 (PGE_2) can be used to induce labour just as oxytocin. PGE_2 and PGF_{2a} increase uterine contractility by Ca^{2+}-influx and moderation of cAMP. Prostaglandins are especially useful in second-trimester abortion.

PGE is a potent vasodilatator, which can be used for intracavernous injection for impotence.

10. Male puberty

Young children have low plasma [gonadotropin] from birth. The gonadotropin releasing hormone is formed in the *hypothalamus* (a decapeptide *GnRH* = LHRH), and stimulates *pituitary gonadotropins* to pulsate secretion of FSH and LH (= ICSH). Through childhood they develop pulsate secretion of pituitary gonadotropins, with a LH peak at night in puberty. The nocturnal LH peak disappears when adult status is reached.

At the onset of puberty a timing device in the brain triggers the gonadotropin producing machinery in the *hypothalamic-pituitary-testicular axis*. Puberty is probably triggered by GnRH in a sufficiently mature CNS. The hypothalamic neurons mature in accordance with a genetic (familial) pattern.

Puberty is a maturation process descending from the programmed brain (hypothalamus) to the pituitary gland, the gonads and eventually to the entire body. Hormones are produced at high rates, and the secondary sex characteristics then develop.

Negative feedback control operates both before and after puberty, but the output of FSH and ICSH from the adenohypophysis is more than 100 times greater in young adults than in boys.

Circulating inhibin is the primary inhibitor of FSH secretion by negative feedback on the pituitary gonadotropins. FSH stimulates Sertoli cells to produce more inhibin at puberty (Fig. 29-5).

Circulating testosterone regulates ICSH secretion by negative feedback primarily on the eminentia mediana hypothalami. The plasma [testosterone] is highest during the night and in the morning (circadian rhythm) but there is virtually no seasonal rhythm with testosterone secretion in humans.

Enlargement of the testes is the first clinical sign of male puberty. The testis consists of Leydig cells that produce testosterone. Gap junctions connect adjacent Leydig cells, and their testosterone has local nourishing effects on germ cells. The seminiferous tubules contain the germ cells (spermatogonia) and Sertoli cells. Each spermatogonium can divide into 64 spermatozoa within 65 days. The Sertoli cells secrete a wide variety of growth factors, activin, inhibin, oestrogens and an androgen-binding protein, all of which nourish the germ cells. The seminiferous tubules drain into rete testis Halleri, which communicates with the epididymis via ductuli efferentes. The epididymis is a maturation chamber for spermatozoa, where they lose their cytosol and become increasingly mobile within a few weeks. The store of mature spermatozoa is emitted into the female genitalia during copulation. The human testes of an adult male are positioned in the scrotum at a temperature around 35^{o}C.

In disease or old age the seminiferous tubules may cease functioning, but the *sexual capacities* (other than fertility) are well maintained as long as testosterone is produced.

Pathophysiology

Aberration of sex development can arise from two different causes. *1.* The *sex chromosomes* can create *genetic sex disturbances*, and *2. hormones* can disturb our *sex differentiation*. This paragraph also deals with *3. Psychosocial sex-deviations*, *4. Cryptorchism*, *5. Castration*, *6. Oral contraception*, *7. Impotence/Prostate disorders*, *8. Menstrual disorders/Occlusion of the fallopian tube*, *9. Menopause*, *10. Osteoporosis*. *11. Breast cancer*, *and* *12. Abortion*.

Sexually related *infections* are gathered at the end of Chapter 33.

1. Genetic sex-disturbances

In 1938 Turner described a syndrome in small persons, retarded in growth and in sexual development. They are apparent females with small or no ovaries and a *XO chromosomal karyotype*. Since they have only one sex chromosome (X), their total chromosome number is 45. The *Turner patient* lacks the inputs from two active X chromosomes and from an Y chromosome. The lack of antimullerian hormone and testosterone leads to Mullerian duct development and female genitals, but the ovary is just a fibrous streak devoid of germ cells. The Turner patients have no sex chromatin and no drum stick (Fig. 29-6).

In 1942 *Klinefelter* described a syndrome in persons appearing as men. These males are tall, have small dysgenetic testes, some have female breasts (*gynaecomastia*), and they are sterile. Their cells contain XXY chromosomes (47 instead of the normal 46). Thus Klinefelter patients must have one sex chromatin and one drumstick just like normal females (Fig. 29-6). These *phenotypic XXY-males* have significantly higher LH & FSH, and lower blood [testosterone] than matched XY controls. The seminiferous tubule development and spermatogenesis are deficient in Klinefelter males. The XXY-males did not show more feminine behaviour than matched controls. A similar group of tall males with *XYY chromosomes* were not extraordinarily masculine. Some XYY-males have significantly higher [testosterone] in their blood than matched XY controls.

Some small *super women* have an extra X chromosome: XXX, making a total of 47 chromosomes. We expected them to have two sex chromatin and two drumsticks, and this has been confirmed. The XXX females have deficient germ cell development and often a short reproductive life.

Fig. 29-6: Intersex syndromes.

Apparent men with XXXY (48) chromosomes have Klinefelter *characteristics* with testes, and also two sex chromatin and two drumsticks (Fig. 29-6).

Individuals with *four* X-chromosomes are extremely rare. They are *apparent females* with XXXX (48), and *apparent males* with XXXXY (49). Cells with 4 X-chromosomes contain a maximum of 3 sex chromatin (Barr bodies) and 3 drumsticks, regardless of whether the cells come from apparent females or males (Fig. 29-6).

A very small number of individuals end up being of *indeterminate gonadal sex* (ie, has both ovarian and testicular tissues present). Some persons have an ovary on one side and a testis on the other - a *true hermaphrodite*. In the Greek mythology *Hermaphrodites* was the child of Hermes and the beautiful Aphrodite. *Pseudo-hermaphrodites* have external genitals from both sexes, but only one gonadal sex. Males have normal XY chromosomes, but small testes with poor sperms (poor spermatogenesis). Some of these genetic (XY) boys are born as apparent girls, but they may change from female to male at puberty if the penis grows. An enzyme defect that blocks the conversion of testosterone to *5-α-dihydrotestosterone* disturbs the development of the external genitals. Female hermaphrodites have ovaries, female ducts, XX chromosomes, and varying degrees of masculine differentiation of the external genitals. Any XY individual with a genetic defect in testosterone synthesis develops testes due to the presence of the Y chromosome, and Mullerian duct regression due to the presence of antimullerian hormone. The Wolffian duct does not develop normally, because of the testosterone deficiency.

Other XY individuals lack the androgen receptor. They develop testes (Y chromosome presence) and the so-called *X-linked testicular feminisation syndrome*. These XY persons show Mullerian duct regression because the antimullerian hormone is present. The lack of androgen receptors and the effects of androgens on the Wolffian ducts prevents masculinization and the external genitals are feminine.

2. Hormonal differentiation disturbances

The virilising effect of testosterone on the *urogenital sinus* in early life causes the *adrenogenital syndrome* in XX individuals. They have ovaries (XX chromosome presence) and the Mullerian duct develops normally, because of the absence of antimullerian hormone. The androgen hypersecretion results in variable development of male external genitalia. The *adrenal hyperplasia* is caused by enzyme defects.

XY individuals with deficient testosterone synthesis ability to convert testosterone to dihydrotestosterone develop testes, but the Wolffian duct structure are underdeveloped to a varying degree ranging from a partial to a complete female pattern.

XY individuals who lack oestrogen receptors or have a mutant gene for aromatase, lack oestrogen effects. The functional lack of oestrogen results in unfused epiphyseal zones, so these males are tall, and they have high plasma concentrations of LH although testosterone is normal.

3. Psycho-social sex-deviations

Sex identity is the individual *perception* of herself or himself as a female or a male. Sex identity is established early, and is not lost by castration. Both psychological and social factors can interfere with normal sexual development on the psychological plane. An imminent urge to change sex (operative sex shifts etc.) characterises *trans-sexual persons*.

The *sex role* is the social behaviour or cultural role played by or forced upon each individual. Some male homosexuals wish to express their femininity while other males clearly signal that they are men. *Transvestites* love to dress like the opposite sex. Transvestites are heterosexual, homosexual or asexual just as others.

4. Cryptorchism

Cryptorchism means *hidden orchids* (testes). The flower orchid (French orchidé) has a root, which is actually shaped like a testis.

If the testes do not descend from the abdominal cavity to the scrotum, heat destroys the sperm-producing seminiferous tubule cells. Heat does not harm the Leydig (testosterone-producing) interstitial cells.

5. Castration

Certain cultures castrate boys to preserve their tenor voices. Puberty and natural sex development does not take place.

Adult males retain their *secondary* sex characteristics and *erection* but they often lose libido. Eunuchs are more or less trustworthy in Harems.

The effects of castration of adult females are surprisingly trivial, as long as the pituitary is working well. Castration, of course, stops their menstrual periodicity (*artificial menopause*), and they are *sterile*.

6. Contraception/ Infertility

Modern contraception is obtained with tablets (pills) containing 20-30 mg ethinyl-oestradiol and variable progesterone. The oestrogen content suppresses the hypophyseal release of gonadotropins, which prevents the maturation of the follicle, the ovulation and the luteinisation. The progesterone content favour the secretion of sperm-hostile mucus in the uterus, inhibits tuba motility and endometrial nidation.

Side effects of the combined tablet are more frequent in smoking female over 35 years, and in all females with cardiovascular risk factors. Side-effects are weight gain, accentuation of cervical and breast cancer, hypertension, acute myocardial infarction, stroke, increased clotting capacity, phlebothrombosis, gallstones, hepatomes, migraine, depression, impaired glucose tolerance (Fig. 27-6), diabetes mellitus, hypercholesterolaemia, and infertility. These serious side-effects are rare but still present. Prescription of even the modern low risk pills necessitates careful control of all risk factors.

Infertility is a diagnosis used on a couple, which have been unable to conceive during one year of unprotected intercourse. The causes are oligospermia, tuba blockage, ovulatory disorders, or combined problems with both persons in the couple.

In some cases ovulation can be elicited by a synthetic *oestrogen receptor antagonist* (clomiphene), which has a high affinity towards hypothalamic oestrogen receptors. Clomifene administration simulates oestrogen deficiency in the infertile patient with a hypothalamic defect, and by negative feedback clomifene increases GnRH and FSH/LH secretion and promote fertility.

7. Impotence/Prostate disorders

Impotence is frequently seen without organic cause such as hypogonadism. Diabetes mellitus, essential hypertension, and neuropathy of the autonomic system cause impotence just as antihypertensive drugs (diuretics, methylDOPA and b-blockers). *Intracavernosal injections* of prostaglandin E, papaverine, phentolamine, and other vasodilatating substances can provide erection for a few hours. The patient can use such a *cocktail* when needed.

Prostate disorders, such as benigh prostatic hypertrophy and prostate cancer, increase in frequency with age above 60. Both disorders interfere with micturition and can obstruct renal function leading to renal insufficiency (Chapter 25). The enzyme *5α-reductase* normally produces dihydrotestosterone from testosterone. Inhibition of this enzyme minimises the hormone conversion and causes the prostate to shrink.

Prostate cancer is frequently present in males with elevated plasma concentrations of *prostate-specific antigen* (PSA). Manifest prostate cancer is removed immediately. In a case where surgery is containdicated, long acting GnRH agonists reduce testosterone secretion and the growth of prostate cancer.

8. Menstrual disorders/Occlusion of the fallopian tube

Amenorrhoea or *oligomenorrhoea* are terms used for absence or irregularity of menstrual periods. Deficient GnRH release prevent FSH secretion from recruiting a dominant follicle, and complete loss of menses (amenorrhoea) may result. In oligomenorrhoea the oestrogen secretion is sufficient for uterine bleeding to occur in an irregular pattern, but often insufficient to induce a midcycle peak of LH and ovulation.

Causes are ovarian disease or absence (*Turners syndrome*, XO), hypothalamic deficiencies, *congenital adrenal hyperplasia* (adrenogenital syndrome), and *starvation amenorrhoea* (anorexia nervosa and excessive exercise), hypothyroid amenorrhoea with increased TRH and prolactin, and withdrawal amenorrhoea (following oral contraception). Starvation amenorrhoea and anovulatory bleeding cycles often occur in female long distance runners and ballet dancers, as well as in *anorexia nervosa* patients (Chapter 7). These females have lost substantial amounts of

fat and suffer from a serious oestrogen deficiency, which even may lead to osteoporosis (Chapter 30).

Occlusion of the fallopian tube

From the start of the menstrual cycle the woman is given FSH to stimulate her ovaries before ovulation. On the 12th day she is given hCG. When ovulation occurs (after 30 to 35 hours), egg cells are *sucked out*, placed in a tissue culture and *exposed* to spermatozoa. After 48 hours some eggs fertilise into the 4-8-cell stage. A few of these *fertilised* eggs are placed in the uterus. One in four of these eggs will nidate.

Therapy is directed towards the cause of the disorder.

9. Menopause

The *menopause* is the event in the life of a female, where the menses stop. The last ovulations are anovulatory and conception is no longer possible. The ovaries become atrophic, the concentrations of pituitary gonadotropins (FSH more than LH) in blood plasma and in urine are the highest in the life of the female, because the follicles become more and more insensitive to gonadotropin stimulation and the oestrogen and inhibin production diminishes. Functional changes in other organs are less definitive, but vascular flushing of the head and neck are typical, probably due to the release of large amounts of hypothalamic gonadotropin releasing hormone (GnRH). Attacks of sweating during the night are classical complains.

Adrenal and ovarian stromal cells secrete androgen precursors that are converted to oestrogens by *aromatase* in adipose tissues. This is why menopausal females with sufficient adipose tissue suffer less from oestradiol deprivation than lean females.

Females with severe complains are treated with oestrogen, which ameliorates the disorders and reduces the rate of *heart disease* and of *postmenopausal osteoporosis*.

10. Osteoporosis

Osteoporosis or *thin bones* is a term used for a marked reduction in all elements of bone mass. Postmenopausal females reduce their bone mass progressively with age up till the age of 70-75 years. This bone reduction also occurs in elderly males, but at a much slower rate. Elderly patients living indoors all year round are less exposed to sunshine and do not synthesise vitamin D in the skin. If their diet simultaneously is poor in vitamin D and Ca^{2+}, it is not surprising that their bones get thin.

Oestrogen therapy is beneficial as a preventive strategy in *postmenopausal* osteoporosis. So is increased dietary Ca^{2+} with *vitamin D*. Walking, jogging, golf are exercises retarding bone mass loss. *Calcitonin* has proven of benefit in some studies. A promising approach is the use of oestrogen-receptor modulators to prevent osteoporosis and thrombo-embolic events, without increasing the risk of breast cancer.

11. Breast cancer

Breast cancer tumours can be treated with synthetic blockers of the oestrogen receptor. The blockers suppress the growth of oestrogen-sensitive breast cancer. – Another therapy principle is to diminish oestrogen production. This is done with the drug, aminoglutethimide, which inhibits the desmolase reaction and thereby reduces adrenal steroid synthesis as a whole.

12. Abortion

Synthetic blockers of the progesterone receptor (mifepristone) induces *early abortion* by removing the positive progesterone effects on the conceptus.

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A. The menarche is the last menstrual bleeding.

B. Pseudo-hermaphrodites have external genitals from both sexes, but only one gonadal sex.

C. HIV means Human Immunodeficiency Virus. HIV is the cause of Acquired Immune Deficiency Syndrome. HIV triggers a progressive and irreversible depletion of T-helper lymphocytes.

D. Transvestites love to dress like the opposite sex. Transvestites are heterosexual, homosexual or asexual.

E. At the onset of puberty a timing device in the brain triggers the Gonadotropin producing machinery in the hypothalamic-pituitary axis.

Case History A

A 24-year-old female is going through her last menstrual cycle before pregnancy.

1. *Summarise schematically the most important hormonal events in her menstrual cycle.*

2. *Summarise schematically the most important hormonal events during continued pregnancy and delivery.*

Case History B

A pregnant woman delivers oxygen to her foetus. Her A-haemoglobin (A = adult) is functionally different from that of her foetus (F-haemoglobin).

1. *Why is this difference important? How are the two dissociation curves related?*

2. *FSH and LH are important for this woman. Describe why. Describe the function of the two hormones in her husband.*

3. *Following birth the mother breastfeed her baby and experience a feeling of sexual pleasure including uterine contractions. Describe the mechanism.*

See answers

Highlights

- *The presence of normal ovaries or testes determines the gonadal sex. Without normal ovaries or testes any genetic sex will develop into an apparent female.*
- *The brain is an important sex organ. The sex desire (libido) is stimulated by a multitude of sense impres (visual, auditive, olfactory, and psychological). Potency refers to the ability to engage in intercourse.*
- *On the first day of the menstrual bleeding, the low progesterone and high prostaglandin level probably releases enough Ca^{2+} to start spontaneous contractions of the myometrial cells. Ca^{2+} -ions enters myometrial cells and stimulates their activity in the secretory (progesterone) phase.*
- *At certain high plasma level of oestradiol can increase FSH output. This is called the positive feedback release ovulation. At lower levels oestradiol is a potent inhibitor of Gonadotropin-RH secretion and thus of FSH output (negative feedback). The negative feedback forms the basis for the ovulation-inhibition by contraceptives.*
- *The primary inhibitor of FSH secretion is the peptide, inhibin that is secreted by the ovary and testis, and blocks the effect of Gonadotropin-RH.*
- *The plasma [oestradiol] increases sharply in the last part of the follicular phase, while the [LH] also increases. The sharp rise in LH and a modest rise in FSH coincide with ovulation. The LH not only causes rupture of the follicle; it continues to act on the follicular cells, turning them into a yellow endocrine organ, the corpus luteum.*

- *The spermatozoa can keep their vitality for more than 4 days if they reach the tube. They lose their protection cover in the uterine tube. The head of the spermatozoa swell and liberates proteolytic enzymes. These enzymes dissolve the corona radiata around the egg (oocyte). The oocyte can only live 14 hours without conception.*

- *Due to the priming effect of oestrogen on progesterone receptors, both hormones stimulate the growth of the endometrial glands, so that they curl like a helix. The progesterone effect in particular provides the endometrial/myometrial tissues with their high secretion and blood perfusion, so the uterus is prepared to receive the fertilised ovum.*

- *The β-group of hCG is specific and found in the blood by specific antibody methods even before the first menstrual bleeding fails to appear. The hCG is detectable in the urine 8-12 days after the first missing vaginal bleeding.*

- *During puberty FSH, LH, growth hormone, and insulin are important for the breast development. The thyroid hormones (T_3/T_4) are permissive. At the end of pregnancy there are other hormonal events. Progesterone secretion reaches a peak and then falls. This fall in progesterone allows the pituitary to release prolactin (LTH).*

- *Relaxin is a pro-insulin-like polypeptide produced by the corpus luteum. The hormone relaxes pelvis articulations and softens the uterine cervix in order to facilitate passage of the foetus. These and several other factors are involved in human labour, but the exact trigger mechanism remains unclear.*

- *Turner described a syndrome in small apparent females, retarded in growth and in sexual development, and with small or no ovaries. Since they have only one sex chromosome (X), their total chromosome number is 45. They have no sex chromatin and no drumstick.*

- *Klinefelter described a syndrome in persons appearing as males. They are tall, have small testes, some have female breasts (gynaecomastia), and they are sterile. Their cells contain XXY chromosomes (47 instead of the normal 46).*

- *Amenorrhoea or oligomenorrhoea are terms used for absence or irregularity of menstrual periods. Causes are ovarian disease or absence (Turners syndrome, XO), hypothalamic deficiencies, congenital adrenal hyperplasia (adrenogenital syndrome), starvation amenorrhoea such as in anorexia nervosa and excessive exercise, hypothyroid amenorrhoea with increased TRH, and withdrawal amenorrhoea (following oral contraception).*

Further Reading

Johnson MH and BJ Everett. *Essential reproduction*. Blackwell Science, Oxford, 1995.

Return to top

Return to Content

Contraction Of Vascular Muscle Cells

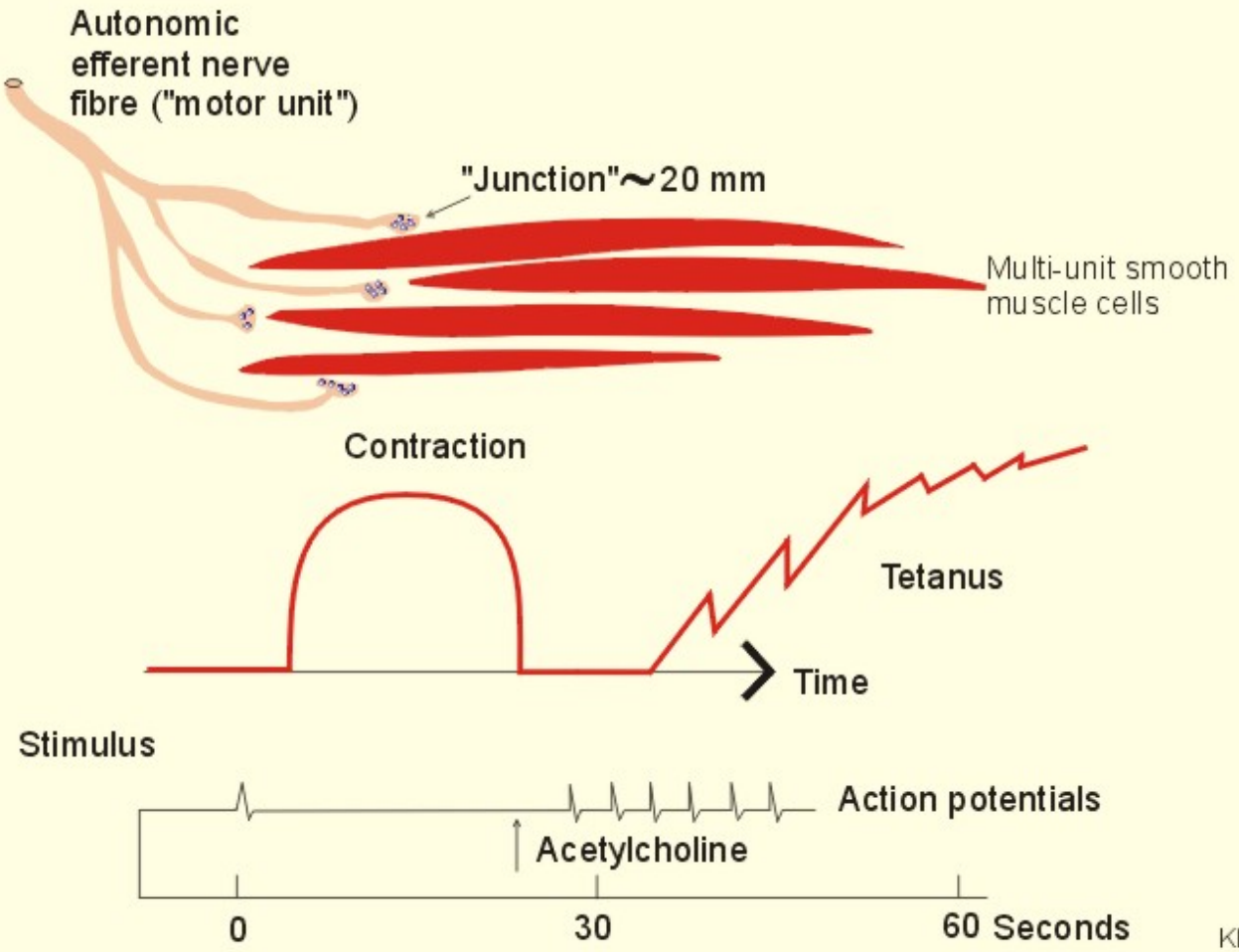
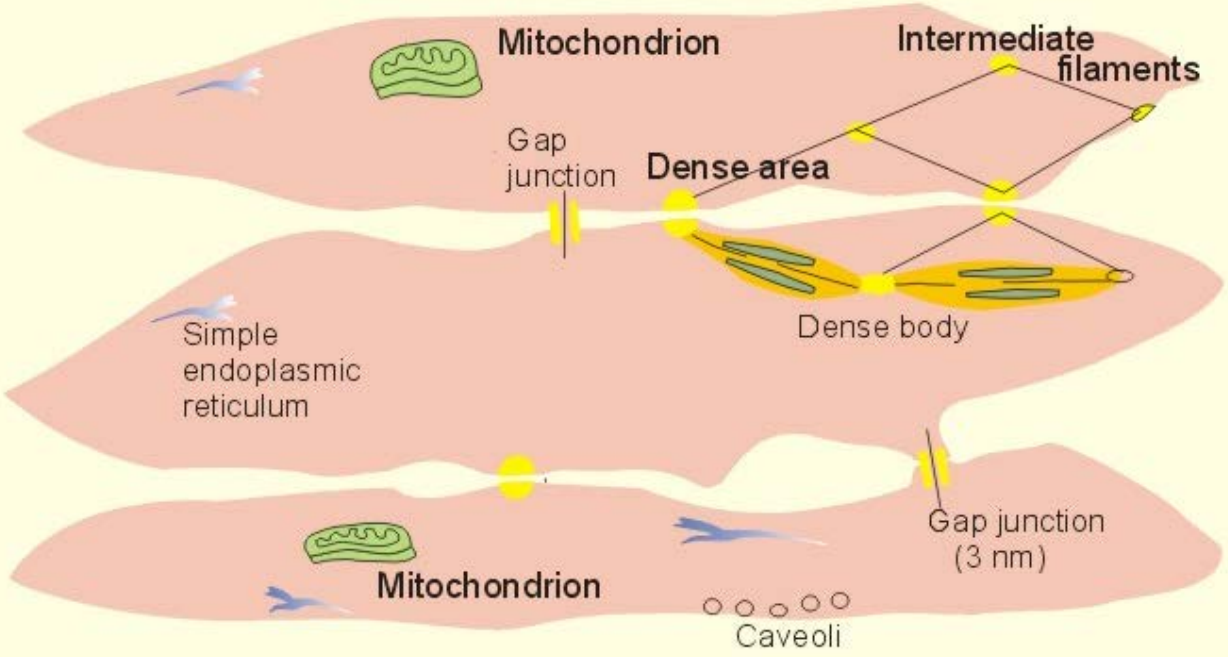


Fig. 2-10

KMc

Structure of Visceral Smooth Muscle Cells



Electrical syncytium by Gap junctions

Fig. 2-11

KMc

Chapter 14.

Gas Exchange And Disorders

Study Objectives

- To define cardiac output, diffusion, diffusion- and perfusion-limited gas exchange, hypercapnia, hypocapnia, hypoxia, respiratory quotient (RQ), ventilatory exchange ratio(R) , and the ventilation-perfusion ratio ($\dot{V}_A / \dot{Q}$ -ratio) .
- To describe Henry's and Dalton's laws, factors of importance for the lung diffusion capacity and pulmonary perfusion, and its measurements, describe the P_{O2}-P_{CO2}-diagram, hypo-and hyperventilation, pulmonary water balance, and mixed venous blood composition.
- To calculate the pulmonary perfusion, use the alveolar gas equation, the alveolar ventilation equation, the final $\dot{V}_A / \dot{Q}$ -equation, and the law of mass balance in calculations.
- To explain the alveolar oxygen uptake and carbon dioxide output, pulmonary vascular resistance and pressures. To explain the alveolar dead space, veno-arterial shunts, uneven regional ventilation-perfusion ratio in health and disease, and peripheral gas exchange.
- To use the concepts in problem solving and case histories.

Principles

- ***Bernoulli's principle or equations (see Eq. 13-8).***

Definitions

- **Alveolar oxygen uptake** per min is the uptake of oxygen molecules into the passing pulmonary blood - into the cardiac output.
- **Cardiac output** is the volume of blood leaving the left (or the right) ventricle each min.
- **Diffusion** is a transport of atoms or molecules caused by their random thermal motion.
- **Diffusion capacity for the lung** (D_L) is defined as the volume of gas diffusing through the lung barrier per min and per unit of pressure gradient ($D_L = \dot{V}_{O2} / \Delta P$).
- **Diffusion-limitation** of *gas exchange* is a condition where equilibration does not occur between the gas tension in the pulmonary capillaries and the alveolar lumen. When the distance between the capillary blood and the cells is large, diffusion becomes a limiting factor even at high bloodflow.
- **Hypercapnia** refers to a resting condition with *hypoventilation,* where P_{aCO2} is higher than 6.4 kPa (48 mmHg).
- **Hypocapnia** is a *hyperventilation* disorder with abnormally reduced P_{aCO2} at rest (below 33 mmHg or 4.4 kPa).
- **Hypoxia** denotes *oxygen deficiency in tissues* due to insufficient delivery of oxygen or inability to utilize oxygen. Hypoxia may be present both with low P_{aO2} and with normal P_{aO2}.
- **Hypotonic** or *hypobaric* **hypoxia** is characterised by a P_{aO2} less than 7.3 kPa (55 mmHg). This is the threshold,

below which the ventilation starts to increase by carotid body stimulation. As the altitude increases, the barometric pressure decreases and the partial pressure of oxygen in the alveolar air falls.

- **Hypoxic pulmonary vasoconstriction** is a compensatory mechanism in alveoli with low ventilation and low oxygen partial pressure. The mechanism is triggered directly by smooth muscle contraction in the vessel walls at a P_{aO2} less than 7.3 kPa (55 mmHg).

- **Multiple inert gas technique** is a procedure, where multiple inert gases of different air—to-blood solubility ratios are infused intravenously until steady state of pulmonary gas elimination is reached. The partial pressure of each gas is measured inthe infused fluid and in the expired air. The $V°_A / Q°$ -equation (Eq. 14-5) and the law of mass balance is used to compute the most likely regional $V°_A / Q°$ -distribution. - Clinically, the alveolar-arterial oxygen tension gradient is measured instead of this complicated reseach procedure.

- **Perfusion-limited** or flow-limited *gas exchange* is limited by the bloodflow. The only limitation to net movement of small molecules across the capillary wall is the rate at which bloodflow transports the molecules to the capillaries.

- **Peripheral Resistance Unit (PRU)** is measured as driving pressure per bloodflow unit (eg, mmHg*s*ml^{-1}).

- **Pulmonary hypertension** is a condition with a mean pulmonary artery pressure above normal - a pressure above 2 kPa or 15 mmHg.

- **Pulmonary vascular resistance (*PVR*)** is the ratio between the pressure gradient and the bloodflow. The basic equation is: *PVR* (PRU) = D*P* / bloodflow (PRU in mmHg*s*ml^{-1}).

- **Pulmonary oedema** is an emergency caused by filtration of fluid out of the pulmonary capillaries into the interstitial space (interstitial oedema), and eventually in the alveolar spaces (alveolar oedema).

- **Respiratory Quotient (RQ)** is a metabolic ratio between the carbon dioxide output and the oxygen uptake of all cells of the body.

- **Standard affinity** is the binding force between two molecules, when half of the binding sites are occupied (at 50% saturation). In the case of oxyhaemoglobin the P_{50} is used. Here, standard affinity is equal to $1/P_{50}$.

- **Ventilatory exchange ratio (*R*)** is the ratio between the carbon dioxide output and the oxygen uptake measurable with gas exchange equipment at the mouth.

Essentials

This paragraph deals with

1. Gas exchange, 2. A key to lung disorders, 3. Uneven distribution of tidal volume and perfusion, 4. Blood gasses, 5. The P_{O2} - P_{CO2} diagram, 6. The $V°_A / Q°$ - curve, 7. Blood-R-curves, 8. Dead space, 9. Anatomic venous-to-arterial shunt, 10. Ficks law of diffusion, 11. Single -breath diffusing capacity, 12. Compensation of $V°_A /Q°$ - mismatch, 13. Pulmonary bloodflow, and 14. Regional ventilation.

1. Gas exchange

Gasses are exchanged between the atmosphere and the alveolar air, and gasses diffuse between the alveolar air and the blood flowing through the pulmonary capillaries.

Oxygen is transported from the atmosphere, via the alveolar ventilation and then carried by the pulmonary bloodflow (equal to the cardiac output), into the cells and their mitochondria for metabolic purposes. Carbon dioxide, the final end-product of metabolism, migrates from the cells to the atmosphere.

A healthy normal person at rest, ventilates his lungs with 5 litres (l) min^{-1} of fresh air ($V°_A$). The *Respiratory Quotient* (RQ) is a metabolic ratio between the carbon dioxide output

($V°_{CO2}$) and the oxygen uptake ($V°_{O2}$) defined for all body cells as a whole. In respiratory steady state, RQ can be measured as the *ventilatory exchange ratio (R)* (Fig. 14-1).

On a diet dominated by carbohydrate the metabolic RQ for all cells of the body is approaching 1, and in a *respiratory steady state*, identical to the *ventilatory exchange ratio, R*, which is measured in the expired air (Fig. 14-1).

Fig. 14-1: The respiratory quotient (RQ) is compared to the measurable ventilatory exchange ratio (*R*).

The normal resting carbon dioxide output is 10 mmol or 224 ml STPD per min from an adult person, and the cardiac output is typically 5 l min^{-1}. The blood volume of 5 l carries each min about 10 mmol (or 224 ml STPD) of oxygen towards the mitochondria. Following passage of the capillary system, the same amount of CO_2 is carried towards the lungs in the venous blood as long as RQ and *R* is 1.

Blood passing the pulmonary capillaries of a healthy person is rapidly equilibrating with the alveolar air. Oxygen from the air diffuses into the blood and binds reversibly with haemoglobin. The *normal oxygen capacity* is 200 ml STPD per l of blood (150 g haemoglobin per l carrying 1.34 ml STPD per g).

The six zones of the *alveolar-capillary barrier* are: 1) a fluid layer containing *surfactant*, 2) the *alveolar epithelium*; 3) a *fluid-filled interstitial space* ; 4) the *capillary endothelium* with *basement membrane*; 5) the *blood plasma;* and 6) the *erythrocyte membrane*. The six zones form an almost ideal gas exchanger for oxygen and carbon dioxide diffusion.

There are 300 million tiny blind end sacs (alveoli) in both lungs together. Fortunately, the alveoli are diluted continuously with fresh air as we breathe.

2. A key to lung disorders

The alveolar ventilation-perfusion ratio is presented as a straight line in Fig. 14-2.

Alveolar ventilation ($V°_A$) and pulmonary bloodflow (equal to the cardiac output, $Q°$) is considered in three extreme situations:

1. The *normal condition* in which $V°_A$ and $Q°$ are matched (ideal $V°_A / Q°$ -ratio = 5/5 = 1), is shown with the typical normal arterial gas tensions (Fig. 14-2).

2. *Pulmonary embolism* creates an *alveolar dead space*. The $V°_A$ is maintained, but there is no bloodflow ($Q°$ regional), so the $V°_A / Q°$ -ratio of the lung region approaches infinity. In the alveolar dead space, alveolar gas pressures approach the levels in inspired air.

3. *Occlusion of the airway* represents an extreme mismatch of venous to arterial shunting of blood, namely perfusion with no ventilation at all (ie, the total ratio approaches zero). The arterial blood gas tensions approach those of venous blood (Fig. 14-2).

The straight line (or $V°_A / Q°$ -axis) of Fig. 14-2 represents an infinite row of ventilation-perfusion-values. Each value refers to an alveolus with equilibrated blood flowing by.

Two well-known equations are relevant here: the *Fick cardiac output equation* (Eq. 14-1) and the *alveolar gas equation* (Eq. 14-3).

The hyperbolic relationship between $V°_A$ and F_{ACO2} is described in the *alveolar ventilation equation* (Eq 14-4).

These three equations can be combined to one equation, which can be expressed in several ways. The calculations are not shown here. The *final equation* reads as Eq. 14-5:

$$V°_A / Q° = R(C_{aO2} - C_{v\text{-}CO2}) / F_{ACO2}$$

Solutions of this equation provide values from zero to infinity for the ventilation-perfusion-ratio. These solutions can be plotted in a P_{O2} - P_{CO2} diagram (Fig. 14-6), where complicated calculations are performed and solved graphically at a glance by looking at the red $V°_A / Q°$ --curve. In the venous point the $V°_A / Q°$ -ratio is zero, and in the I-point on the abscissae the $V°_A / Q°$ -ratio is infinite.

The regional $V°_A / Q°$ -ratio is the *all-important variable*. In any cardio-pulmonary disease, the normal variation of the ratio for the entire system (Fig. 14-2) is exagerated.

Fig. 14-2: Three pulmonary regions or alveoli representing 3 $V°_A / Q°$ ratios from zero to infinity. Normally, the ventilation/perfusion-ratio is 0.8-1.2 for the entire system. Blood gas tensions are given in kPa (133.3 Pa equals 1 mmHg).

3. Uneven distribution of tidal volume and perfusion

can eventuate from *uneven resistance to airflow* within the lung (bronchoconstriction, collapse and compression of airways). Uneven distribution can also be caused by *uneven regional lung compliance* (insufficient surfactant, loss of elastic recoil as in destruction of alveolar tissue, and increase of elastic recoil as in connective tissue scarring or fibrosis with stiff lungs). *Hypoperfusion* can be caused by *compression* of pulmonary vessels, *obliteration* of vessels by fibrosis, or *blockage* by emboli or thrombosis.

Functional shunts arise with any consolidation of alveolar regions that continue to have bloodflow (pneumonia, oedema, haemorrhage, cell necrosis, lack of surfactant).

4. Blood gasses

Blood gases from an arterial blood sample of a healthy person typically show the values of Box 14-1:

Box 14-1: Blood gas values (ranges) from healthy persons at rest. - Normal mean tensions for mixed venous blood and for alveolar air are shown below.

PaO2 :	10-13 kPa (75-95 mmHg)
PaCO2:	4.8-6 kPa (36-45 mmHg),
Base Excess:	Zero (Chapter 17),
pHa :	7.35-7.45 (ie,[H+] = 35-44 nM)
PAO2 :	10 - 13.3 kPa (75-100 mmHg)
PACO2 :	4.8-6 kPa (36-45 mmHg). Mean: 5.3 kPa (40 mmHg),
Mean PvO2 :	6 kPa (45 mmHg)
Mean PvCO2 :	6.1 kPa (46 mmHg)

The blood gasses are essential in the management of severely ill persons with respiratory or circulatory diseases. Any patient - as any healthy person - has some degree of *ventilation-perfusion mismatch*. The P_{aO2} in itself is a good detector of consequential mismatch, but skilled management necessitates interpretation of P_{O2} -P_{CO2} combinations. This is quite easy with the use of Fenn & Rahn´s P_{O2} - P_{CO2} diagram.

Let us develop this excellent clinical tool from simple mathematics and geometry.

5. The P_{O2} - P_{CO2} diagram (Fenn-Rahn)

The general P_{O2} - P_{CO2} - diagram is actually a rectangular triangle with the following corners: (0,0), (P_{O2} , 0) and (0,

P_{CO2}) in Fig. 14-3.

The total tension of all three dry gasses is equal to (P_{O2} + P_{CO2} +P_{N2}) or the barometric pressure (P_B) minus the tension of water vapour in the alveolar air at 37^o C.

The total dry tension at P_B=760 mmHg is thus (760 - 47) = 713 mmHg or (101.3 - 6.3) = 95 kPa. Accordingly, the 713 mmHg on the abscissa refers to pure oxygen (O_2), the 713 on the ordinate refers to pure carbon dioxide (CO_2), and (0,0) represents pure nitrogen (N_2).

Let us assume that analysis of an alveolar air sample results in values shown in point A of the diagram. The diagonal is the hypotenuse of the triangle with a slope of –1. The vertical or the horizontal distance to the diagonal gives the size of the P_{N2}, so in each point all three tensions can be read.

Maintained breathing of pure oxygen leaves all possible expiratory values on the diagonal of Fig. 14-3, namely -1, which is an *R*-value of 1.

In any event, *R* is always equal to one, when one mole of CO_2 is given off to the alveolar air for each mole of O_2 uptake by the blood (Fig. 14-3 and 14-4).

Fig. 14-3: The general P_{O2} - P_{CO2} diagram as developed by Fenn and Rahn. The two axes show the total partial pressure sum of the three dry gasses at one atmosphere. A is a representative alveolar point when breathing air. - Typical alveolar points in the Andes and on the top of Mt. Everest are also shown. – The total dry tension is 713 mmHg or 95 kPa.

Maintained breathing of atmospheric air at sea level implies a P_{IO2} around 150 mmHg (Fig. 14-3). A line from this point to the "CO_2 corner " of Fig. 14-3 represents the situation, where the venous blood delivers CO_2 to the alveolar air without uptake of O_2. Accordingly,

$V°_{O2}$/ $V°_{CO2}$ approach infinity and thus *R* approaches infinity.

Fig. 14-4: The ventilatory exchange ratio, *R*, in different lung regions when breathing air at one atmosphere of pressure (modified from Fenn-Rahn).

When only O_2 is given off to the blood and no CO_2 is removed, the *R*- line falls on the abscissa (Fig. 14-4). All possible *R-values* are easily constructed graphically from any P_B on the P_{O2} - P_{CO2} diagram. The gas-*R* lines fan from the inspiratory or *I* -point (Fig. 14-4).

6. The $V°_A$ / Q°- curve

All alveolar ventilation-perfusion ratios ($V°_A$ / *Q° -ratio*) from zero to infinity were represented by the *straight line* of Fig. 14-2. This line covers all possible combinations of regional ventilation-perfusion-units in the lungs in health and disease. This *line* can be changed to a *curve* by transfer to the P_{O2} - P_{CO2} diagram (Fig. 14-5). Such a curve connects the points for the regional $V°_A$ */Q° -ratio* equal to zero (v$^-$ with all perfusion-no ventilation) and for the regional $V°_A$ */Q° -ratio* equal to infinity (the inspired point I with no perfusion-only ventilation).

Alveolar tensions around **A** refer to the healthy upright lung, where the regional $V°_A$ /*Q° -ratio* is slightly less than *one* in most alveolar units.

Normally, alveolar ventilation ($V\dot{}_A$) and perfusion ($Q\ddot{}$) are matched and the total $V\dot{}_A / Q\ddot{}$ *-ratio* is between *0.8 to 1.2* with normal alveolar and blood gas tensions. In the normal upright lung the regional $V\dot{}_A / Q\ddot{}$ -ratio is approximately 0.6 at the lower and about 3 at the upper lung region.

Fig. 14-5: Three alveolar regions in the upright lung of a healthy person at rest. The upper alveolus and airway is distended and its bloodflow is minimal. The lower alveolus is compressed by gravity and its bloodflow is high.

The *pulmonary bloodflow decreases* from the lower to the upper parts of the lung of a resting person (Fig. 14-5). Likewise, the relative ventilation of the lung also decreases linearly from the base to the apex, but at a slower rate. Thus, the regional ventilation-perfusion ratio varies from *zero* in the lower region, where there is only bloodflow and no ventilation to *infinity* in the upper region, where there is only ventilation and no bloodflow. At the lower lung region, regional $V\dot{}$ approaches zero and at the top of the lung regional perfusion approaches zero. In a P_{O2} - P_{CO2} diagram each point on the curve represents partial pressures at which alveolar air and blood can equilibrate at a certain $V\dot{}_A / Q\dot{}$ -ratio. Thus for any practically obtainable point, a single value exists for blood gas concentrations (see later in Fig. 14-6). - Lung regions at the base with low $V\dot{}_A / Q\ddot{}$ has low P_{AO2} and high P_{ACO2}, relative to normal mean values. Upper lung regions with *high* $V\dot{}_A / Q\ddot{}$ have relatively *high* P_{AO2} and *low* P_{ACO2}.

7. Blood-R-Curves

For a person in respiratory steady state, the *R*- value of the *blood* is equal to the *R* -value of the *alveolar* gas (the gas-*R*). As respiratory gasses are exchanged with a certain *R*- value (eg, *R*=1), the passing blood must do the same. Accordingly, the *blood-R* is equal to the *gas-R*. The blood -*R* curves fan out from the venous point (Fig. 14-6). One green *blood-R* curve is shown. The green curve intersects with its blue *gas- R*-line on the $V\dot{}_A / Q\dot{}$ - curve.

The shape of the blood-*R* curves is dictated by the oxyhaemoglobin dissociation and the carbon dioxide binding curves, which in turn are affected by the Bohr- and the Haldane-shifts.

Fig. 14-6: Gas-*R* and the related blood-*R* curve (0.8) drawn together with the $V\dot{}_A / Q\dot{}$ - curve. An ideal alveolar point is shown (i) together with normal values for arterial (a), alveolar (A), and expired (E) gas tensions. - The symbol t for tissue (co-ordinates 1,47 mmHg) denotes minimal tensions in the peripheral tissues of a healthy person.

The regional $V\dot{}_A / Q\dot{}$ -ratios (Fig. 14-6) show the *lower* lung regions to be *relatively underventilated* (ratio below one), the *middle* lung regions to be *well matched* (ideal regional ratio of 1), and the *upper* lung regions to be relatively *overventilated* (ratio above 1 and approaching infinity). Also *R* approach infinity when we approach the inspired point **I**, Fig. 14-6).

Points **A** and **E** refer to the alveolar and expired air tensions, respectively. Every regional deviation from the average total $V\dot{}_A / Q\dot{}$ -ratio of 1 in healthy subjects, will result in alveolo-arterial gas tension differences (see later in Fig. 14-6).

Underventilated and overperfused alveoli have increased P_{AN2} and thus increased P_{aN2} , whereas overventilated alveoli reduce their P_{ACO2} almost as much as they increase P_{AO2} . Hereby, P_{aN2} becomes greater than P_{AN2} , and a precisely measured difference is used as a measure of mismatch.

8. Dead space and shunt

Ideal lungs have a matched ventilation-perfusion ratio resulting in an ideal composition of the alveolar air throughout the lung. With all alveoli having identical $V\dot{}_A / Q\dot{}$ -ratio s, the alveolar point **A** would be located in the *ideal* point **i** for all alveoli (see Fig. 14-7). Doubling of alveolar ventilation will move the point **A** halfway down the blue diagonal

towards **I,** and as alveolar ventilation approaches infinity the gas concentrations of the alveolar air approaches those of the inspired air (**I**).

Normally, the expired air values are always represented by a point **E** on the diagonal between **A** and **I** (Fig. 14-6).

All the displacement from ideal point **i** to real life point **A** is caused by *alveolar dead space*, and all the displacement from **i** to E is caused by *alveolar* plus *anatomic* dead space. This sum is also termed the *physiological* dead space (Fig. 14-7). The physiological dead space of a healthy adult at rest is approximately *150 ml* out of a tidal volume of 500 ml (30%). During exercise the physiological dead space will rise to perhaps *200 ml* simultaneously with a rise in tidal volume to 2000 ml as an example. This is a relative physiological dead space of only 10%, which is an advantage to the individual during work.

Fig. 14-7: Alveolar (A), expired (E) and arterial (a) gas tensions from a patient with chronic obstructive lung disease. Both a large alveolar dead space and a serious shunt are present. The ideal point (i) is also shown. The different locations of the symbol's t illustrate the tensions in peripheral tissues of a patient and of a healthy person.

In healthy persons, the alveolar gas tensions vary during a respiratory cycle around a mean value, although the oscillations are close to the *ideal point*. These variations are called alveolar gas tension *oscillations* (see Chapter 16, Fig. 16-8).

Patients with lung disorders often have $V\dot{}_A / Q\dot{}$ -mismatch by a combination of areas of *veno-arterial shunting* often in the lower lung regions, and areas of *increased* alveolar dead space often in the upper lung regions (Fig. 14-7). The location of the arterial point **a** (50,40 mmHg) on the green curve indicates that the **i-a** *distance* is larger than 50% of the total **i-v** *distance*, which must be caused by more than 50% veno-arterial shunting. This **i-a** *distance* is an essential clinical concept, called the *ideal alveolar-arterial* P_{O2} *gradient* (Fig.14-7). The closer point **a** is to point $\mathbf{v}^-$**,** the larger is the shunt. The veno-arterial shunt is *total* (100%), when the point **a** is moved to the point $\mathbf{v}^-$

9. Anatomic venous-to-arterial shunts

Normally, up to 5% of the venous return passes directly into the systemic arterial circulation. This shunt-blood includes nutrient bloodflow coming from the upper airways and collected by the bronchial veins. Also the coronary venous blood that drains directly into the left ventricle through the Thebesian veins is shunt-blood.

The classical way to determine the relative size of a shunt is by the law of conservation of matter. Adolph Fick used the naturally occurring indicator oxygen as substance (Fig. 14 -8). The law of mass balance is applied to both bloodflow and oxygen flux in Eq. 14 -7 and 14-8, where $C_{c'02}$ is the oxygen concentration in the *pulmonary end capillary blood* of an ideally functioning alveolus (Fig.14-8).

The flow and flux relations lead to Eq. 14-9, which shows that the classical method, necessitates cardiac catheterisation to get mixed venous blood ($C_{v\text{-}O2}$) for the determination of mixed venous gas tensions.

Fig. 14-8: The classical method of determining the size of a shunt implies cardiac catheterisation and measurements of blood gas concentrations.

The location of point **E**, more than half way down the diagonal to **I**, suggests a *large physiological dead space* - more than 50% of the tidal volume (Fig. 14-7).

As the disease progresses, the venous point (v^-) moves to the left and upward, so that peripheral tissues with the smallest P_{O2} gradient become increasingly hypoxic.

The broken curve shows the tensions in tissues from the mixed venous driving tension to tissue tensions (**t**) of only one mm Hg (Fig. 14-7). The slopes of these *tissue tension curves* are about 1/20, reflecting that CO_2 diffuses 20 times faster than oxygen (23.2 * 0.85= 20). In the final phase of lung disorders also hypercapnia becomes prominent (see the venous point with high P_{CO2} in Fig. 14-7).

10. Fick's law of diffusion

states that the flux of gas transferred across the alveolar-capillary barrier is related to the *solubility* of the gas, the diffusion area (A), the length of the diffusion pathway from the alveoli to the blood (L), and the driving pressure (P_1 - P_2). These factors are all included in the simplified version of Ficks law marked Eq. 14-2. The *solubility* is also called the Bunsen solubility quotient, α**.**

Although the diffusion area at rest is close to the size of half a tennis court, and the diffusion distance (L) is 0.5 - 1 micrometer, it is difficult to predict their variations between individuals. Therefore, Marie Krogh developed the individual *lung diffusion capacity* (D_L) defined as the flux of gas transferred per pressure unit through the lung barrier of a certain person. Since the counter pressure of CO in the blood is virtually zero, a simple measure of P_{ACO} provides us with the pressure gradient in Eq. 14-2. The *standard affinity* of the haemoglobin-CO reaction is very large and *250* times greater than that of O_2 (Eq. 14-10). The standard affinity is measured as the reciprocal value of P_{50}. The P_{50} for haemoglobin-CO is just a fraction of one mmHg, and the haemoglobin-CO dissociation curve is too close to the ordinate of Fig. 14-9 to show - so an enlargement is drawn to the left. - D_L consists of a *barrier-factor* (consequential in lung oedema and in lung fibrosis) and a *haemoglobin-factor* (which reflects the binding rate of oxygen to haemoglobin). The presence of haemoglobin permits blood to absorb 65-fold as much O_2 as the content in plasma at normal P_{aO2}.

Fig. 14-9: Dissociation curves for Oxy- and CO-haemoglobin.

CO competes with O_2 for binding sites on haemoglobin, and thus exposure to CO *reduces* the O_2 binding to haemoglobin. Persons breathing traces of CO occupy a large fraction of all binding sites by CO. The CO binding causes a leftward shift of the oxy-haemoglobin dissociation curve. All the binding sites that are bound to CO, do not respond to falling P_{aO2}. The remaining O_2 molecules on the CO-haemoglobin molecule are much more avidly bound and unload slower than normal.

Diffusion is rapid over short distances. In normal lungs there are *trans-barrier pressure gradients* for diffusion of both O_2 and CO_2. D_{LCO} is measured by measuring the carbon monoxide uptake and the driving pressure (see *single-breath diffusing capacity* below).

11. Single-breath diffusing capacity

The subject takes a deep breath of 0.3% carbon monoxide (CO) and holds the breath for 10 s before exhaling and alveolar sampling. During the breath holding, CO is taken up by the haemoglobin of the passing blood in proportion to its alveolar tension (P_{ACO}).

A simple assumption is that the CO uptake is directly proportional to the mean alveolar P_{CO} (symbolised with P_{ACO}). The *diffusing capacity of the lung* (D_L) is also called the *transfer factor*, because D_L measures not only diffusion, but the barrier thickness and ventilation-perfusion mismatch as well. Patients with lung disease often have abnormal size and thickness of the alveolar barrier or ventilation-perfusion mismatch. In such cases measurement of the *CO transfer* need not be a true measure of the total diffusing capacity.

Box 14-2: Diffusing capacities of the lungs for different gasses in healthy persons

Units	ml STPD s^{-1} kPa^{-1}		ml STPD min^{-1} $mmHg^{-1}$	
	Rest	**Exercise**	**Rest**	**Exercise**
DLCO	3	7.5	25	62.5
DL02	3.6	9	29	73
DLCO2	70	175	565	1412

The *single-breath CO diffusing capacity* is normally *3 ml* STPD s^{-1} kPa^{-1} at rest. The values during rest and exercise - and in two units - are shown in Box 14-2.

The *transfer factor* is reduced by diseases affecting the lung parenchyma, such as emphysema, pneumonectomy and fibrotic diseases (the alveolar barrier is too small in area or too thick or both).

12. Compensation of $V°_A$ /Q°- mismatch

Low P_{AO2} in poorly ventilated alveoli, causes arteriolar *constriction*, which redistributes bloodflow to well-ventilated alveoli.

Low P_{ACO2} exists in alveolar regions with a *high* ventilation-perfusion-ratio. Low values constrict the small airways leading to these alveoli. Their reduced ventilation results in redistribution of gas to alveoli with better bloodflow.

13. Pulmonary bloodflow

Pulmonary vascular *resistance* (*PVR*) is minimal compared to that of the systemic circulation. The pulmonary vascular system is basically a low-pressure, low-resistance, highly compliant vessel system with a bloodflow sensitive to gravity and to P_{AO2}.

The system is meant to accommodate the entire cardiac output - and not to meet special metabolic demands as in the case of the systemic circulation.

Box 14-3: Blood pressures in the pulmonary system of a healthy supine person at rest

Units	**mmHg**	**kPa**
Right ventricle	25/-1	3.3/-0.133
Pulmonary artery	25/8	3.3/1
Mean pulmonary Artery	13	1.7
Pulmonary capillaries	8	1
Left atrium	5	0.7
Driving pressure	8	1

The pressure in the right ventricle is *3.3 kPa* systolic and - *0.133 kPa* diastolic in a healthy, supine person at rest. The pressure in the pulmonary artery is about *3.3 kPa* systolic and *1 kPa* diastolic, with a mean of 1.7 kPa (Box 14-3). The blood flow of the pulmonary capillaries pulsates and its mean pressure is *below 1 kPa*. The pressure in the left atrium is 0.7 kPa. This value implies a pressure drop across the pulmonary circulation of (1.7 - 0.7) = 1 kPa. This *driving pressure* is *less than 1/10* of the systemic driving pressure.

The walls of the pulmonary vessels are thin, hence their pressure must fall at each inspiration, because the intrapulmonic pressure falls.

Change of posture from supine to erect position, will reduce the pressure toward zero in the apical vessels, whereas it increases the pressure in the basal vessels due to gravity.

When the *driving pressure* in the apical blood vessels approaches zero, the blood flow will also approach zero. Apart from its implication for gas exchange, this phenomenon limits the supply of nutrients. Lung disorders often occur in the apical regions.

The *pulmonary vascular resistance (PVR)* is the ratio between the pressure gradient and the bloodflow. A *peripheral resistance unit (PRU)* is measured as driving pressure per bloodflow unit. The basic equation is: *PVR* (PRU) = D*P*/bloodflow (mmHg*s*ml^{-1}).

At rest, the pulmonary driving pressure is 8 mmHg (Box 14-3), and the bloodflow is 5 l per min (83 ml per s). The ratio is 8/80 = 1/10 PRU (normal *PVR* is only 10% of the systemic resistance at rest: *TPVR* = 1 PRU). Calculated in kPa the *PVR* is 1/80 kPa s ml^{-1}. Such low values for *PVR* are only found in the lungs of healthy, non-smokers.

The *PVR* remains low in healthy persons, even when cardiac output increases to 30 l per min, because of distensibility and recruitment of pulmonary vessels. Stretch receptors, found in the left atrium and in the walls of the inlet veins, are believed to be stimulated by distension. Such a distension blocks liberation of *vasopressin* (antidiuretic hormone, ADH) from the *posterior pituitary* and releases atrial natriuretic factor (ANF) from the atrial tissue. Hereby, the urine volume increases and the extracellular volume decreases.

Changes in *pulmonary vascular resistance* are achieved mainly by *passive* factors, but also by *active modification.*

Passive factors: The larger arteries and veins are located outside the alveoli (extra-alveolar); they are tethered to the elastic lung parenchyma, and are exposed to the pleural pressure. The pulmonary capillaries lie between the alveoli and are exposed to the alveolar pressure.

Alveolar capillary volume. The intra-alveolar vessels are wide open at low alveolar volumes, so that their *PVR* must be minimal. With increasing alveolar distension these vessels are compressed. This increases the *intra-alveolar PVR*. However, at low alveolar (lung) volumes, the extra-alveolar vessels are small because of the small transmural vascular pressure gradient, and their *PVR* is high.

With increasing lung distension, the intrathoracic pressure becomes more subatmospheric. This elevates the transmural vascular gradient and is coupled with the radial traction on these vessels by the surrounding lung parenchyma as it expands. Thus, the *extra-alveolar PVR decreases.* The greatest cross-sectional area exists in the many *intra-alveolar* vessels, hence increasing *PVR* in these vessels offsets decreased *extra-alveolar PVR.*

Thus, total pulmonary vascular resistance is increased at higher alveolar volumes when intra-alveolar *PVR* is high. *PVR* is minimal at FRC, where there is air enough to open the extra-alveolar vessels with minimal closure of the intra-alveolar vessels.

Pulmonary artery pressure. A healthy person at rest (FRC) has approximately *half* of the pulmonary capillaries open, but with increasing arterial pressure, the previously closed capillaries open (recruitment). As the arterial pressure continues to rise, the capillaries become distended. The net effect is a rise in the total cross-sectional area of the lung capillaries, leading to decreased *PVR.*

Left atrial pressure. Patients with *high* left atrial pressure have distended capillaries due to the venous backpressure. As a result of the reduced driving pressure their *PVR* is decreased further.

Gravity. The pulmonary bloodflow per unit lung volume is greatest at the lower and decreases towards the upper lung regions. Gravity creates a gradient of vascular pressures from the top to the bottom of the lungs. The intravascular pressure is much lower at the upper than at the lower lung regions, unlike the *alveolar* pressure, which is essentially constant throughout the lung. At the top of the lung all vascular pressures can approach zero (with the alveolar pressure as reference). Under these conditions there is no bloodflow through the upper region, and if it is still ventilated, it is an *alveolar dead space*.

Active modification is essential: Both sympathetic and parasympathetic fibres sparsely innervate the pulmonary blood vessels. *Sympathetic stimulation* constricts the pulmonary vessels, whereas *parasympathetic stimulation* dilatates them. *Vasoconstrictive agents* include: Arachidonic acid, catecholamines, leucotrienes, thromboxane A, prostaglandin F, angiotensin-II, and serotonin. The *vasodilatators* are acetylcholine, bradykinin, nitric oxide (NO) and prostacyclin.

A *decrease in* P_{AO2} in an occluded region of the lung produces hypoxic vasoconstriction of the vessels in that region as mentioned above. The reduced P_{AO2} causes constriction of the precapillary muscular arteries leading to the hypoxic region. The hypoxic effect is *not nerve-mediated*. This reaction shifts blood away from poorly ventilated alveoli to better-ventilated ones. NO seem to dilatate the vessels of the well-ventilated segments of the lung. Perfusion is hereby matched with ventilation.

14. Regional ventilation

Milic-Emili has developed the elegant *onion skin diagram* of the regional ventilation (Fig. 14-10). The first 25% of the lower abscissa is the residual volume or RV, and this axis shows the total lung capacity (TLC) up to 100% TLC (maximal inspiration). The upper abscissa shows the vital capacity, VC, from zero to 100%. The ordinate is the *regional* ventilation volume in % of the *maximal* regional total lung capacity (TLC). The maximal regional TLC is any given lung region totally filled with air by a maximal inspiration (Fig. 14-10).

The slope of the *onion skin-lines* are constant above FRC, thus the fraction of the tidal volume reaching each lung region, must be constant during the whole inspiration from FRC (Fig. 14-10). The slope is larger in the lower than in the upper lung region, because the lower alveoli are the ones most compressed by the gravity-sensitive pleural pressure. Accordingly, they can distend most during inspiration. The upper alveoli are always more expanded than the lower due to the *pull* of gravity. The upper alveoli follow the *first in - last out* principle. During expiration to residual volume (RV) the upper alveoli are the last to empty (Fig. 14-10). - During inspiration from RV, the lower alveoli are closed up to FRC (*closing volume* and *closing capacity* - see the horizontal blue curve in Fig. 14-10). Around FRC the lower alveoli open.

At the start of the inspiration from FRC the lower alveoli are the smallest, so any inspiration will always distend the lower alveoli most.

Fig. 14-10: The relative, regional ventilation (ordinate) depending upon total ventilation from RV to TLC (modified from Milic-Emili).

The upper alveoli are always expanded by gravity. At TLC all alveoli are assumed to be *maximally* distended (Fig. 14-10). The alveoli and small airways are increasingly distended from the lower to the upper lung regions. As a consequence, their compliance must decrease progressively, and the pleural pressure also decreases towards the top of the lung (Fig. 14-5).

Conclusion:

The multiple inert gas technique has confirmed that the major problems in pulmonary disorders are not true shunts, diffusion barriers, and lamination of alveolar gasses, but dominantly ventilation/perfusion inequality *with functional veno-arterial shunts and alveolar deadspace.*

Pathophysiology

This paragraph deals with *1. Hypocapnia*, *2. Acute hypercapnia* and *3. Vascular lung disorders*. - *Hypoxia* is described in Chapter 15.

1. Hypocapnia

Hypocapnia or *hyperventilation* is a disorder with abnormally reduced P_{aCO2}. The hyperventilation reduces P_{aCO2} and produces an *acute respiratory alkalosis*, characterised by increased pH, and normal or unchanged Base Excess (BE = Zero). Changes in *Base Excess* are effected by renal mechanisms, which take hours to develop.

2. Acute hypercapnia (CO_2-poisoning)

Hypercapnia is a condition, where P_{aCO2} is higher than 6.4 kPa (48 mmHg). Patients with a large dead space and $\dot{V}_A/\dot{Q}$ -mismatch develop hypercapnia, due to hypoventilation. Reduced alveolar ventilation increases P_{CO2} and lowers P_{O2}. Since the CO_2 stores are much larger than the O_2 stores, the initial rise of P_{CO2} is lower than the drop in P_{O2}. Thus, the *R*-value must fall, as seen typically in anaesthetic depression of the respiratory centre. The arterial tensions

follow the alveolar. The changes in mixed venous tensions are small, because Q° is maintained and the slope of the oxyhaemoglobin dissociation curve is steep at a mixed venous P_{vO2} around 45 mmHg.

The patient with acute hypercapnia is flushing, nervous, horrified of death, and has increasing dyspnoea. The death-horror and hallucinations are followed by loss of consciousness and respiratory arrest. The blood gasses show increased P_{aCO2} and reduced pH (acute respiratory acidosis, Chapter 17) with a base excess of zero.

For patients with chronic pulmonary disease, the hypoxia increases the 2,3-DPG concentration in the red cells, which - together with the hypercapnia and fever - displaces the oxyhaemoglobin curve to the right. This is beneficial for tissue oxygenation, because it increases the tissue tension gradient during oxygen unloading .

Fig. 14-11: The oxyhaemoglobin dissociation curve.

Abnormal blood gas values are indicators of the severity of the disorder. The first phase is characterised by *normal* blood gasses at *rest*. The second phase is *respiratory insufficiency* with abnormal blood gasses at rest (*hypoxia*: P_{AO2} less than 7.3 kPa or 55 mmHg, and *hypercapnia:* P_{aCO2} higher than 6.4 kPa or 48 mmHg). The term *terminal* respiratory insufficiency refers to the grave prognosis.

Hypoxia is dangerous because its effects are irreversible, while hypercapnia is reversible. The oxygen treatment increases P_{aO2}, which is vital, so oxygen therapy should be administered instantly to patients with hypoxia – irrespective of hypercapnia. A few patients may have adverse effects with respiratory arrest, when the hypoxic drive for the peripheral chemoreceptors is eliminated. The ventilation will fall, which elicits a substantial rise in P_{aCO2} with anaesthetic effect on the respiratory centre.

The advantage of *oxygen enriched air* can be shown by an example. A patient with asthma is hospitalised with a P_{aO2} of 5.5 kPa (41 mmHg) and a S_{aO2} of 0.75 (Fig. 14-11). Oxygen enriched air is valuable to such a patient. Oxygen enriched air is administered with a nasal catheter or accurately with a simple plastic mask using the Venturi or Bernoulli principle (Chapter 13).

A small increase in the oxygen concentration of atmospheric air from 21% to 24% leads to a rise in P_{IO2} (3% of 95 kPa is 2.9 kPa; 3% of 713 mmHg is 21.4 mmHg). The major part of this rise reaches the arterial blood (2.6 kPa or 20 mmHg) and this rise in P_{aO2} from 41 to 61 mmHg is often enough to save the patient, because S_{aO2} increases to *0.94* (Fig.14-11). The oxygen flux to the tissues depends upon a normal haemoglobin concentration and a normal cardiac output.

3. Vascular lung disorders

Diseases of the pulmonary vascular tree are diagnosed as pulmonary oedema, pulmonary embolism, and pulmonary hypertension.

3a. Pulmonary oedema is an emergency caused by filtration of fluid out of the pulmonary capillaries into the interstitial space (*interstitial oedema*), and eventually in the alveolar spaces (*alveolar oedema*) – see Fig. 10-10.

The amount of fluid filtered out of the pulmonary capillaries is determined by the Starling equation (Eq. 8-7). The capillary hydrostatic pressure is the main outward force, and this pressure is larger at the base than at the apex of the upright lung. The main inward force is the colloid osmotic pressure of the proteins of the blood. Normally, the alveoli are kept free of fluid, because a net outflux of fluid from the vasculature is balanced by a small lymph flow to the hilar lymph nodes.

Pulmonary oedema has at least 3 causes:

1. *Increased pressure.* Patients with *left cardiac failure* (acute myocardial infarct, chronic myocardial failure,

mitral stenosis, aortic stenosis, and hypertension) can drown in their own plasma transudates. The increased venous backpressure distends all pulmonary vessels (lung congestion), and as soon as the pulmonary capillary pressure is higher than the *colloid osmotic pressure* (normally 3.3 kPa or 25 mmHg), there is a filtration of plasma water into the pulmonary interstitial tissues and into the alveoli. The pulmonary vascular pressure rises in the supine position causing attacks of lung oedema to occur at night.

2. *Increased capillary permeability.* Pulmonary oedema can be caused by capillary damage with war gas, toxins, pneumonia etc.

3. *Reduced concentration of plasma proteins* increases net filtration at the arteriolar end of the lung capillary and reduces net reabsorption of filtered fluid at the venular end.

Oedema is particularly serious in the lungs, because it widens the diffusion distance between the alveolar air and the erythrocytes. There is not enough time for oxygen to travel from the air to the individual erythrocyte. Thus, the blood leaving the lungs is only partially oxygenated. Both the VC and the compliance are reduced.

Increased pulmonary capillary pressure is caused by any type of left ventricular failure (acute myocardial infarction or chronic heart failure) and by mitral valve stenosis. A pressure above 2.6 kPa (20 mmHg) causes interstitial oedema, and as the pressure rises above 4 kPa, alveolar oedema develops. Interstitial oedema may not be recognised, but *alveolar oedema* is dramatic.

The patient is severely dyspnoeic, with tachypnoea, tachycardia, and coughing up a frothy pink sputum containing red cells. There is basal crepitation by auscultation and often whistling rhonchi.

Since the fluid-filled alveoli are not ventilated with air, any blood passing them does not participate in gas exchange. The effect is a functional veno-arterial shunt with hypoxaemia, although hypoxic vasoconstriction tends to reduce its size. Initially, the non-affected alveoli are overventilated and P_{ACO2} is low. Hypercapnia is a late complication when the gas exchange is severely compromised.

Other causes of pulmonary oedema include *decreased colloid osmotic pressure* (hypoproteinaemia, overtransfusion), *increased capillary permeability* (pulmonary oxygen toxicity, radiation damage), and *high-altitude oedema*.

Therapy keypoints:

- Primarily, it is important to find the cause of pulmonary oedema, such as left cardiac failure, and correct the disorder.
- Patients with *chronic cardiac failure* have reduced contractility, which improved by positive inotropic agents such as digoxin.
- Patients with *lung oedema* must sit up erect in bed with the legs over the side and calm down. This reduces venous return and cardiac output, and the effective filtration pressure is reduced.
- Breathing of *air enriched with oxygen* reduces hypoxia and dilatates the lung vessels. The filtration pressure is reduced.
- Effective diuretics *increase the excretion of* Na^+ and thus of water via the kidneys. The loss of fluid also implies oedema fluid.
- *Positive pressure breathing* is thought to minimise the difference between the central and the peripheral venous pressure, so the *venous return* and thus cardiac output is reduced. The blockade of lung capillary bloodflow in the overpressure-phase, and the fear of the patient (increases cardiac output) does not make this treatment the best of

choice. The effect is probably similar to the earlier application of bloodletting tourniquet to reduce the pressure gradient from the left to the right atrium.

3b. Pulmonary embolism

is caused by detached parts of thrombi from the venous system. The dislodged thrombus is carried with the venous blood to the pulmonary artery, where the lower lobes are frequently affected, due to their relatively high bloodflow.

The lung tissue is ventilated but not perfused, so the gas exchange suffers and hypoxaemia develops. Destruction of lung tissue of the affected area (pulmonary infarction) is rare, due to the continued oxygen supply by the airways and by the bronchial artery.

The condition can develop into *acute cor pulmonale*, which is sudden failure of the right heart.

Immobilisation by prolonged bed rest, local damage of venous walls with thrombophlebitis, and hypercoagulability of the circulating blood are predisposing conditions.

3c. Pulmonary hypertension

is a condition with a mean pulmonary artery pressure above normal (ie. a pressure above 2 kPa or 15 mmHg).

Pulmonary hypertension is caused by *increased* left atrial pressure (left ventricular failure, mitral valve stenosis), *increased* pulmonary bloodflow (congenital heart disease with left-to-right shunting of blood through septal defects or a persistent ductus arteriosus), and by *increased* resistance of the pulmonary vessels (destruction of the capillary bed in emphysema, obstruction in pulmonary embolism, hypoxic vasoconstriction in chronic bronchitis with emphysema and at high altitude).

Persistent pulmonary hypertension leads to right ventricular hypertrophy and finally to *chronic cor pulmonale*. This is often the final stage of not only chronic *obstructive* lung disease in smokers, but also of the late *restrictive* lung disorder.

Equations

- **The Fick cardiac output equation** states that the cardiac output is calculated from the ratio between alveolar oxygen uptake and arteriovenous oxygen content difference:

  Eq. 14-1: $\dot{Q} = \dot{V}_{O2} / (C_{aO2} - C_{v\text{-}O2})$.

- **Fick's law of diffusion** states that the flux of gas transferred across the alveolar-capillary barrier is directly related to the **solubility** (Bunsen's a, Table 13-1) of the gas, the diffusion area (A), the length of the diffusion pathway from the alveoli to the blood (L), and the driving pressure **$(P_1 - P_2)$: $J_{gas} = (D \times a \times A \times 1/L) \times (P_1 - P_2)$.** Marie Krogh incorporated molecular weight (mol. weight), a, A, and L in her lung diffusion capacity (D_L). D_L is equal to a constant, K, multiplied with a, and divided by the square root of the mol. weight. Thus $D_L = K \times a/ \sqrt{\text{mol. weight}}$. This relationship is used on all three gasses: $D_{LC0} = K \times 0.018/\sqrt{28}$; $D_{LCO2} = K \times 0.51/\sqrt{44}$; and $D_{L02} = K \times 0.022/\sqrt{32}$. Thus:

  $D_{L02}/D_{LCO} = [K \times 0.022/\sqrt{32}]/(K \times 0.018/\sqrt{28}) = 1.14.$

  $D_{LCO2}/D_{LCO} = [K \times 0.51/\sqrt{44}] /(K \times 0.018/\sqrt{28}) = 22.6.$

Hereby she eliminated all the unknown variables, and for carbon monoxide, Ficks law of diffusion is simplified to:

Eq. 14-2: **(J_{gas} =)** $V°_{CO}$ = Δ P_{CO} × D_{LC0}

- **The alveolar gas equation** (P_{IO2} -P_{AO2}) = P_{ACO2} ∗[F_{IO2} + (1-F_{IO2})/R] in terms of alveolar gas tensions. We can simplify the *alveolar gas equation* for R=1:

Eq. 14-3: F_{IO2} - F_{AO2} = F_{ACO2} or **P_{IO2} - P_{AO2} = P_{ACO2}.**

- **The alveolar ventilation equation** describes the hyperbolic relationship between alveolar ventilation ($V°_A$) and F_{ACO2} :

Eq. 14-4: **$V°_A$ = $V°_{CO2}$/** F_{ACO2} .

F_{ACO2} is equal to [P_{ACO2}/(101.3 - 6.3) kPa], so P_{ACO2} is easily substituted for F_{ACO2}.

- **The final ventilation-perfusion ($V°_A$ / $Q°$) equation**

Without showing the calculations, one **equation** combines Eq.s 14-1 to 14-4:

Eq. 14-5: **$V°_A$ / $Q°$ = R(C_{aO2} - C_{v-CO2}) / F_{ACO2}.**

The $V°_A$ / $Q°$ -ratio is obviously independent of the metabolic rate or oxygen uptake.

$V°_A$ / $Q°$- ratio is the **key variable**, because we all have a certain degree of ventilation - perfusion mismatch, and in almost all cardiopulmonary patients this mismatch is consequential.

- **The total tension** of all three dry gasses is equal to (P_{O2} + P_{CO2} +P_{N2}) or the barometric pressure (P_B) minus the tension of water vapour in the alveolar air at 37^o C. The total tension at P_B=760 mmHg is thus (760 - 47) = 713 mmHg or (101.3 - 6.3) = 95 kPa.

Eq. 14-6: (P_B - 47) = (P_{O2} + P_{CO2} +P_{N2}).

- The **law of mass balance** is applied to both bloodflow and oxygen flux in the following two equations:

Eq. 14-7: $Q°_{total}$ = $Q°_{shunt}$ + $Q°_{capillary}$

Eq. 14-8: ($Q°_{total}$ · C_{aO2})= ($Q°_{shunt}$ · C_{v-O2})+ ($Q°_{capillary}$ * $C_{c´02}$)

where $C_{c´02}$ is the oxygen concentration in the **pulmonary end capillary blood from ideal lung units** (Fig. 14-8).

- The flow and flux relations implies the following **shunt equation**:

Eq. 14-9: $Q°_{shunt}$/ $Q°_{total}$ = (C_{aO2} - $C_{c´02}$)/(C_{v-O2} - $C_{c´02}$).

- The CO-Oxy-haemoglobin affinity equation:

Eq. 14-10: C_{aCO}/P_{aCO} = 250* C_{aO2}/P_{aO2}.

CO has a standard affinity for haemoglobin 250 times larger than that of oxygen for haemoglobin: C_{aCO}/P_{aCO} : C_{aO2}/P_{aO2} = 250 : 1.

- **Dalton's law** states that the *partial pressure or tension of a single gas in a mixture is equal to the product of the total pressure and the mole fraction (F).* According to Daltons law the fraction of oxygen in the alveolar air (F_{AO2}) is:

  Eq. 14-11: $\boldsymbol{F_{AO2} = P_{AO2}/(101.3 - 6.3) = P_{AO2}/(760 - 47)}$.

With an alveolar partial pressure of oxygen (P_{AO2}) of 13.3 kPa (or 100 mmHg), the F_{AO2} is **0.14**. There is no interaction between gasses.

- **Henry's law** states that the number of gas molecules dissolved in a fluid is directly proportional to the partial pressure of the gas in air above the fluid. According to Henrys law the concentration (C) of dissolved gas is proportional to its partial pressure (*P*) and the solubility (α or Bunsen's solubility coefficient, Box 13-1):

  Eq. 14-12: $\mathbf{C = P * \alpha}$.

With the pressure given in kPa or mmHg it is necessary to divide by 101.3 kPa or 760 mmHg, respectively, because α is defined at 1 atm.abs. pressure.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. The pulmonary vascular pressure and resistance (*PVR*) is only 1/10 of that of the systemic circulation.

B. The *PVR* is highest in intra-alveolar vessels at high lung volumes.

C. The *PVR* increases when pulmonary arterial pressures increase.

D. The pulmonary circulation is dependent on gravity but the pulmonary ventilation is not.

E. The P_{AO2} has a direct effect on pulmonary circulation.

Case History A

A male person, ages 23 and weight 70 kg, is breathing atmospheric air with traces of carbon monoxide (CO) at one atmosphere. The man is at rest, and has an arteriovenous oxygen content difference of 50 ml per l. An arterial blood sample obtained after equilibrium between alveolar air and pulmonary blood is analysed with the following results: P_{aO2} 13.3 kPa (100 mmHg), C_{aO2} 170 ml STPD per l, C_{aCO} (the concentration of CO in the blood) 28.3 ml STPD per l, and the [haemoglobin] 9.18 mM (148 g per l). The standard affinity between haemoglobin and CO is 260 times greater than the standard affinity between haemoglobin and oxygen. The binding capacity for oxygen and CO is 1.34 ml STPD per g of haemoglobin.

1. *Define the concept standard affinity and P_{50}.*
2. *Calculate the dry CO-fraction in the alveolar air (F_{ACO}).*
3. *Calculate the concentration of oxygen in the mixed venous blood of this patient.*
4. *Calculate the concentration of oxygen in the mixed venous blood of a comparable patient with anaemia (haemoglobin concentration 7.78 mM) and with the same arterio-venous oxygen content difference.*

5. *Is the oxygen supply to the tissues at the venous end of the capillaries better for the CO-poisoned person than for the anaemia patient?*

Case History B

A 49-year-old female, body weight 61 kg and height 1.7 m, is hospitalised due to severe, progressive dyspnoea. Six years ago the diagnosis of pulmonary sarcoidosis was established by mediastinal lymph node biopsy. The cause of the disease is unknown, and the patient has no history of previous lung disease. When stair climbing the patient has difficulties in reaching the 2. floor.

The spirometric standard values for a female of this age, height and weight are: forced expiratory volume on 1 s (FEV_1) of 2.9 l, and forced vital capacity (FVC) of 3.7 l. The patient has a FEV_1 of 1.3, and a FVC of 1.48 l. The patient has an unforced VC of 1.6 l, with an ERV of 300 ml, tidal volume of 600 ml and an IRV of 700 ml, as compared to a normal VC of 3.9 l.

The normal specific lung compliance (at FRC) is 2 ml per Pascal (Pa); for this patient it is determined to only 0.4 ml per Pa at FRC. The normal single-breath CO diffusing capacity is 3 ml STPD s^{-1} kPa^{-1}, but this patient has only 0.5 ml STPD.

An arterial blood sample shows a P_{aCO2} of 4 kPa (30 mmHg) and a P_{aO2} of 8 kPa (60 mmHg).

***1**. What are the arguments for the diagnosis of restrictive lung disease?*

***2.** Why is the single-breath CO diffusing capacity seriously reduced?*

***3**. Is there any indication of alveolar ventilation-perfusion mismatch?*

Case History C

Following 3 days of fishing in cold weather, a 30 year old man is brought to hospital with high fever (40.8 Centigrade), coughing with chest pains and red coloured sputum. Rales are heard over both lungs and a chest x-ray show large infiltrates in both lungs. A blood gas analysis on an arterial sample reveals P_{aO2} of 50 mmHg and

P_{aCO2} of 26 mmHg. pH_a is 7.38. The RQ is assumed to be 1, and P_B is 760 mmHg.

1. *Calculate the alveolar P_{O2} (P_{AO2}) using the alveolar gas equation.*

2. *Assume a likely value for an ideal gas composition (mean alveolar) just before the man became ill.*

3. *Calculate the alveolar (ideal) - arterial P_{O2} difference. What does this difference mean?*

4. *Calculate the difference between the alveolar ideal P_{ACO2} and the arterial (P_{aCO2}). What does this difference mean?*

Case History D

A male, 44 years of age, is brought to hospital due to severe dyspnoea. He has been smoking 40 cigarettes per day in 30 years. Over the last 10 years an increasing respiratory distress has developed, and the patient is well known at the medical department. The arterial blood gas tensions are measured:

P_{aO2} is 60 mmHg (8 kPa), P_{aCO2} is 35 mmHg (4.7 kPa), and pH_a is 7.44.

An alveolar gas sample reveals a P_{AO2} of 129 and a P_{ACO2} of 28 mmHg.

1. *Calculate the alveolar-arterial P_{O2} difference assuming that the ideal P_{AO2} is 100 mmHg (13.3 kPa).*

2. *Provide a likely diagnosis, which explains his respiratory distress.*

3. *Is there an abnormally high alveolar dead space?*

Case History E

A female surgeon, 56 years old, has smoked 25 cigarettes a day for almost 40 years. Her dyspnoea from stair climbing has increased substantially over the last three years as has her morning cough with abundant green sputum in big lumps. A chest X-ray shows hyperinflation, bronchial expansions and a distinct vascular pattern. The surgeon is examined at the respiratory laboratory including function tests and arterial blood gasses with the following results:

FEV_1 = 1.1 l (normal 2.6 l) ; Forced Vital Capacity (FVC) = 1.9 l s^{-1} (normal 3.4 l s^{-1}); P_{aCO2} = 56 mmHg or 7.5 kPa; pH_a = 7.21; P_{aO2} = 49 mmHg or 6.5 kPa; Base Excess = - 5 mM.

1. *What is the cause of the disease?*

2. *Characterise the acute condition including the acid-base status.*

3 *From where in the upper airways do the big lumps of green sputum arise?*

Try to solve the problems before looking up the answers.

Highlights

- *Any patient - as well as any healthy person - has some degree of ventilation-perfusion mismatch.*
- *The regional ventilation-perfusion-ratio is the key to understanding cardiopulmonary function.*
- *The regional ventilation-perfusion ratio varies theoretically from zero at the lower lung region (only bloodflow) to infinity at the upper region (only ventilation).*
- *The upper alveoli are always more expanded than those of the lower due to the pull of the gravity are, and they did follow the first in-last out principle: During inspiration the first to fill – during expiration the last to empty.*
- *The regional ventilation-perfusion ratios show the lower lung regions to be relatively underventilated (ratio below one), the middle lung regions to be well matched (ideal ratio of 1), and the upper lung regions to be relatively overventilated (ratio above 1 and approaching infinity.*
- *Pulmonary embolism creates an alveolar dead space. The alveolar ventilation of the region is maintained, but there is no bloodflow, so the $V°_A / Q°$ -ratio of the lung region approaches infinity. In the alveolar dead space, alveolar gas pressures approach the levels of inspired air.*
- *Tracheal occlusion represents an extreme mismatch of venous to arterial shunting of blood, namely perfusion with no ventilation at all (ie, the total ratio for the person approaches zero). The arterial blood gas tensions approach those of venous blood.*
- *The pulmonary vascular system is basically a low-pressure, low-resistance, highly compliant vascular system, which is meant to accommodate the entire cardiac output.*
- *The standard affinity of the haemoglobin-CO reaction is 250 times greater than that of haemoglobin-O_2 .*

- *The single-breath CO diffusing capacity (transfer factor) is normally 3 ml STPD s^{-1} kPa^{-1} at rest and 7.5 during maximal exercise.*
- *Uneven distribution of tidal volume can eventuate from uneven resistance to airflow within the lung (bronchoconstriction, collapse and compression of airways) or from uneven regional lung compliance (insufficient surfactant, loss of elastic recoil as in destruction of alveolar tissue, and increase of elastic recoil as in connective tissue scarring or fibrosis with stiff lungs).*
- *Hypoperfusion can be caused by compression of pulmonary vessels, obliteration of vessels by fibrosis, or blockage by emboli or thrombosis.*
- *Functional shunts arise with any consolidation of alveolar regions that continue to have bloodflow (pneumonia, oedema, haemorrhage, cell necrosis, lack of surfactant).*
- *Patients with lung disorders often have $V\dot{}_A / Q\dot{}$ -mismatch by a combination of serious veno-arterial shunting in the lower lung regions, and increased alveolar dead space in the upper lung regions.*
- *A healthy person at rest (FRC) has approximately half of the pulmonary capillaries open, but with increasing arterial pressure, previously closed capillaries open (recruitment).*
- *Pulmonary oedema is an emergency caused by filtration of fluid out of the pulmonary capillaries into the interstitial space (interstitial oedema), and eventually in the alveolar spaces (alveolar oedema).*
- *Patients with left cardiac failure (acute myocardial infarct, chronic myocardial failure, mitral stenosis, aortic stenosis, and hypertension) can suffocate, when the alveoli are filled with oedema fluid.*
- *The gas exchange of the chronically ill lung patient is reduced over the years, and abnormal arterial blood gas tensions develop already at rest. This late stage of lung disease is called terminal respiratory insufficiency, due to the grave prognosis.*

Further Reading

West, J.B. *Respiratory Physiology*. Williams & Wilkins, Baltimore. USA, 1999.

Änggård, E. "Nitric oxide: mediator, murderer, and medicine." *Lancet* 343: 1199-1206, 1994.

Return to content

Return to chapter 14

Chapter 12

Blood Flow, Distribution And Shock

Study Objectives

- To *define* concepts such as anaphylactic shock, bloodflow, hydrostatic indifference point, hypotension, mean circulatory equilibrium pressure, mean transit time, and shock.
- To *describe* the principle of mass balance (Fick principle) for cardiac output determination, the dilution principle (tracer or indicator bolus), the isotope-wash-out-method, the venous occlusion plethysmography, and the mean transit time method.
- To *draw* an indicator dilution curve, an isotope-wash-out-curve and a plethysmography curve with flow determinations.
- To *calculate* one variable when relevant information is given.
- To *explain* the control of bloodflow to the brain the myocardium, the kidneys, the muscles, the gastrointestinal channel, the skin, and the foetus. To explain the compensatory reactions to shock. To explain the cerebral ischaemic response.
- To *use* the concepts in problem solving and case histories.

Principles

- **The law of conservation of matter** (see Chapter 8). This principle is used to measure bloodflow.
- **Fick´s principle** for determination of cardiac output (see Chapter 10).
- **Poiseuille´s law**. When the radius of a vascular bed is doubled, its bloodflow may increase by as much as 16 times. This is because of Poiseuille´s law (see Eq. 8-3).

Definitions

- **Anaphylactic shock** *(anaphylaxis)* refers to a severe allergic disorder in which the cardiac output and the mean arterial pressure fall rapidly and drastically (see Chapter 32).
- **Hydrostatic indifference point** is the point in the cardiovascular system, in which the pressure does not change with change of body position.
- **Hypotension.** Severe hypotension refers to a condition with a systolic blood pressure below 75 mmHg (10 kPa).
- **Mean circulatory equilibrium pressure** is a pressure of 1 kPa measurable in all divisions of the circulatory system just after cardiac arrest.
- **Mean transit time** for indicator particles in a system is equal to the sum of all transit times for all single particles divided with their number.
- **Shock** is a clinical condition characterized by a gradual fall in arterial blood pressure and rapid heart rate. Respiration is also rapid and the skin is moist, pale or bluish-grey.
- **Vasovagal syncope** or ***emotional fainting*** is a condition, where the fainting is caused by a strong emotional

activation of the parasympathetic nervous system via the hypothalamus with bradycardia, vasodilatation and decreasing arterial pressure.

Essentials

This paragraph deals with *1. The coronary bloodflow, 2. The regulation of coronary bloodflow, 3. Brain bloodflow, 4. Skin & fat bloodflow, 5. The splancnic circulation, 6. The foetal circulation, 7. Fick's principle, 8. The dilution principle, 9. Clearance, 10. The isotope-wash-out-method, 11. The mean transit time, 12. Vascular pressure reference (supine to standing).*

1. The coronary bloodflow

The myocardial metabolism is an exclusively *aerobic* process under normal conditions. It depends on oxidative phosphorylation in order to re-synthesise ATP. The O_2 needs of the myocardium are therefore great, even at rest. Exercise can increase the needs *six-fold;* however, the myocardium cannot extract a greater fraction of the O_2 delivered, since the myocardial O_2 extraction is already close to maximum at rest. Thus the coronary bloodflow must rise importantly during exercise in order to deliver the O_2 needed.

Two main coronary arteries arise from the aorta. The *left main coronary artery* (Fig. 12-1) divides into two major branches: The *left anterior descending artery*, which courses down the interventricular groove towards the apex of the heart, and the *left circumflex artery*, which courses leftward and posteriorly in the atrioventricular groove to the postero-lateral wall of the left ventricle.

The *right coronary artery* (Fig. 12-1) arises from the right aortic sinus and courses rightward and posteriorly in the atrioventricular groove to reach the right atrium, and via the *posterior descending artery* to the posterior wall of the left ventricle and the lower part of the interventricular septum. Later the right coronary artery also gives off branches to the *posterolateral wall* of the left ventricle.

This arrangement of coronary vessels exists in *half* of the population in western countries. In 30% of the population the *posterior descending artery* arises from the *right coronary artery*, and the posterior left ventricular branch arises from the *left circumflex artery*. In another 20% of the population the *right coronary artery* is small and supplies only the right atrium and the right ventricle with blood, and all the blood supply to the left ventricle comes from the *left main coronary artery*.

The main arteries run along the epicardial surface and divide several times on the surface of the heart before they send off small penetrating vessels forming a network of intramural arteries, arterioles and capillaries in their way to the endocardium. The myocardial capillaries feed into a net of intramural venules. They drain eventually into the epicardial collecting veins. *Right ventricular* venous blood drains into the right atrium. *Left ventricular* venous blood drains into the *coronary sinus* that empties in the right atrium, except for a small blood volume, which drains into the left ventricle. The epicardial coronary vessels contain a preponderance of constrictor receptors called *adrenergic α-receptors,* whereas the intramuscular and endocardial coronary receptors have a preponderance of dilatator receptors called *adrenergic β-receptors.*

Fig. 12-1: Coronary bloodflow and receptors.

Due to the contraction of the myocardium in systole, the myocardial bloodflow is blocked and the heart receives its nutrition in the diastolic period (Fig. 10-7). The coronary bloodflow is *phasic*.

2. The regulation of coronary bloodflow

The coronary bloodflow is described before in Chapter 9 (paragraph 3) and in relation to Fig. 10-7.

The coronary bloodflow is mainly controlled by *local metabolic autoregulation,* and sympathetic stimulation does not always cause significant vasoconstriction. Accordingly, a moderate decrease in arterial blood pressure down to 9.3 kPa

(70 mmHg) does not significantly reduce the bloodflow through the myocardium.

Unlike skeletal muscle tissue, the myocardium cannot function anaerobically for extended periods by building up an *oxygen debt.* Thus, oxidative ATP synthesis must continuously match ATP utilisation in the heart. At rest the heart produces 70% of its ATP from oxidation of fatty acids and 30% from oxidation of carbohydrates.

During exercise with lactate production by skeletal muscles, this lactate becomes an important substrate for the myocardial metabolism, entering the tricarboxylic acid cycle after conversion to pyruvate.

Catheterisation of the venous sinus of the heart in healthy subjects at rest reveals a venous haemoglobin saturation fraction of 0.30. Hence, 0.7 parts of the haemoglobin concentration of the venous blood is desaturated. Thus, arterial blood with a normal oxygen concentration (C_{ao2} of 200 ml per l) liberates (200 × 0.7) or 140 ml of oxygen per l to the myocardium. Variations in the arteriovenous O_2 content difference at the *steep* part of the O_2-haemoglobin dissociation curve can only change the myocardial O_2 tension modestly. The extremely high O_2 content difference of the heart at rest implies that a *rise in coronary bloodflow* must be the main source of extra O_2 to the heart during exercise.

Most of the blood entering the coronary circulation is delivered during the diastolic phase. This is because the myocardial tissue pressure increases during systole, and the contraction squeezes the blood/myocardium - in particular in the subendocardial layer. Therefore, the systolic bloodflow through the inner layer of the left ventricular wall approach zero. The duration of each diasole is reduced with increasing heart rate, so the increased oxygen demands during exercise calls for a higher coronary artery pressure in diastole in order to secure the necessary bloodflow.

3. Brain bloodflow

The blood reaches the brain through the internal carotid and the vertebral arteries. The dominant control of cerebral bloodflow (CBF) is metabolic autoregulation but also a pressure dependent myogenic autoregulation is present. The smooth muscle walls of the small cerebral arteries respond immediately to changes in the transmural pressure gradient. Hereby, the CBF is maintained constant despite changes in systolic blood pressure between 80 and 160 mmHg (10.7-21.4 kPa). The small brain vessels are metabolically regulated. Increased P_{aCO_2} and reduced P_{aO_2} dilatates brain vessels and increase CBF. CO_2 (not H^+) passes the blood-brain barrier easily. The mean arterial pressure can double without any appreciable rise in CBF. A neuropeptide released in response to transient hypotension *(calcitonin gene-related peptide)* is probably involved in the autoregulation.

The vertebral arteries join to form the basilar artery, which forms the circle of Willis together with blood from the internal carotids. When brain arterioles dilatate the CBF increases, and since the brain tissue within the cranium is relatively incompressible, the venous outflow must balance.

CBF is normally 55 FU in humans at rest. One FU is one ml of blood per min per 100 g of brain tissue. With a normal brain weight of 1300 g this value corresponds to a total of (55 × 13) = *715 ml* of blood per min. This resting CBF and the oxygen uptake of the brain can double during *cerebral activity* and triple in active brain regions during an *epileptic attack* (Chapter 7).

The sympathetic nervous system plays a secondary role for the CBF. Some brain vessels contract by *sympathetic stimulation.* This neurogenic control only concerns the *larger* cerebral arterioles.

Some degree of autoregulation is found in many other organs including the *skeletal muscle mass, the splancnic area, and the kidneys.*

4. Skin & fat bloodflow

Blood flows through the skin and subcutaneous tissues in order to *nourish* the cells, and to *regulate* shell temperature. Blood flows much faster through the arteriovenous anastomoses in the skin of the face, the fingers and toes in a cold

environment. The sympathetic activity constricts the metarterioles that lead to the skin, so the blood bypasses the cutaneous circulation. Hereby, the skin bloodflow can fall from about 5 FU´s and approach zero. The heat content of the blood returns to the body core, which helps to maintain the core temperature.

In a warm environment, the sympathetic tone is minimal and the arterioles dilatate, so that the skin perfusion can rise to perhaps 70 FU, and much energy is given off to the atmosphere. *Psychological influence* can cause one to blush or to have a white face, by *changing α-adrenergic* constrictor tone and through the effect of local, vasoactive substances normally found in the skin. When a large fat combustion occurs (during hunger and distance running), the fat bloodflow can increase from 3 to 20 FU´s. Cold and warm environments alter the fat perfusion just like the skin perfusion (Chapter 21). The sympathetic regulation of the arteriolar tone in fat and skin tissue is also similar. The sympathetic change in tone is not related to the classical baroreceptors.

5. The splancnic circulation

The splancnic area is drained (1.5 l per min) via the hepatic veins at rest, so *all blood* passes through the liver. The liver receives more than one litre of blood from the portal vein and less than 0.5 l from the hepatic artery each minute. A special characteristic for the splancnic circulation is that two large capillary beds are partially in series with one another forming a *portal system*. The splancnic perfusion increases after meals, and decreases during fasting and duration exercise. The sympathetic nervous system has a tonic activity on splancnic vessels via a-adrenergic nerve fibres. Vagal fibres dilatate the splancnic vessels. Haemorrhagic shock can elicit a fatal splancnic hypoxia.

6. The foetal circulation

The foetus depends completely on the mother and her placenta. The *placental barrier* can be passed by low molecular substances (nutrients, gasses and waste) by diffusion.

The *foetal* haemoglobin (F) has a sigmoid dissociation curve, which is shifted to the left relative to haemoglobin A for adult, because haemoglobin-F has a greater affinity for oxygen than does adult haemoglobin (Chapter 15).

Haemoglobin F is not affected by 2,3-DPG. Foetal blood has a high haemoglobin concentration (200 g l^{-1}), so the foetus takes up large amounts of O_2 in placenta. This occurs even at low P_{O_2}, and the maximal value in the placental blood is only 6.7 kPa or 50 mmHg. Often the foetus achieves an arterial oxygen concentration similar to that of the mother.

Steroid hormones, maternal thyroid hormones, and catecholamines cross the placental barrier to the foetus. Peptide hormones cannot traverse the placental barrier, except small peptides such as thyrotropin-releasing hormones (TRH) and antidiuretic hormone (ADH).

Foetal insulin contributes to anabolism and lipid storage. Human chorionic somatomammotropin, prolactin, and IGF-2 are the most important growth factors during foetal life. Foetal parathyroid hormone stimulates the transport of Ca^{2+} to the foetus.

The *maternal blood* rich in O_2 and nutrients is injected into the *intervillous spaces* of the placenta via spiral arteries, and returns with CO_2 and waste to the mother via veins draining to the uterine veins. The *chorionic villi* dip into the internal sinuses of the placenta. The exchange of nutrients, metabolic waste and gasses across the placental barrier occurs by diffusion. Foetal blood rich in O_2 and nutrients returns to the foetus from the placenta in the umbilical veins.

The placenta has a diameter of 16-20 cm and the placental barrier has an area of about 10 m^2. The blood flowing in the umbilical veins continues in the *ductus venosus* or the blood enters the foetal liver and then all blood is gathered in the inferior caval vein.

The 80%-saturated blood from the umbilical veins flows into the foetus. The saturation is reduced to 2/3 (67%) when passing through the oval communication between the right and left atria (*foramen ovale*) to reach the left ventricle (Fig. 12-2). The blood in the left ventricle is mixed with desaturated blood from the lungs, whereby the oxygen

saturation is reduced to about 65% in the blood passing to the upper-body foetal organs (brain and heart).

The blood in the right ventricle (a mixture of blood mostly from the superior caval vein, but also the coronary sinus and the inferior caval vein to some extent) is only half-saturated. The foetal pulmonary vascular resistance is high due to the compressed inactive lungs (Fig. 12-2). Thus, the major part of the right ventricular output to the pulmonary artery bypasses the pulmonary circulation and flows through a foetal channel (*ductus arteriosus*) between the pulmonary trunk and into the descending aorta (Fig. 12-2).

When the major part of the blood passes the ductus arteriosus and joins the left ventricular output, the resulting oxygen saturation is around 60% in the blood of the descending aorta reaching the lower part of the body and back to the placenta in the umbilical arteries. This shunt delivers well-saturated nutritive blood to upper-body foetal organs (brain and heart). Venous drainage from these essential organs returns to the foetal heart in the superior vena Cava. The right ventricle ejects the venous return.

Much of the blood flowing in the descending aorta of the foetus is directed toward the placenta, so venous drainage from all organs is shunted toward the placenta, where wastes are eliminated from foetal blood, whereas O_2 and nutrients are acquired. Maternal hypoxia and reduced venous return or pressure on the umbilical vessels during birth reduces the oxygen supply to the foetus, which is reflected as bradycardia.

At birth, P_{CO_2} increases, and the first breath reduces the intrathoracic pressure in the new-born, so placental blood is sucked into the baby (placental transfusion). When the bloodflow through the umbilical vein ceases, the muscular sphincter of the ductus venosus contracts. Massive sensory stimuli of the baby caused by labour and delivery, cutaneous cooling after delivery, the falling P_{AO2} and rising P_{ACO2} without the placenta, and withdrawal of a placenta-produced respiratory inhibitor all adds up in activation of the respiratory centre and maintaining breathing in the newborn. The newly established air-liquid interface reduces pulmonary surface tension, which eases the lung expansion.

Distension of the lungs with air also distends pulmonary vessels, so pulmonary vascular resistance (*PVR*) decreases drastically. Hereby, pulmonary bloodflow increases. As a consequence, bloodflow via ductus arteriosus slows, and pulmonary venous return to the heart increases (Fig. 12-2).

Fig. 12-2: The foetal circulation.

The left atrial pressure increases above that in the right atrium and the inferior vena cava by the newly established decrease in pulmonary vascular resistance, because of the large pulmonary bloodflow to the left atrium.

Occlusion of the umbilical vein reduces the bloodflow to the right atrium, and occlusion of the umbilical arteries increases the resistance to the left ventricular output of blood. The resulting elimination of the pressure gradient across the atria, abruptly closes the valve over the *foramen ovale*, and the septal leaflets fuse within a couple of days.

The low pulmonary vascular resistance reduces the pressure in the pulmonary artery, whereas the aortic pressure rises. This reverses the flow of oxygenated blood through the *ductus arteriosus*. Within minutes after lung expansion, the muscular wall of the ductus arteriosus constricts, and its closure is complete within 10 days.

Failure of the foramen ovale or the ductus arteriosus to close gives rise to two congenital cardiac abnormalities (see later in this Chapter).

Normal arterial blood gas tensions are established by 30 minutes of age. Left atrial pressure increases and foramen ovale closes soon after birth. This *reverses* the blood pressure gradient across the foramen ovale, so now the left atrial pressure exceeds the right. When the umbilical cord is closed, and the placental circulation is thus eliminated, the *TPVR* of the newborn increases. The decrease in pulmonary vascular resistance (*PVR*) and increase in *TPVR* means a great difference in the size of the blood pressures in the aorta and in the pulmonary artery.

In conclusion, the *parallel* foetal circulation is transformed into a *series circulation* in the baby. The foramen ovale

and ductus venosus close within 3 days of birth (Fig. 12-2). The *sharp increase* of O_2 content (C_{aO_2}) in the baby's blood is a potent and universal vasodilatator. The dramatic changes in gas exchange affect cardiopulmonary and vascular regulation, probably via local mediators such as *arachidonic acid* and *prostacyclin*.

7. Fick's principle

According to the *law of conservation of matter*, mass or energy can neither be created nor destroyed (the principle of mass balance). *Adolph Fick* applied natural occurring indicators like O_2 and CO_2 when measuring cardiac output. Using O_2 as an indicator and the law of mass balance, he claims that the O_2 flux, taken up by the lung blood, plus the venous O_2 flux to the lung, must be equal to the O_2 flux, which leaves the lung in the oxygenated blood. Thereby, Fick proposed that the cardiac output could be calculated according to Eq. 12-1.

A classical example of the usefulness of Fick's principle is to consider the data of a young healthy male at rest. The typical data for such a person are an O_2 uptake of *250* ml STPD per min and an arteriovenous O_2 content difference of *50* ml STPD per l of blood. According to Fick's principle, this male can only satisfy his O_2 demands, if 5 l of blood is oxygenated in his lungs every minute. Thus, a cardiac bloodflow of 5 l per min is his cardiac output.

The oxygen concentration of mixed venous blood is usually obtained through a venous catheter inserted up the median cubital vein, through the subclavian vein, and finally into the right ventricle or pulmonary artery, where the blood is well mixed. Arterial blood is easily obtained from the radial artery (C_{aO2}). The disappearance rate of oxygen from the respired air can be recorded in a metabolic ratemeter as the oxygen uptake.

The principle of mass balance is valid only for a system in *steady state*. Steady state is a state where the indicator is administered at a constant rate, and is neither stored, mobilised, synthesised nor used by the system, and where no shunts are present.

This method has been used to measure a large increase in cardiac output in different patient groups. For example, patients with anaemia have been found to have higher cardiac output at rest.

8. The dilution principle

When an *indicator* bolus (mass or dose of *tracer* in weight or molar units) is instantaneously injected in the right side of the heart, the indicator and blood will mix. The mixture leaves the right ventricle through a *well-mixed outlet*, passes the pulmonary circulation and then returns to the left side of the heart. The indicator concentration during the first passage of any peripheral artery is recorded continuously or by multiple sampling. The resulting curve is shown in a semilog scale (Fig. 12-3). The indicator concentration (in mol/ml of blood) reaches a peak and then decreases in a few seconds, before it again rises due to indicator recirculation with the blood (Fig. 12-3). The first decrease in concentration is assumed to be mono-exponential. Hence it is easy to extrapolate to the concentration zero, and read the so-called *first passage time*, T_1. In this case the T_1 is 9 s (Fig. 12-3). The mean concentration (c mol/ml) of indicator in the period T_1 seconds is determined by planimetry.

The average amount of indicator (in moles) leaving the left ventricle per second in one ml of blood is c, hence c is given in mol/ml of blood. The volume of blood (V) in which the indicator dose is distributed is dose/c. Since blood carries only c mol of tracer in each ml, the heart needs at least a bloodflow of V ml (dose in mol/ c) in order to carry the entire dose through the aortic orifice in T_1 (9) seconds. Accordingly, the cardiac output per second is dose/($c*T_1$). The product ($c*T_1$) is the area under the curve (Fig. 24-1). Thus the *dose/area ratio* must be equal to the bloodflow (ml/second) leaving the left ventricle.

Put simply, the bloodflow (cardiac output in ml of blood per s or more convenient per min) can be measured by dividing the *dose* of indicator injected upstream by the *area* under

the downstream concentration curve.

Fig. 12-3: The indicator dilution principle.

The bloodflow equation is also called the *dose/area equation.*

An attractive choice of indicator is to use *cold saline*, of known temperature and volume. A flexible catheter, with a thermistor located at its tip, and an opening through which cold saline can pass, is used. The catheter tip is advanced to the pulmonary artery, while the opening supplies the right atrium with saline. The thermistor records the downstream alterations in temperature as the saline bolus passes. This is the *thermodilution technique*. This technique can be frequently repeated without having harmful effects. Moreover, there is negligible recirculation, and the method spares the patient the ordeal of an arterial puncture.

This method is widely used. For example, interesting indicator dilution studies have shown the pump effect of external cardiac massage to be modest.

9. Clearance

Clearance is a theoretical tool for estimating bloodflow in the kidney and other organs. *Clearance* is the volume of blood plasma, which is *totally cleared* each minute of a given indicator by a specific organ (eg, renal clearance). The extraction (E) is the fraction of substance, which is extracted from the total amount transported to the organ per minute (Eq. 12-3).

Clearance for para-amino hippuric acid (PAH) at low plasma concentrations is a measure of the *renal plasma flow* (RPF – see Ch. 25). The high hepatic extraction of bromsulphalein or of indocyanine is used to estimate the *splancnic perfusion.*

10. The isotope-wash-out-method

A lipid-soluble indicator, such as ^{133}Xenon dissolved in saline, is injected in the *tibialis anterior muscle* (* in Fig. 12-4 and Eq. 12-4). At steady state, the tracer concentration in the venous blood (C_v) is assumed to be the *average* blood concentration, and C_{tis} the *mean* tissue concentration.

Fig. 12-4: Isotope (^{133}Xe) wash out from the gastrocnemius muscle before, during and after walk on a treadmill. Upper curve is when the femoral artery is occluded - lower curve is from the healthy leg.

The *fractional fall* in the mean tissue concentration of Xenon (C_{tis}) per time unit (dt) is *constant* during the whole elimination period (a rate constant = ln 2/$T_{½}$). The flow/W_{tis} (weight of tissue) is a perfusion coefficient in FU (ml of blood per min and per 100 g tissue). The fall in mean tissue concentration per time unit (ratio dC_{tis}/C_{tis}) is measurable as $T_{½}$ on the skin surface at the Xenon deposit in muscle tissue with a scintillation detector (Fig. 12-4 and Eq. 12-4).

The method (see Eq. 12-4) is used clinically to detect peripheral vascular diseases. An example is *intermittent claudication* that refers to constricting pain arising during activity of any muscle group but most commonly in the calf muscles. The hypoxic pain and cramp appear after having walked a certain distance and is promptly relieved by rest. The cause is *femoral occlusion* due to arteriosclerosis with insufficient local bloodflow and *ischaemic hypoxia* (Fig. 12-4).

11. The mean transit time

The *mean transit time* (t_{mean}) for indicator particles in a system with the volume, V, is equal to the sum of all transit times for all single particles divided with their number.

This concept is used in a wide variety of indicator methods (Eq. 12-5).

By means of intravascular catheters it is possible to measure the *partial* circulation time through most parts of the circulation. For a healthy adult at rest the normal ranges include the following: arm-ear 8-12 s, arm-lung 5-7s, and

lung-ear 3-5 s.

12. Vascular pressure reference

The heart is not always the correct reference point for blood pressure measurements. The elastic properties of the vascular tree differ throughout the body.

Actually, the point in which the pressure does not change with change of body position is approximately 5 cm beneath the diaphragm during expiratory relaxation (Fig. 12-5). This is called the *hydrostatic indifference point* (HIP). Above this horizontal level, all vascular pressures are lower in the erect than in the recumbent position. The subatmospheric intrathoracic pressure counteracts venous collapse, so the intrathoracic veins remain open and the atrial pressures are zero in the erect position. The veins of the neck and face are collapsed. The venous sinuses of the brain are kept open by attachment to the surrounding tissues, and their pressures are around -1.3 kPa (-10 mmHg) in the erect position.

Fig. 12-5: The hydrostatic indifference point (HIP) in an adult male. The subject changes position from recumbent to erect.

HIP must not be mixed up with the *mean circulatory equilibration pressure*, which is a pressure of 1 kPa (6 mmHg) measurable in all divisions of the circulatory system just after cardiac arrest. This is also called the *mean circulatory filling pressure*, because it is a determinant of the venous return.

When a supine person arise, his TPVR increases, the systolic blood pressure falls andthe diastolic blood pressure rises. Thus, the pressure amplitude falls, but the mean arterial pressure is unchanged. The stroke volume is reduced more than the heart rate rises, so the cardiac output will decrease when attending the standing position.

An elegant way of studying circulatory consequences of standing is by use of *lower-body-negative-pressure* (LBNP). LBNP applied to a recumbent subject simulates the circulatory effects of standing.

The venous return to the heart is dependent upon the body position, and upon the total blood volume. The venous return is also dependent upon the venous compliance and upon the sympathetic tone in the venous system and in the arteriolar system.

When a person is located on a tilt table in a horizontal position, his blood pressures in a superficial vein on the feet is approximately 1.6 kPa (12 mmHg) and in the femoral veins 0.8 kPa (6 mmHg). When the person is turned upright towards the vertical plane, the venous pressure increases by the hydrostatic column up to the *hydrostatic indifference point* (HIP) just below the heart as long as he is not standing and using his skeletal muscle pump. If the tilt table is turned, so the head of the person is downward (*Trendelenburg position*), then the venous pressure increases in neck and head. The Trendelenburg position is rational during neck and head surgery. With the head upward the patient risks *air* embolism, if blood vessels are cut during neck and head surgery due to the subatmospheric pressure in the vessels.

Pathophysiology

This paragraph deals with *1. Shock and 2. Congenital heart disease.*

1. Shock

Shock is defined as a clinical condition characterised by a gradual fall in arterial pressure and a rapid heart rate. Respiration is also rapid and the skin is pale, moist and grey. The general circulatory insufficiency causes the bloodflow to vital tissues to be inadequate, so delivery of oxygen and other nutrients as well as elimination of waste products is insufficient.

In principle the circulatory insufficiency can be caused by disorders in the heart (cardiac insufficiency with imminent or manifest *cardiogenic shock*) or in the vessels (vascular insufficiency developing into vascular shock).

The *cardiogenic shock* can be caused by restricted ventricular filling (bi- or tricuspidal stenosis, pericardial fibrosis, or cardiac *tamponade*); the cause can also be myocardial disorders (infarctions, myocarditis etc) or restricted/ineffective ventricular ejection in in cases wwith semilunar stenosis/insufficiency or shunts.

The *vascularly generated shock* is caused by loss of blood or other fluids (absolute hypovolaemia) or by vasodilatation (relative hypovolaemia). *Absolute hypovolaemia* is caused by blood loss, plasma loss (burns or other denuding conditions, ascites, hydrothorax etc) or dehydration (water deprivation, severe diarhoea or vomiting, excessive sweating, intestinal obstruction with luminal fluid accumulation, urinary loss of proteins/salt/water, excessive use of diuretics, hypoaldosteronism etc). *Relative hypovolaemia,* sometimes with universal vasodilatation, is released by endotoxins (septic shock from viral or bacterial infections), anaphylactic shock (see Ch. 32) or by a neurogenic vasodilatation (neurogenic shock by severe pains or stress, anaesthetics or brain stem lesions close to the vasoconstrictor centre).

The *reduced delivery* of oxygen and nutrients to virtually all cells of the body, is consequential: The mitochondria synthesise less ATP, the Na^+-K^+-pump operates insufficiently, the metabolic processing of nutrients is depressed which profoundly depresses muscular contractions, and finally digestive enzymes destruct the damaged cells. Glucose transport across the cell membranes in the liver and in the skeletal muscles is depressed including a severe inhibition of the actions of insulin and other hormones (Chapters 26 and 30). During progressive shock the metabolism is reduced and thus the heat energy, so the *body temperature* tends to decrease, if the patient is not kept warm.

Compensatory mechanisms in shock are called *negative feedback* mechanisms, because they operate to counteract the fall in blood pressure. Baroreceptor responses and many hormonal control systems, that tend to raise the falling blood pressure, are examples of negative feedback (Fig. 12-6). The gain of a feedback system is defined as the ratio of the response to the stimulus itself.

Decompensatory mechanisms exaggerate the primary fall in blood pressure. This is called *positive feedback*. A positive feedback mechanism can lead to a *vicious cycle* and death, if its gain is above one. Two examples with ischaemic brainstem depression and cardiac depression are shown in Fig. 12-6.

Shock is divided into 3 stages by severity:

1A. *Mild shock* is a condition, where compensatory reactions can cure the patient without external help. A latent shock is produced when a healthy blood donor delivers more than the usual 500 ml of blood for transfusion, but the volume is often replaced within an hour. A number of negative feedback mechanisms oppose the induced changes of shock. The fall in MAP and pulse pressure reduces the stimulation of the high-pressure baroreceptors in the carotid sinus and the aortic arch. The negative stimulation of the cardiovascular control centres in the brainstem enhances the sympathetic tone (and reduces the vagal tone) leading to increased heart rate and contractility as well as to arteriolar and venous constriction mainly in the skin, skeletal muscles and the splancnic area. The bloodflow favours the brain and the heart as long as possible.

An array of other compensatory reactions are given in Fig. 12-6: Increased vascular permeability, reduced capillary pressure with *autotransfusion* from the interstitial fluid, thirst and drinking followed by absorption of fluid from the gastrointestinal tract, and release of powerful vasoconstrictors such as adrenaline, angiotensin II, vasopressin etc.

Catecholamines and enkephalins are released from chromaffine granules in the adrenal medulla. Catecholamines increase the heart rate and the cardiac output by stimulation of the adrenergic b_1-receptors in the myocardium. Catecholamines constrict vessels all over the body by stimulating a_1-receptors located on the surface of vascular smooth muscles.

ADH (*vasopressin*) is secreted from the posterior pituitary gland in response to shock, because the *sinoaortic* baroreceptors are under-stimulated. Vasopressin is a modest vasoconstrictor and a strong antidiuretic hormone. The increased ADH secretion causes increased fluid reabsorption by the kidneys and restores blood pressure and volume.

Renin is secreted from the juxtaglomerular apparatus, when blood pressure and renal perfusion falls drastically. Renin acts on the plasma protein, angiotensinogen, to form inactive angiotensin I, which is transformed to the powerful vasoconstrictor, angiotensin II by angiotensin converting enzyme, ACE. The most likely trigger of the *renin-*

angiotensin-aldosterone cascade is described in Chapter 24. - The rise in normal plasma-[K^+] due to the ischaemia of shock also releases aldosterone.

ACTH and β-endorphins are released into the blood from the anterior pituitary gland Chapter 26) in response to haemorrhage or other forms of stress. ACTH and endorphins both exaggerates and restricts the development of shock. These opioids depress the brainstem control centres that normally mediate autonomic responses to stress. Hence, naloxone (an opioid antagonist) improves the circulation and increases the rate of survival from life-threatening shock. - On the other hand, ACTH has a small aldosterone and a strong cortisol stimulating effect.

Initially, the bleeding patient suffers from *hypercoagulability*. Thromboxane A_2 (TxA_2) aggregates thrombocytes, and the aggregate releases more TxA_2. This positive feedback prolongs the clotting tendency. In this phase anticoagulants (heparin) reduce the mortality from shock.

Fig. 12-6: Development of shock conditions. Effects, effectors and reactions are shown.

1B. *Serious shock* leads to *myocardial damage*, because the arterial pressure is too low to secure a coronary bloodflow adequate for nutrition. Myocardial contractility is depressed, and the ventricular function curve shifts to the right (Fig. 10-5B).

Loss of more than 35% of the total blood volume of a healthy person is a threat, if the loss is unaided by blood transfusion. An arterial blood pressure below 8 kPa, where there is no additional baroreceptor response can stimulate the chemoreceptors of the carotid body and increase ventilation. The sucking effect of the low inspiratory pressure improves the venous return to the heart.

Cerebral ischaemia is consequential at arterial pressures below 5 kPa. The cerebral hypoxia elicts a generalised and powerful sympathetic stimulation with a pronounced arteriolar-and venous-constriction. Further hypoxia in the brainstem activates the vagal centres resulting in bradycardia.

1C. *Irreversible shock* is a terminal condition, where all therapy is frustraneous. Nothing can save the patient. The progressive deterioration becomes irreversible at a blood loss of more than 50% of the total blood volume. The drastic fall in arterial blood pressure reduces the renal glomerular filtration pressure below the critical level, so filtration is diminished or abolished, leading to abolish urine output (anuria). The low cardiac output and bloodflow result in stagnant hypoxia of all mitochondria. Hypoxia increases lactic acid liberation. Renal failure with tubular necrosis prevents excretion of excess H^+. The high H^+-concentration further depresses the myocardium, reduces blood pressure and thus the tissue bloodflow. This aggravates the metabolic acidosis - a classical *vicious* cycle (Fig. 12-6).

In the later stage of haemorrhagic shock, there is fibrinolysis and prolonged coagulation time (*hypocoagulability*). Hence, heparin therapy can be lethal.

The phagocytic activity of the reticulo-endothelial system (RES) is depressed during shock. Endotoxins constantly enter the blood from the bacterial, intestinal flora of a healthy person. The macrophages of the RES (Chapter 32) normally inactivate these endotoxins and release mediators such as hydrolases, proteases, oxygen free radicals, coagulation factors, prostaglandins, thromboxanes and leucotrienes. Some of these mediators modulate the temperature control and hormone secretion.

Following loss of half the total blood volume the shock patient must have lost about 50% of his circulating macrophages, and control substances modulating the phagocytic activity of RES. The depressed defence mechanisms in RES result in an endotoxic shock which aggravates the haemodynamic shock - a vicious cycle.

The patient loose consciousness and falls into a state of *stupor* or of *coma*.

A severe shock becomes irreversible, when the high-energy phosphate stores of the liver and heart are depleted. All of the creatine phosphate is degraded and almost all ATP has been degraded to ADP, AMP and eventually to the even more efficient vasodilatator, *adenosine*. Adenosine diffuses out of the cells and into the circulation, where it is

converted into hypoxanthine and uric acid, a substance that cannot re-enter the cells. Cellular depletion of high-energy phosphate is probably causing the final state of irreversibility.

The cerebral bloodflow is now so low that the function of the cardiovascular brainstem centres is depressed. The loss of sympathetic tone leads to cardiac depression with terminal bradyardia and vasodilatation with falling peripheral resistance. The fall in arterial pressure intensifies the damage, and a vicious cycle is established.

Two types of shock deserve special consideration:

Anaphylactic shock (anaphylaxis with relative hypovolaemia) is a severe allergic disorder in which the cardiac output and the mean arterial pressure fall rapidly and drastically due to relative hypovolaemia. As soon as an antigen to which the patent is sensitive, has entered the blood the *antigen-antibody reaction* (Chapter 32) triggers release of histamine from basophilic cells in the blood and mast cells in the tissues. Histamine dilatates arterioles and most peripheral vessels. This results in falling arterial pressure and increased capillary permeability with rapid loss of plasma water into the interstitial fluid.

Septic shock (relative hypovolaemia). Septic shock or blood poisoning is a widespread bloodborn bacterial infection - often life threatening. Examples are gas gangrene bacilli spreading from a gangrenous limb, colon bacilli with endotoxin spreading into the blood from infected kidneys (pyelonephritis), and *fulminant peritonitis* due to acute abdominal disease. Frequent causes of fulminant peritonitis are rupture of the infected gut or the uterus, and rupture of the uterine tube due to extrauterine pregnancy. Septic shock is characterized by tremendously high fever, high cardiac output, marked vasodilatation, red cell agglutination, disseminated intravascular coagulation withmicroclots spread all over the circulatory system. When the clotting factors are used up, internal haemorrhages occur. Endotoxins produce vasodilatation, induce synthesis of nitric oxide (NO) synthase in the vascular smooth musculature. Overproduction of NO may contribute to the vasodilatation and the depressed myocardial contractility found in septic shock. The high cardiac output is due to a high stroke volume and a high heart rate. The diastolicpressure is low and the systolic pressure is high until endotoxins begin to inhibit myocardial contractility seriously. Now the condition is in a vicious cycle, which is often fatal.

Therapy keypoints are*:*

First of all the *cause* has to be established in order to give the appropriate therapy.

1. *Head-down position* (placing the patient's head below the level of the heart) is the immediate therapy of haemorrhagic and neurogenic shock.

2. *Haemorrhagic arrest* (closing the abdominal aorta with the pressure of a fist, or blocking the bleeding from an artery with a finger) is often life saving when applied without unnecessary delay.

3. *Replacement transfusions*. The best possible therapy of haemorrhagic shock is whole blood transfusion. The best treatment of shock caused by plasma loss is plasma transfusion, and the best therapy of dehydration shock is transfusion with the appropriate solution of electrolytes.

4. *Oxygen breathing* is always helpful in shock with insufficient delivery of oxygen.

5. *Sympathomimetic drugs* (noradrenaline, adrenaline etc) are often beneficial in neurogenic and anaphylactic shock. They are seldom useful in haemorrhagic shock, where the sympathetic nervous system is already activated to its maximum.

2. Congenital heart disease

On a global scale approximately 1% of live births result in *congenital heart disease*. The mother may have suffered from rubella infection, abuse of drugs or alcohol, exposure to influential radiation or other factors causing genetic or

chromosomal abnormalities.

Two pathophysiological phenomena occur: *Right-to-left shunting of blood* results in cyanosis, clubbing of fingers, reduced growth in children, *exertion syncope*, and paradoxical emboli from veins to systemic arteries. Children with Steno-Fallot´s tetralogy (Fig. 12-7) use *squatting*. The advantage of squatting for the children is improved cerebral oxygenation as the position reduces the right-to-left shunt.

Left-to-right shunts in the heart result in *pulmonary hypertension*, because of persistently increased pulmonary bloodflow and vascular resistance (the so-called *Eisenmenger* response - frequently caused by ventricular septal defect).

Three classical congenital heart diseases are Steno-Fallot´s tetralogy, Coarctation of the aorta, and persistent ductus arteriosus.

Steno-Fallots tetralogy

The four elements suggested by the name tetralogy are: 1. Ventricular septal defect, 2. Overriding aorta (ie, the aortic orifice is located above the ventricular septal defect and therefore receives blood from both ventricles), 3. Pulmonary stenosis (right ventricular outflow obstruction), and 4. Right ventricular hypertrophy.

The pulmonary stenosis causes a high right ventricular pressure, and as this pressure supersedes the left ventricular pressure there is a right-to-left shunting of blood through the ventricular septal defect. The mixed blood passes through the overriding aorta, and the patient is cyanotic with all the hypoxic consequences described above (eg, exertion syncope, squatting, small stature, finger clubbing and polycythaemia). The children are tired and dyspnoeic, and growth is retarded although they often demonstrate a surprising appetite, because of the enormous cardiac work. Complete surgical correction is possible.

Fig. 12-7: Three common congenital heart disorders.

Coarctation of the aorta

Coarctation is a narrowing of the aorta distal to the insertion of ductus arteriosus (Fig. 12-7), and often associated with stenosis of the aortic orifice (a bicuspid aortic valve). The obstructed aortic bloodflow forces blood through collateral arteries such as the intercostal and periscapular arteries. The high blood pressure may cause nose bleeds and headaches, and the low distal bloodflow may cause claudication and cold legs.

Turbulent bloodflow through the coarctation is often recognized as a forceful systolic murmur even on the back.

Surgical excision of the coarctation with end-to-end anastomosis must be performed in childhood, because a low renal bloodflow, maintained over years, frequently results in irreversible systemic hypertension

Persistent ductus arteriosus (PDA)

In the foetus the ductus arteriosus leads blood into the systemic circulation instead of through the unexpanded lungs. Hereby, the foetal blood is oxygenated during its passage of the placenta. At birth, the expansion of the lungs with atmospheric air triggers contraction and closure of ductus arteriosus by constriction of its muscular wall. Premature babies and children borne by mothers, who suffered from rubella in the first trimester, are often born with PDA.

A *PDA* shunts blood from the aorta to the pulmonary artery throughout the cardiac cycle (Fig. 12-7). This is because the aortic pressure is much higher (150/80 mmHg) than that of the pulmonary artery (40/20 mmHg or 5.3/2.7 kPa).

The condition may be symptom-less for years, but a large shunt increases the work of the left ventricle and causes left heart failure, which increases the risk of pulmonary congestion and oedema. The person is *easily fatigued* even from

moderately strenuous exercise.

The frequency of *infective endocarditis* is increased in PDA. Infective endocarditis commonly occurs on congenitally or rheumatically damaged valves. The endocardium also suffers from jet lesions located on the endocardial surface opposite to a shunt with a high driving pressure. This is common for all types of congenital heart disease apart from atrial septal defect, where the driving shunt pressure is too small to damage the endocardium.

The continuous shunting of turbulent blood causes a continuous machinery murmur best heard below the left clavicle.

Surgical ligation of the duct should be performed as early as possible.

The general trend today is an increasing survival rate of congenital heart disease. Adults living with congenital heart disease for years may present themselves as cardiac arrhythmias resistant to standard therapy or at autopsy following sudden cardiac death.

Terminal heart failure is now managed by heart-lung transplantation.

Equations

- The Fick cardiac output equation states that the cardiac output is calculated from the ratio between alveolar oxygen uptake and arteriovenous oxygen content difference:

Eq. 12-1: $\dot{Q} = \dot{V}_{O2}/(C_{aO2} - C_{v\,O2})$; [ml STPD*min^{-1}/ml blood*min^{-1}]. The last oxygen concentration is in mixed venous blood.

- The law of mass balance is applied to both bloodflow and oxygen flux in Eq. 14-7 and 14-8. The flow and flux relations implies the following shunt equation:

Eq. 12-2: $\dot{Q}_{shunt}/\dot{Q}_{total} = (C_{aO2} - C_{c'02})/(C_{v\,O2} - C_{c'02})$. The last oxygen concentration is in pulmonary end-capillary blood.

- Clearance is the volume of blood plasma, which is totally cleared each minute of a given indicator by a specific organ (renal clearance). The extraction (E) is the fraction of substance that is extracted from the total amount transported to the organ per minute.

Eq. 12-3: $E = \dot{Q}(C_a - C_v)/(\dot{Q} \times C_a)$; $E = (C_a - C_v)/C_a$

- *The isotope-wash-out-method*. A homogenous muscle tissue of the weight, W_{tis}, is presumed. A lipid-soluble indicator such as ^{133}Xenon dissolved in saline, is injected in the tibialis anterior muscle (* in Fig. 12-4). At steady state, the tracer concentration in the venous blood (C_v) is assumed to be the average blood concentration, and C_{tis} the mean tissue concentration. A distribution coefficient is introduced: $C_{tis}/C_v = 1$, which is known for Xenon. The decrease of the mass of indicator in the tissue per time unit (dt) must be equal to the mass supplied (which is zero) minus the mass of indicator leaving the tissue in the venous blood. The principle of mass balance provides the following equation:

$(W_{tis} \times dC_{tis})$ = [mass supplied minus mass eliminated].

$(W_{tis} \times dC_{tis}) = (- C_v \times Flow \times dt)$ or $W_{tis} \times dC_{tis}/C_{tis} = (- C_v/C_{tis} \times Flow \times dt)$.

Eq. 12-4: $dC_{tis}/C_{tis} = - Flow \times dt/(W_{tis} \times l)$.

The fractional fall in the mean tissue concentration of Xenon (C_{tis}) per time unit (dt) is constant during the whole elimination period (a rate constant = $\ln 2/T_{½}$). $Flow/W_{tis}$ is a perfusion coefficient in FU (ml of blood per min and per 100 g tissue). The ratio dC_{tis}/C_{tis} is measurable as $T_{½}$ on the skin surface above the Xenon deposit in fat or muscle tissue with a scintillation detector.

- The volume equation for a cylindrical system implies that the flow per second ($\dot{Q}_s$) and its volume (V) is related by t_{mean}:

Eq. 12-5: $t_{mean} = V/\dot{Q}_s$.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. The capillaries have the greatest cross-sectional area of the systemic circulation.

B. The systemic arterioles offer the greatest vascular resistance to bloodflow.

C. The sympathetic regulation of the arteriolar tone in fat and skin tissue depends upon the classical baroreceptors.

D. Increased compliance in the venous system means decreased venous return.

E. The lactate produced by skeletal muscles during exercise is not an important substrate for the myocardial metabolism.

II. Each of the following five statements have True/False options:

A. Vasopressin is a strong antidiuretic hormone and a modest vasoconstrictor.

B. ACTH and b-endorphins are released into the blood from the posterior pituitary gland in response to haemorrhage or other forms of stress.

C. Fulminant peritonitis is a frequent form of acute abdominal disease.

D. With the head of the patient downward, the patient risks air embolism if blood vessels are cut during neck and head surgery.

E. Hypotension and ischaemic hypoxia depresses myocardial contractility.

Case History A

20-year old soldiers, body weight 70 kg, is brought to the field hospital with a profusely bleeding gun wound in the left shoulder. The skin is cold and wet, the arterial pressure is 65/ 40 mmHg, and the heart frequency is 144 beats per min (bpm). There is no diuresis.

The bleeding is stopped by ligation of the bleeding arteries, and blood transfusions are given instantly (2 litres of whole blood with a normal packed cell volume, PCV, of 0.45). Hereby, the arterial pressure increases to a steady level of 105/70 mmHg, and the heart rate is reduced to 100 bpm. The condition of the soldier is clearly improved and his cardiac output is measured to 3.5 l per min.

24 hours later the soldier had a relapse, and his PCV was measured to 0.35. The PCV is corrected for trapped plasma, and assumed representative for the body as a whole (ie whole body haematocrit).

1. Describe the initial cardiovascular events leading to the shock condition.

2. *Describe the effect of the blood transfusions.*

3. *Calculate the order of size of the blood loss.*

4. *Was the blood transfusion therapy sufficient?*

Case History B

A male, 21 years old, is located in the supine body position. He has a cardiac output of 5.4 l per min at rest, a circulating blood volume in the pulmonary circulation of 650 ml and in the systemic circulation of 4750 ml. His total mass of skeletal muscle is 35 kg. The muscular perfusion is 3 ml of blood per min per 100 g of muscle tissue (3 FU), and the mean passage time in the muscular capillaries is 5 s.

1. *Calculate the mean transit time for all red blood cells in the total circulatory system.*

2. *Calculate the mean transit time for the pulmonary circulation only.*

3. *Calculate the total perfusion of the skeletal muscle mass at rest (ml/min) and calculate the bloodflow per second.*

4. *Calculate the functioning capillary volume in the muscular capillaries.*

Case History C

A 64-year old male normally has a body weight of 74 kg, a total blood volume of 5 l and a blood [haemoglobin] of 10 mmol per l (mM). One day he suddenly vomits large quantities of fresh blood. For two days his stools have been tarry. The last weeks have been stressful at work. The patient calls his doctor and the emergency ward at the hospital is alerted. Due to an incompetent local ambulance service, the patient is brought to hospital without delay by taxi. Here, the mean arterial blood pressure is below 10 kPa (75 mmHg) and falling. The heart rate is above 150 beats per min and rising. The blood [haemoglobin] is 5 mM measured one hour after the first massive blood loss.

The emergency team immediately institute transfusion of blood. The following 8 days the patient receives three transfusions of blood and at least 10 l of physiological saline. On the second day his [haemoglobin] has increased to 7.2 mM, but on the third day it falls again to 5 mM. On the 4th day at the hospital the patient develops high fever (maximum 40.6 ^{o}C), and a broad-spectrum antibiotic program is started without delay. On the 8th day at hospital the patient has normal temperature, but he develops watery swellings of legs and lower abdomen, in spite of pronounced urination. The body weight is now 80 kg.

1. *What is the most likely cause of the haemorrhage?*

***2**. Estimate the size of his blood loss.*

***3**. Why did the patient develop high fever?*

***4**. Why did the patient accumulate water?*

Case History D

A small girl, borne with ventricular septal defect, pulmonary stenosis, overriding aorta and a right to left shunt through the septal defect is examined with cardiac catheterisation. The oxygen concentrations in her blood are 138, 195, and 220 ml per l in the pulmonary artery, brachial artery and pulmonary veins, respectively. Her oxygen uptake at rest is 164 ml per min, and she has polycythaemia with a blood haemoglobin of 164.2 g per l. The girl has a 30% right-to-left shunt through the septal defect and directly to the overriding aorta.

1.What is the diagnosis?

2. Calculate the oxygen concentration in the pulmonary veins and estimate the saturation degree.

3. Calculate the bloodflow through the lungs.

4. Calculate the total bloodflow through the aorta assuming that the lung bloodflow ($Q°_{lung}$) and the shunt bloodflow ($Q°_{shunt}$) equals the total cardiac output, $Q°_{total}$.

5. Estimate the size of the venous return.

6. Is it likely that this patient develops cyanosis? Calculate the concentration of reduced haemoglobin in mean capillary blood.

Case History E

A male, 18 years old, suspect of congenital heart disease, is examined at the hospital. His body weight is only 60 kg and he has always abstained from exercise. The patient is tall and slim, although he is always hungry and is actually eating more than normal. Cardiac catheterisation is performed with the patient resting in the supine position, and reveals the following: Mixed venous blood from the right ventricle and from the pulmonary artery (C_{vO2}): 160, arterial blood from the aorta (C_{aO2}) 195, and blood from the right atrium 130 ml STPD l^{-1}. The oxygen uptake ($V°_{O2}$) at rest is 310 ml STPD per min.

1. *Calculate the cardiac output (Q°) from the right ventricle.*
2. *Calculate the Q° from the left ventricle.*
3. *Provide arguments for a certain cardiac abnormality, which explains the findings.*
4. *Why has the patient always avoided exercise?*
5. *Why is the patient slim although he is eating a lot and not performing exercise?*

Try to solve the problems before looking up the answers.

Highlights

- *The coronary bloodflow is mainly controlled by local metabolic autoregulation. Accordingly, a moderate*

decrease in arterial blood pressure down to 9.3 kPa (70 mmHg) does not significantly reduce the bloodflow through the myocardium.

- *The blood reaches the brain through the internal carotid and the vertebral arteries. The dominant control of cerebral bloodflow (CBF) is autoregulation. The sympathetic nervous system plays a secondary role.*
- *Some degree of autoregulation is found in many other organs including the skeletal muscle mass, the splancnic area, and the kidneys.*
- *CBF is normally 55 FU in humans at rest. One Flow Unit (FU) is one ml of blood per min per 100 g of brain tissue. This resting CBF and the oxygen uptake of the brain can double during cerebral activity and triple in active brain regions during an epileptic attack. Brain vessels are metabolically regulated. Increased P_{aCO_2}, and reduced P_{aO_2} dilatate brain vessels and increase CBF.*
- *Blood flows through the skin and subcutaneous tissues in order to nourish the cells, and to regulate shell temperature. Blood flows much faster through the arteriovenous anastomoses in the skin of the face, the fingers and toes in a cold environment. The sympathetic activity constricts the metarterioles that lead to the skin, so the blood bypasses the cutaneous circulation.*
- *Psychological influence can cause one to blush or to have a white face, by changing α-adrenergic constrictor tone and via the effect of local, vasoactive substances normally found in the skin.*
- *When a large fat combustion is occurring (eg, hunger and distance running), the fat bloodflow can increase from 3 FU to 20 FU. Cold and warm environments alter the fat perfusion just like the skin perfusion.*
- *The placental barrier has an area of 10 m^2, and can be passed by low molecular substances (nutrients, gasses, and waste) by diffusion. The foetal haemoglobin (F) has a dissociation curve that is shifted to the left relative to adult haemoglobin, and F is not affected by 2,3-DPG..*
- *Foetal blood has a high haemoglobin concentration, so the foetus takes up large amounts of O_2 in placenta.*
- *Much of the blood flowing in the descending aorta of the foetus is directed toward the placenta, so venous drainage from all organs is shunted toward the placenta, where wastes are eliminated from foetal blood, whereas O_2 and nutrients are acquired.*
- *The parallel foetal circulation is transformed into a series circulation in the baby. The foramen ovale and ductus venosus close within 3 days of birth, and the ductus arteriosus closes within 10 days. The sharp increase of O_2 content (C_{aO_2}) in the baby's blood is a potent and universal vasodilatator. The dramatic changes in gas exchange affect cardiopulmonary and vascular regulation, probably via local mediators such as arachidonic acid and prostacyclin.*
- *When a supine person arise, his TPVR increases, the systolic blood pressure falls and the diastolic blood pressure rises. Thus the pulse pressure amplitude falls, the MAP is unchanged. The stroke volume is reduced more than the heart rate rises, so the cardiac output will decrease in the standing position.*
- *The bloodflow (cardiac output in ml of blood per s or more convenient per min) can be measured by dividing the dose of indicator injected upstream by the area under the downstream concentration curve.*
- *Clearance is the volume of blood plasma. Which is totally cleared each minute of a given indicator by a*

specific organ.

- *Catecholamines constrict vessels all over the body by stimulating α_1-receptors located on the surface of vascular smooth muscles.*
- *ADH (vasopressin) is secreted from the posterior pituitary gland in response to shock, because the sinoaortic baroreceptors are under-stimulated. Vasopressin is a modest vasoconstrictor and a strong antidiuretic hormone.*
- *The enzyme renin is secreted from the juxtaglomerular apparatus, when blood pressure and renal perfusion falls drastically. Renin acts on the plasma protein, angiotensinogen, to form inactive angiotensin I, which is converted to the powerful vasoconstrictor, angiotensin II by ACE in the lungs.*
- *Angiotensin II is a powerful stimulator of the aldosterone secretion from the renal cortex. Aldosterone promotes the reabsorption of Na^+ and increases the secretion of K^+ and H^+ in the distal tubular system of the kidneys. Water follows by osmosis, so the extracellular volume is increased.*
- *ACTH and β-endorphins are released into the blood from the anterior pituitary gland in response to haemorrhage or other forms of stress.*
- *Septic shock or blood poisoning is a widespread bloodborn bacterial infection - often life-threatening. Examples are gas gangrene bacilli spreading from a gangrenous limb, colon bacilli with endotoxin spreading into the blood from infected kidneys, and fulminant peritonitis due to acute abdominal disease.*
- *Haemorrhagic arrest (haemostasis) is often life saving in shock, when applied without unnecessary delay.*
- *Replacement transfusions. The best possible therapy of haemorrhagic shock is whole blood transfusion, of shock caused by plasma loss plasma transfusion, and of dehydration shock transfusion with the appropriate solution of electrolytes.*
- *Oxygen breathing is helpful in shock with insufficient delivery of oxygen.*
- *Sympathomimetic drugs (noradrenaline, adrenaline etc) are often beneficial in neurogenic and anaphylactic shock. They are seldom useful in haemorrhagic shock.*

Further Reading

Calver, A., J. Collier and P. Vallance. "Nitric oxide and cardiovascular control." *Experimental Physiology* 78: 303-326, 1993.

Return to chapter 12

Return to content

Chapter 10

Cardiac Performance And Disorders

Study Objectives

- To *define* concepts such as cardiac insufficiency, central venous pressure, compliance, contractility, venous pump, venous return and ventricular stroke work.
- To *describe* four capacities of the cardiac pump called bathmotropic, chronotropic, inotropic, and dromotropic states. To describe the pressures of the left and the right half of the heart. To *describe* atherosclerosis, arteriosclerosis and risk factors involved.
- To *draw* pressure-volume curves and contractility lines for the heart at different conditions, and pressure variations in the left side of the heart and the aorta.
- To *calculate* the external work on the blood by the right and left ventricle, and the kinetic energy.
- To *explain* the autonomic innervation of the heart, venous return, cardiac contraction, Starling´s law of the heart, cardiac performance and oxygen demand, and cardiac filling pressure. To explain ischaemic heart disease, rheumatic heart disease, Buerger´s disease, and Raynaud´s disease.
- To *use* the above variables and concepts in problems and case histories.

Principles

- *Starling´s law of the heart: With an increased ventricular filling during diastole (venous return), the ventricular fibre length increases, so the ventricular contraction and stroke volume increases. - This is an intrinsic adaptation of the pumping capacity to the venous return. - Starling's law of the heart is also called the Frank-Starling relationship, because it was described independently by Otto Frank and Starling.*
- *The Fick principle (cardiac output) states that the volume of oxygen taken up by the blood in the lungs, divided by the arteriovenous oxygen content difference, is equal to the cardiac output. This example utilises the law of conservation of matter.*
- *The law of Laplace - Eq. 8-6*
- *Poiseuille´s law - Eq. 8-3.*

Definitions

- **Angina pectoris** (chest pains) is pain felt beneath the sternum or in the precordial area. The hypoxic pains are typically provoked by exercise or cold and submitted by subendocardially situated nerve fibres. The pains are relieved rapidly by nitro-glycerine and rest.
- **Arteriosclerosis** refers to atherosclerosis (and further changes) of the peripheral arteries.
- **Atherosclerosis** is a process of progressive lipid accumulation (*atheromatosis*) and calcification of the inner arterial walls in the abdominal aorta, lower extremities and the arteries of the heart, brain and kidneys.
- **Bathmotropic state** refers to the irritability of the myocardium.
- **Cardiac insufficiency** is a disorder, where the heart cannot pump enough blood to satisfy the nutritive needs of the body.
- **Central venous pressure** (CVP) is the pressure measured in the caval veins at the level of the heart or in the right atrium.
- **Chronotropic state** refers to the cardiac frequency.
- **Compliance** of the resting cardiac chambers refers to d*V*/d*P* (chamber compliance) - the reverse of the elastance (d*P*/d*V***)** of relaxed tissue.
- **Cardiac contractility** is the d*P*/d*V* of the contracting ventricle. The contractility is depicted on the pressure-

volume loop of the cardiac ventricle. Contractility refers to the change in slope of the pressure and volume increase from isovolumetric rest to contraction. Contractility is a function of contraction by crossbridge cycling.

- **Ejection fraction** refers to the stroke volume of blood as a fraction of the end-diastolic ventricular volume. This is a useful index of contractility.
- **Inotropic state** is another term for the myocardial contractility.
- **Intermittent claudication** refers to chronic ischaemia of the legs with hypoxic pains while walking.
- **Dromotropic state** refers to the conduction velocity of the myocardial conduction system.
- **Maximum oxygen uptake** is the oxygen uptake during maximum exercise. This is a measure of *endurance capacity*, and when expressed per kg of body weight it is also called the *fitness number* (ml min^{-1} kg^{-1}).
- **Mean circulatory equilibration pressure** (MCEP) is the filling pressure everywhere in the circulatory system following cardiac arrest.
- **Venous pump** refers to all local external forces acting on valvular veins and facilitating venous return.
- **Venous return** is the bloodflow reaching the right atrium (in steady state a similar bloodflow reaches the left atrium).
- **Ventricular stroke work** is the work applied to the blood at each ejection from the ventricle.

Essentials

This paragraph deals with *1. Cardiac electro-mechanics, 2. Pressure-volume work, 3. Venous return, 4. The venous pump, 5. The lipoprotein metabolism, and 6. Risk factors*.

1. Cardiac electro-mechanics

The heart is a four-chambered double pump. Every 24 hours the heart is ejecting more than 10^4 l of blood, and contracting more than 10^5 times. The total amount of work performed over the lifetime of a person is enormous. The order of size is calculated at the end of this chapter.

The cardiac cycle describes volume, pressure, and electric phenomena in the left ventricle as a function of time, and one heartbeat is shown below (Fig. 10-1). The red clock-shaped curve is the *intraventricular pressure.* The left ventricle is closed to the aorta in diastole and blood flows from the atrium to the left ventricle (Fig. 10-1). The ECG is explained in Chapter 11. Contraction or *ventricular systole* results in closure of the mitral valve (Fig. 10-1).

Fig. 10-1: Electro-mechanical events in the cardiac cycle. The aortic pressure is shown with a green curve. The blue atrial pressure curve has a, c, and v waves. For the ECG see Chapter 11.

The systole is *isovolumetric* (ie, the volume of blood in the ventricle is unchanged) until the intraventricular pressure exceeds the aortic pressure. Then the aortic valve opens and ventricular ejection occurs (see thick upward arrow in Fig. 10-1). *Bulging* of the cuspidal mitral valve into the left atrium during isovolumetric contraction causes a rise in left atrial pressure (see the *c-wave* in the thin *a-c-v-curve* of Fig. 10-1). The intraventricular pressure reaches a plateau around 15-16 kPa and then begins to decrease. The aortic valve closes when the intraventricular pressure falls below aortic pressure (see small arrow indicating *retrograde* flow in aorta in Fig. 10-1). This is the end of the ejection phase or the *left ventricular ejection time* (LVET).

The *electrocardiogram* (ECG) does not reflect the mechanical performance of the heart, but *closure* of the aortic valve corresponds in time with the *end of the T-wave* in ECG (Chapter 11).

The *a-wave* occurs during the contraction of the right *atrium* by which it squeezes out extra blood just before ventricular systole (early atrial contraction). As the atrium relaxes, the CVP is reduced (blue atrial curve in Fig. 10-1). The next atrial wave is the *c-wave*, and this wave is produced by closure of the *cuspidal* valves and by the right ventricular contraction, because the increased ventricular pressure is transmitted backwards to the right atrium and large veins.

Filling of the atrial chambers with blood is aided by the large *atrial compliance*. This is why the atrial pressure only rises modestly. The third atrial wave is the *v-wave* for *venous return*. Throughout ventricular systole and isovolumetric relaxation, *venous blood* returns to the heart, but the tricuspid valve is closed, so that the central veins and right atrium

are distended. The coinciding pressure build-up is relieved, when the tricuspid valve opens at the start of diastole, and the pressure is reduced (Fig. 10-1). The increase in *right atrial pressure* is transmitted backwards to the large veins near the heart. Prominent waves can often be seen in the neck veins when supine.

The periods just described for the left heart can be shown to be the same in the right heart, except for the fact that the systolic pressures are considerably lower in the right ventricle and pulmonary artery. The stroke volumes of the two ventricles are the same. Contraction of the right ventricle begins just after that of the left side and lasts for a shorter time (the peak pressures obtained are much less).

2. Pressure- volume work

At the end of the ventricular diastole the left atrial pressure increases, because the atrial systole begins just before the ventricular systole.

The intraventricular *pressure-volume loop* is a time-independent representation of the cardiac cycle, where the instantaneous intraventricular pressure and volume is plotted (Fig. 10-2). In diastole from B to C, the ventricle receives blood from the left atrium. The small increase in ventricular pressure reflects passive expansion and elastance of the myocardial wall. Pressure and volume increase with a slope that is related to *contractility*. The green line represents minimal contractility. The line starts from the ventricular *dead volume***,** that is a virtual minimal volume of blood, which can never be ejected (50 ml in Fig. 10-2). A steep rise in pressure occurs from C to D, with no change in ventricular volume (the isovolumetric contraction). At D the aortic orifice opens, because the end-diastolic pressure in the aorta is passed. During the rapid ejection phase, the fall in ventricular blood volume, is accompanied by a continuous increase in pressure. During ejection the volume falls by a size equal to stroke volume, pressure rises and falls until the *residual* ventricular volume is attained (about 80 ml in Fig. 10-2). The last ejection phase is slow, because the pressure decreases towards A, where the aortic orifice closes. The last event from A to B is the *isovolumetric relaxation* with a sharp drop in pressure at constant volume (Fig. 10-2). The steep slope of the curve refers to optimal contractility at a given condition (Fig. 10-2). Actually, the precise contractility concept is the *change* in slope from isovolumetric rest to the highest slope of the curve.

Fig. 10-2: The left ventricular pressure-volume loop from a healthy person at rest. – The pressure-volume loop is a widely applicable pathophysiological tool.

The *area* of the loop represents the pressure-volume work on the stroke blood volume performed by the ventricular contractile elements during ejection.

The pressure in C is the *end-diastolic* intraventricular pressure or the so-called *preload*. The force against which the ventricle contract is termed *afterload*. A good index of the afterload is the *peak aortic* (or intraventricular) pressure during systole, equal to the highest pressure shown in Fig. 10-2. The afterload is almost equal to the *peak* systolic pressure in the arterial tree. When the afterload is increased at constant end-diastolic pressure and volume, a greater ventricular pressure develops in order to expel blood (the dashed curve from D in Fig. 10-2). The result is a smaller stroke volume (and hence a greater residual ventricular volume), because of the high aortic and intraventricular pressure (Fig. 10-2).

The Frank-Starling relationship states that increasing left ventricular end-diastolic volume increases the stroke volume of the next heart beat. During isovolumetric contraction, the *end-diastolic* intraventricular pressure (EDIP) or *preload* increases. Increasing end-diastolic volume increases the ventricular filling pressure and thus stroke volume and stroke work. An increase in afterload occurs when the aortic pressure increases. Such an increase causes a decrease in stroke volume. Hereby, the EDIP and the end-diastolic volume increase, so the cardiac *stroke work* increases concomitantly (Fig. 10-2). The *ventricular* stroke work is the sum of the *pressure-volume work* and the *kinetic work* and calculated according to Eq. 10-1. *Preload* is the end-diastolic filling pressure of the ventricle just before contraction (C and E in Fig. 10-3).

Fig. 10-3: Left ventricular pressure-volume loops at rest (red area) and following an increase in end-diastolic volume by increased diastolic filling (blue loop with E-F).

When the left ventricle expands from C to E, by receiving more blood than before, the end-diastolic volume is increased. The greater diastolic filling results in a larger stroke volume according to *Starling's law* of the heart. The *Frank-Starling relationship* can be formulated as follows: Any increase of preload (ventricular filling) invokes a progressive increase in the blood volume ejected by the ventricles beat-by-beat until the cardiac output equals the input.The ventricular pressure increases with the rise in systolic aortic pressure (increased *afterload* in Fig. 10-3). The

stroke work of the heart on the blood is the area A, B, E, F, A. Accordingly, the stroke work is increased. A sympathetically mediated increase in contractility without a change in end-diastolic pressure (*preload*) results in an increased intraventricular pressure (Fig. 10-4). The slope of the line through G illustrates the *increased contractility*. The larger stroke volume, smaller residual volume, and larger stroke work on the blood (area C, D, G, H) is also shown in Fig. 10-4.

Fig. 10-4: Left ventricular pressure-volume loops from a healthy person at rest and following an increase in contractility during exercise (G-H).

The transmural pressure rises during ventricular contraction even at constant fibre tension. This is because the short ventricular radius is reduced to the same extent (see the law of Laplace, Eq. 8-5).

Echocardiography shows that the short ventricular radius is reduced by 5-6 mm during systole in healthy people at rest. The Laplace law is acceptable in such a situation.

If a cardiac chamber of a heart patient increase its diameter to double, and a spherical chamber is assumed, it implies a *two-fold* greater fibre tension (Fig. 8-9B). To the patient, this means an enormous energy requirement in the myocardium to maintain the necessary pressure. This disorder is called *cardiac failure* or insufficiency (see later).

3. Venous return

Veins are highly distensible or *compliant* vessels (ie, they have a large volume/pressure ratio, d*V*/d*P*) that have *one-way* valves. The venous system (veins, venules, and venous sinuses) controls the amount of blood that is translocated from the venous to the arterial side of the circulation.

In this paragraph the circulatory system is simplified to a venous system connected by a heart pump to an arterial system (Fig. 10-5A). The *central venous pressure* (CVP) is the pressure measured in the caval veins at the level of the heart or in the right atrium.

When the heart pump stops, the pressure is the same in all compartments of the circulatory system, (ie, the mean circulatory equilibration pressure, MCEP). MCEP depends upon the blood volume and the compliance of the vessels (Fig. 10-5A). As the heart pump starts and moves blood from the venous system, the pressure here will fall and the arterial pressure will rise. A further rise in pumping activity (cardiac output) reduces the central venous pressure (CVP), and finally the CVP is negative, so the central vessels collapse. This impedes the bloodflow into the atrium (the venous return), so that the cardiac output cannot rise any longer (Fig. 10-5A).

Fig. 10-5: The central venous pressure (CVP) as a function of cardiac output (A). – The cardiac output as a function of CVP (B). – Combined venous return and cardiac output curves as a function of CVP (C).

The cardiac output must be a function of CVP at a given steady state. Increasing the venous return will increase CVP from –1 kPa towards zero, and increase cardiac output as well (Fig. 10-5B). As CVP becomes increasingly positive, it exerts a backpressure on the venous system to impede venous return. The rise in cardiac output levels off, and there is no further rise at values around MCEP.

The third step is to combine the curves in Fig. 10-5 A and B. At the normal CVP (or right atrial pressure) the venous return curve crosses the cardiac output curve, and both flows are 5 l per min. If the atrial pressure is suddenly increased to the *mean circulatory equilibration pressure*, then all flow of blood is stopped (Fig. 10-5C). The low pressure during arrested circulation is mainly due to the very distensible venous system. The right atrial pressure has only increased slightly, but enough to decrease the venous return to zero and thus the cardiac output is zero (Fig. 10-5C). The more the atrial pressure falls below the venous pressure, the more the venous return will rise up to a certain level at an atrial pressure of -0.2 kPa (almost zero). Negative atrial pressures have the same venous return. This is because negative transmural pressures in the central, thoracic veins imply collapse. The venous return is therefore constant, regardless of a further fall in right atrial pressure (Fig. 10-5C).

The cardiac output must equal venous return in the steady state. Thus the cardiac output- curve for the left ventricle must cross the venous return curve in one point (Fig. 10-5).

The steep part of the cardiac output curve shows that the cardiac output can double following a small rise in pressure.

The *driving pressure* for the systemic circulation is the mean aortic pressure (MAP) minus CVP. The relationship to cardiac output and total peripheral vascular resistance (*TPVR*) is given by Eq. 10-2. A small fraction of *TPVR* is found in the venous system.

Eq 10-3 expresses the venous return. Small variations in CVP alter the volume of blood considerably in the venous

system. A normal value for venule pressure is 1.3 kPa (10 mmHg) and for CVP about zero. Since venous return must equal cardiac output in steady state, the venous resistance is only about 10 % of *TPVR*.

4. The venous pump

The venous pump is defined as all local external forces that facilitate venous return to the heart. Two important pump mechanisms are involved:

1. *The skeletal muscle- pump*. The deep veins of the arms and legs are affected by pressure exerted by exercising skeletal muscles. The veins are compressed by muscle contraction, and the one-way valves prevent the blood from flowing backward, and thus secure the transfer of blood toward the heart. Even the superficial veins are compressed during contraction.

   As soon as a venous segment is emptied of blood, its *transmural* pressure is so low that the *filling* pressure from more peripheral veins can fill the empty segment with blood.

   The skeletal muscular venous pump is also called the *peripheral venous heart.* In the erect position the peripheral venous heart must overcome the force of gravity and prevent overpressure in the dependent limb. During muscle rest there is an added hydrostatic pressure load of 13.3 kPa (100 mmHg) in the dependent limb. With a MAP of also 13.3 kPa the total pressure in a foot artery is 26.6 kPa, and in the dependent vein just above 13.3 kPa. During muscle contraction the venous pressure rises driving blood into the heart and just after muscle contraction the venous pressure falls again.

2. *The thoraco-abdominal pump.* The large veins are also affected by the positive intra-abdominal pressure and by the negative pressure in the thoracic cavity. The inferior caval vein returns blood from lower regions to the heart. During *inspiration* the intrathoracic pressure becomes more negative, and blood is *sucked* into the caval veins facilitating venous return to the heart. The inspiratory contraction of the diaphragm increases abdominal pressure favouring venous return. During *expiration* the intra-abdominal pressure decreases and intrathoracic pressure increases but remains negative, so that the venous return is maintained. *Intrinsic* cardiac mechanisms, including the *length-tension relation* (Chapter 2), allow the heart to increase stroke volume *beat-by-beat*, when the venous return increases. – Straining against a closed glottis is called Valsalva´s manoeuvre, and it is part of our every day life during coughing, defaecation, urination and lifting of heavy weights. The intrathoracic pressure increases abruptly, whereby the venous return is inhibited and the cardiac output is reduced (Starling´s law). Normally, fainting is prevented by a strong arteriolar and venous constriction released by the baroreceptor reflex.

5. The lipoprotein metabolism

This paragraph is inserted here in order to help the reader understand the pathophysiology of the most common cardiovascular disorders.

Lipoprotein particles are build up by a non-polar core containing triglycerides (TG) and cholesterol esters (E in Fig. 10-6). The polar shell of each particle consists of phospholipids, apoproteins and non-esterified cholesterol. These substances provide the particle with a negative electrical charge, which allow it to remain soluble in plasma (Fig. 10-6 right).

Hepatic synthesis of cholesterol varies inversely with the dietary intake.

Fig. 10-6: Lipoprotein metabolism (left). - The lipoprotein particles are found both in the plasma of healthy persons and of patients with atherosclerosis (right).

Four different lipoprotein particles (VLDL, IDL, LDL and HDL) control the transport of cholesterol to the cells.

a) *VLDL, IDL and LDLVery low-density lipoproteins* (VLDLs), which contain mainly the liver-produced TG, are synthesised in the liver and carries a characteristic surface protein called apoproteins B-100 (Fig. 10-6 right). VLDL is liberated from the liver in the postabsorptive phase

Chylomicrons are formed in the enterocyte from dietary fat after a meal and absorbed from the intestine into the blood (Fig. 22-13). Cylomicrons contain all the dietary lipids including lipophilic vitamins and have a half-life of 5 min.

VLDL from the liver, and chylomicrons absorbed from the intestine, are hydrolysed by the enzyme lipoprotein lipase (LPL) on the endothelial surfaces in the capillaries into glycerol, chylomicron remnants and free fatty acids (FFAs). From the FFAs the TG molecules are resynthesised and used or stored in adipocytes, heart and striated muscle cells (Fig. 10-6).

As more and more TG is removed from VLDL, the density of the particles becomes greater, and they are now termed *intermediate* density lipoproteins (IDLs). Normally, the liver cells take up half of the IDL particles, because they have receptors for the apoprotein B-100 on the IDL surface. The hepatic lipase removes TG from the IDLs to produce *low-density* lipoproteins (LDLs) still maintaining their apoprotein B-100. This apoprotein is recognised by the LDL receptors of all cells. LDL is the largest cholesterol fraction in blood plasma, and has a half-life of 24 hours. LDL delivers cholesterol and other lipids from the liver to the cells for metabolic and structural purposes (forward transport). An increasing concentration of cholesterol inside the cell automatically down-regulates LDL receptors and thus regulates the receptor-mediated endocytosis of additional LDL. The circulating LDL concentration is controlled by the number of hepatic LDL receptors and by enzyme activity in the cholesterol synthetic pathway.

Genetic LDL receptor *deficiency* elevates the ratio of LDL to HDL in blood plasma, and a ratio greater than 4 is a serious risk factor for cardiovascular disease.

b) HDL The remains of the chylomicrons co-operate with IDL and *high-density* lipoproteins (HDLs) to form cholesterol esters. Cholesterol esters are then exchanged for TG in VLDL and chylomicrons by the *cholesterol ester-transfer protein,* whereby HDL_3 changes to the less dense HDL_2.

HDL is the substrate for *lecithin-cholesterol acyltransferase* (LCAT). This enzyme catalyses the conversion of free cholesterol to cholesterol ester. LCAT is reduced in severe liver disease.

HDLs in plasma are disk formed particles, mainly produced in the liver. They contain an entirely different apoprotein called apoprotein Apo-I or Apo-II, and also Apo-E (Fig. 10-6 right). In the fasting state, the HDL concentration in the blood plasma is generally increased in females, by oestrogens, by exercise, and by moderate alcohol intake.

Similarly, fasting HDL concentrations are reduced (and LDL increased) in males, by androgens, by smoking, by obesity, and by an inactive sedentary life-style.

The cell membranes contain specific HDL receptors. HDL absorbs cholesterol in peripheral tissues and thereby matures from the nascent state (Fig. 10-6 right). LCAT is activated by the apoprotein A on the HDL surface. Mature HDL facilitates the transport of cholesterol back to the liver ((backward transport or reverse cholesterol uptake), where it binds to Apo-E.

Normally, HDL particles carry 30% of the total quantity of cholesterol in the blood. HDL protects against development of atherosclerosis, and a high HDL/LDL ratio reduces the risk of cardiovascular disease.

Population groups at risk are advised to eat a low fat diet with unsaturated lipids and a low cholesterol content, in order to prevent or delay the development of atherosclerosis.

Persons with extremely low total cholesterol in their plasma demonstrate a higher mortality than persons with values around 5 mM. The reason for the increased mortality (mainly death of cancer and gastrointestinal diseases), is probably insufficient immune defence and genetic factors.

6. Risk factors

A *risk factor* for a disease is a factor showing a statistical co-variance with the disease. A risk factor is not identical with a definite *disease factor*, where all causal steps are clearly understood. Nevertheless, risk factors may obtain an increasing degree of causal relationship to the disease. Two or more risk factors present frequently potentiates the risk.

Major risk factors that predispose to atherosclerosis and ischaemic heart disease (IHD is atherosclerosis if the coronary arteries) are consequences of the Western World lifestyle. These consequences are often notified as age changes. The fact remains that the western life style is characterised by unhealthy fast-food, a high dietary fat fraction, obesity, years with lack of exercise, low fitness, smoking, hypertension, hyperlipidaemia, diabetes, gout, oral contraceptives (synthetic steroids), drugs and doping (testosterone or other steroids in excess).

Consideration of risk factors must be supplied with other relevant information in order to provide a whole patient status, including genetic and immunological factors as mentioned above.

The inactive lifestyle of the Western World is documented by measurement of a low maximal oxygen uptake (endurance capacity) as an average for large population groups. Values below 34 ml per kg and per min are unhealthy for any age. Such a low endurance capacity is related to high mortality (Chapter 18), especially from IHD (males and postmenopausal females). Regular exercise seems to protect against IHD.

The following risk factors for IHD are dealt with below: Obesity, male sex, smoking, hyperlipidaemia, familial

hypercholesterolaemia, diabetes mellitus, gout and asymptomatic hyperuricaemia, oral contraceptives or synthetic steroids for doping.

Obesity is clearly associated with IHD, but probably not linked independently to IHD. The obese patient is characterised by an inactive lifestyle, low fitness, and preferring a fatty diet.

Sex. Males are more frequently affected by coronary artery disease (IHD) than fertile females. The smaller incidence in females is obviously related to the presence of natural female sex hormones (oestradiol in natural dosage). After the menopause, the female incidence approaches that of males. Female sex hormones in natural dosage may be protective, and male sex hormones *atherogenic*. Testosterone stimulates hepatic cholesterol synthesis.

Smoking a certain number of cigarettes per day is directly related to the incidence if IHD, and following 10 years of abstention, the risk declines towards the normal level.

Hypertension is associated with an increased risk of IHD (Chapter 9). Most forms of hypotensive therapy only reduce the risk of cardiac events to some extent, but clearly reduce the risk of *stroke*. Some hypotensive drugs reduce the arterial blood pressure but still have unwanted side effects. Light exercise reduces the blood pressure and implies other beneficial effects.

Hyperlipidaemia. Total cholesterol, HDL with calculation of LDL, and triglyceride

concentrations should be measured in all patients. High total serum cholesterol combined with a low HDL, and also high triglycerides are associated with an increased risk of IHD (Fig. 10-6).

Heterozygous familial hypercholesterolaemia is a relatively common genetic defect caused by mutations in the gene coding for the LDL receptor. With defective genes there is malproduction of LDL receptors in the liver. Some patients are without physical signs – others have cholesterol deposition around the eyes (xanthelasma) or in the tendons (xanthomas). These patients require diet with fibres and reduction of the cholesterol intake. Alcohol consumption must be reduced. The body weight must be kept close to ideal with exercise. Usually the patients require treatment with lipid-lowering drugs.

Homozygous familial hypercholesterolaemia is extremely rare. These patients are without LDL receptors in the liver, so they accumulate cholesterol and other lipids in the aorta, arteries, organs and skin. The HDL/LDL ratio in blood plasma is greatly reduced (below 1:4), and the fasting total cholesterol increases towards 30 mM. Drugs are needed, but the patients usually die young from ischaemic heart disease.

Diabetes mellitus. Increased blood glucose after fasting and an abnormal glucose tolerance test is associated with increased risk of atherosclerosis and a high LDL (Chapter 27).

Gout and asymptomatic *hyperuricaemia* (Chapter 20) is associated with an increased risk of ischaemic heart disease and atherosclerosis.

Intake of several types of oral contraceptives or synthetic steroids for doping (Chapter 18) increases the risk of atherosclerosis and thrombo-embolic phenomena.

Pathophysiology

This paragraph deals with heart disorders and atherosclerosis, which cause most people of the Western Hemisphere to die. *Coronary or Ischaemic Heart Disease* (CHD or IHD) is the most widespread. The remaining cases are caused by arteriosclerotic damage of the brain (strokes) and other organs (liver, kidney etc).

The two typical cardiovascular disorders are I. Atherosclerosis and II. Rheumatic heart disease. - Atherosclerosis is involved in several widespread disorders (ischaemic heart disease, peripheral arterial disease and hypertension). These diseases are related to the typical life style of the Western Hemisphere, whereas rheumatic heart disease is more frequent in poor countries with high frequency of infections and malnutrition.

I. Atherosclerosis

Atherosclerosis is a process of progressive lipid accumulation (*atheromatosis*) and calcification of the inner arterial walls in the abdominal aorta, lower extremities and the arteries of the heart, brain and kidneys.

Atherosclerotic plaques contain cholesterol, and the most important single factor for their development is a high plasma concentration of total cholesterol, in particular a high concentration of LDL.

Atheromas are yellow streaks or lesions found in arteries at autopsy. They are formed in the intima (lamina intima) by

lipid accumulation in macrophages and monocyte adhesion. As more and more cholesterol crystals are deposited, the atheromas grow and the surrounding fibrous and smooth muscle tissue is involved. Finally – as the subendothelial distortion leads to platelet aggregation - large arteriosclerotic plaques are formed. They consist of cholesterol and other lipids, dead cells, collagenous fibres, and there is excessive proliferation of the smooth muscle cells. The fibrosis or sclerosis makes the arterial wall stiff, which lead to systolic hypertension. Later Ca^{2+} salts precipitate and a factual calcification of the arterial wall may occur.

Typical for atherosclerosis patients are a high total cholesterol concentration in the blood plasma (total cholesterol above 6.2 mM), a dangerously high LDL and a low HDL fraction in fasting plasma (below 20% of the total). Often the atherosclerotic patient also has a high total triglyceride concentration (above 2 mM). In a fasted patient the triglyceride concentration depicts the precursor concentration of dangerous cholesterol: Very Low Density Lipoprotein (VLDL).

Large atherosclerotic plaques narrow the arterial lumen and produce arterial stenosis with reduced bloodflow. Insufficient oxygen delivery to the tissue is called ischaemic hypoxia, and hypoxic pains develop as in angina pectoris and intermittent claudication. Total occlusion of the arterial lumen is caused by a thrombus or an embolus in the lumen, or by wall bleeding. Disruption of the endothelium results in accumulation of thrombocytes and fibrin with thrombus formation and a complex atheromatous lesion is produced.

Arteriosclerosis (atherosclerosis) manifests itself in the coronary arteries as Ia. ischaemic heart disease and in peripheral arteries as Ib. peripheral arteriosclerosis.

Ia. Ischaemic heart disease (IHD)

Atherosclerotic coronary artery disease remains a leading cause of death, and is manifested as focal narrowing in the epicardial coronary arteries. The gradually narrowed vessel segment can be abruptly occluded by clot formation (thrombus) or by vasoconstriction at the atherosclerotic lesion. When a thrombus flows along the arterial tree with the blood and occludes the vessel, it is called an embolus.

Ischaemic heart disease is caused by reduced bloodflow to a region of the myocardium. Myocardial ischaemia diminish delivery of oxygen and nutrients, and accumulate potentially toxic substances such as lactic acid and K^+ around the cardiac cells, whereby necrosis may result. The causes are *atherosclerosis* with atheromas, thrombosis, emboli, or spasms in the coronary arteries.

Fig. 10-7: Bloodflow through the left coronary artery at rest and during exercise in a healthy person (upper) and in a patient with coronary obstruction (lower) and angina pectoris.

Coronary bloodflow is restricted in the systole by strong myocardial contractions and in diastole by the high heart rate of exercise. Normally, the coronary bloodflow in healthy persons is small during systole and increases during diastole (Fig. 10-7A). The high heart rate at exercise implies a short diastolic duration, but the rise in pressure secures a great *diastolic* bloodflow (Fig. 10-7B).

Four clinical manifestations of IHD are treated below:

1) Angina pectoris (chest pains) is pain felt beneath the sternum or in the precordial area - often referred to the left arm-shoulder-neck-jaw etc. Exercise and cold bring on hypoxia pains in the substernal or precordial area. Hypoxic pains are transmitted by subendocardially situated nerve fibres. The coronary resistance vessels contain a-adrenergic constrictor receptors and badrenergic dilatator receptors. The vasodilatatory capacity of the coronary resistance vessels can be maximally mobilised already at rest (Fig. 10-7C). Exercise shortens the diastolic duration and restricts the rise in diastolic bloodflow further (Fig. 10-7D). The aggravated myocardial ischaemia results in a lactate acidosis.

Following sublingual administration of nitro-glycerine, peak concentrations are achieved in the plasma within 1 min. Organic nitrates dilatate constricted coronary vessels, improve the bloodflow to the subendocardial (pain sensitive) part of the myocardium, and dilatate resistance & capacitance vessels. This dilatation reduces the venous return to the heart and the arterial pressure (reduced preload and afterload).

The beneficial effect of drugs such as glycerol- trinitrate in angina have been known for more than a century. Recently it was realised that the drugs act by releasing nitric oxide (NO) in the vascular wall (see Chapter 5).

Ca^{2+}-channel blockers block the Ca^{2+} flux into the smooth muscle cells of the coronary arteries, so they relax. The Ca^{2+} -channel blockers also reduce the force of contraction and thus the oxygen demand of the myocardium.

Coronary angioplasty is a method by which atheromatous obstructions are dilatated by an inflated balloon.

The arterial oxygen concentration is also reduced in anaemia, CO poisoning and in shock. Patients with hyperthyroidism or hypertension may have increased coronary oxygen demand and all these paitients may experience *chest pains* caused by myocardial hypoxia.

Another manifestation of ischaemic heart disease is myocardial infarction.

2) Acute myocardial infarction (AMI) is due to a sudden coronary thrombosis from an atheromatous plaque causing cellular death (*infarct*) of a myocardial area. Distal to the coronary occlusion the blood pressure is low. The thin-walled subendocardial vessels are squeezed most and receive the smallest bloodflow, often leading to *subendocardial* infarcts. The myocardial infarcts are sometimes *silent* (which means *without pains*; the pain relief is due to destruction of subendocardial nerve fibres). The typical infarct causes severe and long lasting pain.

Acute myocardial infarction renders the heart incapable of pumping the minimal blood volume required to transfer sufficient oxygen to the mitochondria.

The patient experiences a sudden chest pain and the pain is lasting for hours in contrast to *angina*. The patient may develop signs of shock. Necrotic myocardial cells liberate cellular enzymes such as *creatine kinase*, which peaks in the blood plasma within 24 hours. The total enzyme release depicts the size of the infarction. Lactic dehydrogenase (LDH) isoenzymes peaks a few days later, and *LDH 1* is rather specific for myocardial necrosis.

Read Chapter 11 before this paragraph:

Non-Q wave infarction: When only part of the wall is necrotic there are deeply inverted, symmetrical T-waves (*coronary T- waves*) and mostly ST depression in the ECG. These signs of ischaemia are often transient - only during the acute attack - and found in all precordial leads located above the infarcted area. Such a subendocardial infarct does not show deep Q waves, and epicardial involvement implies ST segment elevation.

Fig. 10-8: Myocardial infarction of the left ventricular wall with lack of movement of infarcted tissue during systole. The ECG changes are typical for the anterolateral location of the infarct (see Ch. 11).

Q-wave infarction: A wide and deep Q wave in the ECG is a lesion wave, and the sign of transmural myocardial infarction with necrosis through the whole of the myocardial wall. The deep Q wave is maintained for years after the event. A typical *anterior wall infarction* shows changes in lead I and in V_2-V_6 dependent upon the localisation (Fig. 10-8). During systole the infarcted area does not contract or move due to cell necrosis (paradoxical movement). There is always a danger of rupture of necrotic tissue.

A typical *posterior wall infarction* is diagnosed by a *mirror image* with changes in V_1-V_2 (*reciprocal* changes) and ST- depression in lead I. ST-depression or ST-elevation is indicative of myocardial ischaemia.

Chapter 11 is a prerequisite for the understanding of the above paragraph!

Fig. 10-9: Left ventricular pressure-volume loops in a healthy person (red) and for persons with acute (blue area) or chronic, congestive (green area) cardiac failure.

The AMI patient is extremely tired. Even the work of breathing is a heavy task. The condition is often a case of general ischaemic hypoxia (Fig. 10-9) and can develop into cardiogenic shock (see below).

Interaction of platelets with collagen in the vessel wall is the first step in platelet aggregation leading to thrombosis. The activated platelets release thromboxanes A2 (TxA2) from arachidonic acid in the phospholipids of the platelet membrane. Platelet aggregation is inhibited by cAMP and by acetyl-salicylic acid, which inhibits platelet cyclo-oxygenase.

Eicosapentaenoic acid (EPA) in the diet reduces the frequency of thrombosis by reduction of the TxA2 production.

3) Cardiac failure or *cardiac insufficiency* is a disorder, where the heart cannot pump enough blood to satisfy the nutritive needs of the body. Cardiac insufficiency is manifest by a consequential decrease in cardiac output (*lower output failure*) or by an increase in cardiac output (*higher output failure*). The cardiac failure can be acute or chronic.

Damming of blood in the vessels behind the insufficient heart pump is typical.

Acute cardiac failure is caused by AMI, acute intoxications, anaesthesia etc. Occlusion of the coronary artery to the left ventricle impairs contractility, and left ventricular failure develops due to the diminished cardiac output from the left ventricle. Initially, the right ventricular output is maintained, whereby the left atrial pressure (and pulmonary venous pressure) is increased beat-by-beat. As a consequence, the left ventricular output will increase until the cardiac

outputs of the two ventricles are equal. The increased pulmonary venous pressure leads to reduced lung compliance (dV/dP) and increased respiratory elastic resistance with increased respiratory work and distress. Eventually, plasma fluid flows into the alveoli and *pulmonary oedema* is developed (Fig. 10-10).

Chronic or congestive cardiac failure occurs in conditions such as IHD and following severe hypertension. In *chronic cardiac failure* blood is accumulated and expands the venous system and the left ventricle (Fig. 10-9).

Cardiac oedema develops during congestive cardiac failure, because the kidneys retain NaCl and water. The accumulated fluid increases venous return, which in turn elevates the right atrial pressure. The rising atrial pressure elevates the venous and the capillary pressure. This causes loss of fluid into the interstitial fluid volume. Accumulation of abnormal volumes of interstitial fluid is the definition of *oedema*.

The low cardiac output and blood pressure causes an increased sympathetic tone with constriction of the afferent renal arterioles to the glomeruli. As a consequence, the renal bloodflow (RBF), and the glomerular filtration rate (GFR) decrease (Fig. 10-10). Also the NaCl concentration decreases in the renal macula densa (Fig. 25-17). The *renin-angiotensin-aldosterone cascade* is activated, which enhances salt-and water-retention. Angiotensin II is a strong vasoconstrictor, which further decreases the renal bloodflow, and aldosterone promotes the reabsorption of NaCl and water from the distal renal system. A certain salt-water retention is beneficial in the early stages of cardiac failure, because of improved venous return and thus improved cardiac output according to Starlings law of the heart. However, prolonged activation of the renin-angiotensin-aldosterone cascade and the *sympathetic nervous system*, damage the heart muscle further and reduce its contractility. This is because circulating vasoconstrictors, such as catecholamines, vasopressin and angiotensin II, imply an extra workload on the damaged myocardium.

When the salt and water retention results in even a small rise in the osmolarity of plasma, there is a stimulus of osmoreceptors, located close to the neurosecretory cells in the hypothalamus. The osmoreceptors stimulate both production and secretion of vasopressin (antidiuretic hormone, ADH) in the neurosecretory cells. ADH eases the renal reabsorption of water in the outer cortical collecting ducts leading to a low urine flow (antidiuresis).

Vasopressin is also an universal vasoconstrictor.

Fig. 10-10: Formation of pulmonary oedema in left ventricular failure (mitral stenosis) and in congestive cardiac failure with ankle oedema.

Increased venous pressure with stasis of blood dilatates the central vessels and the heart chambers. The distended atrial wall liberates *atrial natriuretic peptide* (ANP), which increases Na^+-excretion and dilatates peripheral vessels (Chapter 24). This is a partial compensation of the increased *preload* (the water loss by high urine flow reduces venous return) and *afterload* (the vasodilatation reduces outflow resistance).

Venoconstriction shifts significant quantities of blood from the peripheral to the central circulation. Since *central venous pressure* (CVP) varies inversely with *TPVR*, it is possible to maintain cardiac output in resting patients with congestive heart failure (insufficient contractile force) at the expense of increased CVP, by reduction of *TPVR* (Eq. 10-2).

When cardiac output decreases more and more during development of congestive heart failure, the compensation fails, and both CVP and end-diastolic ventricular pressure (preload) and volume rises further (Fig. 10-9). The superficial meck veins are expanded, when CVP is abnormally elevated. Eventually, large volumes of plasma water flow from the liver into the peritoneal cavity due to the elevated CVP. Fluid accumulation in the abdomen is called *ascites*.

Patients with a cardiac output much *higher* than normal can develop cardiac failure. The venous return is much too high, and after some time with an overexpansion of the heart, the cardiac pump fails to eject the same blood volume as it receives, and an increasing blood volume is accumulated behind the insufficient ventricle. The rise in left atrial pressure leads to pulmonary oedema and eventually the right ventricle fails so peripheral oedema develops.

Examples of this condition are cardiovascular disorders with a drastic reduction of the *TPVR*. The low opening pressure in the left ventricle being equal to a low end-diastolic aortic pressure (Fig. 10-11) illustrates the low *TPVR*.

Fig. 10-11: Left ventricular pressure-volume loops from a healthy person, and from a person with high metabolic rate (hyperthyroidism) - or a person with arteriovenous shunts.

In *hyperthyroidism* (a disease with an abnormally high metabolic rate described in Chapter 28), all the vessels in the systemic circulation dilatate, and the venous return overloads the heart. On the other hand, short-term administration of L-thyroxin to patients with chronic heart failure improves cardiac and exercise performance.

Any major *arteriovenous shunt* leads a large fraction of the arterial blood directly into the veins. This greatly increases venous return and overloads the heart (Fig. 10-11).

4) Cardiogenic shock - terminal pump failure - is such a severe reduction of cardiac output that the peripheral tissues suffer seriously from lack of oxygen, the cells deteriorate and within hours or days the patient die. The pulmonary capillary wedge pressure is normal or elevated in contrast to other types of shock (blood loss or vasodilatation).

During insufficient pumping capacity and cardiac arrest, the cardiac pump do not get rid of the blood volume received and a large blood volume is therefore accumulated in the distensible venous system, pulmonary system and the thin-walled chambers of the heart. This is why the lower part of a newly diseased body is filled with blood in distensible vessels (*livores*), and the upper part of the body is pale.

I b. Peripheral arteriosclerosis

Peripheral arteriosclerosis refers to atherosclerosis and other changes of the large and medium large peripheral arteries. The muscular lamina media grows and becomes fibrotic often with atheromas in small arteries. The walls of the elastic arteries become thick of hyaline and the lumen narrows, which causes systolic hypertension. The narrowing also leads to ischaemia, which further promotes systolic hypertension.

The most frequent of these disorders is chronic ischaemia of the legs, also called *intermittent claudication* from its prominent symptom. Claudication is a cramp-like pain, which occurs during exercise and subsides at rest. Occlusive atheromatous lesions between the common iliac and the common femoral artery lead to claudication of the thigh and calf. Lower femoral artery disease usually causes claudication of the calf, and occlusive lesions of the popliteal artery causes claudication of the calf or the foot.

If possible, regular relaxed exercise should be undertaken in an attempt to develop anastomoses. In severe ischaemia there is pain at rest. Balloon dilatation is often useful, and amputation may become necessary.

II. Rheumatic Heart Disease

Rheumatic fever is caused by repeated pharyngeal infections with group A streptococcus. An autoimmune reaction is triggered by the streptococci and the patient develops fever, joint pains, diastolic mitral murmur caused by mitral valve inflammation, cardiac enlargement with pericardial effusion and pericarditis (raised ST-segment in ECG – see Chapter 11) or myocarditis (inverted T-waves). More than 50% of those who suffer from *acute rheumatic fever* with carditis develop rheumatic heart disease many years later.

The rheumatic valvular disease mainly affects the mitral and the aortic valves.

Mitral stenosis and regurgitation

Practically all cases of mitral stenosis are caused by rheumatic heart disease. Severe mitral valve stenosis is present, when the mitral valve orifice is reduced to 10^{-4} m^2 as compared to the normal area of ($5 * 10^{-4}$ m^2). The left atrium dilatates and its walls hypertrophy in order to maintain sufficient bloodflow to the left ventricle. Obviously, the pressures also increase in the entire pulmonary vascular bed: veins, capillaries, arteries and right ventricle.

The stenotic mitral valve and the resistance of the pulmonary arteries determine the pressure in the pulmonary capillary bed. Up to a certain point, pulmonary arteriolar vasoconstriction protects the patient from pulmonary oedema.

Patients with severe mitral stenosis have cyanotic cheeks and ears (mitral faces) caused by stasis of the blood. Auscultation at the apex, lying on the left side just following exercise reveals a split second heart sound with a *mitral snap* (second component) as the mitral valves open, then a mid-diastolic rumbling murmur - like a sack of potatoes falling on a floor - caused by turbulence through the narrowed orifice. The murmur ends in a loud first heart sound, because the cusps are kept open until the start of the ventricular systole (Fig. 10-12). As the left atrium grows it becomes activated later than the right, and the P-wave is bifid (P- mitral). The large left atrium favours the development of atrial fibrillation with thrombo-embolism and emboli to the brain, kidneys and gastrointestinal tract.

Mitral stenosis is often combined with regurgitation. Regurgitation is recognised by a systolic murmur without any 1.heart sound. The 1.heart sound is caused by the closure of the cusps, and in mitral stenosis they do not close. Usually the condition is asymptomatic for decades following rheumatic fever. Light pulmonary oedema presents itself as coughing and exertion dyspnoea.

Replacement of the mitral valve is performed with artificial valves, which may work for decades with adequate

anticoagulant therapy.

Fig. 10-12: Mitral stenosis, aortic stenosis and aortic regurgitation.

Aortic stenosis and regurgitation

Disorders of the normal tricuspid aortic valve are mainly caused by rheumatic fever or by atherosclerosis. Almost half of all cases of rheumatic heart disease include the aortic orifice, usually associated with the mitral orifice.

Symptoms and signs are characteristic: Exercise-induced syncope and angina. This occurs when the disease is severe and the area of the aortic orifice is reduced to 1/3 of normal. The left ventricular pressure rises and the left ventricle hypertrophies. The increased oxygen demand of the myocardium leads to ischaemia with angina pectoris, arrhythmias and left ventricular failure. Healthy persons can increase their cardiac output by a factor of 5 or more, but this is not possible for patients with severe aortic stenosis. The arterial blood pressure falls, the patient is pale (*aortic face*), chest pains worsen, and the patient may lose consciousness. There is a strong systolic murmur in the aortic area and over the carotid arteries. The echocardiogram demonstrates thick aortic valve cusps and left ventricular hypertrophy. A ventricular-aortic pressure gradient above 6.7 kPa (50 mmHg) measured by cardiac catheterisation, is indicative of surgery. Without surgical intervention death frequently ensues within a few years from the occurrence of the first serious signs.

Tricuspid stenosis and regurgitation

This is an uncommon complication, which is related to rheumatic heart disease or is congenital. Regurgitation is more frequent than stenosis, but often the two conditions are combined.

Isolated tricuspid stenosis dilatates the right atrium and the caval veins, and the liver swell just like the condition with constriction of the heart. Atrial fibrillation is frequent from the dilatated chamber.

Tricuspid regurgitation is a condition where the right ventricle delivers blood to both the pulmonary artery and the right atrium at each systole. The right heart is dilatated, whereas the lung vessels are not. Pulsation of the neck veins and a large tender liver are typical signs. There is often a blowing systolic murmur over the sternum.

Replacement of the tricuspid valve is performed with artificial valves.

Left ventricular hypertrophy

Left ventricular hypertrophy is an abnormal increase of the left ventricular mass caused by increased demands of cardiac work. If due to pressure overload it is called *concentric* hypertrophy in which new contractile elements are lined up in parallel, and if due to volume overload it is called *eccentric* hypertrophy, in which new contractile elements elongate the myocardial cell. - This disease is described in Chapter 11.

Other cardiovascular disorders

Thrombo-angiitis obliterans (Buergers disease). *Buerger´s disease* occurs in small arteries of the limbs of young smokers – typically males. The vessel wall is inflamed, but many lesions look like atherosclerosis. The patient is invalid by intermittent claudication, and the only choice of treatment is to stop smoking.

Raynaud´s disease is a condition with cold precipitated attacks of spasms of the small arteries and arterioles, supplying the fingers and toes. The disease is usually bilateral and affects predominantly young girls and female smokers.

First the skin becomes pale and white from vasoconstriction, due to slow bloodflow, and finally red because of hyperaemia. The vasoconstriction occurs in the digital arteries, arterioles and skin capillaries. A few minutes later the capillary smooth muscle spasm is released due to local vasodilatators, and the capillaries are filled with oxygen poor blood (the skin becomes blue and is still cold). Finally, the arterial and arteriolar constriction is released and the classical physiological reactive hyperaemia occurs, with red, warm fingers and *paresthesia* (numbness). Centrally controlled vasoconstrictor tone, sensitive to cold signals, is probably implicated in cases with symmetrical spasms. A centrally increased vasoconstrictor tone involving the coronary bloodflow is consistent with the fact that some of the Raynaud patients also suffer from *chest pains* (angina pectoris) and *migraine*.

Raynaud´s disease occurs in a primary and a secondary form. Primary Raynaud´s disease is a condition where the cause is unknown (ie essential). There is a benign familiar occurrence of so-called *dead fingers* (ie, digiti mortui familiaris), and a malignant form with symmetrical gangrene (ie, symmetric gangrene).

Secondary Raynaud´s disease (*Raynaud´s phenomenon*) occurs together with connective tissue disorders (dematomyositis, polymyositis, systemic lupus erythematosus and systemic sclerosis). Raynaud´s phenomenon is also a

side effect to treatment with b-blockers, in which case they must be withdrawn. The patient has to wear warm clothes in order to protect both the shell and the core temperature. *Nifedipine* is sometimes beneficial.

Varicose veins have incomplete valves. Normally, the muscular venous pump maintains the venous bloodflow towards the heart. Patients with defective valves can develop venous pooling or stasis and ankle oedema. This is because the contracting leg muscles squeeze the blood in the retrograde as well as in the anterograde direction.

Equations

Ventricular stroke work rate is the sum of the pressure-volume work and the kinetic work:

$$\textit{Eq. 10-1: Stroke work rate} = [(P \times V) + \tfrac{1}{2} m * v^2].$$

Both the pressure-volume work and the kinetic work are work per stroke duration or time unit that is comparable to work-rate or effect in Watts.

The driving pressure for the systemic circulation is the mean aortic pressure (MAP) minus CVP. The relationship to cardiac output and total peripheral vascular resistance (TPVR) is given by:

$$\textit{Eq. 10-2: Cardiac output} = (MAP - CVP)/TPVR.$$

A small fraction of TPVR is found in the venous system. The venous return is expressed by the approximative equation:

$$\text{Eq. 10-3: Venous return} = (\text{venule pressure} - CVP)/\ \text{venous resistance}.$$

Self-Assessment

Multiple Choice Questions

I. The following five statements have True/False options:

A. Smoking has no effect on Buerger´s disease.

B. When the brain is deprived of new blood, gray out occurs 4.5 seconds following blockage of its bloodflow.

C. There is a mid-systolic rumbling murmur at the apex in severe mitral stenosis.

D. The central venous pressure varied inversely with the total peripheral vascular resistance.

E. Organic nitrates dilatate constricted coronary arteries, improve the bloodflow to the subendocardial part of the myocardium especially, and dilatate resistance and capacitance vessels.

II. The following five statements have True/False options:

A. Exercise-induced syncope and angina is characteristic of aortic stenosis and regurgitation.

B. Concentric ventricular hypertrophy is due to volume overload, whereas eccentric hypertrophy is due to pressure overload of the ventricles.

C. IDL particles are further hydrolysed by LPL, resulting in slightly higher density of the so-called low-density lipoprotein still maintaining its *apoprotein* B-100.

D. Echocardiography shows that the short ventricular radius is reduced by 15-20 mm during systole in healthy persons at rest.

E. The residual ventricular volume is approximately 150 ml of blood in healthy persons at rest.

Case History A

Three years following heart-lung transplantation, a patient is examined at the hospital for the cause of frequent exertion syncope. Heart catheterisation reveals that the systolic/diastolic pressure in the left ventricle is 26.7/0 kPa (200/0 mmHg), and in the aorta 10.7/6.7 kPa (80/50 mmHg).

1. *What is the cause of exertion syncope?*
2. *What is a likely diagnosis?*
3. *Argue for the size of the left ventricular cavity and wall thickness.*

Case History B

A 24-year old sporty male consults the doctor because of syncope while playing handball.

The examination reveals a systolic murmur to the right in the second intercostal space aortic site). The systolic murmur is audible also over the neck. The arterial blood pressure is 145/85 mmHg (19.3/11.3 kPa). There is a history of rheumatic fever at the age of 11.

ECG shows a deep S-wave in V_1 and a high R-wave in V_6 (the sum is 5 mV), and there are asymmetrically negative T-waves in V_4-V_6.

1. *What is the most likely diagnosis?*
2. *What is the ECG diagnosis*
3. *Describe the prognosis and the therapy.*

Case History C

A 55-year old female complains of headache and an arterial blood pressure of 190/100 mm Hg (25.3/13.3 kPa) is found. Her ECG shows deep negative S-waves in V_1 and V_2, and high R-waves in V_5 and V6. The sum of one S- and one R-wave is above 4 mV.

1. *Calculate the mean arterial pressure and compare the result to a normal value.*
2. *What is the diagnosis?*
3. *Is the patient suffering from any cardiac disease?*
4. *Why is the arterial pressure amplitude (eg, systolic minus diastolic pressure) much larger than normal?*

Case History D

A girl, 17 years of age, with frequent episodes of acute tonsillitis and articular pain during her childhood, developed a cardiac disease. The diagnosis was made as the doctor heard a diastolic murmur over the precordial area. The main complaint of the girl was dyspnoea at exertion. Cardiac catheterisation revealed a mean pressure in the pulmonary artery of 58 mmHg (7.7 kPa), and in the left atrium 28 mmHg (3.7 kPa), whereas the pressure in the left ventricle was only 2 mmHg (0.27 kPa) in early diastole. The pressure in the aorta was 105/80 mmHg (14/10.6 kPa). The cardiac output was measured at rest with the Fick principle to be 2.5 l min^{-1}. Her circulating blood volume is somewhat less than her total blood volume, and exactly 5000 ml. The blood volume of the systemic capillaries is 3% of 5000 ml.

1. *What is the name of the condition with frequent episodes of tonsillitis?*
2. *Calculate the pulmonary vascular resistance.*
3. *In what orifice is the cardiac disorder located?*
4. *Calculate the mean passage time in the systemic capillaries.*

Case History E

A female, 33 years old, complains of attacks of severe pain in the fingers of both hands, when she is outdoors in cold weather. She is a heavy smoker since the age of 16. The patient also suffers from another pain disorder. She has half-sided headache with visual disturbances at least twice a month in the cold season. The patient explains the pain attacks in the fingers as follows. First the skin of both hands and the fingers (not the thumb) becomes pale and white, and they feel like dead. A few minutes later the skin is blue, and the pain is severe. After 5-10 minutes the skin suddenly becomes red and it is very painful.

1. *What is the diagnosis of the finger disease?*
2. *Are the finger disease and the headache related?*
3. *Describe the pathophysiology of the finger disease.*

Try to solve the problems before looking up the answers.

Highlights

- *The ventricular stroke work is the sum of the pressure-volume work and the kinetic work.*

- *A good index of the afterload is the peak aortic pressure during systole.*
- *A ventricular-aortic pressure gradient above 6.7 kPa (50 mmHg) is indicative of surgery in aortic stenosis.*
- *The venous pump is defined as all local external forces that facilitate venous return to the heart.*
- *The skeletal muscular venous pump is also called the peripheral venous heart, because its force must be equal to or larger than that of the heart in order to return the blood in the upright position.*
- *During inspiration, the intrathoracic pressure becomes more negative, and blood is sucked into the large thoracic veins facilitating venous return to the heart.*
- *Atherosclerosis is a process of progressive lipid accumulation (atheromatosis) and calcification of the inner arterial walls in the abdominal aorta, lower extremities and the arteries of the heart, brain and kidneys.*
- *High-density lipoproteins (HDLs) in plasma are disk formed particles, mainly produced in the liver for cholesterol transport. HDLs protects against the development of atherosclerosis as they transport the cholesterol back to the liver, where the elimination begins.*
- *Acute myocardial infarction renders the heart incapable of pumping the minimal blood volume required to transfer sufficient oxygen to the mitochondria at rest.*
- *Cardiogenic shock - or terminal pump failure - is such a severe reduction of cardiac output that the peripheral tissues suffer seriously from lack of oxygen, the cells deteriorate and within hours or days the patient die. The pulmonary capillary wedge pressure is normal or elevated in contrast to other types of shock.*
- *Left ventricular hypertrophy is an abnormal increase of the left ventricular mass caused by increased demands of cardiac work. If due to pressure overload it is called concentric hypertrophy in which new contractile elements are lined up in parallel, and if due to volume overload it is called eccentric hypertrophy, in which new contractile elements elongate the myocardial cell.*
- *Cardiac oedema develops during chronic cardiac failure, because the kidneys retain fluid.*
- *Buerger´s disease occurs in small arteries of the lower limbs of young male smokers. The vessel wall is inflamed, but many lesions look like atherosclerosis. The patient is invalid by intermittent claudication, and the only choice of treatment is to stop smoking.*
- *Raynaud´s disease is a condition with cold precipitated attacks of spasms of the small arteries and arterioles, supplying the fingers and toes. The disease is usually bilateral and affects predominantly young girls and female smokers.*
- *Rheumatic fever is caused by repeated pharyngeal infections with group A streptococcus. The streptococci trigger an autoimmune reaction and the patient develops fever, joint pains, diastolic mitral murmur caused by mitral valvulitis, cardiac enlargement with pericardial effusion and pericarditis or myocarditis.*
- *More than 50% of those who suffer from acute rheumatic fever with carditis develop rheumatic heart disease many years later.*
- *The rheumatic valvular disease mainly affects the mitral and the aortic valves.*
- *Practically all cases of mitral stenosis are caused by rheumatic heart disease. Severe mitral valve stenosis is present, when the mitral valve orifice is reduced to 10^{-4} m^2 as compared to the normal area of ($5 * 10^{-4}$ m^2).*
- *Disorders of the normal tricuspid aortic valve are mainly caused by rheumatic fever or by atherosclerosis. Almost half of all cases of rheumatic heart disease include the aortic orifice, usually associated with the mitral orifice.*

Further Reading

von Spiegel T, G Wietash and A Hoeft (1998). Basics of myocardial pump function. *Thorac Cardiovasc Surgery* 46, Suppl 2, 237-41.

Return to content

Return to chapter 10

Chapter 11

Cardiac Action Potentials and Arrhythmias

Study Objectives

- To *define* asystolia, atrial fibrillation, atrial flutter, axis-deviation, bradycardia, heart block, tachycardia, sinus rhythm, ventricular fibrillation, ventricular tachycardia and vulnerable period.
- To *describe* cells with pacemaker-function, the normal impulse conduction in the heart, the recording of the electrocardiogram (ECG), and different types of heart block.
- To *draw* membrane potentials of sinus cells, atrial cells, and myocardial cells during a cardiac cycle, and draw a normal ECG.
- To *estimate* the electrical QRS-axis from the R-waves of the standard limb leads.
- To *explain* the effect of the autonomic nervous system on the membrane potential of the sinus node. To *explain* the genesis of the ECG in health and disease. To *explain* the Adam-Stokes syndrome.
- To *use* these concepts in problem solving and case histories.

Principles

- *The body conducts electrical signals uniformly in all directions (an almost uniform volume conductor increasing the potential electrical field around the heart generator).*
- *Einthoven's law states that any two of the three bipolar standard limb leads determine the third one with mathematical precision.*

Definitions

- **An electrocardiogram (ECG)** is a curve showing the potential variations against time in the whole body stemming from the heart, which is an electrochemical generator suspended in a conductive medium.
- **Asystolia** refers to cardiac arrest.
- **Atrial fibrillation** is a continuous atrial activation with 400 or more contractions per min. Contractions spread through the atrial tissue almost without mechanical effect and only few electrical signals are conducted to the ventricles.

Atrial flutter is an atrial contraction rate around 300 per min, often with every second contraction conducted to the ventricles. Sawtooth-like flutter waves characterise the ECG.

- **Bradycardia** is an unduly slow heart rate. ***Sinus bradycardia*** is a sinus rhythm at rest below 60 beats per min during the day or less than 50 at night.
- **Calcium concentration** in plasma (total): Normal range is 2-2.5 mM.
- **Heart block** is a blockage somewhere along the pathway for impulse conduction in the heart.
- **Mean QRS- axis** of the ventricles or the *mean cardiac vector* is the net force in the frontal plane during ventricular depolarisation and repolarisation. Many electrical potentials are propagated in different directions and most of these cancel each other out. The main direction of the mean cardiac vector is from the base of the ventricles towards the

apex.

- **Left-sided axis deviation** is characterised by a positive R_I and a negative R_{III} (Fig. 11-7). Cardiologists use the net area of the QRS-complex for precise diagnosis.
- **Pacemaker cells** are small pale cells located in the sinus node of the heart. The sinus node is the primary determinator of the cardiac rhythm, because its cells have the highest spontaneous frequency.
- **Potassium concentration** in plasma: Normal range is 3.5-5 mM.
- **Right-sided axis deviation** is characterised by a negative R_I and a positive R_{III} (Fig. 11-7). Cardiologists use the net area of the QRS-complex for precise diagnosis.
- **Sinus rhythm** refers to the normal cardiac pacemaker rhythm from the sinus node. The spontaneous discharge at rest is usually 100 beats per min, but the parasympathetic inhibitory tone predominates in healthy individuals resulting in a resting heart rate around 75 beats per min.
- **Tachycardia** refers to a cardiac rate above 100 beats per min. *Sinus tachycardia* is a sinus rhythm above 100, which can be caused by anaemia, cardiac failure, catecholamines, emotion, exercise, fever, pregnancy, pulmonary embolism or thyrotoxicosis.
- **Ventricular tachycardia** is defined as three or more ventricular beats occurring at a rate of 120 beats per min or more.
- **Ventricular fibrillation** is an extremely rapid ventricular activation without pumping effect. Electrical defibrillation is the only effective therapy.
- **Vulnerable period** is a dangerous period in cardiac cycle just at the end of the contraction (simultaneous with the T-wave in the ECG). Electrical conversion (an electrical shock) given during this period may in itself initiate ventricular fibrillation. Refractory areas of cardiac muscle are spread among non-refractory areas.

Essentials

This paragraph deals with *1. Neural regulation of the heart function, 2. The action potential, 3. The role of calcium, 4. Spontaneous depolarisation, 5. Development of electrocardiography, and 6. The normal ECG.*

1. Neural regulation of the heart function

A large number of sympathetic (S) and vagal (X) motor nerve fibres end close to the sinoatrial node (SA in Fig. 11-1).

Fig. 11-1: The neural control of the heart.

Sympathetic stimulation speeds up the sinus node (*sinus tachycardia*) and vagal activity slows the node (*sinus bradycardia*). Increased concentration of the sympathetic transmitter noradrenaline, and of adrenaline from the adrenal glands, cause *positive inotropic state* (increased contractility), *positive chronotropic state* (increased frequency), *positive dromotropic state* (increased conduction velocity), and *positive bathmotropic state* (increased irritability) on the heart. Noradrenaline activates a-adrenergic constrictor receptors in the coronary vessels, whereas adrenaline activates b-adrenergic vasodilatator receptors.

The neurotransmitter acetylcholine, activating muscarinic receptors, and vagal stimulation causes *reduced contractility* (negative inotropy), *reduced frequency* (negative chronotropy), *reduced conduction velocity* (negative dromotropy), and *reduced irritability* (negative bathmotropy).

2. The action potential

Across the ventricular cell membrane there is a steady potential difference of almost the same size as the equilibrium potential for K^+ (-94 mV), that is -*90* mV (Fig. 11-2). This negative potential is referred to as the *resting membrane potential* (RMP), because it represents the potential difference across the cell membrane (inside negative) at rest between successive action potentials.

Fig. 11-2: Recordings of ECG (above), intracellular membrane potential (red curve) and contraction (blue curve) of one heart cycle in a ventricular fibre.

Any process that reduce the absolute size of the RMP (ie, depolarise the membrane) tends to activate (open) *fast* Na^+-channels. These channels contain fast opening and fast closing gates (inactivation gates). Electrochemical forces favour the abrupt influx of Na^+ from neighbouring regions. Hereby, the potential is further diminished and more and more Na^+-channels are activated or opened. The threshold potential for release of an action potential is a rise of 25 mV from -90 mV. The cardiac action potential is an *all-or-none response,* which can be divided into *five phases:*

The fast depolarization (phase 0) is shown by the abrupt upstroke, which is related to the rapid entry of Na^+ into the cell through the *fast* Na^+-channels, suddenly allowing the electrostatic and chemical forces to work (Fig. 11-2). The *fast* Na^+-influx causes phase 0 of atrial, ventricular and Purkinje action potentials (Fig. 11-2). The fast Na^+-channels are both *voltage*- and *time*-dependent. Phase 0 stops at about +30 mV, because the fast Na^+-channels become voltage-inactivated by closure of inactivation gates. The potential difference approaches the equilibrium potential for Na^+ (+ 60 mV), but only reaches +30 mV. The conduction velocity along the fast response fibre increases with the AP-amplitude and especially with the slope of phase 0.

Phase 1 is the early repolarization from the upstroke. This is related to K^+-outflux.

Phase 2 is the plateau of the action potential, where the *slow* Ca^{2+}-Na^+-channels remain open for a long period - up to 300 ms. The net influx of Ca^{2+} and Na^+ is almost balanced by a net efflux of K^+, so the balance is forming the *plateau* (Fig. 11-2). Ca^{2+} activates the muscle contractile process. When the slow Ca^{2+}-Na^+-channels close at the end of the plateau, the voltage-gated K^+-channels are activated, and the permeability for K^+ increases rapidly.

Phase 3 is the terminal repolarization. With *all* the K^+-channels open, large amounts of K^+ diffuse out of the ventricular fibre. The equilibrium potential for K^+ (-94 mV) and the RMP is rapidly approached.

Phase 4 is recognized by the RMP of - 90 mV (Fig. 11-2). The Na^+-K^+ pump restores ionic concentrations by exchanging Na^+ for K^+ in a ratio of 3:2.

Phase 5 covers the *relative* refractory period (RR), and the T-wave in the ECG. The long *absolute* refractory period (AR) of the ventricular cells covers the whole shortening phase of the contraction (Fig. 11-2 blue curve). In the absolute refractory period all fast Na^+-channels are voltage-inactivated and closed, which prevents sustained tetanus. As a consequence, no stimulus is sufficient to trigger contraction regardless of size.

In the relative refractory period, enough of the fast Na^+-channels are recovered, so that a sufficiently large stimulus can break through and produce an action potential although smaller than normal.

The long absolute refractory period protects the cardiac pump, as it is not possible to bring ventricles into smooth tetanus.

As described above the cardiac muscle fibre has a horizontal plateau, because of the *slow* Ca^{2+}-Na^+-channels (phase 2). - Skeletal muscle fibres have no plateau, because they do not open slow Ca^{2+}-Na^+-channels for such a long time.

3. The role of Ca^{2+}

Cardiac fibres contain many mitochondria and intercalated desks with gap junctions, transverse tubules of invaginated

sarcolemma, and sarcoplasmic reticulum (Fig. 11-3). The fibres require a continuous supply of oxygen and they are provided with a rich capillary supply, about one capillary for each cardiac fibre.

The spontaneous firing from the pacemaker spreads over the entire heart as a propagating wave.

The wave of excitation is conducted rapidly along the long axis of a cardiac fibre, and spreads along the myocardial sarcolemma from cell to cell via electrically conducting *gap junctions.*

The excitation also reaches the interior of the cell through the large T-tubules filled with mucopolysaccharides. During the phase 2 plateau of the action potential, Ca^{2+} permeability increases, and Ca^{2+} flows down its electrochemical gradient into the cell through the *slow* Ca^{2+}-channels.

The *channel proteins* are phosphorylated by a *cAMP-dependent protein kinase A*. The small Ca^{2+} influx is called *trigger*-Ca^{2+}, because it releases large amounts of Ca^{2+} from the sarcoplasmic reticulum (Fig. 11-3). Hence the cytoplasmic $[Ca^{2+}]$ increases from the resting level of 10^{-7} molar by a factor of 10-100 during excitation. The free Ca^{2+} binds to troponin C, just as in striated muscle cells, and the complex interacts with tropomyosin to activate sites between the actin and the myosin filaments. This process starts *crossbridge cycling* and thus contraction of the myofibrils.

Fig. 11-3: Excitation-contraction coupling in a cardiac fibre.

When the Ca^{2+}-influx ceases at the end of systole, the Ca^{2+} -movement is reversed. Now, Ca^{2+} is pumped into the sarcoplasmic reticulum by a Ca^{2+} -pump. The binding of Ca^{2+} to *troponin C* is inhibited by phosphorylation of troponin I, and the binding sites between actin and myosin are blocked resulting in diastole. In diastole the Ca^{2+} surplus is removed by a 3 Na^+- 1 Ca^{2+}-exchanger, and by an electrogenic Ca^{2+}-pump (Fig. 11-3).

Catecholamines and increasing extracellular $[Ca^{2+}]$, raise the cytoplasmic $[Ca^{2+}]$ and thus the developed force of contraction. This is accomplished in the following way. Adrenaline and noradrenaline activate adenylcyclase, whereby cAMP is formed and the dependent *proteinkinase A* phosphorylates and activates Ca^{2+}-channel proteins (Fig. 11-3). Hereby, the cytoplasmic Ca^{2+} is increased, and thus the force of contraction.

The excitability of cardiac fibres, striated muscle cells and neurons is reduced by a *reduction* of the cellular Ca^{2+}-gradient (hypocalcaemia), a *rise* in the Na^+-gradient across the cell membrane, or the *administration* of Ca^{2+}-blockers that prevent Ca^{2+} from entering the cell.

Cardiac glycosides, such as digoxin, are beneficial for patients with chronic cardiac failure. Digoxin inhibits and reduces the number of Na^+-K^+-pumps in the membranes of cardiac cells, hereby producing high intracellular levels of Na^+. Some of the intracellular Na^+ is exchanged for extracellular Ca^{2+}, and cytoplasmic Ca^{2+} enhances the force of cardiac contraction (positive inotropic effect). Exactly the necessary minimum dose of digoxin must be used for each patient with cardiac failure.

In myocardial cells, as in nerve and skeletal muscle cells, K^+ plays a major role in determining the RMP**.** At physiologic concentrations of K^+ outside the myocardial cell ($[K^+]^o$ about 4 mM), the RMP is determined by a dynamic balance between the membrane conductance to K^+ and to Na^+. As $[K^+]^o$ is increased (hyperkalaemia), the membrane depolarises. Depolarisation inactivates voltage-dependent K^+ -channels and activates Na^+ -channels, allowing Na^+ to make a proportionally larger contribution to the RMP. Increased $[K^+]$ in the extracellular fluid reduces the force of contraction, so the heart becomes dilatated and flaccid.

Because the equilibrium potential for Na^+ is positive (+ 60 mV), it tends to depolarise the RMP. Since RMP is -90 mV at normal $[K^+]$, and the equilibrium potential of K^+ is always more negative (-94 mV), there is a small outflux of K^+

from the cell at equilibrium. The K^+ and Na^+ gradients are maintained by the efficient Na^+-K^+-pump.

4. Spontaneous depolarisation

The rhythm of the heart is initiated by a complex flow of electrical signals, which are called *action potentials*. The action potentials generated in the sinus node (SN) display *automaticity* (ie, they undergo spontaneous and rhythmic depolarization without external stimuli). The sinus node is the *primary pacemaker,* because it has the highest frequency. The automaticity (Fig. 11-4) is associated with small *pale round cells* in the SN.

The SN and the AV-node also contain elongated cells that react with a special AP, a so-called *slow response*. The AP is smaller than the fast response, phase 0 is long and caused by slow Ca^{2+}-influx, phase 1 is absent, phase 2 is slow repolarisation, and the real repolarisation in phase 3 is due to inactivation of the slow Ca^{2+}-channels and increased K^+–outflux. Phase 4 is horizontal, and the relative refractory period is long, extending well into phase 4. The slow response propagates slowly and tends to be blocked, which causes cardiac rhythm disturbances.

The *membrane potential* of pacemaker cells (about -55 to -60 mV) is *never* constant (Fig. 11-4). The diastolic depolarization of pacemaker cells (producing the very special slope in phase 4 of these cells) is ascribed to an influx of Na^+ through special channels in their cell membrane. Towards the end of phase 4 there is also a certain influx of Ca^{2+} through voltage-activated Ca^{+2} -channels. Hypocalcaemia diminishes the amplitude of the action potential and the slope of the pacemaker potential. The diastolic pacemaker depolarization is opposed by the K^+-outflux.

Fig. 11-4: Pacemaker potentials from a sinus nodal fibre: a: Sympathetic stimulation; b. Normal heart rate; c. Vagal stimulation. There is a constant Na^+-influx between two heartbeats. Reduction in the slope of the pacemaker potential and increase of the threshold both diminish the cardiac frequency.

The K^+-channels remain open during the repolarization, so the membrane potential approaches -60 mV at the end of the action potential. The rising pacemaker potential between two heartbeats, is caused by the inherent leaks of the pacemaker membrane to Na^+. The *constant* Na^+-influx causes the membrane potential to approach the threshold potential. When the membrane potential reaches the threshold potential (about -40 mV), the slow Ca^{2+}-Na^+-channels open and the cycle starts again. This process continues for a lifetime (Fig. 11-4).

The heartbeat is self-initiating, and normally the propagating wave (impulse) originates in the sinus node (see above). The impulse propagates from the sinus node via *three* bundles of internodal syncytial cells, through the left and right atrial wall to the *atrioventricular* (AV) *node*. This point is typically reached within 40 ms. After passing through the AV node (with a so-called AV- *delay* of 100 ms), the propagating wave (excitation) reaches the *bundle of His*. Here, specialised conduction fibres activate almost synchronously all the ventricular tissue and thus impart maximal thrust to the blood. The propagation velocity in these large *Purkinje fibres* is 1-4 m per s, which is the fastest velocity possible in the heart. The Purkinje system terminates just under the endocardial surface on *gap junctions* in the myocardial cells. The AV-delay provides time for the atrial systole to pass extra blood to the ventricles before the ventricular systole occurs.

Adrenergic transmitters and sympathetic stimulation increase the slope of the diastolic pacemaker depolarization.

Acetylcholine and vagal stimulation increase the K^+-efflux, so the slope is reduced and thus the cardiac frequency is reduced.

Some cardiac fibres show a slow response comparable to that of the pacemaker cells, but with a constant phase 4. Ischaemia may activate *ectopic* pacemaker cells.

5. Development of electrocardiography

In 1903 Einthoven began a systematic study, with a string galvanometer of the potential differences between electrodes placed on the skin surface during heart beats. The string galvanometer was developed to become the first electrocardiograph. The fluids of the body conduct electricity quite easily. The body conducts electrical signals uniformly in all directions.

An electrocardiogram (ECG) is a curve showing the potential variations in time in the body stemming from the heart, which is an *electrochemical generator* suspended in a conductive medium acting as a volume conductor. The myocardial cell membranes have separate charges, and small ion gradients produce electrochemical gradients. Small ion fluxes across the cell membrane only occur during depolarization and repolarization, where potential differences are produced between polarised and depolarised tissue regions in the heart. Each cardiac fibre behaves as a dipole, the magnitude and direction of which is symbolised by an arrow or a vector. The dipole vector points from minus to plus by definition.

Originally Einthoven assumed that the sum of all electrical activity in the heart resulted in an electromotive force - a *main cardiac vector* originating in the middle of the heart.

Einthoven used three standard bipolar limb leads forming a triangle in the frontal plane (Einthoven's triangle in Fig. 11-5 and 11-7). Lead I records the potential difference between the right and left arms, lead II between the right arm and left leg; and lead III between the left arm and left leg. The right leg is used to ground the patient (Fig. 11-7; observe the positive and negative signs). The connections were arbitrarily chosen so most healthy individuals had dominating upright (positive) QRS-complexes and T-waves in their leads.

The actual recording sites are at the junctions between limbs and trunk, because the arms and legs act as *extended electrodes*. Einthoven actually placed the limbs of the patient in bathing tubes and used these as the first electrodes. Einthoven arranged the ECG equipment so that the *direction* of the mean QRS-axis towards the *positive pole* of a bipolar lead produces an *upright* deflection (*positive* R-waves in the 3 leads shown in Fig. 11-5). When directed towards the *negative pole* a *downward* deflection (a *negative* RIII-wave in left-sided axis deviation, Fig. 11-7) was recorded.

Einthoven's law states that any two of the three bipolar limb leads determine the third one with mathematical precision. The potential differences recorded over time in an ECG can be estimated using the rules of *vectorial* projection with force parallelograms in the frontal plane (Fig. 11-5).

Fig. 11-5: Einthoven's triangle. To the left is shown the main direction of the conduction system in the frontal plane.

The R-waves (and QRS deflections) in two of the three standard leads are drawn graphically in Einthoven's triangle and their resultant is the *mean QRS-axis* of the heart or more exact the mean ventricular axis in the frontal plane (Fig. 11-5). This is a mathematical concept representing the integral cardiac vector operating in Einthovens triangle.

The *mean QRS-axis* of the heart is usually located around 60 degrees in Einthoven's triangle.

Right- sided axis deviation is found in children or high thin individuals with a vertical located heart, and in persons with right ventricular hypertrophy. The axis is now located to the right and the condition is diagnosed by a positive R_{III} combined with a negative R_I (between +90 to +180 and even +90 to –90 degrees in Fig. 11-7).

Left-sided axis deviation is found with left ventricular hypertrophy, in fat individuals and in late pregnancy, when the heart is pressed upwards to the left. The condition is diagnosed by a negative R_{III} combined with a positive R_I (Fig. 11-7).- Actually cardiologists operate with the *net areas* of the QRS-complexes in order to diagnose significant deviations.

6. The normal ECG

The ECG is recorded from the surface of the body, and used to demonstrate the presence of AV-blocks, ectopic foci, premature beats, sinoatrial arhythmias, atrial fibrillation and ventricular fibrillation etc.

Each heart cycle has a fixed pattern of ECG waves (P,Q,R,S and T). The sinus node is a minimal muscle mass, and there is no potential difference (wave in the ECG) before the atria depolarise with a P-wave. When the propagating impulse wave is directed towards the *positive* electrode (as in lead II) the atrial depolarization will produce a *positive* P-wave (Fig. 11-6). The P-waves correspond to the impulse distribution in the atria. Retrograde excitation of the atria creates a *negative* P-wave. When atrial excitation coincides with the QRS, the P-wave is often hidden. The PR-interval (normally 0.12-0.2 s) measures the impulse speed through the supraventricular tissue from the sinus node to the

bifurcation of the Hiss bundle. Prolonged PR-interval is caused by disturbances of AV conduction. The QRS-complex measures the passage through the left and right bundle branch followed by depolarization of the strong ventricular myocardium (Fig. 11-6). The QRS-complex is prolonged (> 0.12 s), when the left or right bundle branch is blocked by disease. When excitation originates in the ventricles, it spreads slowly and the QRS-complex is severely deformed.

Since the activation of the septum is mainly from left towards right, the propagating wave moves away from the exploring electrode, and the *Q-wave* becomes *negative*, whereas the dominating ventricular transfer of the propagating wave towards the apical electrodes provides the large, *positive* R-wave in almost all leads. The small propagating wave moving away from the electrode at the apex and to the right to reach the thin-walled right ventricle, is responsible for the small, *negative* S-wave. The T-wave represents ventricular repolarization and has the same direction as the QRS-complex in most normal leads. When the QRS-complex is positive and the T-wave is negative, it indicates that the repolarization proceeds in a wrong direction. Abnormal T-waves are due to cardiac hypertrophy, myocardial damage or electrical disturbances.

The isoelectric ST-segment represents the period, where the entire ventricular myocardium is depolarised (therefore isoelectric). Any deviation (up or down from the isoelectric level) indicates *anoxic damage* of the myocardium.

The QRS- plus ST-intervals correspond to the duration of the ventricular systole.

Fig. 11-6: Normal ECG (II. lead). The action potentials from an atrial (green curve) and a ventricular fibre (blue curve) are shown above. – To the right is shown the direction of propagating waves in the frontal plane and their relation to the ECG waves.

The *T-wave* is positive in most leads, and due to the apical directed repolarization of the ventricular action potential (Fig. 11-6).

Unipolar, precordial leads (Wilson) have *one* exploring electrode as the actual recording electrode, detecting changes in the local potential relative to zero. The exploring electrode is defined relative to the three standard extremity leads, which are connected to form one indifferent *reference electrode*. Conventionally, the *6 Wilson leads* are recorded from specific locations on the precordium (V_1 to V_6).

The exploring electrodes of leads V_1 to V_3 are located to the right and are looking at the right side of the heart. The QRS complexes of the normal heart are mainly negative in these leads. The exploring electrodes of leads V_4 to V_6 are located to the left and are looking at the left side of the heart, where the QRS complexes are typically positive.

Fig. 11-7: Standard limb leads (Einthoven's triangle) and precordial ECG leads.

In unipolar recordings, upright deflection indicates movement of the propagating wave towards the positive, exploring electrode, and downward deflection (a negative ECG wave) indicates that the propagation wave is moving away (Fig. 11-7).

Pathophysiology

This paragraph deals with *cardiac arrhythmias.*

Arrhythmia is any cardiac rhythm different from the normal sinus rhythm. Sinus rhythm is the automaticity elicited from the normal cardiac pacemaker, the sinus node. The sinus node depolarises spontaneously with a frequency between 60 and 100 beats per min (bpm) at rest. Normally the parasympathetic tone predominates, whereby the heart rate is reduced from 100 to 60-70 bpm at rest. Any reduction in parasympathetic tone or increase in sympathetic tone results in tachycardia (ie, a ventricular pumping frequency above 100 bpm). Reduction in sympathetic tone or increase of the parasympathetic tone leads to bradycardia (ie, a ventricular pumping rate below 60 bpm).

Cardiac arrhythmias are divided into two groups: I. Pacemaker abnormalities, and II. Conduction abnormalities (cardiac block).

I. Pacemaker abnormalities

arise in the sinus node (ie, sinus tachycardia, sinus bradycardia or sinus arrhythmia) or outside the node (ie, ectopic beats, tachycardia, and fibrillation and shifting pacemaker). These arrhythmias are disorders of rhythmogenesis.

Sinus node disease is a mixture of conditions with insufficient discharge of signals from the sinus node. The *sick sinus syndrome* is caused by damage of the nodal tissue. Sinus pauses lead to sinus bradycardia, tachycardia, tachy-brady-cardia syndrome, ectopic beats or atrial arrest (sinus arrest).

Sinus tachycardia (heart rate above 100 bpm) is caused by psychological (panic, anxiety) or physical stress (anaemia, hypoxia, exercise, fever, intoxication, shock etc). Any condition with increased sympathetic tone results in tachycardia. Adrenergic b-blockers are effective in slowing down the heart rate.

Sinus bradycardia (heart rate below 60 bpm) is a normal phenomenon in well-trained people and in elderly persons. Athletes at rest have a high stroke volume and a low pulse, because they are dominated by vagal tone in this condition. All persons react with sinus bradycardia, during hypothermia, hypothyroidism, jaundice, increased intracranial pressure, treatment with digoxin or b-blockers, and in sinus node disease.

Sinus arrhythmia is a normal phenomenon. There is a rise in heart rate during inspiration followed by a fall during expiration. During inspiration, the low intrathoracic pressure improves the venous return, which – with a delay through the pulmonary circulation - increases the stroke volume of the left ventricle. The resulting increase in aortic intravascular pressure combined with the low intrathoracic pressure around the aorta leads to an increased transmural pressure over the aortic wall and thus to a strong stimulation of the aortic baroreceptors. This is why the vagal tone falls and the sinus node discharge quickens. The reflex and circulation delay explains why the fall in heart rate manifests itself during expiration and not during inspiration.

Ectopic beats originate in pacemaker cells outside the sinus node. The abnormal pacemaker tissue is triggered by ischaemia, mechanical or chemical stimuli. Cardiac catheterisation can trigger ectopic beats. Ectopic beats are either of atrial or ventricular origin.

1) *Atrial ectopic beats* appear as early (premature extrasystoles) and abnormal P-waves in the ECG; they are usually followed by normal QRS-complexes (Fig. 11-8). Following the premature beat there is often a compensatory interval. A premature beat in the left ventricle is weak because of inadequate venous return, but after the long compensatory interval, the post-extrasystolic contraction (following a long venous return period) is strong due the Starling´s law of the heart. - Adrenergic b-blockers are sometimes necessary.

2) *Ventricular ectopic beats* (extrasystoles) are recognized in the ECG by their wide QRS-complex (above 0.12 s), since they originate in the ventricular tissue and slowly spread throughout the two ventricles without passing the Purkinje system. The ventricular ectopic beat is recognized by a double R-wave (Fig. 11-8). The classical tradition of simultaneous cardiac auscultation and radial artery pulse palpation eases the diagnosis. Now and then a pulsation is not felt, and an early frustraneous beat is heard together with a prolonged interval. A beat initiated in the vulnerable period may release lethal ventricular tachycardia, since the tissue is no longer refractory.

Fig. 11-8: Atrial (left) and ventricular (right) ectopic beats.

After a period with high cardiac frequency from activity of an ectopic focus (ie, overdrive), there often follows a period with a remarkable fall in frequency (ie, overdrive suppression). During the overdrive, the Na^+-K^+-pump is extremely active in order to extrude Na^+ from the myocardial cells in the short phase 4 periods, and the Na-influx exceeds the K-influx (Na:K= 3:2). Hereby, the cell becomes hyperpolarised in the end, which may stop the high frequency and turn it into suppression.

Tachycardia occurs in *paroxysms* and is either of atrial or ventricular origin.

1) *Atrial tachycardia* is elicited in the atrial tissue outside the SN as an atrial frequency around 200 bpm. Often only every second impulse passes the AV-node to the ventricles, so a 2.1 AV-block is found in the ECG (Fig. 11-9).

2) *Ventricular tachycardia* is elicited from one focus in the ventricular tissue with a frequency around 200 bpm (more than 120 bpm) and abnormal intraventricular impulse conduction (disturbed QRS complexes). Of course, there are no P-waves in the ECG, and the QRS-complexes are broad and irregular (Fig. 11-9).

Fig. 11-9: Left: Atrial tachycardia with a QRS-frequency of 100 bpm. - Right: Ventricular tachycardia with a

QRS-frequency of 200 bpm following 2 sinus beats.

Fibrillation is either atrial or ventricular in origin.

1) *Atrial fibrillation* is a condition in which the sinus node no longer controls the rhythm and the atrial muscle fibres undergo a tumultuous rapid twitching. A total irregularity of ventricular contractions characterise the fibrillation. An excitation wave with 400-600 cycles per min, courses continuously through the atrial wall over a circular pathway about the origin of the great veins (the *circus motion theory*). There is a continuous activation with more than 400 P-waves per min, where regular atrial contraction is impossible. It is difficult to see and count the P-waves of the ECG. Because of the refractoriness of the AV-bundle, only some of the excitation waves result in ventricular beats. The pulse of the patient is therefore irregular as the occurrence of QRS-complexes in the ECG. The many P-waves (also called f-waves for fluctuations) are characteristic for atrial fibrillation. Untreated atrial fibrillation has a QRS-frequency of 150-180 bpm (Fig. 11-10). Old patients with chronic heart disease often show the so-called *slow atrial fibrillation* with a QRS-frequency below 60 bpm.

Most cardiac disorders can lead to atrial fibrillation or flutter.

Atrial flutter is related to atrial fibrillation, but the atrial frequency - counted from the P-waves - is much lower than 400 bpm - usually around 300 bpm and the AV-conduction is more regular. The consequences to the patient depend upon the number of impulses conducted from the atria through the AV-node to the ventricles (recorded as QRS-complexes). Often every second impulse reaches the ventricles, so the ratio of AV-blocks is 2:1, but the ratio can also be 3:1, 4:1 etc. Atrial flutter is recognized in the ECG as sawtooth-like P-waves (Fig. 11- 10).

2) *Ventricular fibrillation* is a tumultuous twitching of ventricular muscle fibres, which are ineffectual in expelling blood. The condition is lethal without effective resuscitation. The irregular ventricular rate is 200-600 twitches/min. Without contractile co-ordination the force is used frustraneous. Actually, the heart does not pump blood, so within 5 s unconsciousness occurs, because of lack of blood to the brain. In patients with coronary artery disease, ventricular fibrillation is a cause of sudden death. The trigger is *anoxia* (with an ineffective Na^+-K^+-pump) and the impulses arise from several foci in the ventricular tissue. There is no regular pattern in the ECG. Ventricular fibrillation is initiated when a premature signal arrives during the downslope of the T-wave (*vulnerable period*). Electrical shock (electrocution) also triggers ventricular fibrillation.

Fig. 11-10: Ventricular defibrillation of a patient with cardiac arrest.

Ventricular fibrillation is the most serious cardiac arrhythmia. It must be converted to sinus rhythm at once by the application of a large *electrical shock* to the heart (ventricular defibrillation) or the patient will die. Alternating current is applied for 100 ms or 1000 volts direct current is applied for a few milliseconds. The *vulnerable period* (VP in Fig. 11-7 is actually phase 3 and represented in the ECG as the T-wave) is dangerous, because an electrical shock, when given during this period, will cause in itself *ventricular fibrillation*.

Here is shown sinus rhythm and one ectopic beat followed by ventricular fibrillation. The only effective treatment is rapid institution of electrical defibrillation.

Fig. 11-11: Left: Atrial fibrillation with a QRS-frequency of 180 bpm. - Right: Ventricular fibrillation following 2 beats of sinus rhythm and one ectopic beat.

Shifting pacemaker is a condition where the impulse originates in shifting locations inside the SN, or the pacemaker shifts from the SN to the AV-node. In the first case the P-wave change size from beat to beat, and in the second case the P-wave is found either in front of the QRS-complex or behind.

II. Conduction abnormalities

are 1) sino-atrial block, 2) atrioventricular block, 3) bundle branch block, 4) WPW- syndrome, and 5) the long QT-syndrome.

1) Sino-atrial block (see SN disease) is characterized by long intervals between consecutive P-waves, and caused by blockage of the formation or conduction of the stimulus from the SN to the atrial tissue (ischaemia or infarction of the SN).

2) Atrioventricular block is blockage of the conduction from the atria to the AV-node.

The first-degree AV block is a prolongation of the PQ (PR)-interval (above 0.2 s) implying a delay of the conduction - not a real block. All beats are conducted, so there is a QRS-complex following each P-wave, although with delay (Fig. 11-12).

The second-degree AV block occurs when some signals are not conducted to the AV-node, so some of the P-waves are not followed by QRS-complexes. The ventricles actually drop some beats. A typical example is Mobitz type I block or Wenchebach block, which is a predictive loss of a QRS-complex. The PQ-interval is increased progressively until a P-wave is not followed by a QRS-complex. Mobitz type II block occurs without warning. Suddenly, a QRS- complex falls out (Fig. 11-12).

The third degree AV block (complete AV-block) is a total block of the conduction between the SN and the ventricles. Also blocked Hiss bundle conduction results in an AV-block (Fig. 11-12). An AV- or ventricular pacemaker maintains life with a spontaneous escape rhythm around 40-50 bpm, or cardiac arrest occurs with the fainting paroxysms of Adam-Stokes syndrome.

The Adam-Stokes syndrome is a clinical disorder caused by a partial AV-block, with a long P-Q interval and a wide QRS complex in the ECG, suddenly becoming a total bundle block. The condition results in unconsciousness and cramps caused by brain hypoxia and sometimes resulting in universal cramps (*grand mal*) due to violent activity in the motor cortex. The keyhole is the AV node and the bundle of His. Disease processes here elicit the Adam-Stokes syndrome. Therapy is to provoke sinus rhythm by a few forceful strokes in the precardial area of the thorax, accompanied by mouth-to-nose-resuscitation and external heart massage. A sympathomimetic drug can be injected if necessary even in the heart directly. The patient must be immediately brought to hospital for special intensive care. Permanent pacemaker treatment may become necessary.

Fig. 11-12: Four types of atrio-ventricular (AV)-block. From above downwards: First-degree AV-block, Second-degree Mobitz I block (Wenchebach), Second-degree Mobitz II block, and Complete AV-block.

3) Bundle branch block is a block of the right or the left bundle branches. The signal is conducted first through the healthy branch and then it is distributed to the damaged side. This distribution takes more time than usual, so the QRS-complex is wider than normal (more than 0.12 s in Fig. 11-13).

In *right* bundle branch block, the right ventricle is activated late, which is shown by a tall double R-wave in V1 (ie, the second late R-wave is from the right side), and a deep wide S-wave in leads I and V6 .

The *left* bundle branch block is characterized by a late activation of the left ventricle from apex towards basis. This results in a solid R-wave in the left precardial leads (V5 andV6), whereas there is a deep broad S-wave in V1 and III (Fig. 11-13).

Fig. 11-13: Right and left bundle branch block.

4) WPW-syndrome or *Wolf-Parkinson-White block* is not a direct block of the conduction through the Hiss bundle and branches, but is caused by a short cut through an extra conduction pathway from the atria to the ventricles. This abnormal conduction pathway is congenital and called the *bundle of Kent* (Fig. 11-14). Due to this short-cut, the slow conduction through the AV-node is bypassed and the ventricles are depolarised faster than normal. The WPW-syndrome is recognized in most ECG leads as a short PQ (PR)-interval followed by a wide QRS-complex with a delta wave (Fig. 11-14). The patients often have paroxysmal tachycardia or they may develop atrial fibrillation.

Some patients are treated with ablation of the *bundle of Kent*. Other patients are asymptomatic and in good physical condition.

Fig. 11-14: The WPW-syndrome and the long QT-syndrome.

5) The long QT-syndrome. This is frequently a genetic condition, where fast repolarised cells are restimulated by cells that have not repolarised. When acquired the condition is caused by myocardial ischaemia, by drugs or by a low serum $[Ca^{2+}]$ - below 2 mM. Normally, the QT-interval is less than 50% of the preceding RR-interval (Fig. 11-14). The long QT-interval symbolises a long ventricular systole. Actually, the ST-interval is simultaneous with the phase 2

plateau of the ventricular membrane action potential. Here, the slow Ca^{2+} -Na^{+} - channels remain open for more than 300 ms as normally. The net influx of Ca^{2+} and Na^{+} is almost balanced by a net outflux of K^{+}. Hereby, a long phase 2 plateau or isoelectric segment is formed.

Cardiac pacemakers

Implanted cardiac pacemakers are successful in keeping heart patients alive. This is often a beneficial treatment of Adam Stokes syndrome or ventricular tachycardia. An electrical pacemaker is a small stimulator with battery planted underneath the skin. The electrodes are connected to the right ventricular muscle tissue, whose contraction rate is controlled by the stimulator.

Cardiopulmonary resuscitation

Cardiac arrest is cessation of all spontaneous cardiac rhythmicity. Cardiac arrest is most often caused by anoxia. The cause of anoxia is inadequate respiration due to terminal lung disease, thoracic trauma, and shock or deep anaesthesia.

Cardiopulmonary resuscitation is important in keeping the heart alive until electrical defibrillation can be performed with a large electrical shock. Alternating current is applied for 100 ms, or 1000 mV direct current is applied for a few ms.

Ventricular hypertrophy

The consequences of the rise in cardiac mass are the same for the locally recorded ventricular fibre action potentials. The action potentials increase in magnitude, with a parallel increase of QRS voltage observed in the ECG. The heart partially adapts to the increase in workload by an increase in muscle mass (hypertrophy). The degree of hypertrophy is roughly proportional to the increase in load. As the electrically active surface area is increased, there is an increase in ventricular fibre action potential, and thus in the amplitude of the R wave in the left precordial leads.

The *ECG criterion* of left ventricular hypertrophy is that the sum amplitude of S in V_1 and R in V_6 is larger than 3.5 mV (3.5 cm).

With a delay in conduction through the large left ventricle - or with left bundle branch block - there is a wide QRS complex.

The normal positive T-wave is due to the apical directed repolarization of the ventricular AP. Therefore, asymmetrical T-inversion or bi-phasic T-waves and downward-sloping ST-segment signal abnormal repolarization with the propagating wave moving away from the apical electrode in most leads (so-called strain pattern in Fig. 11-15).

Left axis deviation is often found in the standard leads of left ventricular hypertrophy patients (augmented R-wave in I and S-wave in III).

Fig. 11-15: ECG from a patient with left ventricular hypertrophy.

The disorders causing *left ventricular hypertrophy* are:

Heart diseases: Myocardial disorders, pericarditis, valvular disorders, congenital heart disease.

Vascular disorders: Atherosclerosis, systemic hypertension (Chapter 12) , aortic stenosis, renal disorders, arteriovenous shunts, and aneurysms.

Thoracic diseases: Diseases of the lungs & pleura, and kyphoscoliosis.

Pumping of increased volume load: Acromegaly, anaemia, obesity and excessive alcohol intake, thyrotoxicosis, severe manual work and sports.

Left ventricular hypertrophy is often demonstrated by echocardiography or found on the ECG.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

Stem statement: The ventricular action potential is

A. initiated by rapid entry of Na^+.

B. characterised by slow Ca^{2+} -Na^+- channels.

C. characterised by closed K^+- channels in phase 3.

D. dependent upon Ca^{2+}-influx.

E. independent of the Na^+-K^+ -pump in phase 4.

II. Each of the following five statements have True/False options:

A. In myocardial cells, as in nerve and skeletal muscle cells, K^+ plays a minor role in determining the resting membrane potential.

B. The impulse propagates from the sinus node via five bundles of internodal syncytial cells through the left and right atrial wall to the atrioventricular node.

C. The long absolute refractory period of the ventricular cells, covers the whole shortening phase of the contraction, where all the fast Na^+-channels are voltage-inactivated. As a consequence, no stimulus is sufficient regardless of size.

D. The fast Na^+-influx causes phase 0 of atrial- , ventricular- , and Purkinje- action potentials. The fast Na^+-channels are both voltage- and time-dependent.

E. Noradrenaline activates a-adrenergic constrictor receptors in the coronary vessels, whereas adrenaline activates b-adrenergic vasodilatator receptors.

III. The following five statements have True/False options.

A. WPW-syndrome or Wolf-Parkinson-White block is caused by a short cut through an extra conduction pathway from the atria to the ventricles.

B. Atrial fibrillation is more malignant than ventricular fibrillation.

C. All pacemaker abnormalities arise in the sinus node.

D. Premature beats are also called atrial ectopic beats.

E. Only few cardiac arrhythmias can lead to atrial fibrillation and flutter.

Try to solve the problems before looking up the answers.

Highlights

- *Heart rate is controlled mainly by the autonomic nervous system. Sympathetic stimulation speeds up the sinus node*

(sinus tachycardia) and vagal activity slows the node (sinus bradycardia).

- *The autonomic nervous system controls myocardial contraction by varying the Ca^{2+} -permeability of the sarcolemma via hormones and the adenylcyclase system.*
- *Increased concentration of the sympathetic transmitter noradrenaline, and of adrenaline from the adrenal glands, cause increased contractility, increased frequency, increased conduction velocity, and increased irritability of the heart.*
- *Catecholamines and increasing extracellular $[Ca^{2+}]$, raise the cytoplasmic $[Ca^{2+}]$ and thus the developed force of contraction.*
- *Cardiac digitalis glycosides (digoxin) block the Na^{+}-K^{+}-pump. This blockage increases the internal $[Na^{+}]$ to the extent that less Ca^{2+} is removed from the cell. This - and any - form of elevated cytoplasmic $[Ca^{2+}]$ enhances contractile force. Increased $[K^{+}]$ in the extracellular fluid reduces the force of contraction, so the heart becomes dilatated and flaccid.*
- *Noradrenaline activates α-adrenergic constrictor receptors in the coronary vessels, whereas adrenaline activates β-adrenergic vasodilatator receptors.*
- *The neurotransmitter acetylcholine, activating muscarinic receptors, and vagal stimulation cause reduced contractility (negative inotropic state), reduced frequency (negative chronotropic state), reduced conduction velocity (negative dromotropic state), and reduced irritability (negative bathmotropic state).*
- *Changes in blood concentrations of gasses and protons affect the cardiac function directly and indirectly via chemoreceptors.*
- *The long absolute refractory period of the ventricular cells, covers the whole shortening phase of the contraction, where all the fast Na^{+}-channels are voltage-inactivated. As a consequence, no stimulus is sufficient regardless of size.*
- *The electrocardiogram (ECG) is a surface recording of the electrical field generated in the entire body by the heart.*
- *The sinus node is a minimal muscle mass, and there is no potential difference (wave in the ECG) before the atria depolarise with a P-wave. When the propagating wave is directed towards the electrode (as in lead II) the atrial depolarization will produce a positive P-wave.*
- *The P-waves correspond to the impulse distribution in the atria, and the QRS-complex origin from depolarisation of the strong ventricular myocardium.*
- *The QRS deflections in two of the three standard leads can be drawn graphically in a triangle and their resultant is the mean QRS-axis of the heart.*
- *The T-wave is caused by the spread of repolarization over the ventricles.*
- *The small propagating wave moving away from the electrode at the apex and to the right to reach the right ventricle, is responsible for the small, negative S-wave.*
- *The Adam-Stokes syndrome is a clinical disorder caused by a partial AV-block, with a long P-Q interval and a*

wide QRS complex in the ECG, suddenly becoming a total bundle block. The condition results in unconsciousness and cramps caused by brain hypoxia and sometimes resulting in universal cramps (grand mal) due to violent activity in the motor cortex. Disease processes in the AV node and the bundle of His elicit the Adam-Stokes syndrome.

- *Ventricular tachycardia is defined as three or more ventricular beats occurring at a rate of 120 beats per min or more.*
- *Ventricular fibrillation is an extremely rapid ventricular activation without pumping effect. Electrical defibrillation is the only effective therapy.*
- *Vulnerable period is a dangerous period in cardiac cycle represented in the ECG as the downslope of the T-wave. Electrical conversion (an electrical shock) given during this period causes in itself ventricular fibrillation.*

Further Reading

Cardiovascular Reviews & Reports. Monthly journal published by Le Jacq Communications Inc, 777 West Putnam Av., Greenwich CT, 06830, USA.

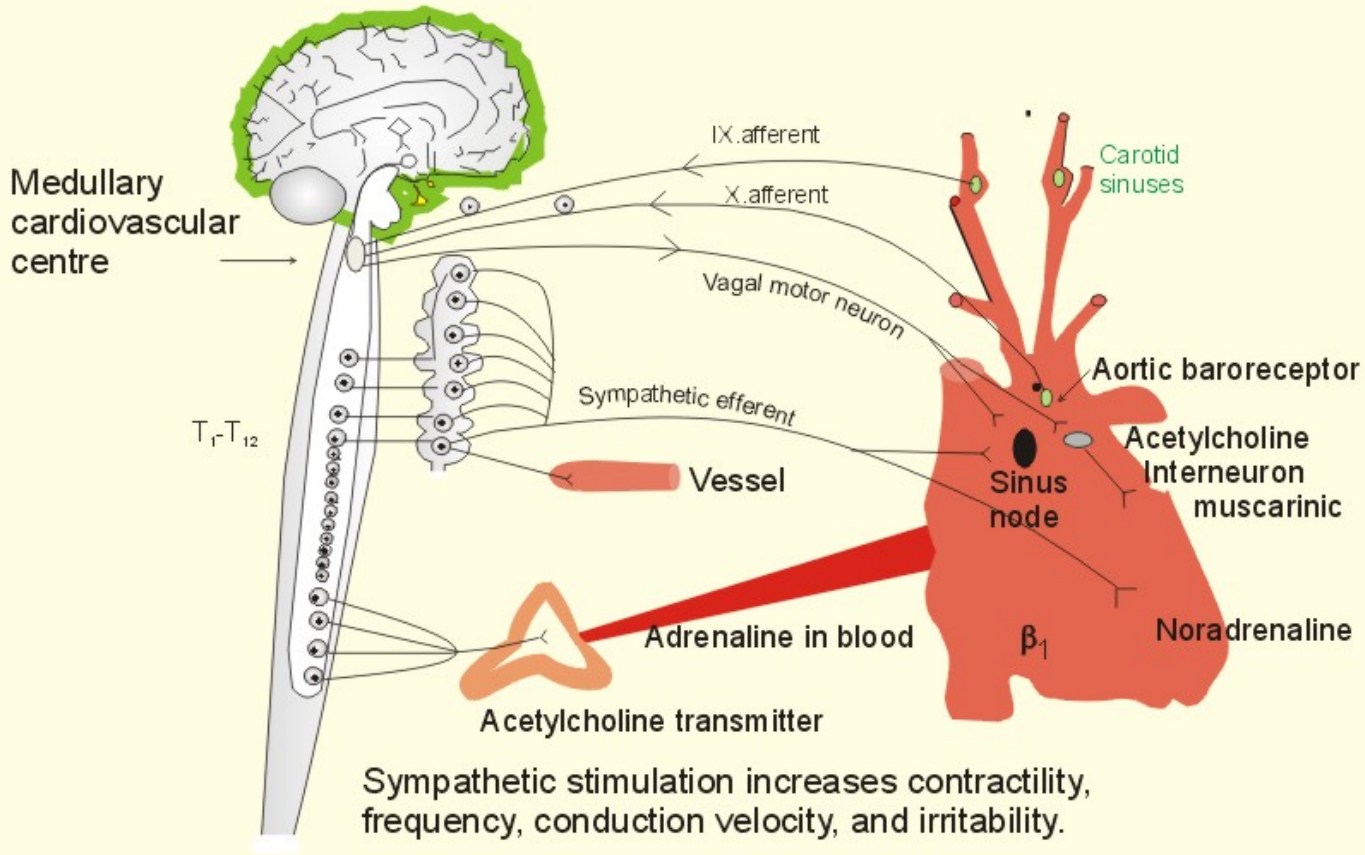
Kastor, JA (1994) *Arrhythmias*. Philadelphia: W.B.Saunders Co.

Surawicz, B (1995) *Electrophysiological Basis of ECG and Cardiac Arrhythmias*. Baltimore: Williams and Wilkins.

Return to content

Return to chapter 11

Control Of Cardiac Function



Sympathetic stimulation increases contractility, frequency, conduction velocity, and irritability.

Fig. 11-1

KMc

Chapter 2.

Multiple Choice Questions

I. Answers **B, D**, and **E** are true statements, whereas A and C are false.

II. Answers A, B, D, and E are false statements, whereas **C** is true.

Return to chapter 2

Return to content

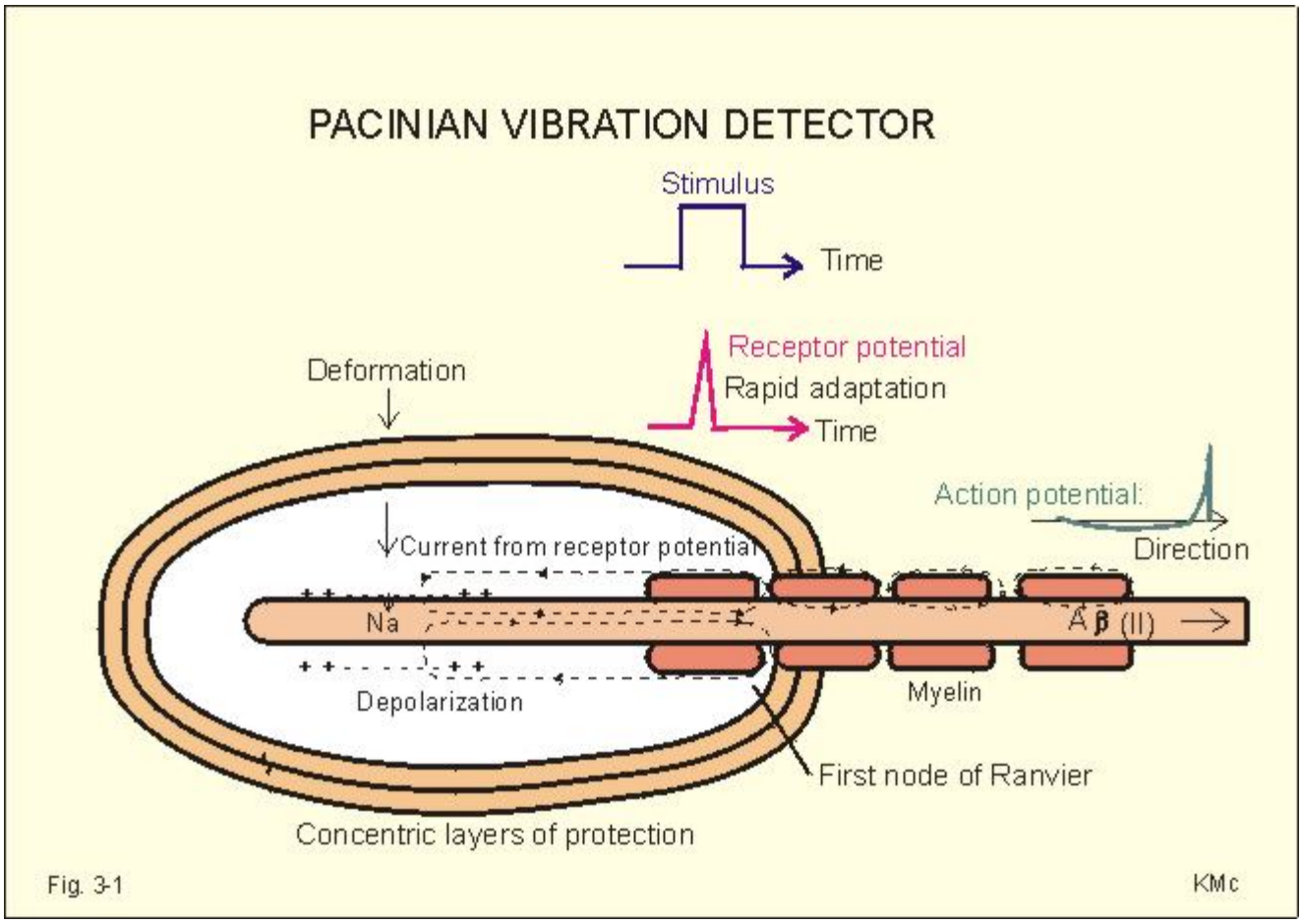

PACINIAN VIBRATION DETECTOR
Stimulus
Time
Receptor potential
Rapid adaptation
Time
Deformation
Action potential:
Direction
Current from receptor potential
Na
Aβ (II)
Depolarization
Myelin
First node of Ranvier
Concentric layers of protection

Fig. 3-1

KMc

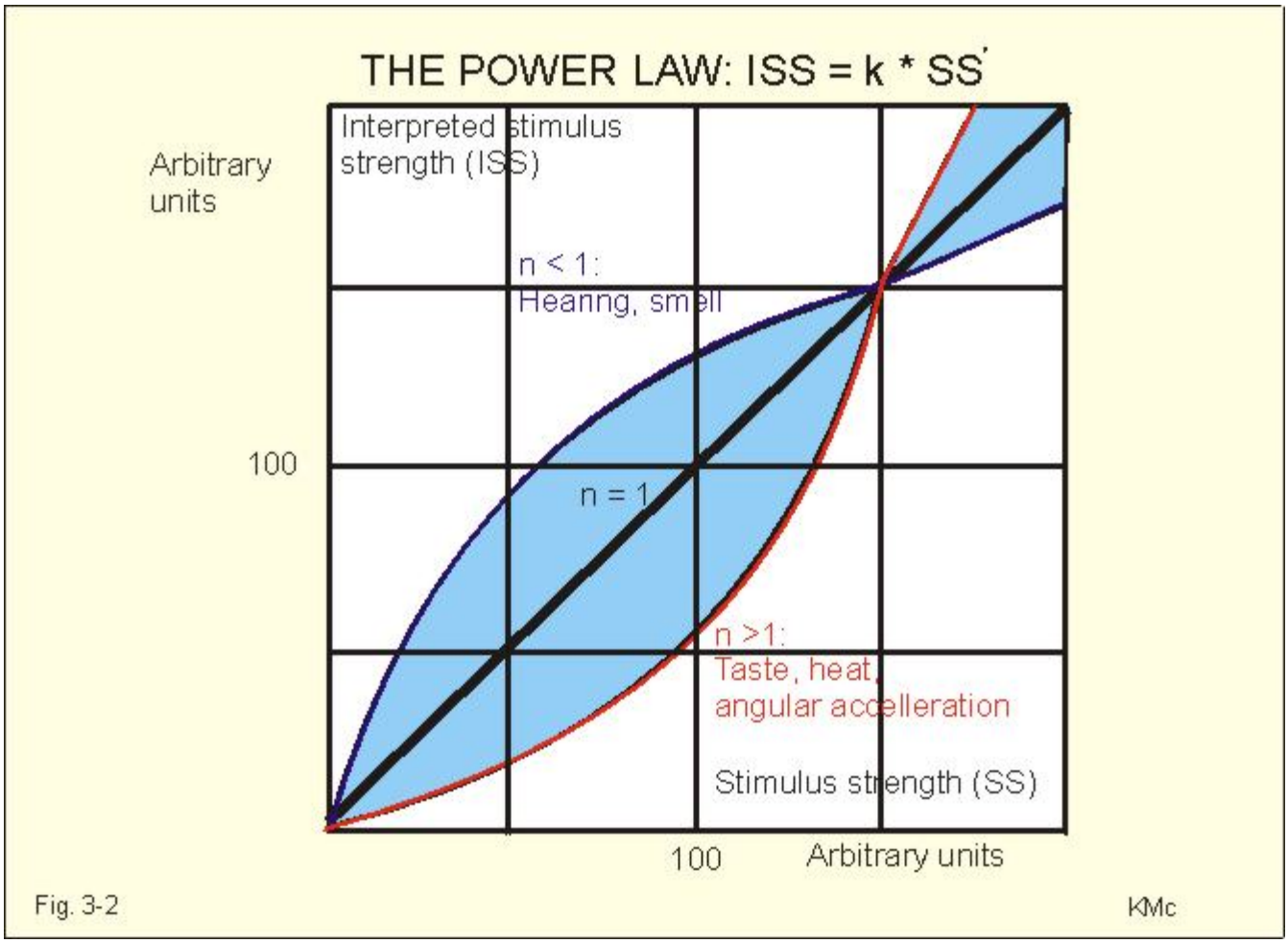
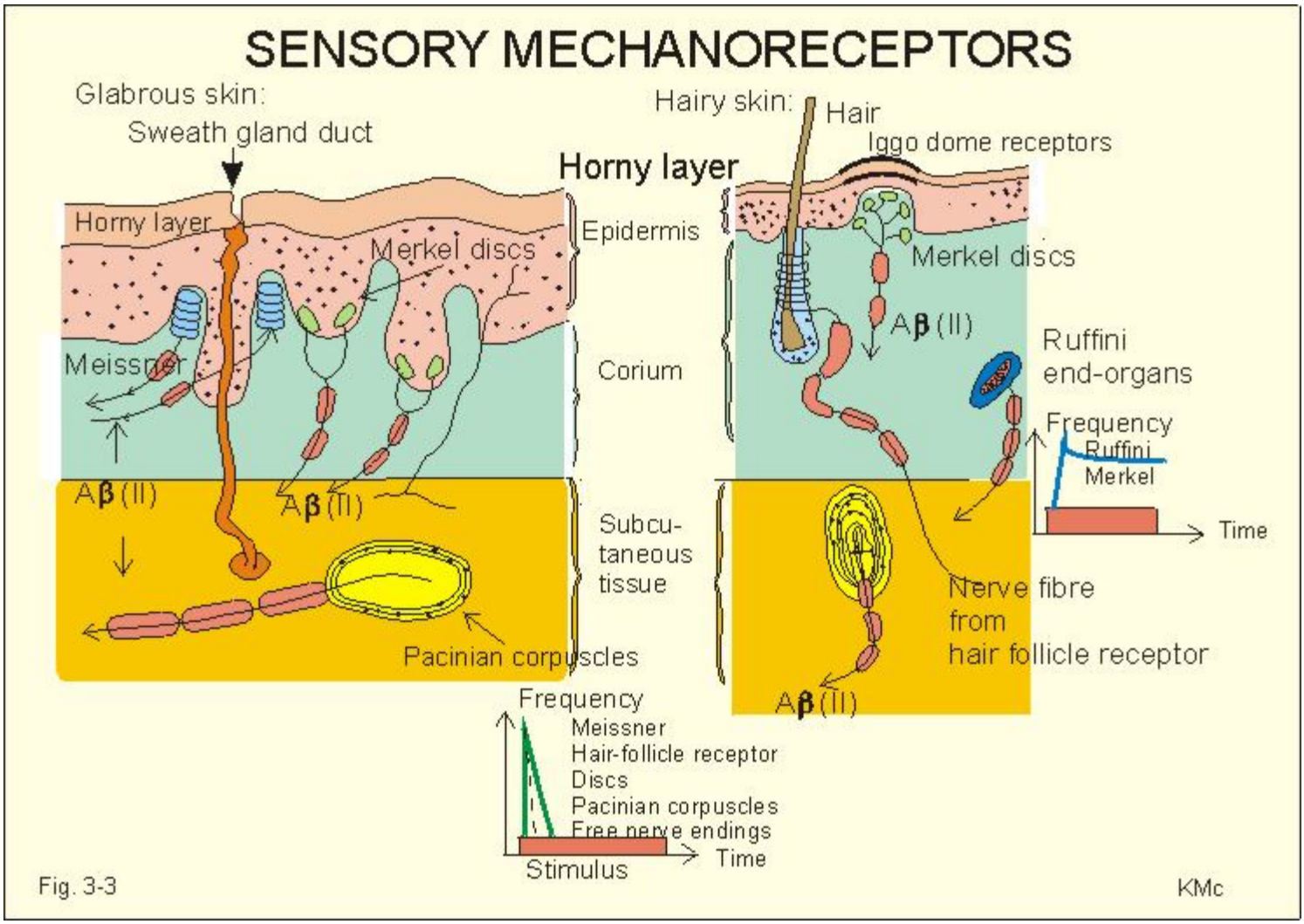

THE POWER LAW: ISS = k * SS'
Arbitrary units
Interpreted stimulus strength (ISS)
n < 1:
Hearing, smell
100
n = 1
n >1:
Taste, heat, angular accelleration
Stimulus strength (SS)
100
Arbitrary units

Fig. 3-2

KMc


SENSORY MECHANORECEPTORS
Glabrous skin:
Sweath gland duct
Horny layer
Epidermis
Merkel discs
Meissner
Corium
Aβ (II)
Aβ (II)
Subcu-
taneous
tissue
Pacinian corpuscles
Hairy skin:
Hair
Iggo dome receptors
Horny layer
Merkel discs
Aβ (II)
Ruffini
end-organs
Frequency
Ruffini
Merkel
Time
Nerve fibre
from
hair follicle receptor
Aβ (II)
Frequency
Meissner
Hair-follicle receptor
Discs
Pacinian corpuscles
Free nerve endings
Stimulus
Time
Fig. 3-3
KMc

ADAPTATION CURVES FOR SENSORY RECEPTORS

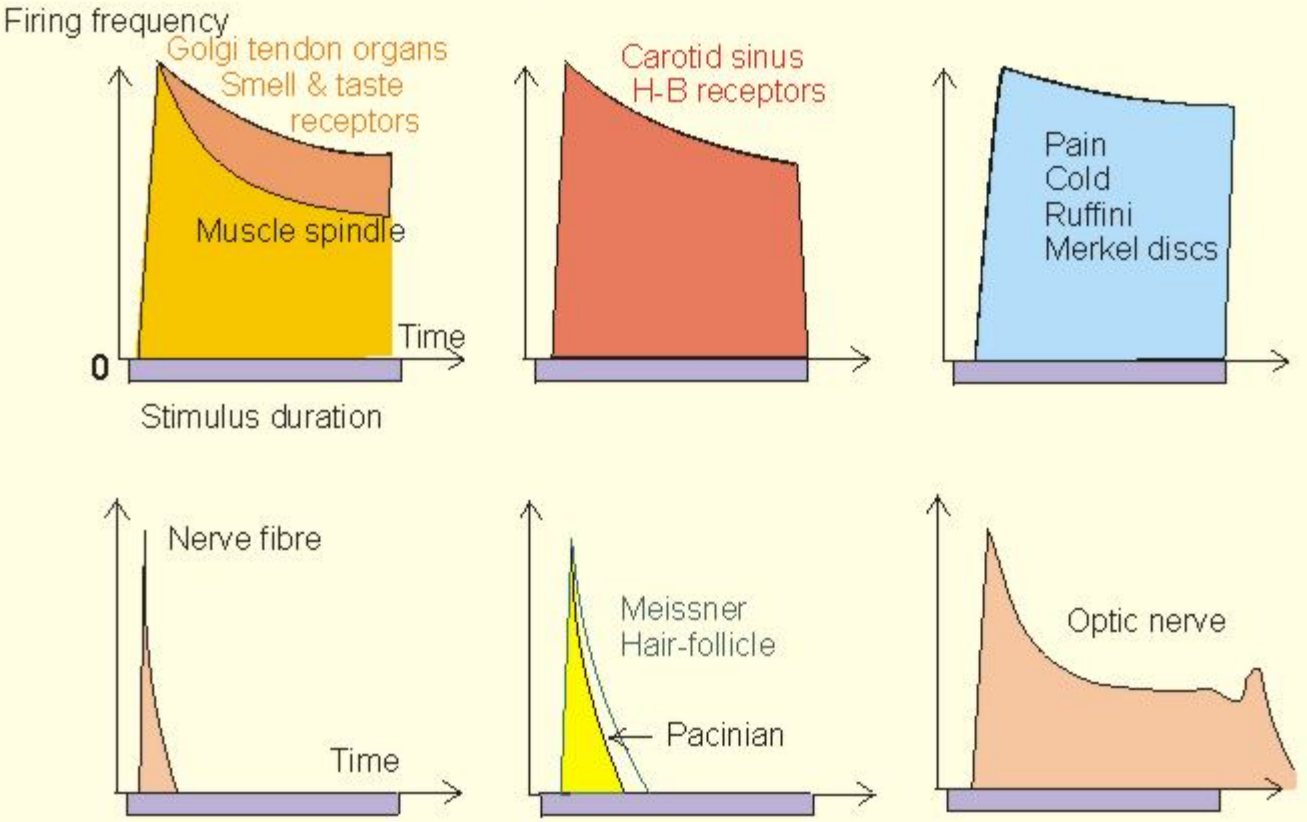


Fig. 3-4

KMc

NOCICEPTORS CAUSING HYPER ALGESIA

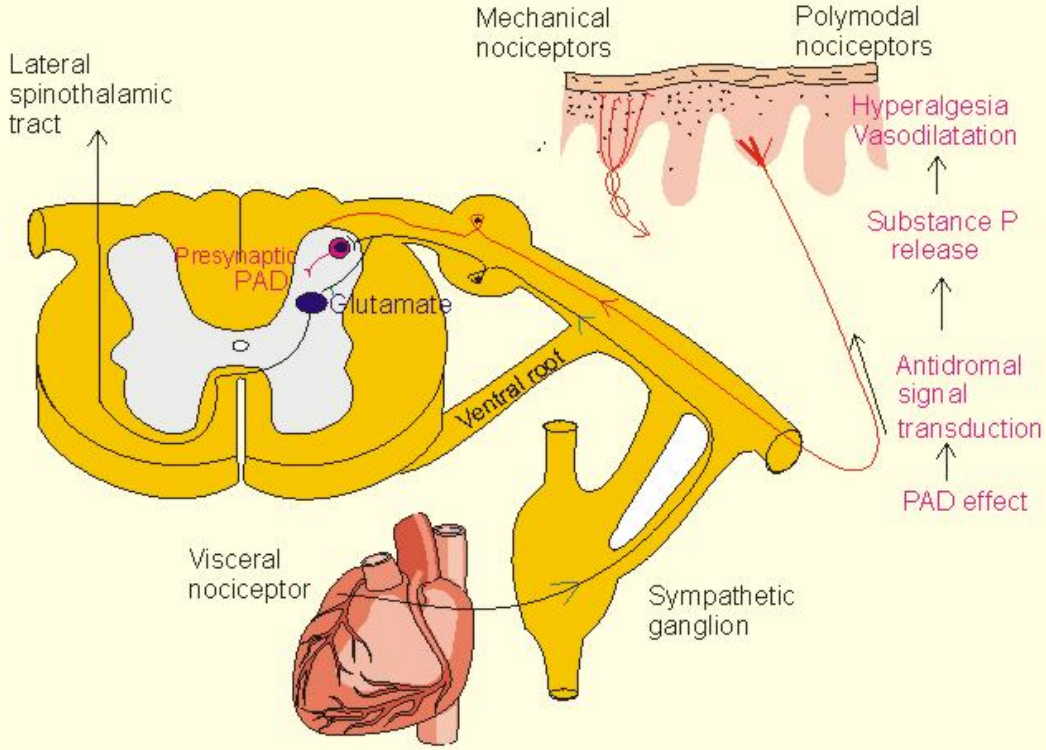


Fig. 3-5

KMc

THE BLOOD-BRAIN & BLOOD-CSF BARRIERS

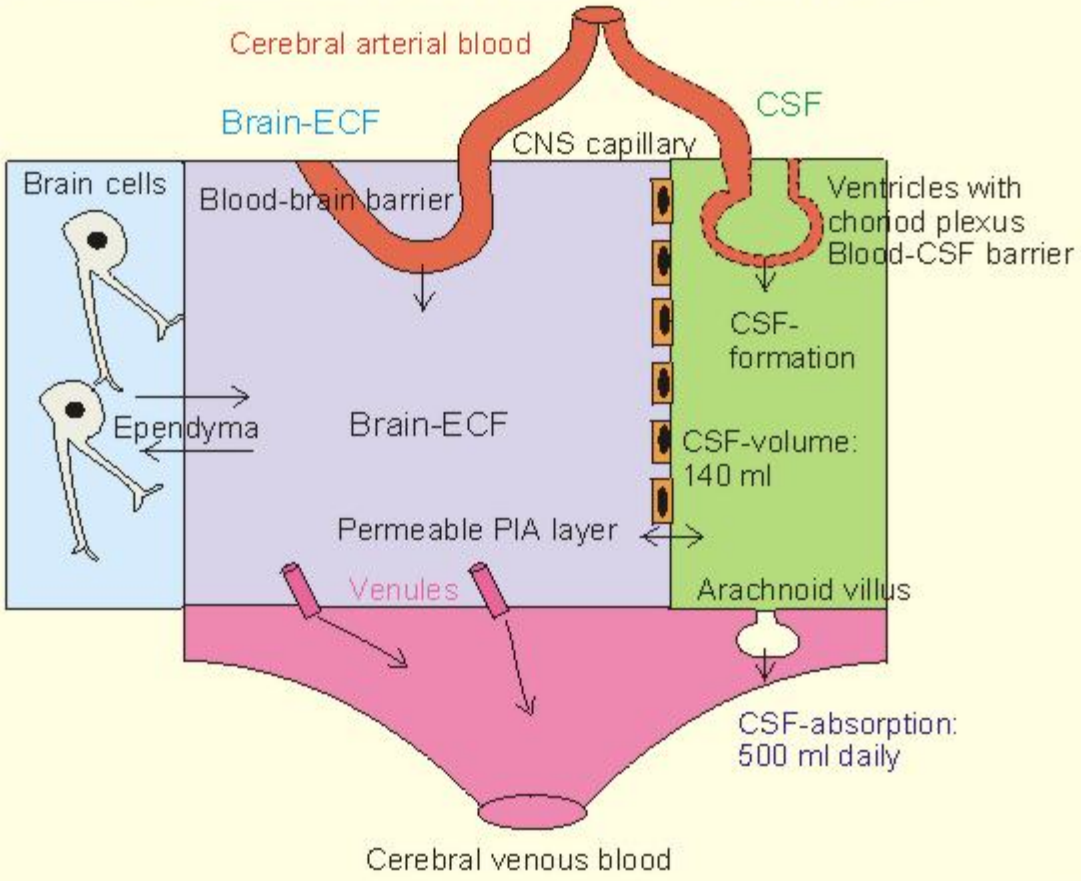
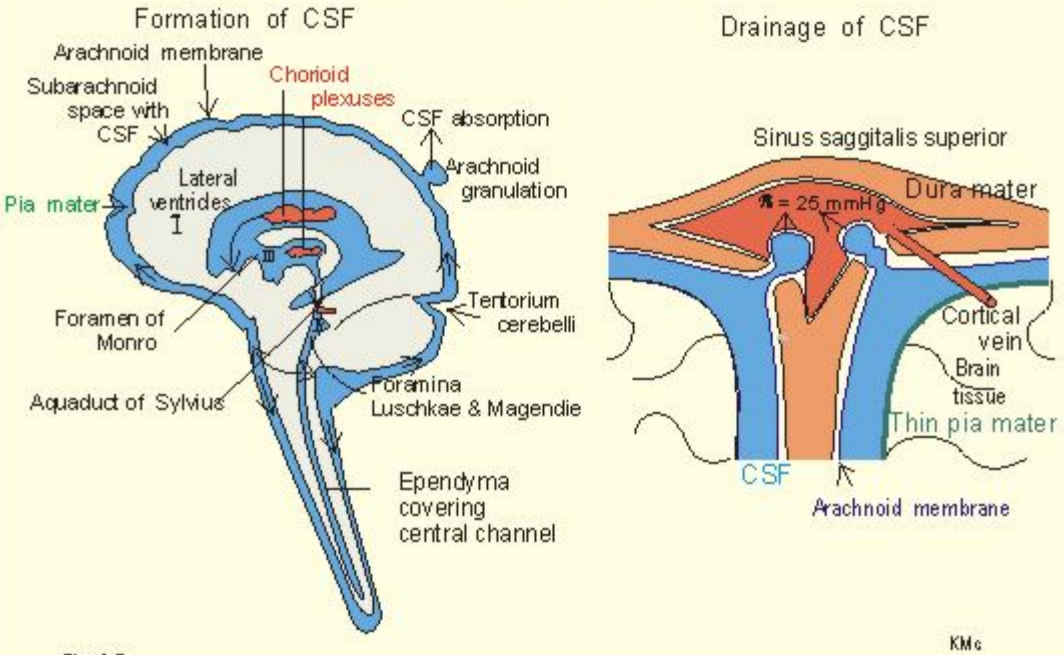


Fig. 3-6

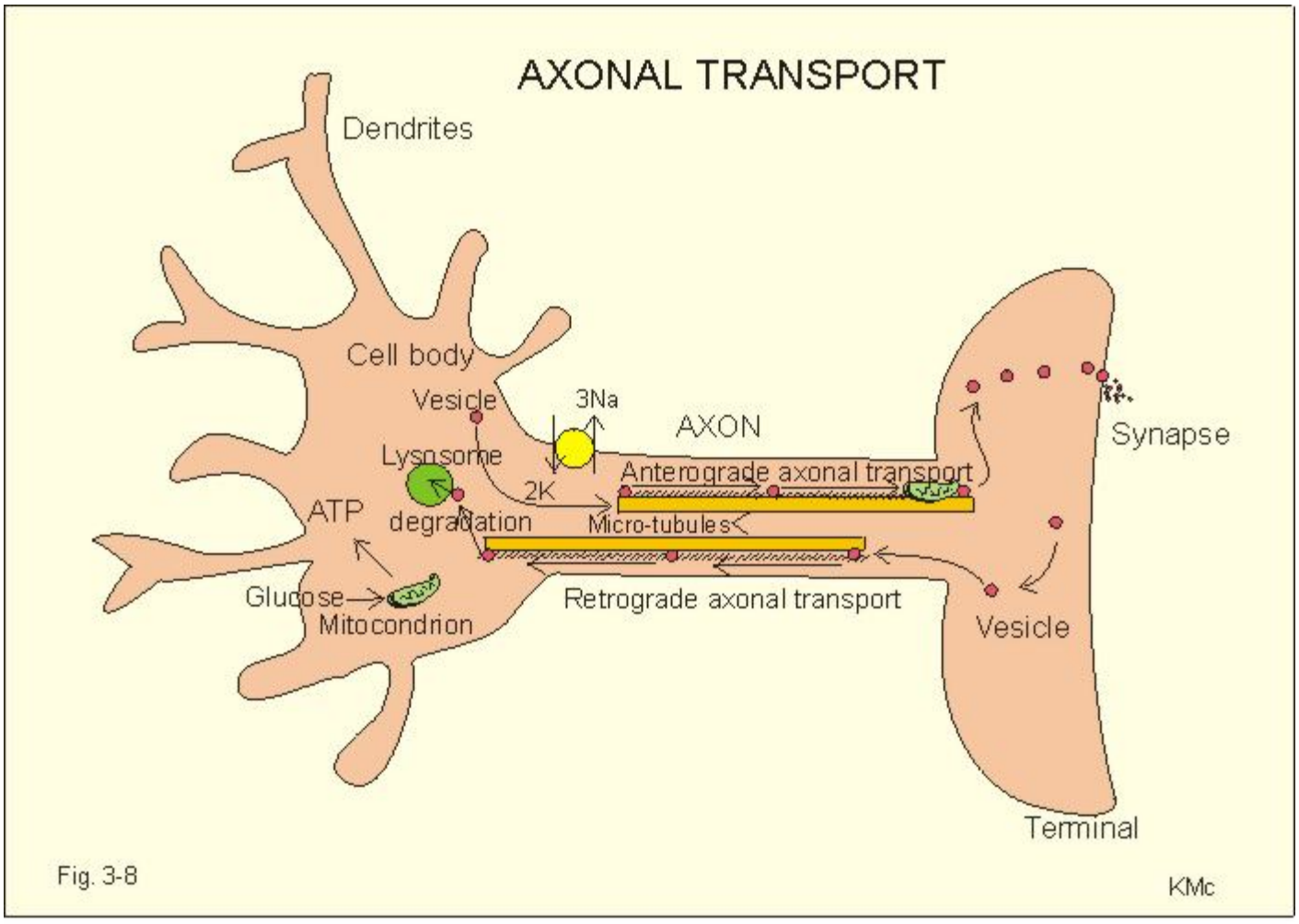
KMc

CSF FORMATION AND ABSORPTION



KMc

Fig. 3-7


AXONAL TRANSPORT
Dendrites
Cell body
Vesicle
3Na
AXON
Synapse
Lysosome
2K
Anterograde axonal transport
ATP
degradation
Micro-tubules
Glucose
Mitocondrion
Retrograde axonal transport
Vesicle
Terminal
Fig. 3-8
KMc

PAIN TRANSMISSION IN SPINOTHALAMIC TRACT

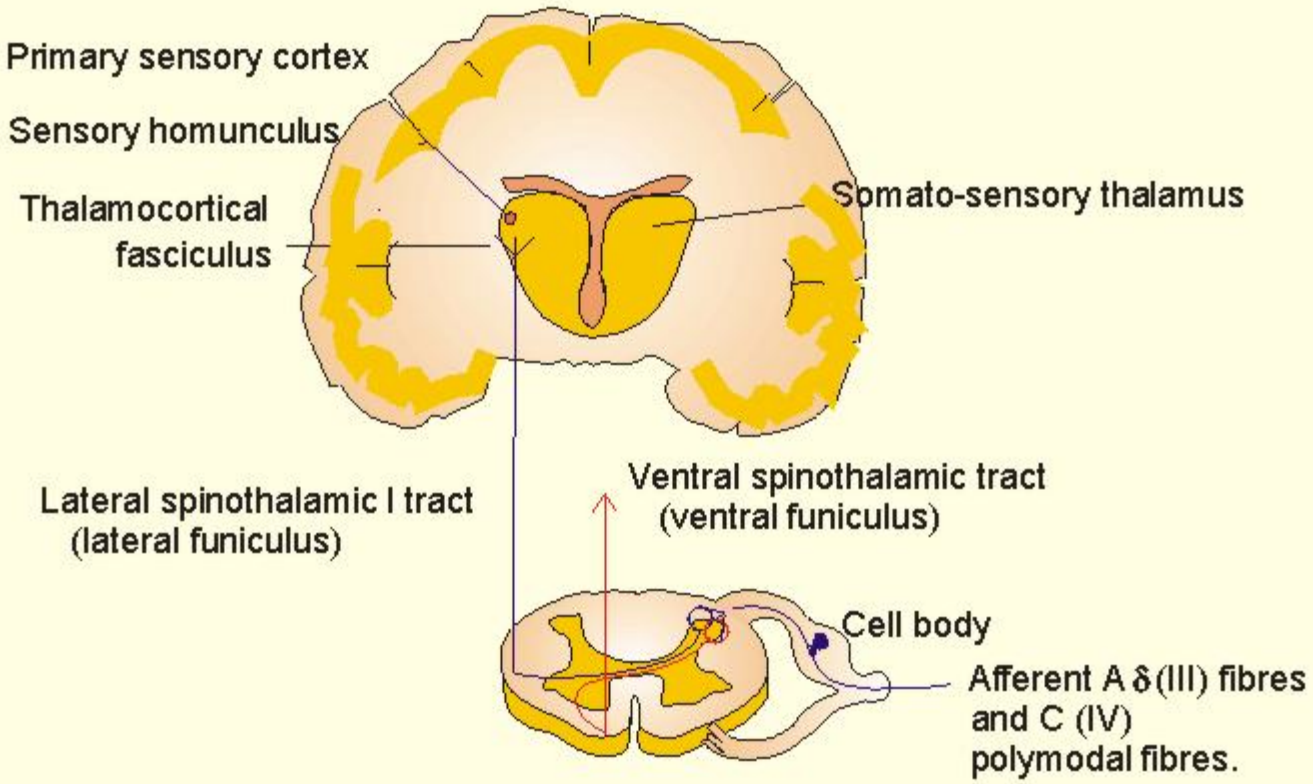
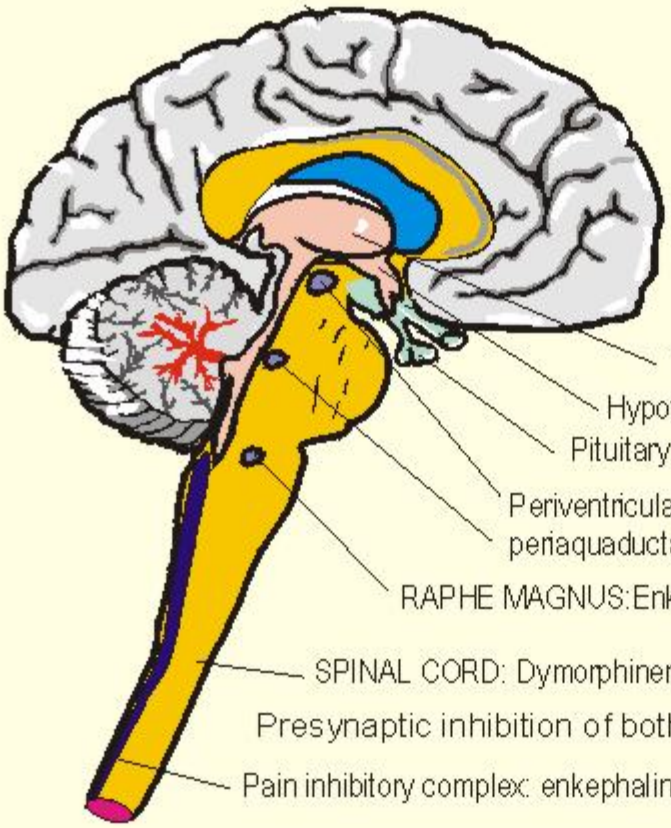


Fig. 3-9

KMc

OPIATE RECEPTORS IN THE CNS



Thalamus

Hypothalamus

Pituitary glands

} β - endorphin (μ-& δ)

Periventricular nuclei

periaquaductal grey

} Morphine Dynorphin (μ)

RAPHE MAGNUS:Enkephalin (δ)

SPINAL CORD: Dymorphinergic neuron with κ receptors

Presynaptic inhibition of both III and IV fibres by enkephalins

Pain inhibitory complex: enkephalin (δ)

Fig. 3-10

KMc

TASTE BUDS AND NEURAL TASTE PATHWAYS

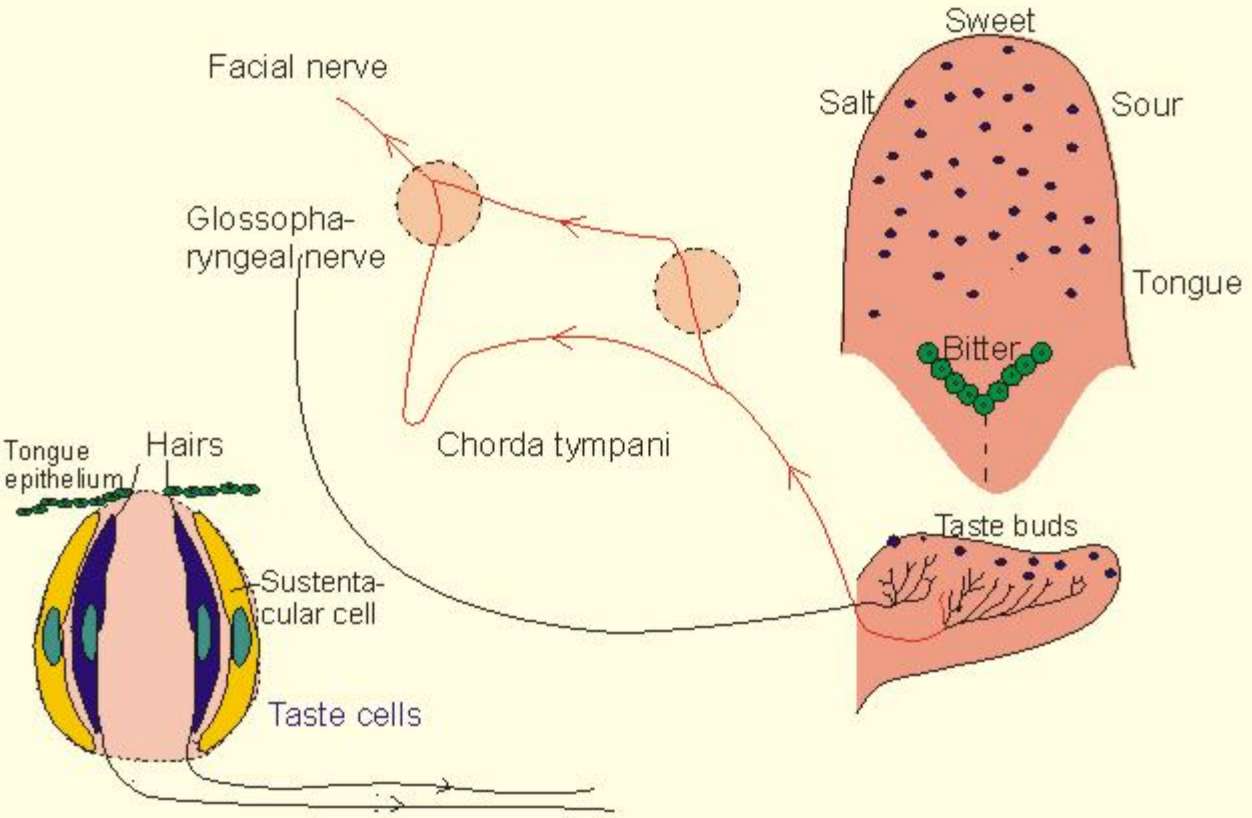


Fig. 3-11

KMc

PATHWAYS FOR SMELL SENSATION

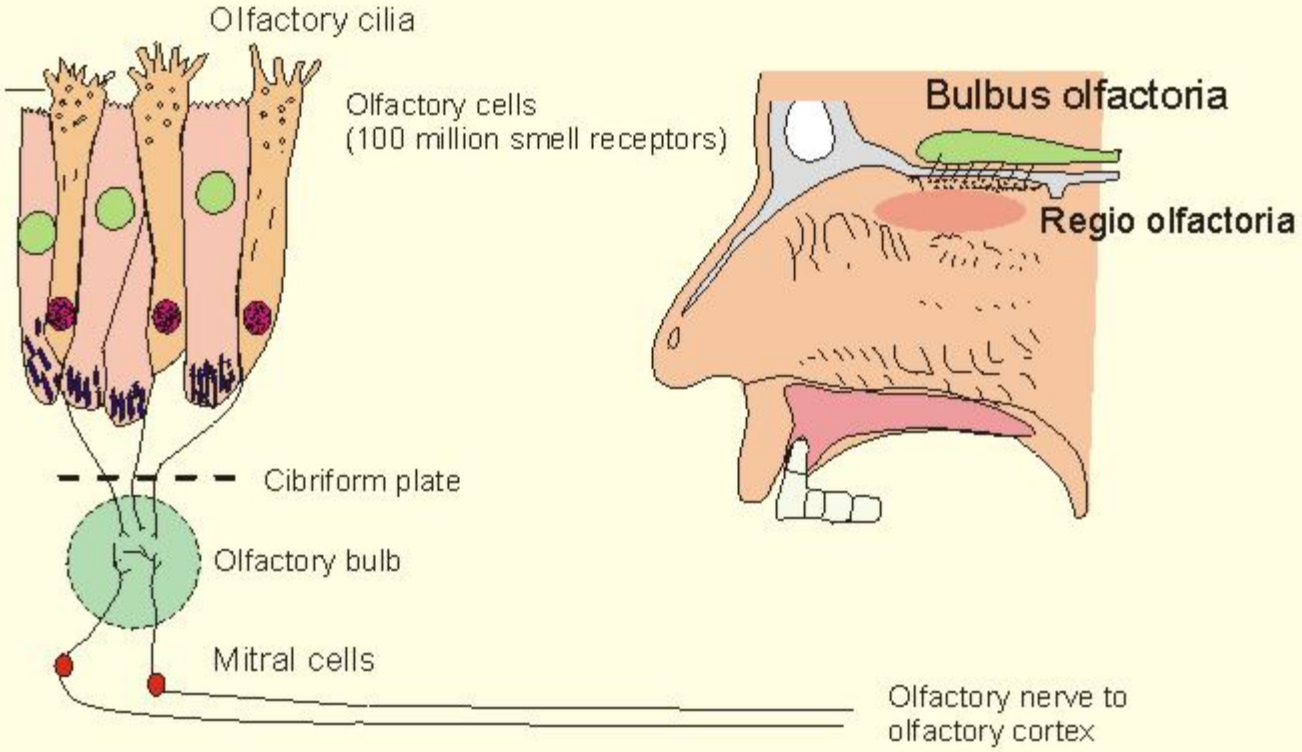
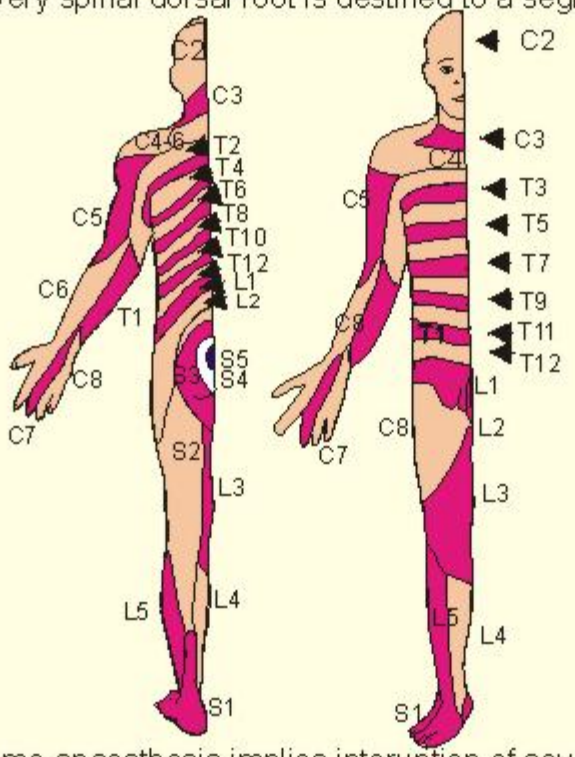


Fig. 3-12

KMc

DERMATOMES:

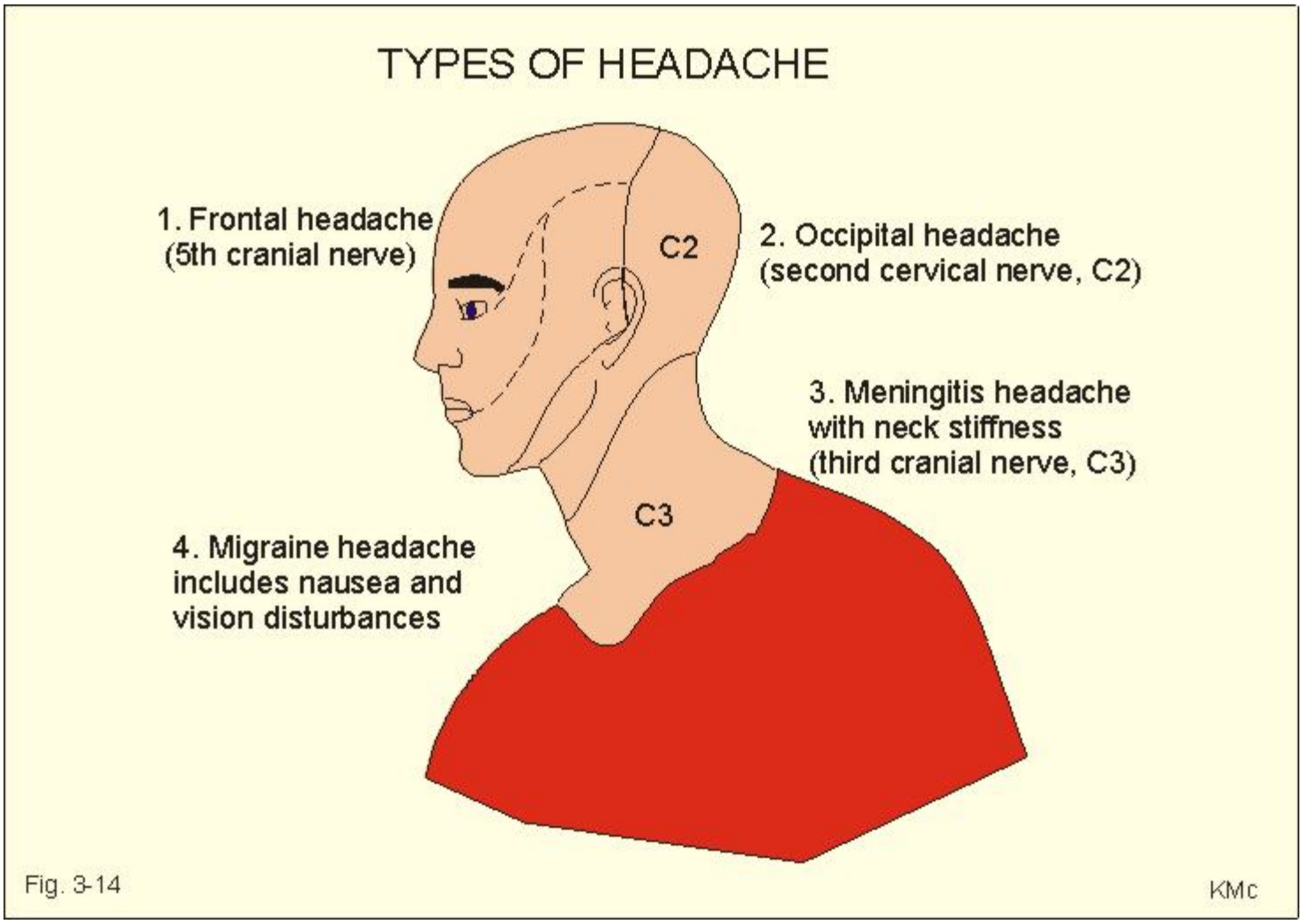
Every spinal dorsal root is destined to a segment of the skin



Dermatome anaesthesia implies interuption of several dorsal roots

Fig. 3-13

KMc


TYPES OF HEADACHE
1. Frontal headache
(5th cranial nerve)
C2
2. Occipital headache
(second cervical nerve, C2)
3. Meningitis headache
with neck stiffness
(third cranial nerve, C3)
C3
4. Migraine headache
includes nausea and
vision disturbances
Fig. 3-14
KMc

Chapter 3

Answers Multiple Choice Questions

Answers **A, C, D,** and E are true statements, whereas **B** is false.

Answers **A, C,** and **D** are true statements, whereas **B** and **E** is false.

Case History A

*The lesion of this patient must be localised in the **lateral spinothalamic tract** of the spinal cord on the right side. The precise location is the thoracic segment, which has the upper affected skin area as dermatome. This is probably **the 10th thoracic segment to the navel***

Return to chapter 3

Return to content

Chapter 5.
Special Senses And Disorders

Study Objectives

- To *define* adequate stimuli for vision, hearing and balance. To define astigmatism, cataract, glaucoma, hypermetropia, myopia, presbyopia, colour blindness, hemianopsia, scotoma, strabismus, and visual acuity and agnosia. To define nerve deafness, conduction deafness, presbyacusis, nystagmus, and transportation sickness.
- To *describe* the anatomy of the eye, including the retina and sensory pathways and cortical visual organisation. To describe the inner ear including hair cells, and the vestibular system. To *describe* the eye movements, receptive fields, the colour vision, the sound transfer, the mechanical-electrical transduction in hair cells.
- To *calculate* the correction of refractive disorders, and the hearing loss from relevant variables given.
- To *explain* the function of photoreceptors, dark adaptation, the pupillary reflex, the sensitivity to sounds, the travelling wave and the frequency theory. To explain the diagnosis and pathophysiology of the following disorders: astigmatism, hypermetropia, myopia, presbyopia, colour blindness, hemianopsia, visual agnosia, strabismus, nerve deafness, conduction deafness, presbyacusis, nystagmus, and transportation sickness.
- To *use* the concepts in problem solving and case histories.

Principles

- *The human eye as light perceives electromagnetic radiation at wavelengths between 400 and 725 nm. Electromagnetic radiation ("waves") do not require a medium for propagation. This spectrum of wavelengths is seen in a rainbow. Light reflected from a star traverses the empty space. Electromagnetic radiation in any narrow band within this spectrum is termed monochromatic light.*
- *The camera obscura is a simple model of the eye. The camera obscura is a device in which a small aperture creates a reversed image on the receptive surface. The aperture can be extended if a convex lens is mounted in front of it. The image then produced on the receptive surface (retina) is reversed and reduced (see textbooks in physics).*
- *A mechanical wave is a wave that propagates by disturbing the particles of a medium. Sound waves are such mechanical or matter waves. The laws of quantum mechanics determine their behaviour.*

Definitions

- **Anopsia** is a visual field defect reaching the periphery of the field (see glaucoma).
- **Astigmatism** is a refractive disorder of the eye in which the curvatures of the cornea or lens are different along different meridians. The different meridians therefore have different focal distances.
- **Cataract** is an eye disease, where the vision is blurred by an opaque lens.
- **Colour blindness** is a group of recessive inherited sex-linked (X chromosome) disorders. Colour blindness is inherited from the father - with the daughter as a carrier - to her son.
- **Conduction deafness** is caused by impairment of the mechanical conduction of sound into the cochlea.
- **Far-point** (F or *punctum remotum*) for the eye is the fixation point in the un-accommodated eye.
- **Glaucoma** is a term used for eye disorders with loss of optic nerve fibres and of the visual field. Frequently, the cause is an increased intraocular pressure (above 25 mmHg) due to reduced fluid outflow at the irido-corneal junction.
- **Hypermetropia** or *far-sightedness* is a refractive disorder with insufficient refractive power, whereby the far point is always located behind the eye.
- **Hemianopsia** means loss of vision in half of the visual field of both eyes. The loss of vision refers to the visual filed, and thus to the contralateral half of each retina.
- **Myopia** or *near-sighted* is a refractive disorder, where the patient can only foveate diverging light waves (both from the near- and the far point). The patients usually have elongated eyeballs.
- **Near-point** (N or *punctum proximum*) of the eye is the fixation point for the maximally accommodated eye (ie when the lens is in its most spherical configuration).

- **Nerve deafness** is caused by damage of the cochlea, the auditory nerve or nucleus. There is hearing loss by both air and bone conduction.
- **Nystagmus** is a disorder with abnormal involuntary movements of the eyeballs.
- **Presbyacusis** refers to the decline with age in the capacity of hearing high tones.
- **Presbyopia** is called old mans sight, because the lens loses its elasticity and hence its ability to assume a spherical shape, so the patient cannot accommodate for near vision.
- **Receptive field** is an area of the visual field from which light is perceived through a certain ganglion cell in the retina.
- **Scotoma** or *localised blindness* is an island-formed visual defect caused by a lesion of the retina in one eye, or by partial interruption of the optic nerve (see glaucoma). The visual field defect may be absolute or relative.
- **Strabismus** *(squint)* or *cross-eyed* is an eye disease, where the visual axes of the two eyes do not converge on the fixation point of the object simultaneously.
- **Transportation sickness** or *kinetosis* is a disorder with vertigo, nausea and vomiting due to rapid changes in the direction of motion.
- **Visual accommodation** is the rise in refractive power of the lens, obtained as the lens rounds up because of contraction of the ciliary muscle and relaxation of zonule fibres.
- **Visual acuity** is the resolution capacity of the eye. Cones have a high-resolution capacity and hence a high visual acuity, because the light is focused on the fovea, where the cones are concentrated.
- **Visual agnosia** or mental blindness, is lack of the ability to combine the seen object into a concept.

Essentials

This paragraph deals with *1. The visual system, and 2. The auditory and vestibular system.*

1. The Visual System

This system detects, transmits and interprets photic stimuli. Photic stimuli are electromagnetic waves with wavelengths between 400 and 725 nm. This is *visible light* or the adequate (effective) stimulus for the eye.

The eyes can distinguish brightness and colour. The photoreceptors are rods and cones located in a specialised epithelium called the retina. In each eye the *retina* contains about 6 million cones and 120 million rods. In the peripheral region of the retina both rods and cones converge on *bipolar cells*. The bipolar cells converge on ganglion cells giving rise to the one million nerve fibres in each optic nerve. In addition, there are horizontal cells and amacrine cells in the retina. They conduct impulses laterally.

Rods and cones

Rods are most sensitive in the dark (*scotopic vision*). More than hundred rods converge on each ganglion cell. There are no rods at all in the fovea.

Cones operate best in light (*photopic vision*). Cones have a *high-resolution capacity* and hence a *high visual acuity*, because the light is focused on the fovea, where the cones are concentrated. The high resolution is also due to the small *convergence* of cones to bipolar cells in the fovea (approximately a 1:1 relationship). Cones are responsible for colour vision. Cones are surrounded by pigment, except where the light enters.

The eye contains chamber fluid, which is produced by filtration and secretion in the ciliary processes. The intraocular pressure is normally 1.3-2.6 kPa (10-20 mmHg). Increased resistance to fluid outflow at the iridocorneal junction leads to *increased intraocular pressure* with loss of optic nerve fibres- or *glaucoma*. In this condition, the retinal artery is compressed at the optic disc, where it enters the eye. This causes retinal and optic nerve atrophy which eventually results in blindness.

A *diopter* is the unit for the refractive power of a lens. The *diopter* (D) equals the reciprocal value of the focal length of the lens in metre (m).

Visual accommodation is the *rise in the refractive power of the lens*, obtained as the lens rounds up, because of contraction of the ciliary muscle and relaxation of the zonule fibres.

Each object we look at has a special target point (the *fixation point*), from which light passes un-refracted through the nodal point of the eye and focuses on the fovea, creating the sharpest possible image. The nodal point in the eye is precisely the point through which a light beam passes un-refracted. The *far point* (F or punctum remotum) for the eye is the fixation point in the un-accomodated eye (Fig. 5-1). The *near point* (N or punctum proximum) of the eye is the

fixation point for a maximally accommodated eye (ie, when the lens is in its most spherical configuration). The refractive power of the lens can vary between 12 and 26 D. The *accommodative power of the eye* is the rise in refractive power from the un-accomodated to the maximally accommodated condition (see Eq. 5-1). A child of 10 years has 12-14 D, a 20 year old person 10 D, and a 60 year old person only 1 D in accommodative power.

The optical distance convention defines all distances measured from a light source to the eye, to be positive. Thus, all distances from the eye to the light source are negative. Hence, the distance from the *nodal point* of the eye to a point in front of the eye is negative. Convex refractive media bend (convergence) in-falling light behind the media and thus have a positive diopter. Concave lenses have refractive powers with negative Diopters, because the focal point is in front of the lens.

Convergence or near vision occurs when the eye focuses on an object closer than 6 m from the eye. Near vision - even with only one eye - triggers *accommodation* and *pupillary constriction*. The ciliary muscle and the pupillary sphincter muscle are innervated by the parasympathetic oculomotor nerve, and the two muscles contract simultaneously for near vision.

The visual fields of both eyes are perceived as only one continuous visual fields (the Cyclops eye effect). This is *fusion* or the illusion that we are looking at the world with only one eye.

In a healthy eye, the light from an object in the visual scenario is focused sharply on the retina by the cornea and the lens. Both of these refract (bend) light. The cornea has a refractive power of *43 D*, and the healthy lens has a refractive power that varies between *12-26 D*. Thus the total refractive power is *56-69 D*. The lens allows the eye to accommodate, so that both near and distant objects can be focused on the retina and thus clearly seen. When we look at distant objects with normal eyes and relaxed ciliary muscles, the object foveates automatically. However, when we look at nearby objects, the light is initially focused behind the retina. The lens then rounds up, by contraction of the ciliary muscles and relaxation of the zonule fibres (i.e., accommodation), to focus the image on the fovea.

The normotropic eye has the ideal refractive power. *Parallel light* from the far point (F in the upper part of Fig. 5-1) *foveates* on the retina in the un-accomodated eye. Light from the near point (N in Fig. 5-1) in the totally accommodated eye also foveates.

Fig. 5-1: Hypothetical light rays for emmetropic, myopic, facultative and absolute hypermetropic eyes.

The coloured space in front of each eye is the fraction of the three dimensional space, which can be focused on the retina for a given visual axis (Fig. 5-1).

Eye movements

Conjugate movements are movements of both eyes in the same direction and magnitude, so that the relation between the visual axes is maintained. When focusing on far away objects, the parallel axes are maintained during conjugate movements. Likewise, conjugate eye movements maintain the convergence angles of the eye required for focusing on nearby objects.

Saccadian or *jumping movements* are rapid eye movements. Saccadian eye movement is an instantaneous reposition of the eye that occurs when reading or when focusing on a flash of light in the peripheral visual field. The velocity of the movement is up to 500^{o} per s. The latency period is 250 ms, and the contraction time is 50 ms. The compensatory eye movement involving the vestibular system, occurs when the head rotates. This is also an example of Saccadian eye movement.

In contrast, *pursuit movements* are smooth eye movements that allow the eye to track a moving object. They have a velocity of up to 30^{o} per s.

These two movements work together in *optokinetic nystagmus*. This is a shift between smooth pursuit movements and correcting jumps. The direction of nystagmus is by convention indicated by the rapid correcting phase.

Even during foveation of an object the eyes are not totally still. The eyes are continuously performing *miniature eye movements*, which occur at a rate of 3 microsaccades per s, with mean amplitude of 0.1^{o}.

Photoreception

The number of photoreceptors in a human eye is estimated to be 110-130 million rods and 5-7 million cones.

Each *photoreceptor cell* includes an outer and an inner segment, which are united by a thin cilium. The outer segments are directed towards the pigment epithelium of the peripheral retina, and contain stacks of *disks* that are rich in photo-pigment molecules. The inner segments contain the cell nucleus and numerous mitochondria. The *rods* are

predominant outside the fovea, and they contain much more pigment (10 rhodopsin or molecules per rod) than do cones. Rods are so sensitive that a single photon can trigger a rod response. Rods are therefore well suited for night vision. Rhodopsin or *visual purple* has two absorption maxims: 350 and 500 nm. The spectral extinction curve for rods corresponds to that of rhodopsin, suggesting that rhodopsin is the chemopigment in rods. Rhodopsin consists of a glycoprotein (opsin) and a chromophore group (11-cis-retinal). Retinal is the aldehyde of vitamin A_1 (retinol).

The fovea only contains *cones*. Cones function in the daytime with maximal visual acuity and colour vision. The human eye possesses *three* types of cones, each with a specific pigment related to the three basic colours: red (erythrolab), green (chlorolab) and blue (cyanolab). The cones in the fovea do not contain cyanolab.

When the human eye is fully adapted to darkness, its rods have open Na^+-channels, and the resulting influx of Na^+ maintains depolarised rods with a resting membrane potential of *-40 mV*. The rod cell synapses with bipolar and horizontal cells, and releases *glutamate* as long as the dark depolarisation is maintained. Na^+ is continuously removed from the rod by the Na^+-K^+ pump.

Inside the rod a special amplification takes place. Light absorption by a *single rhodopsin molecule* activates thousands of G-protein molecules (transducin), which then activate large quantities of cGMP phosphodiesterase in the discs. Each of these enzyme molecules catalyses the hydrolysis of cGMP to 5'-GMP at a rate of thousands per second. The reduction in [cGMP] closes the Na^+-channels, and hyperpolarises the cell. The amplification mechanism is probably why the eye is capable of detecting a single photon.

A similar cascade of reactions takes place in cones, when they are stimulated. Cones are so small that the hyperpolarization occurs rapidly.

Each ganglion cell has a receptive field in the retina that is comprised of a number of photoreceptors.

The fraction of a receptive field belonging to each photoreceptor is added to neighbour areas in order to obtain the receptive area of a bipolar or a horizontal cell. An on-bipolar cell is depolarised by white light, whereas an off-bipolar cell is hyperpolarised. Signals are transmitted from the photoreceptors to the ganglion cells as a graded response. These small receptive areas are summated to form a circular receptive field for each ganglion cell (Fig. 5-2). Ganglion cells can generate action potentials and transmit signals to the brain.

1. One type of ganglion cell has a centrally located excitatory area, surrounded by an inhibitory annular area (Fig. 5-2). Together these form an on-centre off-surround receptive field. Here, an on-response is triggered in the bipolar cell that is connected to the on-ganglion cell.
2. Another type of ganglion cell has a centrally located inhibitory area (inhibited by light), surrounded by an excitatory annular area (Fig. 5-2). These form an off-centre on-surrounding receptive field.
3. A third type of ganglion cell is connected to both on- and off-bipolar cells, so its centre is both stimulated and inhibited by white light.

Fig. 5-2: Ganglion cell receptor fields in the retina.

The ganglion cells can also produce *transient* or *sustained* reactions. These reactions are due to adaptation to light (decreased sensitivity with exposure) and lack of adaptation, respectively (Fig. 5-2).

Ganglion cells in the fovea are connected to few or only one cone. Some ganglion cells are excited by blue light and inhibited by its opponent colour yellow. Other cells are excited by green and inhibited by the opponent colour red. This mechanism is the so-called *colour contrast analysis* of the retinal ganglion cells. Colour opponent neurons are found not only in the ganglion cells but also in the lateral geniculate nuclei.

Retinal signals pass through the main visual pathway: the optic nerve, the lateral geniculate nucleus, the optic radiations (the geniculostriate tracts), the primary visual cortex, the pretectal area, the Edinger-Westphal nucleus, the oculomotor nerve and the ciliary muscle.

Each point of the retina has a corresponding location in the dorsal lateral geniculate nucleus and in the visual, striate cortex (area 17). The nerve fibres in the optic nerve run so the upper quadrants of the retina are represented in the upper half of the nerve, and the lower quadrants in the lower half.

Such a *retinotopic map* is present in the *lateral geniculate nucleus* and maintained throughout the visual pathways and in the visual cortex. The receptive field in the retina is maintained all the way to the cortex. This is the basis of fusion. The consequence is that the *right striate cortex* receives information about objects located in the left side of the visual

field, and the striate cortex in the left hemisphere receives information about the right side of the field of vision. In general, each hemisphere of the brain is connected to sensory and motor activity of the opposite side.

The lateral geniculate nucleus has three different pairs of neuronal layers (1-2, 3-4, 5-6). Ganglion cells from the ipsilateral (same side) eye projects to layers 2, 4 and 6, whereas ganglion cells from the contralateral eye projects to layers 1, 3, 5. The lateral geniculate nucleus is involved in integration and registration of pictures formed in corresponding areas of the retinal surfaces. Some neurons react to white light (with circular receptor fields), while other neurons react to opponent colours. When we jump from one highlight to another in the visual field, each jump is called a saccade. Selection of visual stimuli may be located in the lateral geniculate neurons (possibly performing gate control).

Most of the neurons in the geniculate nucleus projects to the striate cortex by way of the optic radiations (geniculostriate tract). Neurons in a certain column of the lateral geniculate nucleus project to precisely the same part of the striate cortex (area 17). The lateral geniculate nucleus also receives information from the cortex (in particular the visual cortex) that is essential for selection of signals of particular interest.

The striate cortex (area striata, area 17) is located around the calcarine fissure on the medial side of each occipital lobe. The optic radiation ends mainly in synaptic contact with simple cells in layer 4 of the striate cortex. Simple cells have on- and off-fields. Complex cells receive inputs from several simple cells, and hypercomplex cells receive inputs from several complex cells.

Axons from one eye terminate in millions of functional units called ocular dominance columns consisting of about 10^3 neurons. Cortical neurons are arranged in orientation columns showing orientation selectivity for lines edges or bars. Other cortical neurons are arranged in direction columns showing direction selectivity. Colour blobs are interspersed among the other columns (see later).

A large area at the occipital pole represents the macula, and the upper and lower half of the visual field is represented below and above the calcarine fissure. The upper layers of the superior colliculus perform visual processing. The deep layers produce eye movements.

The cortical area V4 contains colour-sensitive neurons, and the visual association areas 18 & 19 (Fig. 4-4) contain many cells with complex functions.

The absolute sensitivity depends upon the adaptive condition of the retina, the pupillary diameter, and the source of light (spectral composition, exposure time, and light source dimensions). The threshold for the completely dark-adapted eye is (7×10^{-11}) Watts/m^2.

The Trichromatic Theory

Light adaptation is a decrease in visual sensitivity during constant stimulation. This occurs rapidly because the rhodopsin bleaches readily. Hence, in daylight (photopic cone vision) we are dependent on cones for vision. Night vision (scotopic rod vision) is extremely sensitive to light, because of dark adaptation. It takes at least 20 min in dark surroundings before the rods become fully adapted. In a dark movie theatre, we have scotopic vision with low visual acuity and colour blindness. As soon as the film is projected we experience partial light adaptation, so that the photopic cone vision is resumed.

The trichromacy theory postulates that an appropriate mixture of the three basic colours can produce any colour: red, green and blue. The three types of cone pigments have different opsins, and opsins that differ from that in rhodopsin. Groups of cortical neurons called cortical *colour blobs* respond specifically to colour signals, and also receive signals from adjacent columns of the visual cortex. Cortical *colour blobs* are probably the primary stations for perception of colour, and they are found both in the primary and the secondary visual cortex areas. Perception of spectral opponent colour pairs is located in discrete colour blobs of the visual cortex.

The three cone pigments are Erythrolab for red (maximal sensitivity at 555 nm), chlorolab for green (525 nm), and cyanolab for blue (450 nm). The absorption spectra of the photopigments overlap considerably. The three cone types are uniformly distributed in the retina, except in the fovea. *Fovea has no cyanolab cones* and no rods. This gives the fovea partial physiologic scotoma (ie, no blue vision and no scotopic dark vision). The real *physiologic scotoma* is the *dark spot* corresponding to the *optic papilla*. Inhibition of neighbour ganglion cells from on centre field ganglion cells is called lateral inhibition; it occurs also in the lateral geniculate nucleus or in the visual cortex. Lateral inhibition provides simultaneous contrasts and enhancement. Each colour-contrast neuron is excited by one colour and inhibited by the opponent colour. Opponent colours are red-green, yellow-blue, and green-purple.

Contrast analysis begins already in the retina and is elaborated centrally in the lateral geniculate nucleus, the thalamus and the visual cortex. If there is a multilevel neural system for the analysis of colour mixing, we also need to assume the existence of a neural system for colour brightness, depending upon the intensity of the light. Healthy people are trichromats, because they have all three cone pigments.

Spatial resolution or minimum separabile is the capacity of the eye to see two stimulated retinal areas as separated. In healthy young humans the spatial resolution is about 1/60 degree, depending upon luminosity, exposure time, patterns and opponent colours in the visual scenario. The most important factor limiting this capacity is the cerebral integration.

Temporal, visual resolution is the capacity of the eye to see consecutive light stimuli as separate. Intensity is directly related to duration of perception of light. Contrast further decreases temporal resolution, a flash of light in the dark is perceived for longer than in bright surroundings. Temporal resolution is also determined by the wavelength of light. The eye is maximally sensitive at the absorption maxims of the three cone pigments and rhodopsin.

The positive after-picture is a visual impression lasting longer than the stimulus. It is visible on a dark background following exposure of the eye to intense light. The negative after-picture follows the positive afterpicture as a dark shadow or as the opponent colour. The negative after-picture is due to adaptation of the area in the retina related to the picture.

A flickering source of light liberates successive flashes so rapidly that they fuse, and appear to be continuous. In the darkness of a movie theatre we do not sense the flickering frequency of 24-48 frames each s, or those of a television screen with 50-60 frames per s. With increasing intensity of illumination the critical fusion frequency increases abruptly. This is why young persons can look directly into a neon light and see its flickering character even with 60-100 flashes each s. Accordingly, the cones of the healthy human eye have a critical fusion frequency around 60-100 flashes per s with optimal illumination. The photopic cones are much more sensitive to rapid alterations of light intensity than the rods.

Movements in the visual scenery are depicted as opposite movements on the retina. Convergent inputs from the eyes result in depth perception (ie, stereopsis or stereoscopic vision). Stereopsis depends upon the medial, longitudinal fasciculus and the corpus callosum. These structures co-ordinate the movements of the two eyes. The two eyes are 7-8 cm apart, which causes slight disparities between their retinal images. Disparate receptive fields and thus excitation of specific cells in the secondary visual cortex probably exhibit the perception of depth.

Distance evaluation requires high visual acuity and experience with objects of known size.

Essential for the development of the baby’s brain is human milk proteins and long chain fatty acids in the mother milk. Protein deficiency from birth reduces formation of brain neurons and thus limits brain development including the development of visual capacity. Many vitamins and key proteins have hormonal and transmitter function in the brain, and lack of such substances in the critical growth period just after birth, results in irreversible damage. The action of endogenous nerve growth factor is necessary for the normal functional and anatomical development of the visual system. In the critical period of visual development, which is the first two years of life, the child must be exposed to a multitude of visual stimuli. This is necessary for the development of neurons and key substances that can record future visual stimuli. The ability to fuse the two optic fields is a process that has to be practised. This fact is an important basis for the treatment of cross-eyedness (strabismus).

Cross-eyedness or squint (strabismus) is an eye disease, where the visual axes of the two eyes do not converge on the fixation point of the object simultaneously. Thus the retinal images do not fuse on corresponding areas on the two retinas. Since the fixation line only foveate in one eye, the patient can learn to suppress the other picture in the brain. Hereby, double vision is avoided at the expense of visual acuity.

2. The Auditory And The Vestibular System

The two systems share the labyrinth, and transmit signals to the brain through the 8^{th} cranial nerve. The two systems record fluid movements and use the so-called hair cells as mechanical transducers.

Sounds are sense impressions that consist of complex mixtures of compression and decompression waves that can be broken down to pure tones by Fourier analysis. Pure tones are sinusoidal waves of a specific frequency (cycles per s or Herz = Hz) and amplitude.

Sinusoidal waves can change phases. The normal human ear is sensitive to pure tones with frequencies between 10 and 30 000 Hz, in a young person.

As people age, their capacity to hear high tones declines. This condition is termed presbyacusis.

Sound propagates at 343 m/s in air at 20^{o}C, although each single air molecule only moves a few mm in the direction of propagation. The unit of sound pressure (p) is Pascal (Pa). According to international convention the sound pressure level (SPL) is expressed in decibel (dB) - see Eq. 5-2.

Any rise in the SPL of 10 dB implies a rise in sound pressure by a factor of 3, since the log of 3 is 0.5: 10 dB = 20 log 3 (Eq. 5-2).

Speech has an intensity of 60-65 dB, and sounds that exceed 100 dB can damage the ear. A constant sound stimulation only results in minor adaptation. The human ear has the largest sensitivity around 1000-4000 Hz, the range for normal speech.

The sound pressure waves in air are converted into sound pressure waves in the fluid column within the cochlea. The pressure wave in the air is transmitted via the tympanic membrane and the ossicles (malleus, incus and stapes), to the fluid of the cochlea. The foot plates of the stapes inserts in the oval window, and separates the middle ears from the fluid of the cochlea. The ratio of the effective surface area of the tympanic membrane to that of the oval window is 14:1, and the pressure is increased further by the differing lengths of the lever arms in the chain of ossicles. By this area-pressure amplification, hearing is improved by more than 25 dB.

When the external ear is filled with water during diving, hearing is seriously reduced.

Two muscles are found in the middle ear. They dampen movements of the ossicular chain when the ear is exposed to extremely high pitch sounds that can be anticipated. These muscles are the tensor tympani muscle supplied by the trigeminal nerve, and the stapedius muscle supplied by the facial nerve. Exposure to sounds above 90 dB elicits reflex contractions.

The cochlea is composed of three tube systems coiled together to form a pyramid: scala vestibuli, scala media and scala tympani (Fig. 5-3). The part of cochlea beneath the oval window is called scala vestibuli, and it is filled with a fluid column termed perilymph.

Fig. 5-3: A cross section through one of the turns of the cochlea.

The perilymph conducts the pressure wave to the basilar membrane, which is displaced within the endolymph together with the whole organ of Corti, which contains the hair cells.

Each hair cell has 40-100 hairs (*stereocilia*). The hairs have different heights, and when the pressure wave displaces the hairs towards the tallest hair, the hair cells are depolarised. When the basilar membrane moves upward towards the scala media, the reticular lamina shifts upward and inward (Fig. 5-3), causing the hair cells to depolarise. Downward movement of the basilar membrane towards the scala tympani moves the reticular lamina downward and outward (Fig. 5-3). This movement hyperpolarises the hair cell membrane.

The endolymph in the scala media has a potential difference of +80 mV with the perilymph as reference. The inside of the hair cell is -60 mV compared to the perilymph; this is a resting membrane potential about the same size as in most neurons. Thus the total potential difference between the inside of the hair cell and the endolymph in the scala media is -140 mV. This resting membrane potential is maintained by Na^{+}-K^{+}-pumps in the Stria vascularis (Fig. 5-3).

Bending of the hair change the conductance of K^{+}-ions through the apical hairy membrane, and this is how the resting membrane potential is changed. A current flow is produced through the hair cell from apex to base, which is resting on the basilar membrane (Fig. 5-3). This current flow or receptor potential can be recorded extracellularly with microelectrodes as the cochlear microphone potential (ie, the sum of receptor potentials from many hair cells). This potential has the same frequency as the acoustic stimulus, and the potential is analogous to the output voltage of a microphone. The cochlear microphone potential follows the sound stimulus without latency, without measurable threshold, and without fatigue in contrast to neuronal action potentials.

Stimulated the hair cells release neurotransmitters (glutamate, aspartate) that excite the cochlear nerve fibres. Thus, the propagating action potentials are generated in the cochlear nerve fibres.

A high frequency tone produces travelling waves along the basilar membrane. High tones travel only a short distance from the stapes along the basilar membrane to their resonant point, where the displacement amplitude of the basilar membrane is maximal (Fig. 5-4). Low frequency tones travel all the way to the apex of the cochlea (Fig. 5-4). The higher the tone frequency, the more basal located in the cochlea is the resonant point and its potential.

Fig. 5-4: Displacement of the basilar membrane illustrates the travelling wave theory (von Bekesy).

The existence of such a maximum of the travelling wave is termed *frequency dispersion.* Since different frequencies

excite differently located hair cells the argument is called the place analysis theory. The brain also utilises the temporal structure of the sound stimulus. This is the so-called periodicity analysis.

The receptor potentials generate action potentials in the cochlear nerve (8^{th} cranial nerve) that travel to the cochlear nuclei. Secondary neurons transfer the signals from here to the superior olivary nuclei that co-ordinates the two ears, or directly to the inferior colliculus through the lateral lemniscus (representing both ears). Axons from the inferior colliculus ascend to the medial geniculate nucleus of the thalamus. Axons from this thalamic nucleus form the auditory radiation, which terminates in the auditory cortex in the superior temporal gyrus (areas 41 and 42 in Fig. 4-4). High frequencies are projected to the rostral auditory cortex, and low tones to the caudal section.

The duration of a sound stimulus is encoded in the duration of the neural signal, and its intensity by the level of neural activity.

Projections from the auditory cortex also descend to the medial geniculate nucleus and the inferior colliculus. The oligocochlear bundle controls several sound impressions. Efferent stimulation through these pathways inhibits the sensitivity of these nuclei for sounds, while increasing their tone selectivity. This phenomenon, and a high degree of motivation, explains how a mother can hear her baby cry in spite of noise, and also how we can hear an individual in a crowd (the cocktail party effect).

Localisation of a sound source depends upon the difference in time between the arrival of a low frequency sound signal to the left and right ears (time delay). The sources of low frequency sounds (below 2000 Hz) are localised by this time delay. The source of high frequency sounds is localised by the difference in sound amplitude arriving at each ear caused by the dampening of the sound intensity.

Sounds in the region of 2000 Hz cannot be detected by either mechanism. On average, the distance between the two organ's of Corti is about 0.16 m. Then, a sinusoidal wave or pure tone with exactly the same wavelength coming from one side of the head, is in phase when they reach the ears. This wavelength corresponds to the frequency of 2144 Hz (343 m/s divided by 0.16 m). In this instance the subject will be unable to determine the source of the sound.

The vestibular system detects if the body is in balance. The sensory unit of the auditory-vestibular system is the membranous labyrinth, located in the petrous portion of the bony labyrinth. The membranous labyrinth contains endolymph and is surrounded by perilymph; it is composed of the auditory cochlear duct or scala media, and the balance regulating the Vestibular system. The vestibular system consists of three semicircular ducts and two otolith chambers (the utricle and the saccule). Each semicircular duct has a swelling termed an ampulla (Fig. 5-5).

Fig. 5-5: The spatial orientation of the three semicircular ducts in the upright person (left). The horizontal duct is not drawn. The membranous labyrinth is shown to the right.

The semicircular ducts consist of a horizontal duct, a superior and a posterior duct at right angles to each other, so that they cover all three planes in space. The semicircular ducts all communicate with the utricle. The utricle joins the saccule, which receives new endolymph from the cochlear duct.

The sensory organ of each utricle and saccule is called a macula. The sensory organ of each semicircular ampulla is the crista ampularis.

Each macula contains thousands of hair cells. Vestibular hair cells each have many stereocilia (hairs) on their apical surface just as cochlear hair cells do; however, they also have a large stereocilium called kinocilium. The hairs are imbedded in a gelatinous substance, the otolithic membrane that also contains earstones or otoliths. These otoliths increase the specific gravity of the otolithic membrane to twice that of the endolymph. Thus their hair cells are sensitive to linear acceleration such as gravity and to static equilibrium control, but not to angular accelerations of the head. The macula of the utricle is located in the horizontal plane, and the macula of the saccule in the vertical plane.

Each crista ampularis consists of many hair cells. Here the hairs are imbedded in a large gelatinous substance termed a cupula. The cupula occludes the lumen of the ampulla completely, and its material has the same specific gravity as the endolymph. The cupula is concerned with equilibrium control during motion and with angular acceleration (rotation of the head), but is unaffected by linear acceleration.

When the stereocilia are bent toward the kinocilium, the conductance of the apical cell membrane increases for positive ions, and the hair cell becomes depolarised. Bending the stereocilia in the opposite direction hyperpolarizes the cell. The depolarised hair cell releases glutamate or aspartate and increases the discharge rate of the nerve fibre with which it synapses.

The utricles and saccules are sensitive to linear accelerations. When we suddenly thrust our body forward, the otolithic membranes fall backwards on the cilia of the hair cells until the thrust stops. Then, the otolithic membranes fall

forwards. The signals to the brain make us feel as if we were falling backwards. Therefore, we lean forward until the otolithic membranes are in balance.

Pathophysiology

This paragraph deals with 1. Refractive disorders, 2. Colour blindness, 3. Visual field defects, 4. Mental blindness, 5. Deafness (hypacusis), 6. Nystagmus, and 7. Kinetosis.

1. Refractive Disorders

(myopia, hypermetropia, astigmatism, presbyopia, and cataract).

Near-sighted (*myopic*) patients usually have *elongated eyeballs*. More rarely, myopia can be caused by too high refractive power in the lens system. Myopic persons can only foveate diverging light waves - both from F and N (Fig. 5-1). The images of distant objects are focused in front of the retina, and the image is blurred on the retina. Both F and N are located in front of the eye. Concave lenses (-D) accomplish correction. The weakest concave lens compatible with optimal visual acuity is the best correction, as the accommodation is eliminated.

Hypermetropic or far-sighted persons usually have *shortened eyeballs*, and F is always behind the eye. In rare cases hypermetropia can also be caused by insufficient refractive power in the un-accomodated eye. The *absolute hypermetropic eye* can only focus images of distant objects behind the retina (Fig. 5-1). The *facultative hypermetropic eye* can focus converging light beams on the retina without accommodation (rest in Fig. 5-1). This patient can read the Snellen letters without problems; they also foveate diverging light beams by accommodation, but then the patient gets eyestrain due to fatigued ciliary muscles. Convex lenses (+D) correct hypermetropia. The strongest convex lens compatible with optimal visual acuity is the best correction, as the accommodation is eliminated.

Astigmatism is a refractive disorder of the eye, in which the *curvatures of the cornea or lens are different along different meridians*. The different meridians therefore have different focal distances. Therefore, astigmatism can be corrected with *cylinder lenses* that correct the curvature differences.

Presbyopia is called *old man's sight*. The far point remains where it is, so the un-accomodated refraction is unaltered. The ability to accommodate is changed, so that N approaches F.

With age, the lenses of most people loses its elasticity, and hence their ability to assume spherical shape. The lens is the organ in our body with the highest protein concentration. Alterations of lens proteins probably cause *progressively increasing stiffness of the lens*. The accommodative power decreases from 14 D in a child to less than 2 D at the age of 50. The patient's eye becomes incapable of accommodation for near vision and reading. Convex lenses correct presbyopia.

Cataract is an eye disease, where the *vision is blurred by an opaque lens.* Precipitation of lens proteins can occur in several ways. It is often due to oxidative processes. The lens needs oxygen, but strong sun light or radiation can *oxidise* lens proteins in unprotected eyes. The oxidation is enhanced by hyperbaric oxygen therapy and by high blood [glucose] in diabetics. Oxidants in the food may be the cause in some patients. Antioxidants, such as vitamin A and D, seem to protect against the loss of transparency in long-term studies. Today, excellent surgical techniques are used to eliminate the opaque lens and re-establish normal refraction.

2. Colour blindness

The three colour genes are located on an *X chromosome*. Females have two X-chromosomes, and colour blindness is rare among females. Colour blindness is inherited from the *father* - via the daughter - to her *son*. The trait is recessive and sex linked. The total incidence is about 8% of the male and 0.5% of the female population. *Monochromats* lack all three or two cone pigments, an extremely rare disorder. *Dichromats* lack one of the three cone pigments. Proteus is the first or red component, so *protanopic people* are blind for the red part of the spectrum. They cannot separate red and yellow signals in traffic. *Deuteranopic patients* are blind for the second or green colour, and *tritanopics* are blind for blue - the third basic colour. *Abnormal trichromats* have a reduced amount of one cone pigment: *Protanomalous trichromats* lack erythrolab, *deuteranomalous* (the most frequent type) lack chlorolab, and *tritanomalous* lack cyanolab.

3. Visual field defects

Visual field defects are caused by interruptions of the visual pathways. *Hemianopsia* means loss of vision in half of the visual field of both eyes. The loss of vision refers to the visual field, and thus to the contralateral half of each retina (shown with two colours in Fig. 5-6).

Homonymous hemianopsia means that the same side of the visual field for each eye is defective. Corresponding halves

of each retina has lost vision (black in the illustration). Homonymous hemianopsia occurs from lesions of the entire optic tract, the lateral geniculate body, the optic radiation, or the entire visual cortex of the contralateral hemisphere (Fig. 5-6). A lesion of the striate cortex often spares the large macular area at the occipital pole. This results in a disorder termed *homonymous hemianopsia* with macular sparing (Fig. 5-6). Partial lesions may cause *quadrant-anopsia.*

Heteronymous hemianopsia can be bitemporal or binasal (Fig. 5-6). Bitemporal hemianopsia results from damage of the optic nerve fibres as they cross the optic chiasm (Fig. 5-6).

An expansively growing pituitary tumour, perhaps related to acromegaly, can damage crossing fibres, originating from ganglion cells in the nasal halves of each retina. Expansion of the tissues surrounding both carotid arteries is a rarity, which can damage nerve fibres from the temporal halves of each retina, and cause binasal hemianopsia.

Fig. 5-6: Visual field defects. Lesions are shown with bars.

As long as the patient sees with both eyes, he may not experience any visual defect caused by damage to non-corresponding areas of the retina.

Localised blindness or *scotoma* is caused by a lesion of the retina in one eye, or by partial interruption of the optic nerve. Interruption of the entire optic nerve results in complete blindness or anopsia (Fig. 5-6).

Ophtalmoscopy is an important diagnostic tool able to establish both eye disorders and systemic diseases.

Fig. 5-7: The eye background (fundus) of the right eye in a healthy person. – A typical hypertensive eye background is shown in Fig. 9-6.

The normal ophtalmoscopic picture of the fundus is seen in Fig. 5-7. The papilla is clearly visible with the central artery and vein, and the cone-filled fovea is located to the left. The papillo-macular nerve bundle connects to the cones of the macula, but this bundle is invisible. This person has a small pigmented area along the lateral side of the papilla (Fig. 5-7).

4. Mental blindness

Bilateral temporal lobe lesions can lead to the Klüver-Bucy syndrome. In this condition, the temporal cortex, hippocampus and the amygdaloid body are damaged. The Klüver-Bucy syndrome includes *mental blindness* (visual agnosia). Mental blindness is the inability to recognise objects seen. Besides mental blindness, the syndrome consists of loss of short-term memory, and hypersexual behaviour incompatible with normal social adaptation. This hypersexual behaviour is related to the visual agnosia.

Damage to visual areas of the temporal cortex alone causes isolated *visual agnosia*. Visual agnosia or mental blindness, is lack of the ability to combine the seen object into a concept. This visual agnosia can be *colour*-agnosia (acromat-agnosia) or *face*-agnosia (prosop-agnosia).

5. Deafness (hypacusis)

Nerve deafness is caused by impairment of the cochlea, the auditory nerve or the nucleus. Chloramphenicol, kinin and streptomycin can damage the cochlea. These drugs can cause hearing loss or deafness for *all* sound frequencies. Deafness to specific frequencies is caused by localised damage of the basilar membrane. This is typical for rock and beat musicians, soldiers, and airline pilots. The nerve deaf patient has a hearing loss when tested both by *air conduction* through the middle ear, and *bone conduction* through surrounding bone structures. A certain type of nerve deafness for high tones develops among older persons (*presbyacusis*).

Conduction deafness is caused by impairment of the mechanical conduction of sound into the cochlea. A hereditary disease called otosclerosis is due to fixation of the faceplate of the stapes to the oval window. Otosclerosis, blockade of the external ear with ear wax, otitis media, damage of the tympanic membrane, and of the ossicles all cause conduction damage to hearing. Persons with conduction damage have normal bone conduction.

6. Nystagmus

Nystagmus is a disorder with *abnormal involuntary movements of the eyeballs*. Opto-kinetic nystagmus occurs when travelling in a train or a car. The eyes remain fixed on an object long enough in order to gain a clear image. The semicircular ducts cause the eyes to rotate in the direction opposite to the direction of travel. Optokinetic nystagmus involves the vestibular nuclei, the *medial longitudinal fasciculus*, and the oculomotor nuclei.

Post-rotatory nystagmus is observed in a person sitting in a rotating chair. This is the physiologic adequate stimulus for nystagmus.

Caloric nystagmus refers to the horizontal reflex movement of the eye when the external ear is flushed with hot or cold

water. The fast phase of the nystagmus is directed away from the ear flushed with cold water, and towards the ear flushed with hot water. The caloric nystagmus test is preferable to the post-rotatory test for testing the nystagmus reflexes, because it examines one ear at a time, and is more convenient.

7. Kinetosis or transportation sickness

Many types of transportation, which subjects passengers to rapid changes in the direction of motion, elicit *kinetosis*. Kinetosis is a disorder with vertigo, nausea and vomiting. The disorder is triggered from the *vestibular system*, provided that the cerebellar function (the flocculo-nodular lobe) is intact. The flocculo-nodular lobes are linked to the equilibrium control of the semicircular system. Persons with destroyed semicircular canals or with destroyed flocculo-nodular lobes can be completely protected from kinetosis at the expense of lost equilibrium during motion.

Equations

The **accommodative power of the eye** is the rise in refractive power from un-accomodated to the maximally accommodated condition:

Eq. 5-1: Accommodative power = 1/F - 1/N.

F and N are the far- and the near point, respectively.

A child of 10 years has the high accommodative power of 12-14 D, a 20-year-old person 10 D, and a 60-year-old person 1 D.

According to international convention the *sound pressure level* (SPL) is expressed in decibel (dB):

Eq. 5-2: Sound pressure level (dB) = 20 log p/p_o.

The actual pressure is p, and the threshold for sound pressure is p_o . The threshold for sound pressure is *20 μPa* in air (p_o) at 1000 Hz in a sound tight chamber for a healthy person. This pressure corresponds to a *sound effect* of 10^{-12} *Watts/m*^{2}.

Self-Assessment

Multiple Choice Questions

I. The following five statements have True/False options:

- **A**. The fovea has no cyanolab cones and no rods.
- **B.** Nystagmus is a disorder with abnormal voluntary movements of the eyeballs.
- **C.** The cornea has a refractive power of 43 D, and the healthy lens has a refractive power which varies between 12-26 D. Thus the total refractive power is 56-69 D.
- **D.** Light absorption by a single rhodopsin molecule activates thousands of G-protein molecules (transducin), which then activate large quantities of cGMP phosphodiesterase in the discs. Each of these enzyme molecules catalyses the hydrolysis of cGMP to 5'-GMP at a rate of thousands per second.
- **E.** Parallel light from the far point foveates on the retina in the fully accommodated eye.

II. The following five statements have True/False options:

A. Nerve deafness is caused by damage of the cochlea, the auditory nerve or nucleus. There is hearing loss by air conduction only.
B. Kinetosis is a disorder with vertigo, nausea and vomiting due to rapid changes in the direction of motion.
C. Drugs, such as chloramphenicol, kinin and streptomycin, can cause hearing loss for all sound frequencies.
D. The organ of Corti contains hair cells, each with 40-100 stereocilia. When the pressure wave displaces the hairs towards the tallest stereocilia, the hair cell is depolarised.
E. Low tones travel only a short distance from the stapes along the basilar membrane to their resonant point, where the displacement amplitude is maximal.

5. Case History A

A 30-year old female complains of eyestrain and frontal headache during reading - sometimes followed by nausea and

vomiting.

The patient is placed 6 m (20 feet) from the Snellen test chart. She is able to read line 6, which is the letter size read by a normotropic eye. Now thin convex lenses are placed in front of her eyes, but she can still read line 6. The diopter of the strongest convex lens with which she can still read line 6 is +4 D for both eyes. These converging light rays must be directed against the far-point (F). F must be located behind the eyes at a distance of ¼ m. Examination with concave lenses reveals the strongest concave lens by which she can read line 6 to be -3 D or N = -1/3 m (in front of each eye).

1. *1. What is the refractive anomaly of the patient?*
2. *2. A normotropic person of 30 years has an accommodative power of 7 D. Compare the accommodative power of the female patient to that of the normal person?*
3. *3. Is it possible for the patient to read a fine text 0.2-m in front of her?*
4. *4. The saggital diameter of the patient's eyes is typical. Describe its characteristics.*

- 5. *This patient has a reduced outflow of chamber fluid at the iridocorneal junction. She has an increased risk of developing an eye disease, which is the most common course of blindness in the world. Describe the most likely condition and the relation to her eye anatomy.*

5. Case History B

A patient with a hearing loss of 26 dB is working in a power station, where the daily sound intensity is 100 dB and the air temperature is 20^{o}C.

- ***1**. Calculate the ratio between the sound pressure in the powerhouse, and the sound threshold pressure for a healthy person.*

- ***2.** Calculate the threshold pressure for the patient.*

5. Case History C

A 9-year-old girl suffers from facultative hypermetropia. She is placed 6 m from the Snellen test chart and asked to read line 6, which is the letter size read by a normotropic eye. When thin convex lenses are placed in front of her eyes, she can still read line 6. The diopter of the strongest convex lens with which she can still read line 6 is +5 D for both eyes. Examination with concave lenses reveals the strongest concave lens by which she can read line 6 to be -4 D.

1. *Where are the far point (F) and the near point (N) located?*
2. *Calculate her accommodative power and compare it to the normal value of 14 D.*
3. *Calculate the correction needed.*

5. Case History D

A 40 year old male diabetic has an accommodative power of 10 D. His near point (N) is located 0.05 m in front of the eye (- 0.05 m).

- *1. Calculate the location of F and the necessary correction.*
- *2. What is the name of the refractive disorder?*
- *3. Is the patient capable of driving a car without corrective glasses? Is the vision mildly or seriously reduced?*

Try to solve the problems before looking up the answers.

Highlights

- *Near vision- even with only one eye - triggers accommodation and pupillary constriction.*
- *Rods are most sensitive in the dark (scotopic vision). More than hundred rods converge on each ganglion cell. There are no rods at all in the fovea.*
- *Cones operate best in light (photopic vision). Cones have a high-resolution capacity and hence a high visual*

acuity, because the light is focused on the fovea, where the cones are concentrated.

- *The cornea has a refractive power of 43 D, and the healthy lens varies between 12 and 26 D.*
- *The accommodative power of the eye is the rise in refractive power from the un-accomodated to the maximally accommodated condition.*
- *Conjugate movements are movements of both eyes in the same direction and magnitude, so that the relation between the visual axes is maintained.*
- *Saccadian or jumping movements are rapid eye movements. Saccadian eye movement is an instantaneous reposition of the eye that occurs. when reading or when focusing on a flash of light in the peripheral visual field. The velocity of the movement is up to 500^{o} per s. The latency period is 250 ms, and the contraction time is 50 ms. The compensatory eye movement involving the vestibular system, occurs when the head rotates. This is also an example of Saccadian eye movement.*
- *The fovea only contains cones. Cones function in the daytime, with maximal visual acuity, and colour vision. The human eye possesses three types of cones, each with a specific pigment related to the three basic colours: red (erythrolab), green (chlorolab) and blue (cyanolab).*
- *The 3 cone types are uniformly distributed in the retina except in the fovea. The fovea has no cyanolab cones and no rods.*
- *Abnormal trichromats have a reduced amount of one cone pigment: Protanomalous trichromats lack erythrolab, deuteranomalous (the most frequent type) lack chlorolab, and tritanomalous lack cyanolab.*
- *Movements in the visual scenery are depicted as opposite movements on the retina. Convergent inputs from the two eyes result in depth perception (ie, stereopsis or stereoscopic vision). Stereopsis depends upon the medial, longitudinal fasciculus and the corpus callosum. These structures co-ordinate the movements of the two eyes.*
- *A flickering source of light liberates successive flashes so rapidly that they fuse and appear continuous.*
- *The normal human ear is sensitive to pure tones with frequencies between 10 and 30 000 Hz in a young person.*
- *Speech has an intensity of 60-65 dB and sounds that exceed 100 dB can damage the ear.*
- *A high frequency tone produces travelling waves along the basilar membrane. High tones travel only a short distance from the stapes along the basilar membrane to their resonant point, where the displacement amplitude of the basilar membrane is maximal.*
- *Low frequency tones travel all the way to the apex of the cochlea. The higher the tone frequency, the more basal located in the cochlea is the resonant point and its potential.*
- *The auditory and the vestibular systems share the labyrinth, and transmit signals to the brain through the 8^{th} cranial nerve. The two systems record fluid movements and use the so-called hair cells as mechanical receptors.*
- *The medial geniculate nucleus and the inferior collicle can increase its tone selectivity by dampening other sound signals. This explains the cocktail party effect.*
- *The cupula is concerned with equilibrium control during motion and with angular acceleration (rotation of the head).*
- *The utricles and saccules are sensitive to linear acceleration.*

Further Reading

- Berardi, N., A. Cellerino, L. Dominici, M. Fagiolini, T. Pizzorusso, A. Cattaneo, and L. Maffei. "Monoclonal antibodies to nerve growth factor affect the postnatal development of the visual system." *Proc. Natl. Sci.* , USA, 91 (2): 684-688, 1994.
- Stryer, L. "Cyclic GMP cascade of vision." *Annual Rev. Neuroscience* 9: 87, 1986.
- Brodal, A. "Neurological anatomy in relation to clinical medicine." *Edition 3.* New York, 1981, Oxford University Press.
- Von Bekesy, G. "Experiments in hearing." *New York, 1960.* Mc-Graw-Hill.

Return to content

Return to chapter 5

Chapter 4. Multiple Choice Questions

I. Answers **A, C, D,** and **E** are true statements, whereas **B** is false.

II. Answer **B** is true, whereas **A, C, D**, and **E** are false.

III. Answers **A, D**, and **E** are true statements, whereas **B** and **C** are false.

4. Case History A

1. The defect in language function is aphasia (word deafness).
2. The difficulties in understanding spoken and written language is called Sensory Aphasia. ***Damage to Wernicke's area in the dominant hemisphere*** *(left hemisphere in 95% of all persons) explains the findings. The brain lesion of the composer is localised here.*

4. Case History B

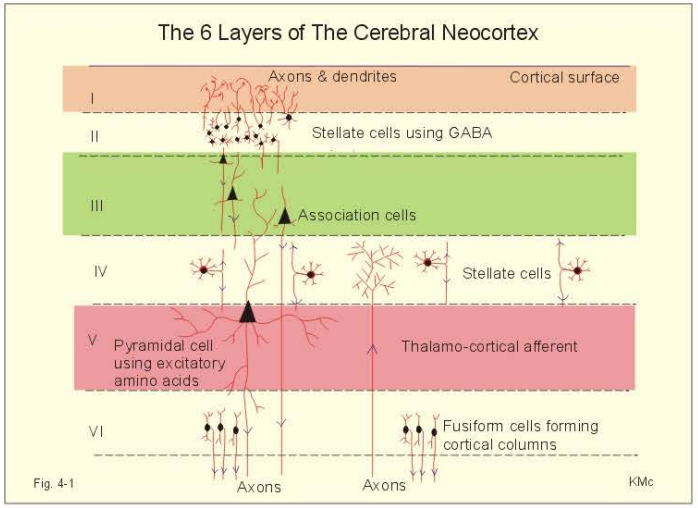
- *1. Bleeding has interrupted the corticospinal tract as it traverses the* ***internal capsule****. The block of the excitatory pathways to the spinal cord result in* ***severe contralateral paresis****.*

- *2.* ***Spasticity*** *means a motor condition dominated by increased tonic and phasic stretch reflexes, and it involves damage of the reticulospinal tract.* ***Foot clonus*** *is a repetitive pattern of violent contractions when the reflex hammer taps the Achilles tendon.*

- *3. The innervation of the frontalis and the orbicularis oculi muscles is* ***bilateral****. Accordingly, these muscles function normally in a patient with a* ***unilateral central paresis of the facial nerve****. A patient with a* ***peripheral paresis*** *of the facial nerve can only knit his brows and screw up his eyes at one side.*

Return to chapter 4

Return to content


The 6 Layers of The Cerebral Neocortex
Axons & dendrites
Cortical surface
I
II
Stellate cells using GABA
III
Association cells
IV
Stellate cells
V
Pyramidal cell
using excitatory
amino acids
Thalamo-cortical afferent
VI
Fusiform cells forming
cortical columns
Axons
Axons
Fig. 4-1
KMc

RAS and the Thalamo-Cortical System

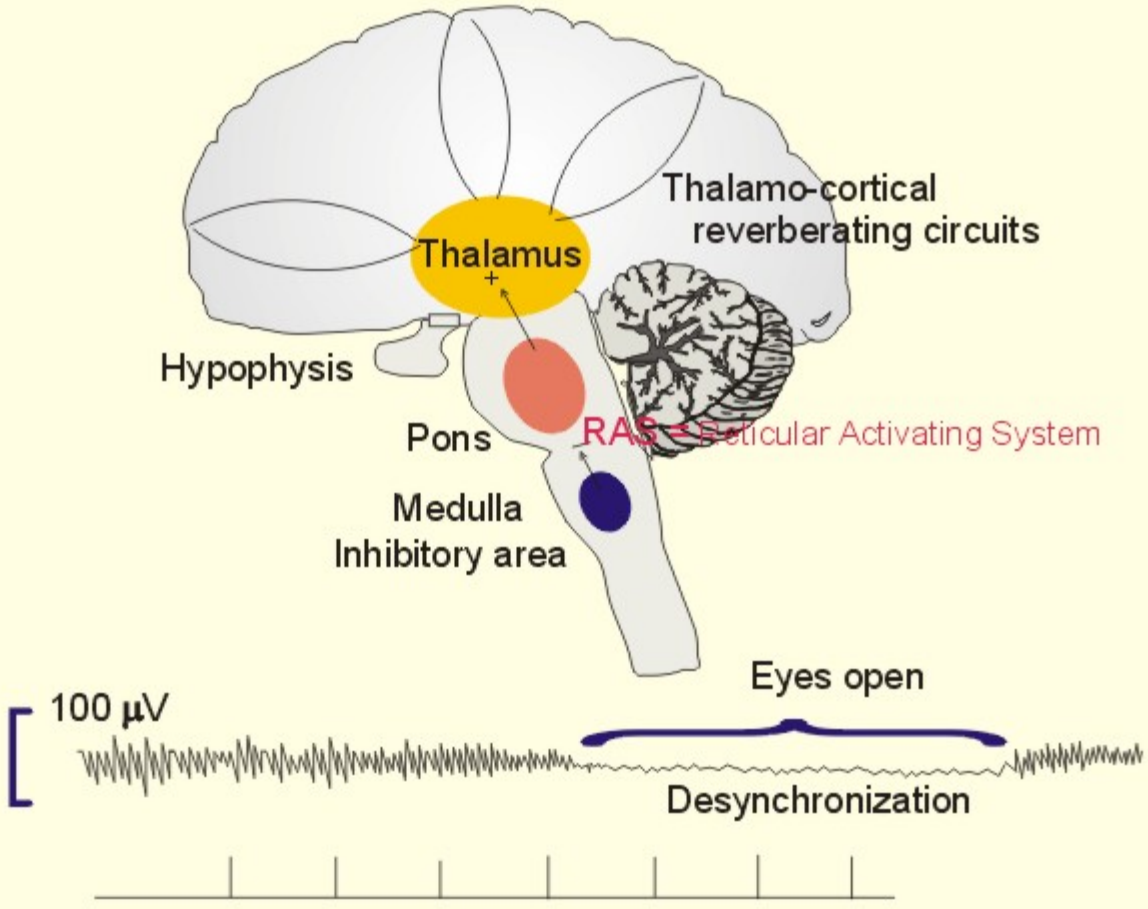


Fig. 4-2

KMc

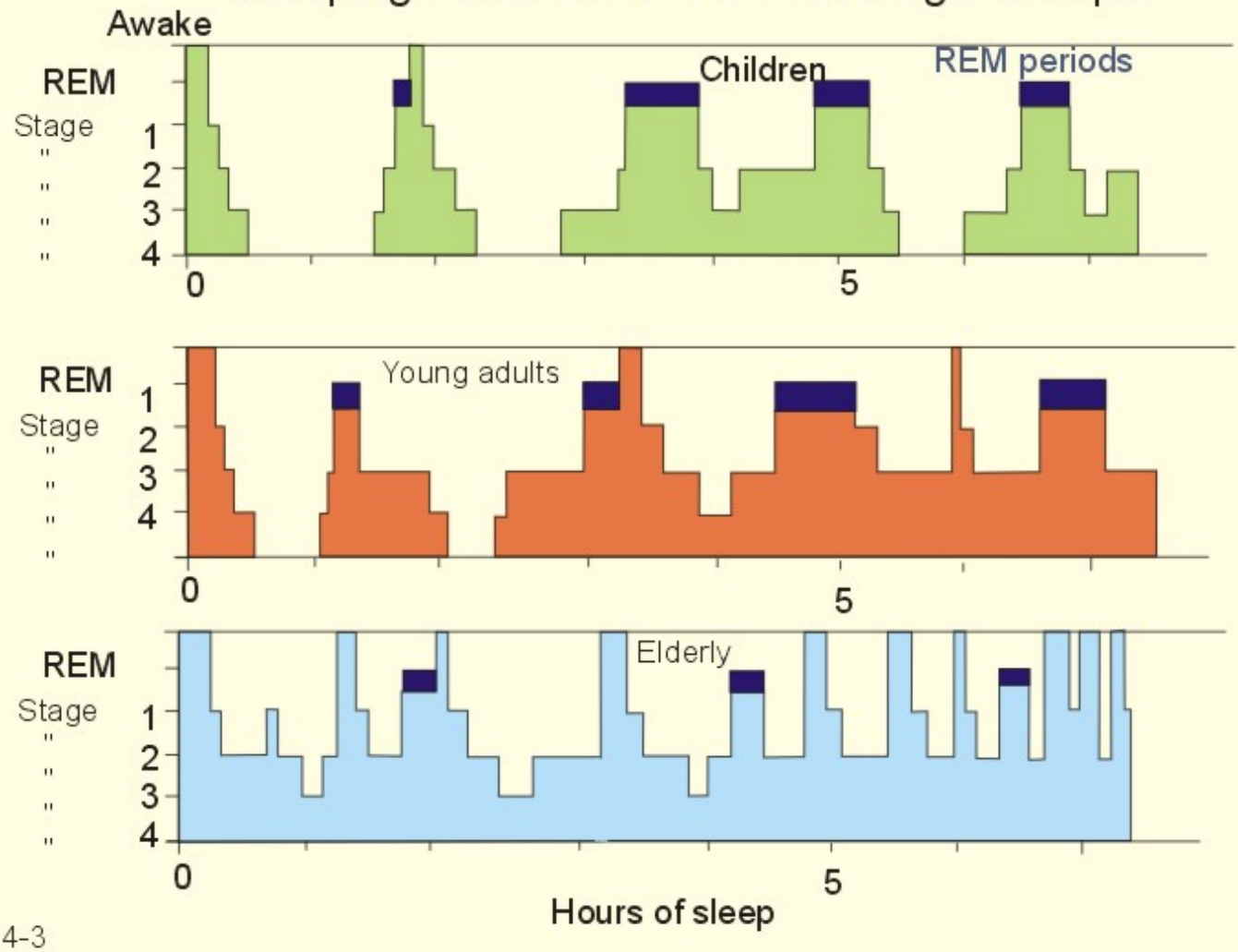

Sleeping Patterns In The Three Age Groups
Awake
REM
Stage
1
2
3
4
0
5
Children
REM periods
Young adults
Elderly
Hours of sleep

Fig. 4-3

KMc

The Human Cerebral Cortex

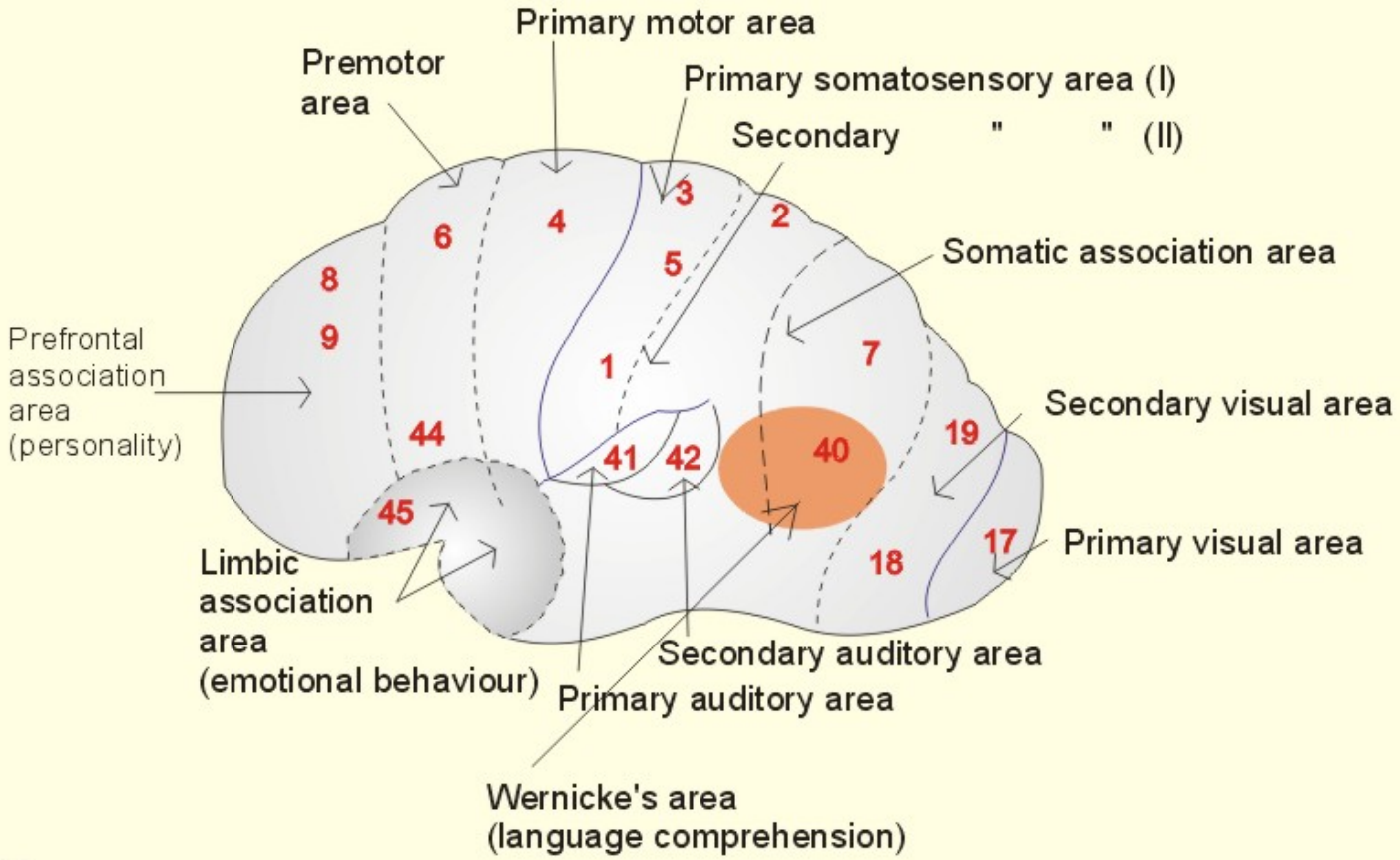


Fig. 4-4

KMc

The Limbic System

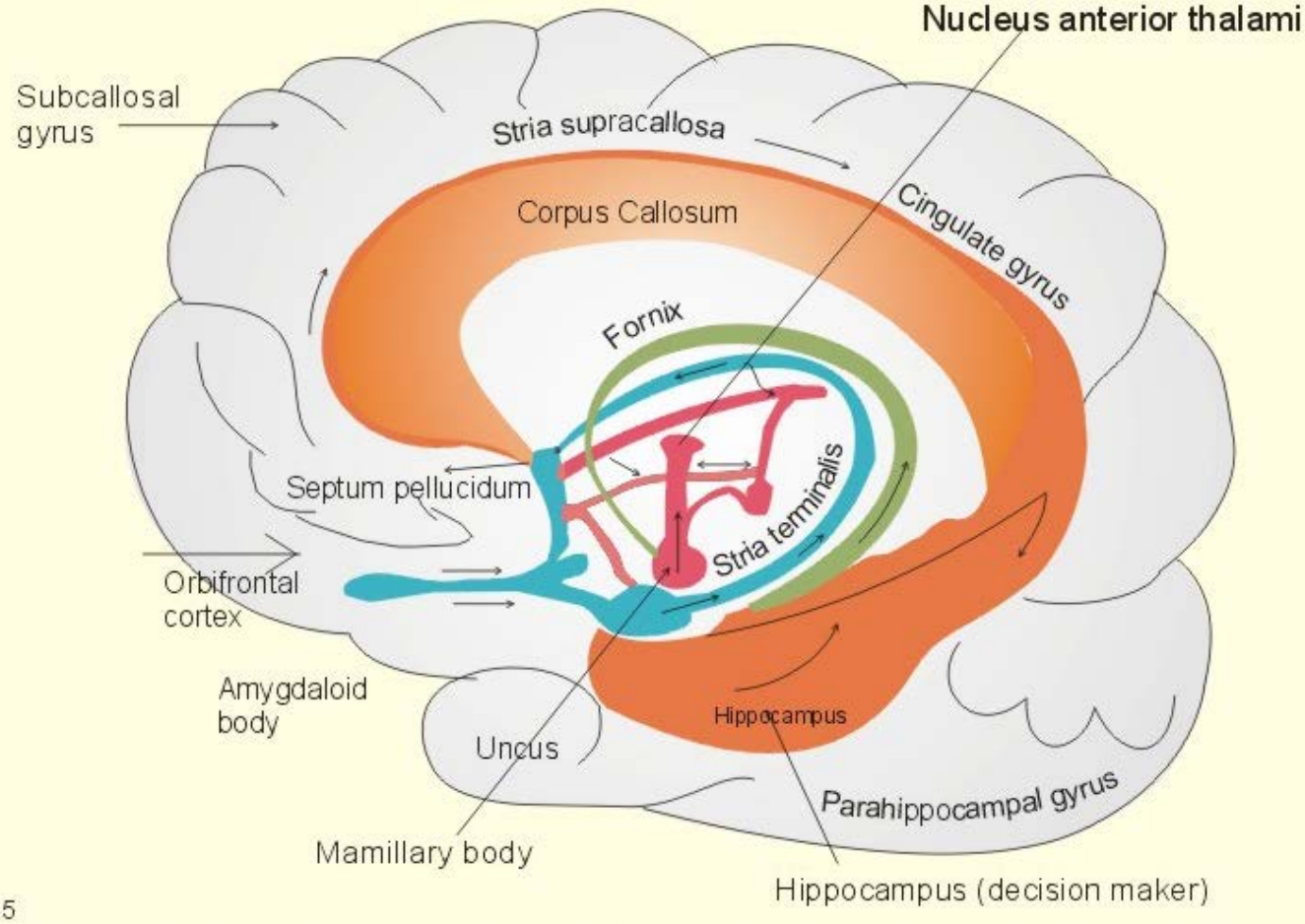
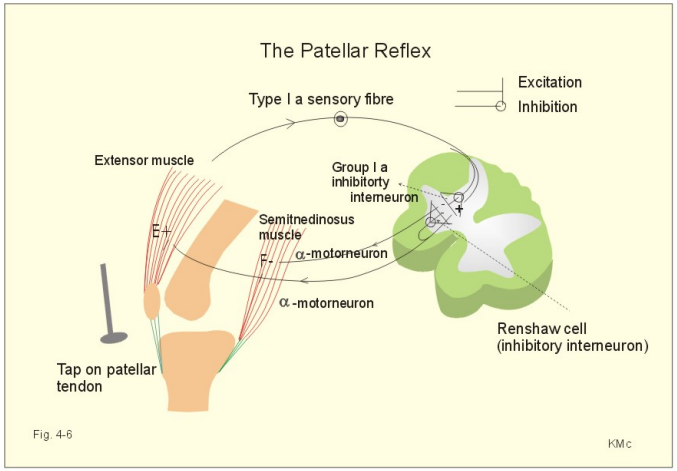


Fig. 4-5

KMc


The Patellar Reflex
Excitation
Inhibition
Type I a sensory fibre
Extensor muscle
Group I a inhibitorty interneuron
Semitnedinosus muscle
E+
F-
α-motorneuron
α-motorneuron
Renshaw cell (inhibitory interneuron)
Tap on patellar tendon
Fig. 4-6
KMc

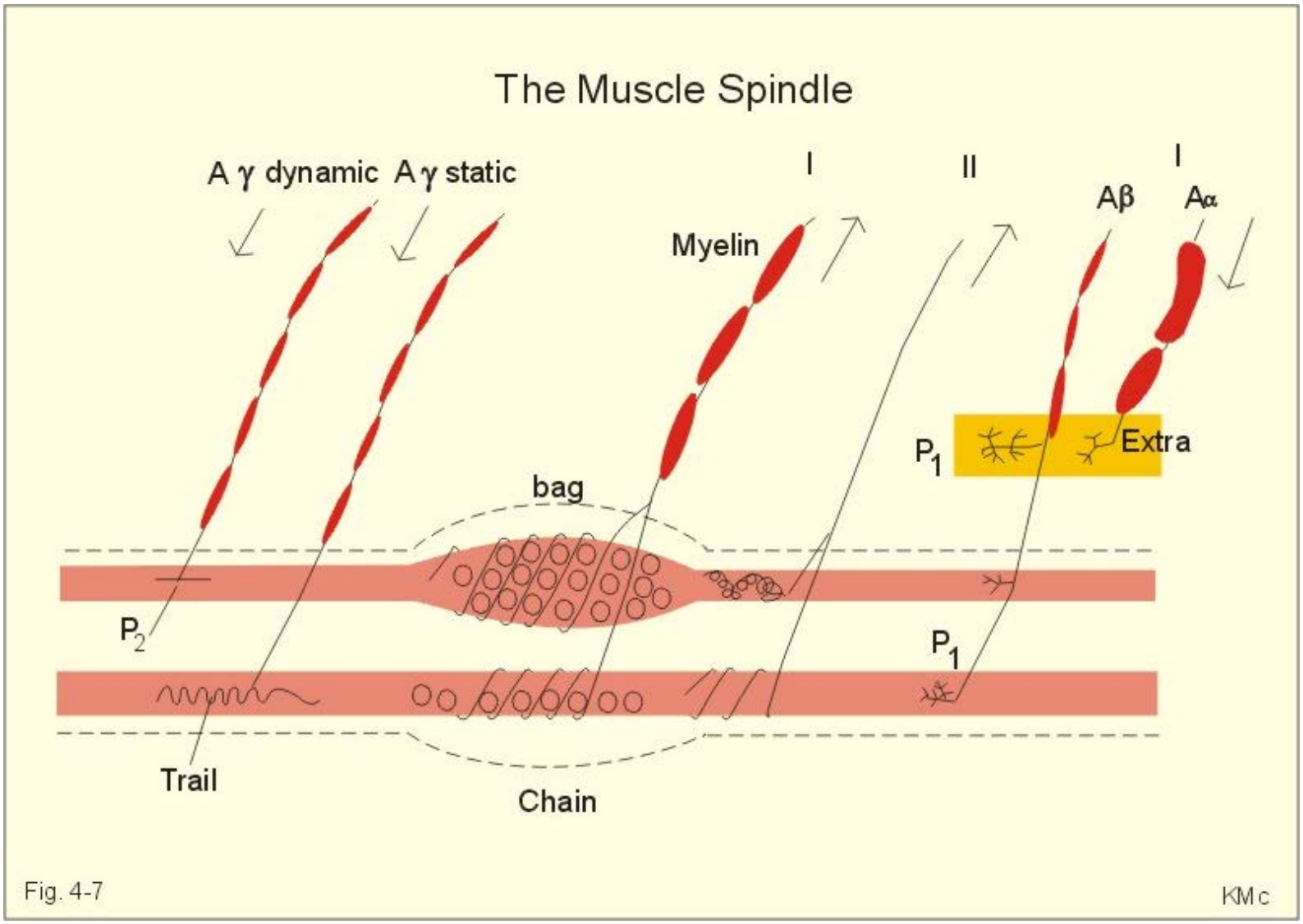
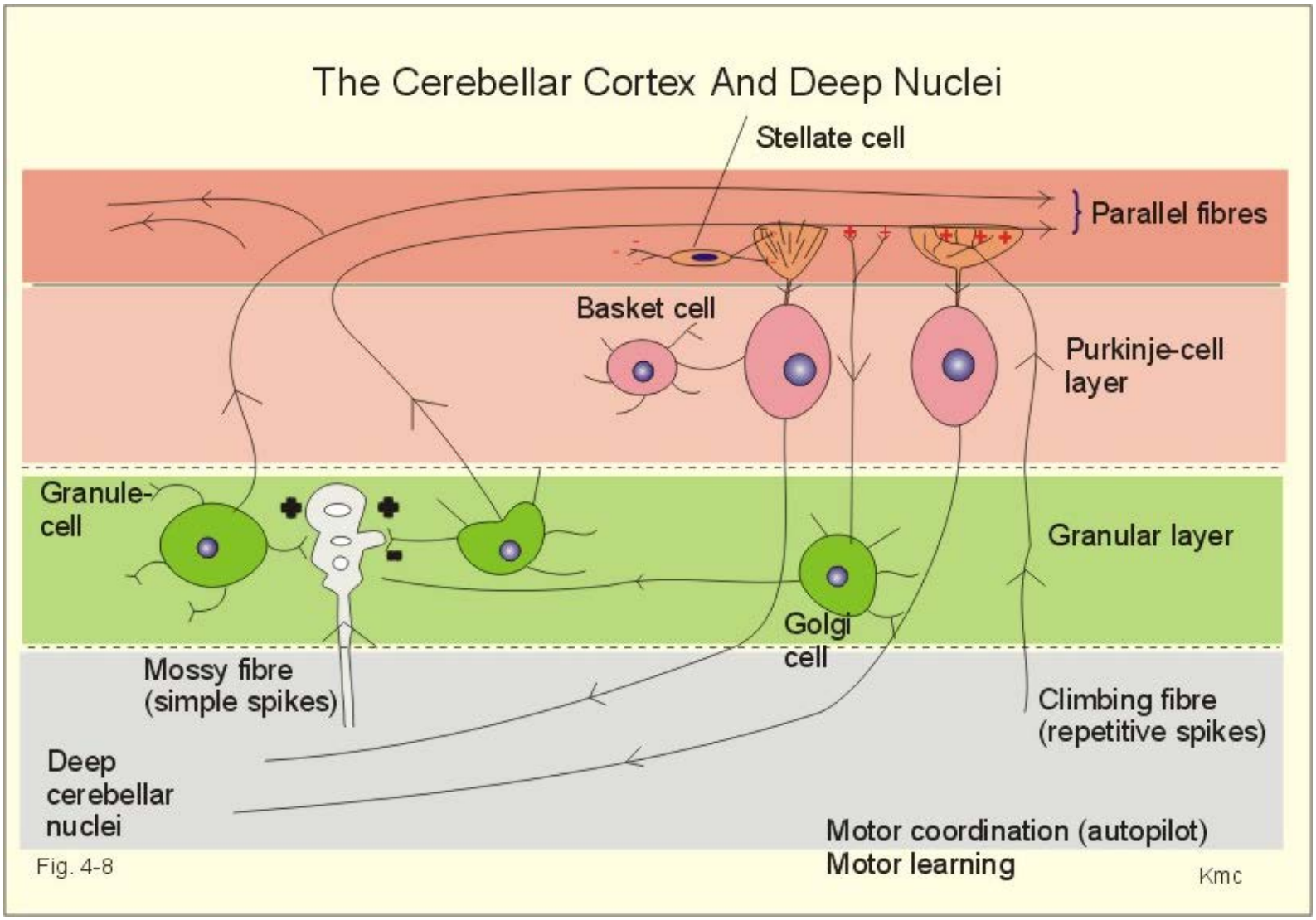
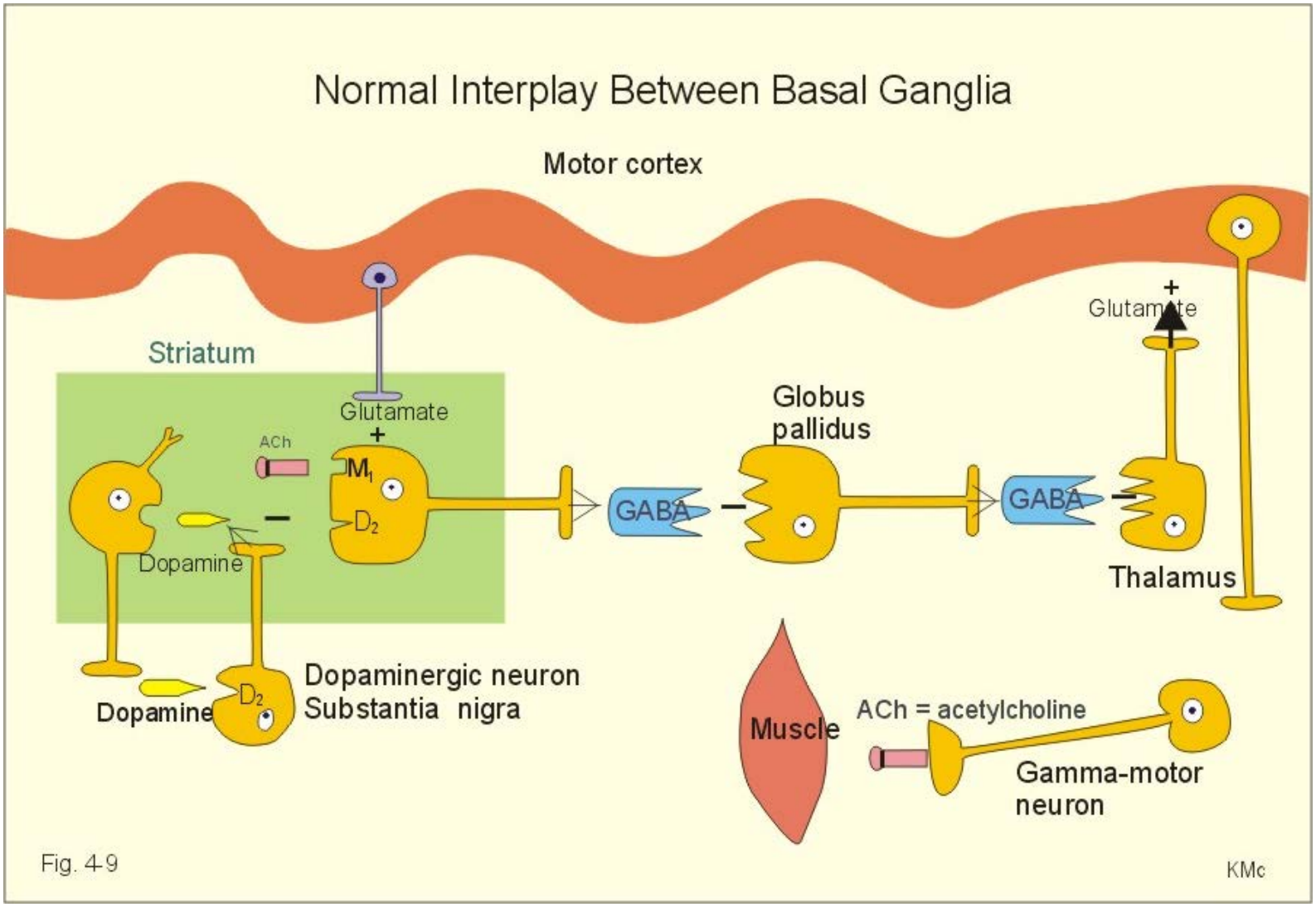

The Muscle Spindle
A γ dynamic
A γ static
I
II
I
Aβ
Aα
Myelin
bag
P1
Extra
P2
P1
Trail
Chain

Fig. 4-7

KM c


The Cerebellar Cortex And Deep Nuclei
Stellate cell
Parallel fibres
Basket cell
Purkinje-cell layer
Granule-cell
Granular layer
Golgi cell
Mossy fibre (simple spikes)
Climbing fibre (repetitive spikes)
Deep cerebellar nuclei
Motor coordination (autopilot)
Motor learning
Fig. 4-8
Kmc


Normal Interplay Between Basal Ganglia
Motor cortex
Striatum
Glutamate
+
ACh
M1
D2
Dopamine
–
GABA
–
Globus
pallidus
GABA
–
Glutamate
+
Thalamus
Dopaminergic neuron
Substantia nigra
D2
Dopamine
Muscle
ACh = acetylcholine
Gamma-motor
neuron
Fig. 4-9
KMc

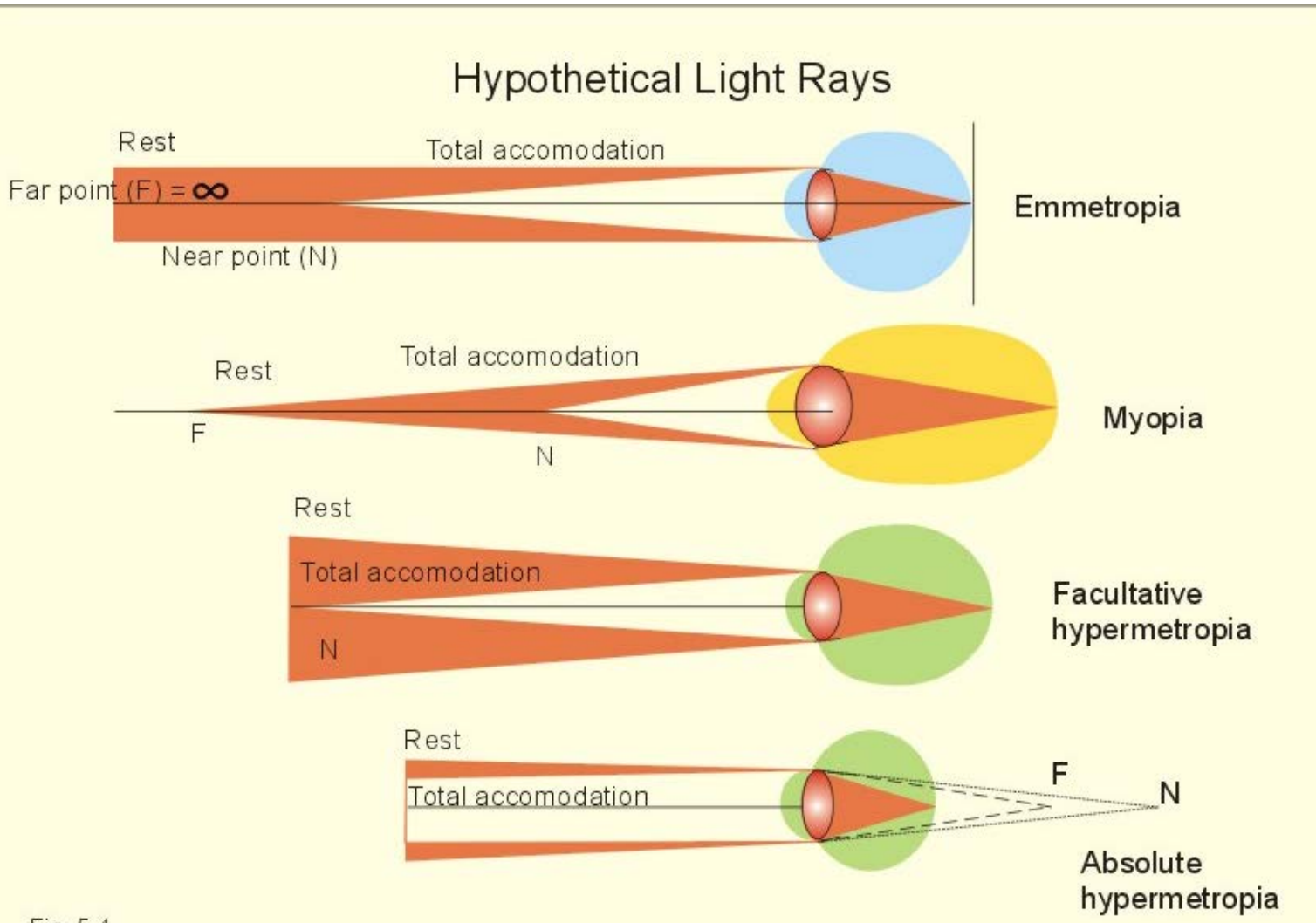

Hypothetical Light Rays
Rest
Total accomodation
Far point (F) = ∞
Near point (N)
Emmetropia
Rest
Total accomodation
F
N
Myopia
Rest
Total accomodation
N
Facultative hypermetropia
Rest
Total accomodation
F
N
Absolute hypermetropia
KMc

Fig. 5-1

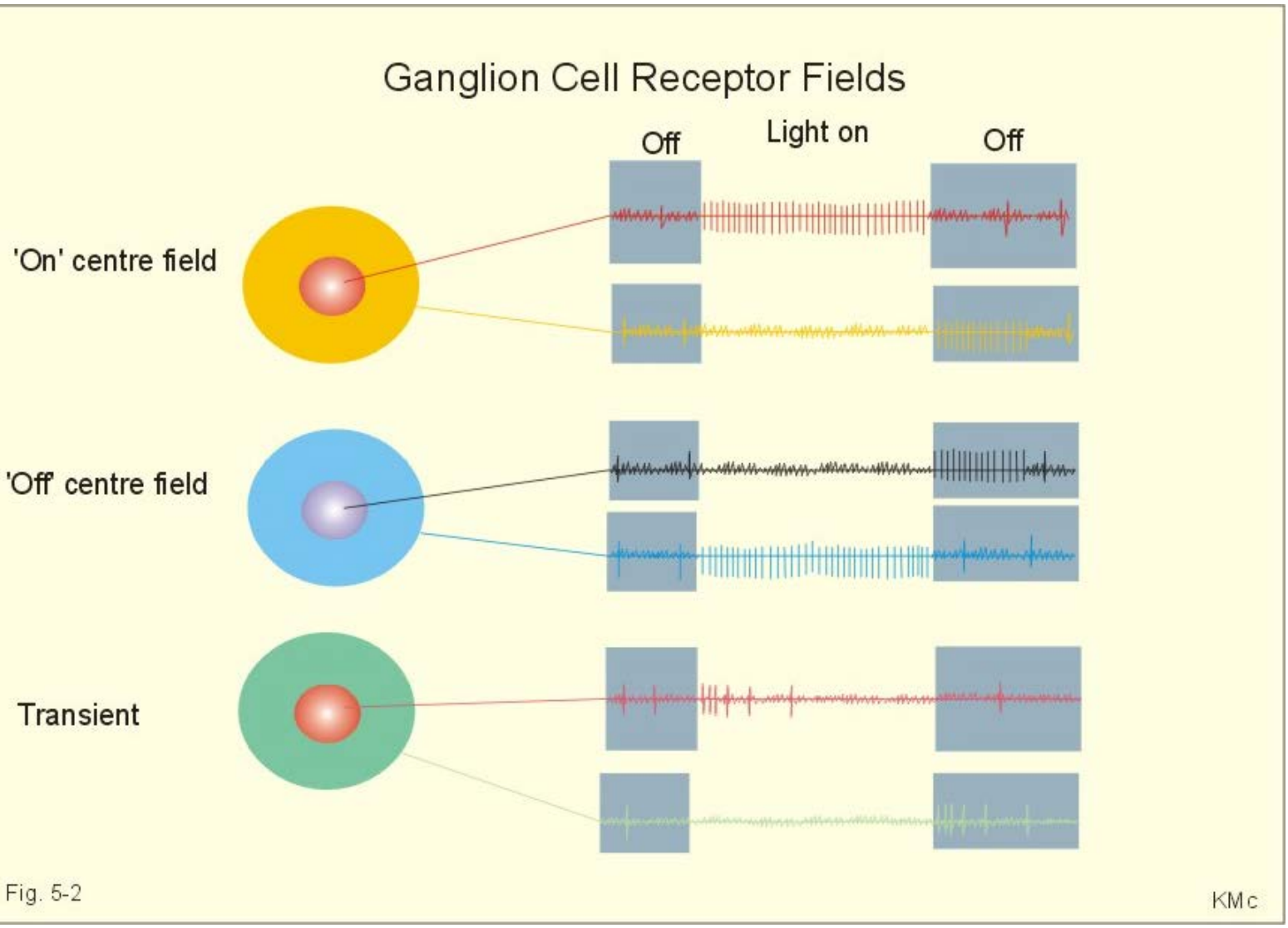

Ganglion Cell Receptor Fields
Off
Light on
Off
'On' centre field
'Off' centre field
Transient

Fig. 5-2

KMc

Cross - Section Of Cochlea

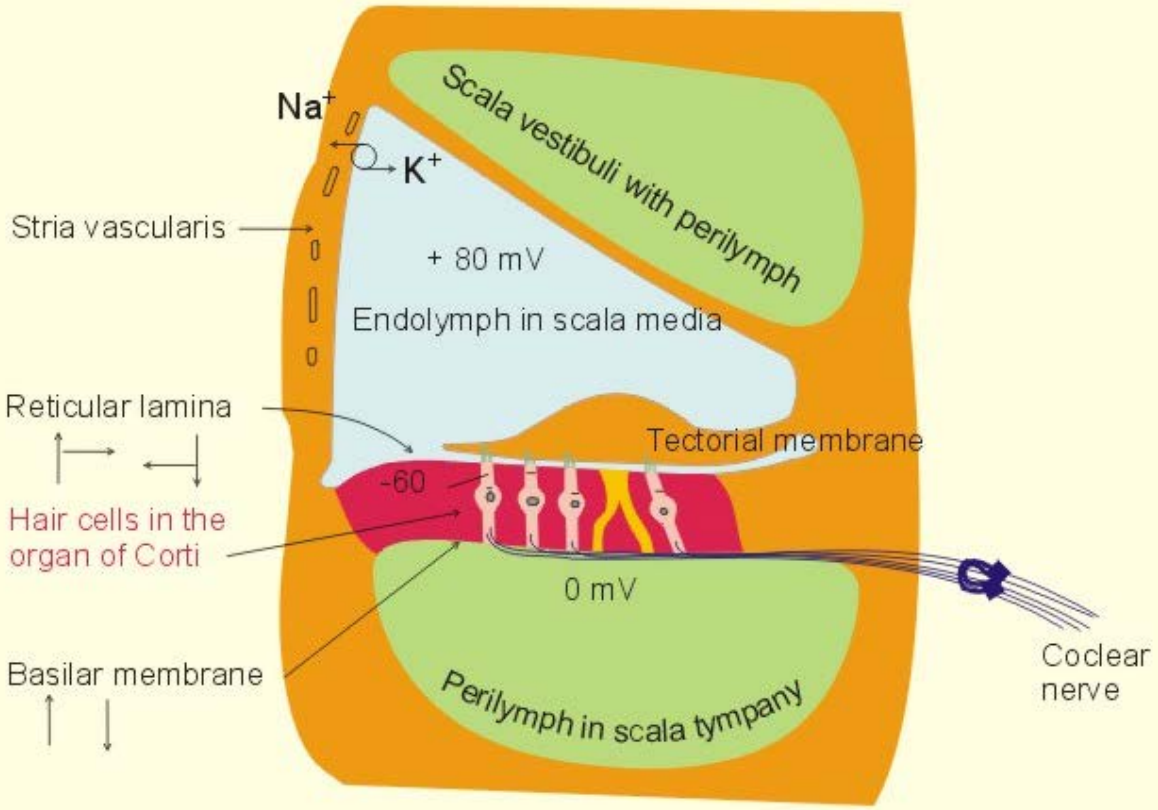
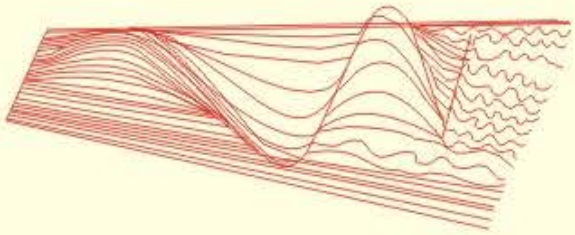


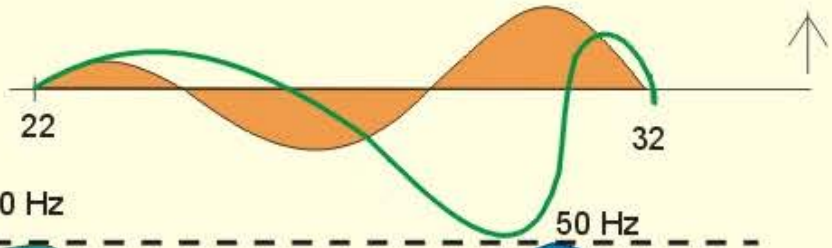
Fig. 5-3

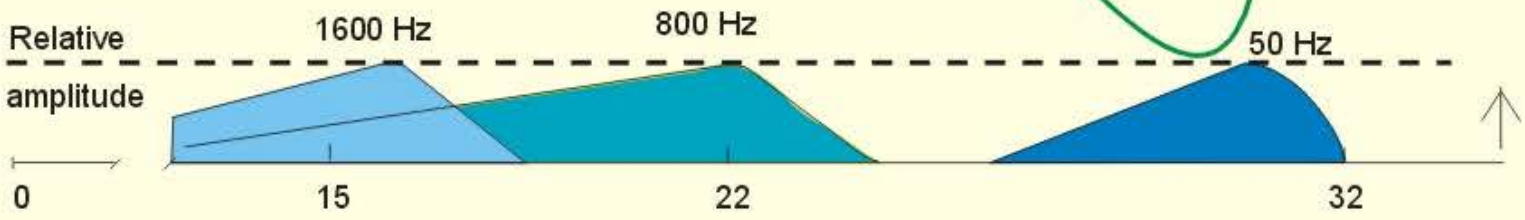
KMc

The Travelling Wave Theory



Displacement of basilar membrane





mm from stapes

Fig. 5-4

KMc

Function of The Vestibular System

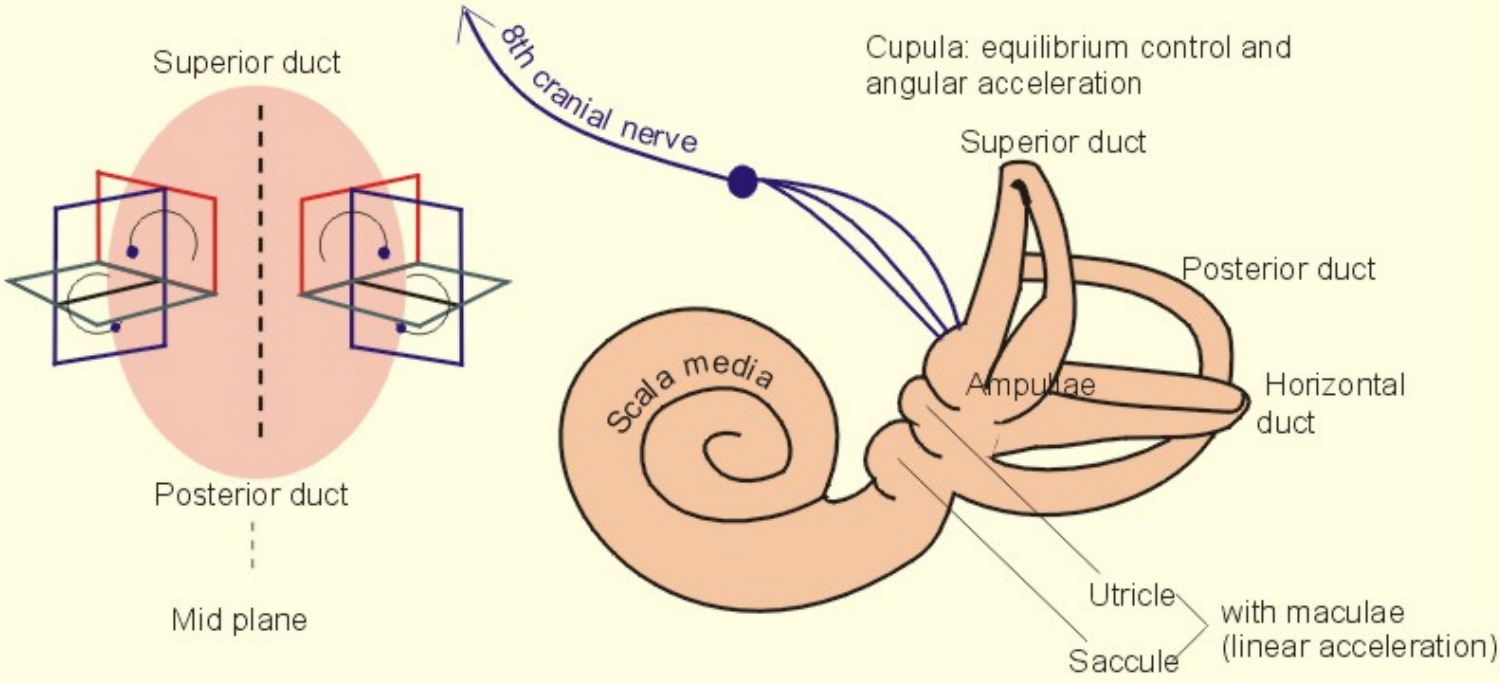


Fig. 5-5

KMc

Visual Field Defect

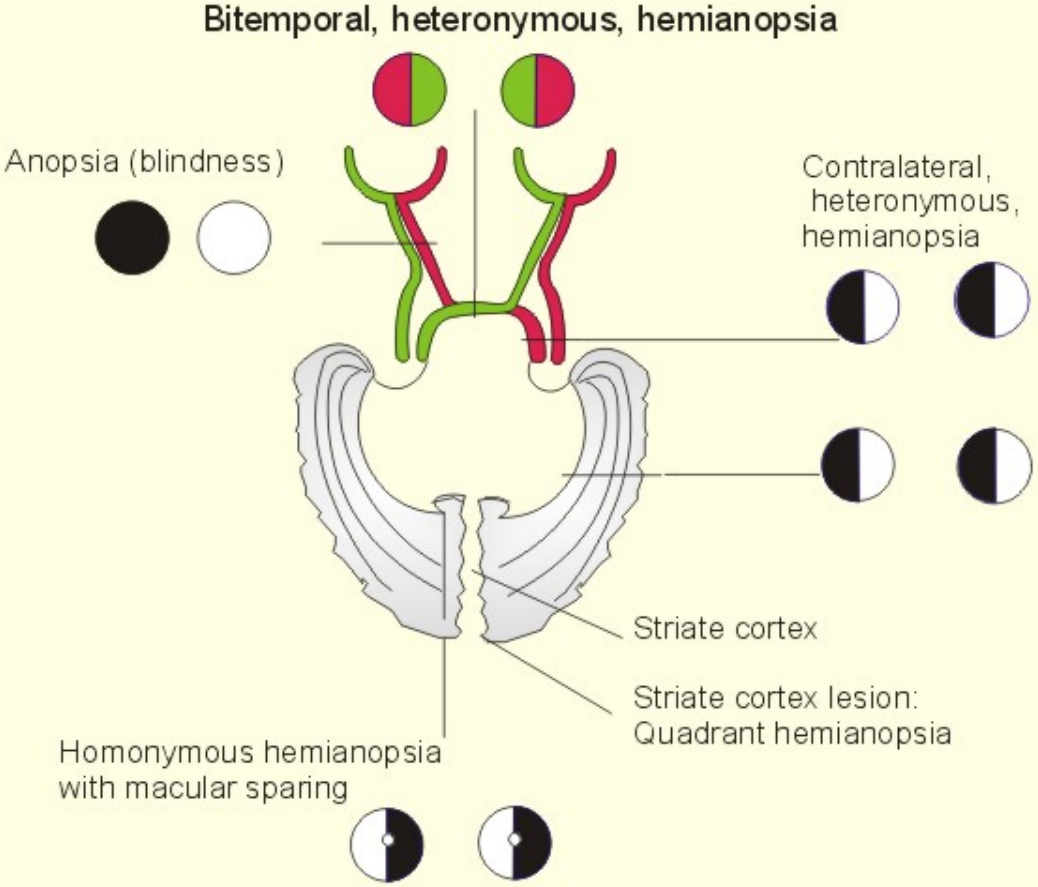


Fig. 5-6

KMc

Normal Ophthalmoscopic Picture

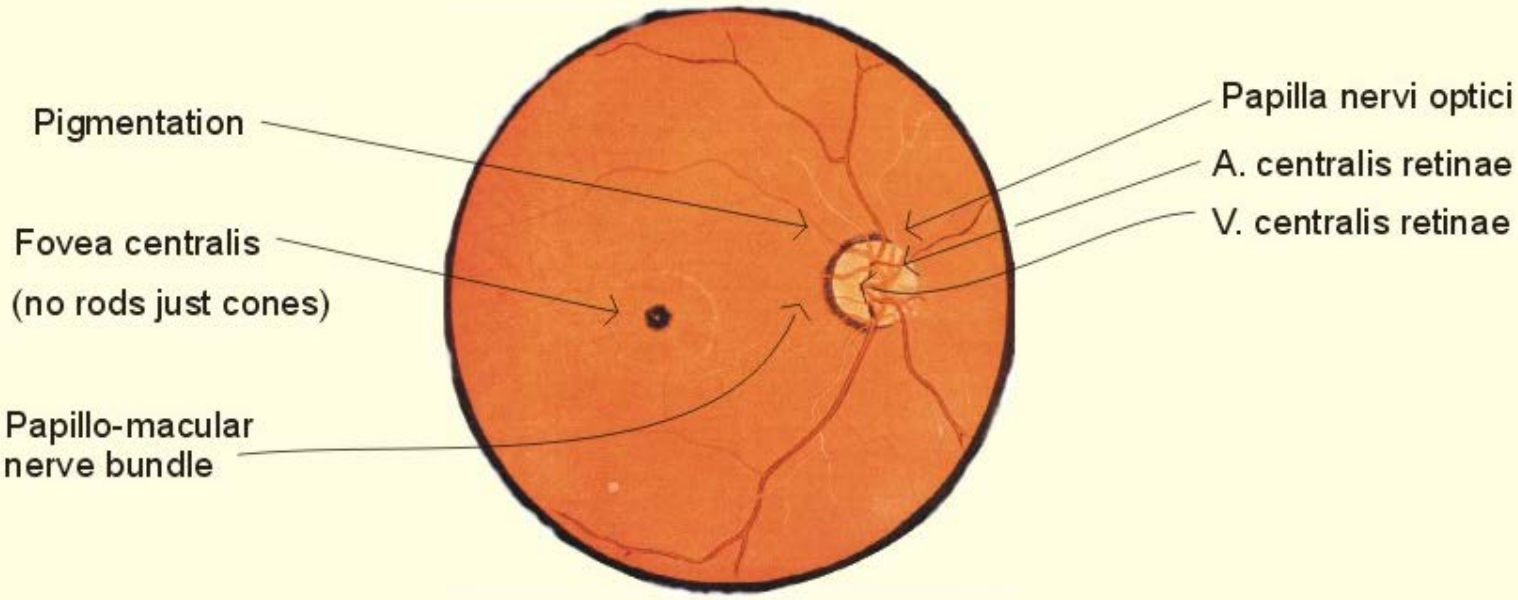
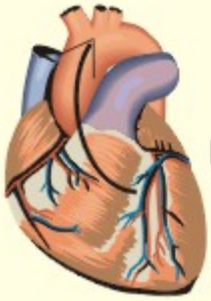


Fig. 5-7

KMc

Causes of Hypertension

1. Chronic renal disease
2. Cushing's Syndrome
3. Hyperaldosteronism



Increased Na^+-retention ← Increased renin-aldosterone-cascade

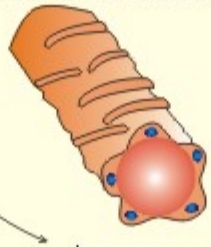
Increased ECV and blood volume

Increased $\dot{Q}$

4. Phaeocromocytoma

Arteriolar constriction (α-receptors) ← Increased sympathetic activity (genetic/environmental)

Resistance vessel



Increased venous return

Wall stretch

Increased ANP

Increased TPVR

Cardiac output

Blood pressure

Reduced Na^+ retention → Reduced ECV and blood volume

Fig. 9-6

KMc

Chapter 5.

Multiple Choice Questions

I. Answers A, B, C and D are true statements, whereas E is false.

II. Answers B, C, and D are true statements, whereas A is false.

Chapter 5. Case History A

- *1. The patient has **facultative** hypermetropic eyes. In order to foveate diverging light she is forced to accommodate relatively more than a normal eye during near vision. Thus, the ciliary muscles are fatigued, and she gets **eyestrain**.*

- *2. The accommodative power is calculated as follows: 1/F – 1/N = 4 – (-3) = **7 D**. This result is exactly the same as that of a normotropic person of 30 years.*

- *3. A fine text can be read further away and until the near point, 1/3 m in front of the eyes. Within the range of 0.2 m, she cannot obtain a sharp picture on the retina (foveation).*

- *4. Hypermetropic persons usually have **too short eyeballs** – the saggital diameter is too short.*

- *5. The fluid outflow at the iridocorneal junction is reduced, because such flat eyes implies a **small chamber angle**. This leads to increased intraocular pressure or **glaucoma**, a frequent disorder in hypermetropic patients*

5. Case History B

1. The threshold pressure for a healthy person is 20 μPa or 2×10^{-5} Pa (Po). Thus, the relation is: dB, SPL = 20 log (P/Po).

- *dB, SPL = 20 log (P/(2 10^{-5})). Accordingly the sound pressure in the power plant is **2 Pa**, and the **ratio is 10^5**.*

- *2. 26 dB, SPL = 20 log (P/(2 10^{-5})).*

The threshold pressure for the patient is (3.99 10^{-4}) Pa. This pressure is 20 times as high as that of the normally hearing person.

5. Case History C

- *1. 1/N = - 4 D or N = -0.25 m. The nearpoint **(N) is located 0.25 m in front of her eye** (- 0.25 m). The far point (F) is calculated as follows: 1/F = 5 D.*

  ***F is located 0.20 m behind her eye**.*

- *2. The accommodative power is: 1/F - 1/N = 1/0.20 - (1/-0.25) = **9 D**. This is almost as good as normal although the girl has facultative hypermetropia.*
- *3. The patient needs correction with the strongest convex lens for optimal visual acuity related to her far point: **+ 5 D**. Lenses with slightly lesser refractive power would allow her to see sharply as well, but she would have to use her accommodation all the time and probably develop eyestrain.*

5. Case History D

- *1.The accommodative power is 10 D equal to 1/F - 1/N. Accordingly, 1/F = 10 D - 1/0.05m or 10 D - 20 D = -*

10 D. This is a concave lens with a refractive power of 10 D. This lens is the weakest concave lens that would refract parallel light beams; so they are directed as if they came from a point 0.1 m in front of the eye (F).

- *2. A refractive disorder with both N and F located in a finite distance in front of the eye must be **myopia**.*

The patient is not allowed to drive a car without spectacles. The disorder is severe. The patient with - 10 D must be regarded as functionally blind, with a far point 10 cm in front of his eyes

Return to chapter 5

Return to content

Autonomic Cardiovascular Control

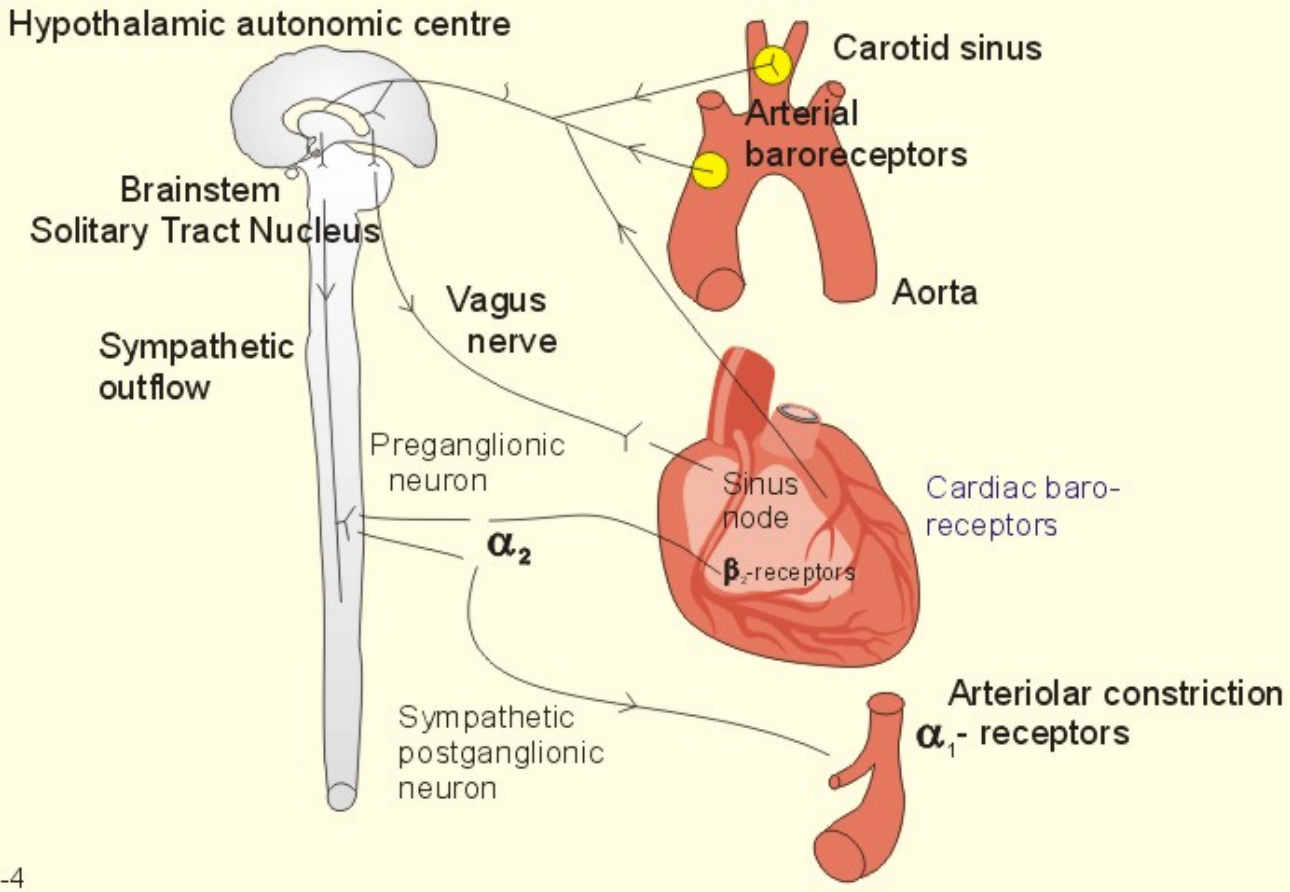


Fig. 6-4

KMc

Chapter 27.

Blood Glucose And Diabetes

Study Objectives

- To *define* glicentin, gluconeogenesis, glycogenolysis, glycolysis, hyperglycaemia, hypoglycaemia, incretins, insulin antagonists, paracrine secretion, primary and secondary diabetes mellitus.
- To *describe* the structural and functional characteristics of the Langerhans islets with cell types and hormones produced. To describe the incretin effect, insulin effects, the insulin receptor, and the glucose transporter. To describe disorders of the different cell types and their clinical picture. To describe methods for evaluation of the glucose combustion.
- To *draw* oral and intravenous glucose tolerance curves with linear and logarithmic ordinates for blood glucose concentration.
- To *explain* the biosynthesis and the effects of insulin, glucagon, pancreatic polypeptide (PP) and somatostatin. To explain the glucose metabolism and the control of blood glucose in the fed and the fasting state. To explain the consequences and the therapy of high and low blood glucose disorders.
- To *use* the above concepts in problem solving and in case histories.

Principles

- *The physiological principle in treatment of diabetes is to inject a fast-acting insulin three times a day just before meals and a slow-acting insulin at night.*
- *Insulin promotes the storage of energy, the synthesis of glycogen, mRNA and proteins.*
- *Certain major tissues (kidney, brain and intestine) are insensitive to the direct action of insulin.*

Definitions

- **Glicentin** is intestinal glucagon. Glicentin is built up from 69 amino acids in contrast to pancreatic glucagon, which consists of 29 amino acid moieties.
- **Gluconeogenesis** refers to formation of new glucose from glycogenic amino acids, lactate, glycerol and pyruvate.
- **Glycogenolysis** refers to glycogen breakdown to glucose in the liver.
- **Glycolysis** refers to anaerobic breakdown of glycogen.
- **Hyperglycaemia** is a condition, where the blood glucose is above 6.7 mM.
- **Hypoglycaemia** refers to a serious condition, where the blood glucose is below 2 mM.
- **Incretins** are hormones, which strongly potentiates the insulin secretion induced by the rising blood [glucose]. The incretins cause a much larger insulin secretion than the iv. administration of glucose, even at the same rise in blood [glucose]. This extra insulin secretion is called the *incretin effect*.
- **Ketogenesis** refers to accelerated lipolysis with liberation of free fatty acids to the blood. Free fatty acids are broken down to fatty acyl carnitine within the liver cells, and this molecule is converted into acetyl CoA, which in turn reach the mitochondria, where ketone bodies are formed.

- **Paracrine secretion** is a release of signal molecules to neighbour cells.
- **Insulin antagonists** are hormones opposing the effect of insulin: Pancreatic and intestinal (glicentin) glucagon, ACTH, growth hormone.
- **Primary diabetes mellitus** refers to all cases, where the cause is not fully explained.
- **Secondary diabetes mellitus** is caused by hypersecretion of one or more of the many catabolic hormones with hyperglycaemic effect (adrenaline, noradrenaline, glucagon, glucocorticoids and growth hormone) or by total destruction of the pancreas from pancreatitis or carcinoma. The hormone disorders are phaeochromocytoma, glucagonoma, Cushing's syndrome and acromegaly.
- **Somatostatin** (GHIH) is a multipotent hormone inhibitor consisting of a disulphide bridge and 14 amino acid units.

Essentials

This paragraph covers *1. the blood glucose regulation in the fed state*, as well as in *2. the fasting states*. Also *3. The endocrine pancreas*, and *4. Pancreatic exocrine control* is dealt with.

1. Glucose regulation in the fed state

In the *absorptive state* after a balanced meal, nutrients enter the blood and lymph from the gastrointestinal tract (as monosaccharides, triglycerides, and amino acids). All the blood passes directly to the liver, which converts most of the other monosaccharides into *glucose*. Much of the absorbed carbohydrate enters the liver cells, but little of it is oxidised; instead most is stored as glycogen. Absorbed glucose, which did not enter hepatocytes but remained in the blood, is stored as glycogen by muscle cells, or it may enter into adipose tissue. A large fraction is oxidised to CO_2 and water in the various cells of the body. Glucose is the major source of energy during the absorptive state. Homeostatic mechanisms maintain the plasma [glucose] within narrow limits in healthy humans, so that the energy needs during the postabsorptive state can be met by stored fuel.

A high glucose intake results in a high blood [glucose] or extracellular *hyperglycaemia*. Hyperglycaemia increases insulin secretion from the b-cells and inhibits glucagon secretion from the a-cells of the pancreatic islets. These hormones block *hepatic glucose production* by glycogenolysis and gluconeogenesis. *Insulin secretion* dominates over all *insulin-antagonists* (growth hormone, glucagon, cortisol and some catecholamines).

The sight and the smell of a meal triggers *cephalic* insulin secretion. When the meal reaches the intestine, several peptides of the *incretin family* are released; this is the *intestinal* secretion phase. Typical representatives of the *incretin family* are Gastric Inhibitory Peptide (GIP), glicentin (intestinal glucagon), and glucagon-like peptides (GLP-1 and -2). Incretins strongly potentate the insulin secretion induced by the rising blood [glucose]. The *incretins* cause a much larger insulin secretion than the iv. administration of glucose, even at the same rise in blood [glucose]. This extra insulin secretion is called the *incretin effect*. The insulin released following a meal increases the storage rate of glucose-related energy in the liver, muscles and fat tissues. The storage effect is much larger than when glucose is administered intravenously.

Glucose is absorbed through the luminal membrane of the intestinal cells in *glucose-Na^+ transporter proteins*. The two substances pass through the basolateral membrane via separate routes: Glucose passes in a special glucose-transporter, and Na^+ is transferred by the Na^+-K^+-pump. Glucose transport proteins and insulin receptors are described in Chapter 1.

The filtration flux for glucose (mmol per min) increases proportionally to the concentration in the blood (as for all other filtered substances). Normally, all glucose is reabsorbed in the first part of the proximal renal tubules with a T_{max} of 1.8 mmol per min or 320 mg per min.

In other words, the passage fraction falls from one to zero already halfway through the proximal tubules. The excretion flux for glucose is *zero* in healthy humans.

Glucose appears in the urine of diabetics, who have a blood [glucose] exceeding the *appearance threshold* (10 mM).

Reabsorption of glucose over the luminal membrane of the proximal tubule cell takes place through the *glucose-Na^+ transporter*.

2. Glucose regulation in the fasting state (rest and exercise)

We keep our blood [glucose] surprisingly constant around the fasting level, considering the wide variety of daily activities.

Glucose production (gluconeogenesis and glycogenolysis) must equal *glucose combustion* in the fasting state. Thus, a precise relation must exist between the secretion of insulin and glucagon from the pancreatic islets.

In the fasting state hepatic *glycogenolysis* produces most of the glucose and the remaining glucose is produced by *gluconeogenesis*. Hepatic glucose is produced by *glycogenolysis* (glycogen breakdown to glucose) and by *gluconeogenesis* (glucose formation from glycogenic amino acids, lactate, glycerol, and small amounts of pyruvate. Muscle glycogen cannot deliver glucose to the blood, since muscle tissues lack *glucose-6-phosphatase*.

In the fasting condition a healthy adult has a blood [glucose] of *4.5-5.6* mM. With an average [glucose] of 4 mM in 15 l of extracellular fluid volume (ECV), therefore, the total glucose content in ECV is 60 mmol or 10.8 g (2 teaspoons of sugar). This amount is equal to the glucose eliminated in one hour at rest (60 mmol each hour), but during maximum exercise that same person may use *five* times more glucose.

The CNS and the erythrocytes neither synthesise nor store glucose, which is their primary fuel. Any surplus of glucose is deposited as liver and muscle *glycogen*. The liver cells contain an especially *efficient glucose transporter* (GLUT 2), and its *glucose uptake rate* cannot be increased further by insulin or by other hormones.

Any small *fall* in blood glucose releases more glucagon. During fasting *glycogenolysis, gluconeogenesis, and lipolysis* are dominant. If a normal person does not eat for 24-48 hours the CNS cells revert to combustion of ketone bodies, and a reversible condition that is similar to diabetes develops (*hunger diabetes*).

Exercise

The exercise stress on the hypothalamus increases the activity of the *sympathoadrenergic system,* which includes increased secretion of adrenaline from the adrenal medulla. Sympathoadrenergic activity inhibits the *insulin secre* from the b-cells. Sympathetic activity also increases hepatic glucose production.

We tend to increase glucagon secretion only if the blood [glucose] falls. A slight fall in blood [glucose] can occur both during exercise bouts and during prolonged exercise.

During exercise the blood [glucose] is maintained rather constant by bihormonal interplay between insulin and glucagon.

Generally, an increasing demand of more energy elicits increased glycogenolysis, lipolysis, and increased gluconeogenesis caused by insulin-antagonistic hormones (catecholamines, glucagon, cortisol, and growth hormone).

Adrenaline inhibits the insulin and stimulates the glucagon secretion, so the blood [glucose] increases.

Somatotropin - human growth hormone (GH) - is an *insulin-antagonist,* but together with insulin probably the most important *anabolic hormone.*

Conditions where energy sources are lacking are hypoglycaemia, hunger, fasting state, exhaustion, and stress. These conditions trigger the release of GRH from the hypothalamus, which in turn stimulates the release of GH from the hypophysis. This hormone has a *tropic effect* on the a -cells of the pancreatic islets. GH releases glucagon from these cells, just as sympathetic stimulation from the hypothalamus does.

GH increases blood [glucose] by increasing *hepatic glucose production* (glycogenolysis but not its gluconeogenesis) and by inhibiting the insulin sensitivity of the muscle cells and thus reduces their glucose uptake. GH also has the same effect on fat cells, mobilising fatty acids and glycerol. GH stimulates protein synthesis, mitoses, chondrogenesis, ossification, and phosphate balance, while increasing glycolysis (ie, anaerobic breakdown of glycogen).

Glucocorticoids are insulin-antagonists. They stimulate the *hepatic glucose production* (glycogenolysis and gluconeogenesis) but inhibit the cellular glucose uptake. Glucocorticoids are *permissive* and *potentiating* for catecholamines and glucagon.

Catecholamines (adrenaline & noradrenaline) are *insulin-antagonists.* Adrenaline *stimulates hepatic glucose production* (glycogenolysis). Catecholamines also stimulate lipolysis. The increase in mitochondrial oxygen uptake by T_3/T_4 is potentates by catecholamines.

The glucostat

Glucose sensitive neurons in the hypothalamus (the *glucostatic centre*) react to hypoglycaemia by releasing glucagon from the pancreatic a-cells and catecholamines from the adrenal medulla by action of the sympathetic system.

The glucostatic centre also reacts to hyperglycaemia to release insulin from pan
glycogen synthesis by vagal stimuli. Insulin promotes the entry of glucose into tissues, including the neurons of the hypothalamic glucostatic centre (but in no other brain neurons). A balanced blood [glucose] is achieved by sympathetic signals stimulating hepatic glucose production. This balance theory is called the glucostatic theory. In the glucostatic theory the hypothalamus is considered a glucostat and the liver is a unique glucose exchanger, due to the portal system and the hepatic glucose-6-phosphatase. Leptin is dealt with in Chapter 20.

Since the hypophysis hormones ACTH and GH are insulin-antagonists the net effect of the hypophysis, when not balanced by a normal pancreatic insulin secretion, is a reduced glucose tolerance.

3. The endocrine pancreas

The endocrine pancreas or the pancreatic islets are synonyms for the produc
Glucagon, insulin, somatostatin, and pancreatic polypeptide (PP).

The one million islets of Langerhans are discrete structures scattered throughout the pancreas, but which only comprise 1% of its total weight. The islets are arranged along fenestrated capillaries, so that the hormones can pass easily to the portal blood. The islets of Langerhans receive both sympathetic (adrenergic) and parasympathetic (cholinergic) fibres.

The membranes of the islet cells contain gap junctions between neighbour cells, so hormones from one cell can act on its neighbour (paracrine action). Gap junctions allow passage of small molecules from one islet cell to its neighbour. In many pancreatic lobules, the a - b - and d - cells form a paracrine syncytium.

3 a. Glucagon

The a-cells of the pancreatic islets is the source of pancreatic glucagon. Glucagon stimulates adenylcyclase in the hepatocytes. This enzyme activates phosphorylase that breaks down glycogen. Actually, glucagon triggers a glycogenolytic cascade, so those considerable amounts of glucose are released in response to the fall in blood glucose. In addition, glucagon stimulates the hepatic production of glucose (gluconeogenesis) from glycerol, alanine and lactate. Glucagon is a direct antagonist to insulin, being catabolic in its actions (gluconeogenetic, glycogenolytic, lipolytic & ketogenic, and deaminating amino acids).

Intestinal glucagon (glicentin) is built up from 69 amino acids. The glucagon from the a-cells of the pancreatic islets only contains 29 of the 160 amino acid residues in pro-glucagon. Conditions where there is intracellular lack of glucose (hunger, insulin deficiency, protein rich meals, and amino acid infusion) liberate glucagon from the a-cells of the pancreatic islets to the pancreatic vein and then to the portal vein. Glucagon stimulates ketogenesis (formation of ketone bodies). High blood [glucose] and [FFA] inhibit glucagon secretion.

Pancreatic and intestinal (glicentin) glucagon are hepatic insulin-antagonist. Glucagon stimulates hepatic glucose production by glycogenolysis in the hepatocytes and thus increases the blood [glucose].

Glucagon also stimulates gluconeogenesis from glycogenic amino acids in the liver and thus increases urea-genesis. Glucagon stimulates ketogenesis (formation of ketone bodies). In addition to the ketogenic effect, intestinal glucagon is a potent stimulator of insulin secretion - as are other members of the incretin family. Incretins act by increasing cAMP in the b-cells.

3 b. Insulin

Banting shared the Nobel Prize with Macleod in 1923 for their work in identifying the role of insulin in the carbohydrate metabolism. Their research led to the practice of insulin therapy for diabetes.

Pre-proinsulin is the precursor of insulin. When pre-proinsulin reaches the endoplasmic reticulum, enzymes separate

the molecule from the signal molecule, to form proinsulin. In the Golgi apparatus enzymes cleave proinsulin to insulin (51 amino acids in two chains: A and B) and the C peptide (Connecting peptide). Insulin and C peptide are wrapped in the same secretion granule. The content of these secretion granules is expelled from the cell by exocytosis.

When the secretory granules release proinsulin to the portal blood and later the extracellular fluid volume (ECV), connecting peptide (C-peptide) and two amino acids breaks off. The split products are carried to the liver, where half of the insulin molecules are degraded and extracted. The degradation products are broken down and eliminated by the kidney. The kidneys only eliminate c-peptide and its rate of production reflects the rate of insulin secretion. Insulin contains 51 amino acid residues in two chains (m.w. 5734).

In healthy persons the blood glucose concentration, B-[glucose], is controlled exactly. The fasting value is within the range of 4-7 mM, with minimum individual variance from day to day, despite varying life conditions with food and exercise.

The liver is a glucose exchanger, because it absorbs glucose from the intestine, stores glucose as glycogen, and produce glucose from fat and protein residues (gluconeogenesis). The liver releases glucose to the ECV in exact proportion to the peripheral rate of glucose utilisation in the postabsorptive state (180-200 g per day).

The brain metabolism of a healthy standard person requires 100 g of glucose per day. The brain glucose is totally oxidised, ir-regardless of the insulin status.

Each meal elicits a peak of insulin secretion, because of the rise in blood [glucose] (Fig. 27-1). The blood [glucose] increases after a meal. Increasing [glucose] is a strong stimulus to the b-cells of the pancreas. Glucose enters these cells through GLUT 2. The cells empty their granules into the ECV, and the granule dissolve immediately after entering the blood. This sequence of events supplies the blood with insulin, C-peptide and proinsulin in the ratio 19:19:1.

Insulin reduces the blood glucose for the following reasons: insulin increases the cell uptake of glucose (and potassium) in most tissues (adipocytes, heart and other muscle tissue). The exceptions are the brain, kidney and erythrocytes. The uptake capacity for glucose in hepatocytes is so large, that any insulin effect is immaterial.

Insulin promotes the formation of tissue stores from circulating nutrients (the actions are all anabolic). The insulin receptor is a tetrameric protein complex with two a-units extracellularly, and two b-units traversing the membrane of target cells (ie, skeletal muscle fibres, cardiac myocytes and adipocytes). The target cells contain the glucose transporter, GLUT-4. In the absence of insulin, all the GLUT-4 units are located in the intracellular vesicles. Insulin binding to the insulin-receptor activates the tyrosine kinase, which resides in the b-units, and promotes the transport of GLUT-4 vesicles towards the surface of the cell, where they melt together with the membrane. This phenomenon increases the number of glucose-channels through the membrane and thus promotes glucose uptake into target cells. Finally, the insulin-receptor complex is internalised by the cell, insulin is broken down and the insulin-receptor recycles to the cell surface for further use.

Glucose is stored in the muscle cells as glycogen, used in the Krebs cycle or it is broken down to lactate.

In the fat cells glucose is utilised as a substrate for triglycerides synthesis. In the post-prandial phase, lipolysis liberates fatty acids (FFA) from triglyceride together with glycerol. Glycerol and lactate are substrates for hepatic gluconeogenesis.

Fig. 27-1: Effects of insulin on target cells.

Falling blood [glucose] is called hypoglycaemia, which activates the sympathoadrenergic system and deplete the glucose-dependent brain for its only fuel. Accordingly, the hypoglycaemia causes sympathoadrenergic (sweating, hunger, tremor, tachycardia), and cerebral manifestations (anxiety, disorientation, cramps, and unconsciousness). The clinical picture is that of hypoglycaemic shock.

Insulin also increases the rate of glycogen synthesis in the liver and muscles and inhibits the rate of gluconeogenesis. Insulin is a direct antagonist to glucagon being anabolic in its actions (increased glucose entry to cells, increased glycogen and lipid synthesis, decreased protein catabolism and ketogenesis).

Glucose-evoked insulin secretion is the result of a chain of events in the pancreatic b-cell (Fig. 27-2).

Fig. 27-2: Insulin release from pancreatic β-cell.

1. The glucose uptake takes place through a specific transporter protein (GLUT-2). The pancreatic b-cell membrane contains several K^+ channels, of which two are directly involved. This is the K^+-ATP channel and the maxi-K^+ channel (Fig. 27-2).

2. The hyperglycaemia accelerates the glucose uptake and metabolism and thus increases the ATP/ADP ratio.

3. Increased ATP closes the K^+-ATP channels, so the cell depolarises (hypopolarises). During hypopolarisation from the normal resting membrane potential of -70 mV, a threshold is reached at - 50 mV, where the voltage dependent Ca^{2+} channels open.

4. The Ca^{2+} influx triggers exocytosis of insulin and C-peptide containing granules following vesicular fusion with the cell membrane.

5. Normally, the maxi-K+ channel and other K+ channels stop depolarisation. When intracellular [Ca^{2+}] and [K^+] has increased, it opens the maxi-K^+ channel. The K^+ efflux restores the resting membrane potential (- 70 mV) towards the equilibrium potential of K^+ (-100 mV).

Insulin is a vital hormone. Blood from the pancreas passes through the liver, where insulin promotes the produc glycogen from the recently absorbed glucose. The liver destroys a substantial amount of the insulin, whereas the C-peptide passes the liver undisturbed. The plasma [C-peptide] is thus a good estimate of insulin secretion. Insulin can now be synthesized from genetically modified micro-organisms.

Insulin is an anabolic hormone.

Insulin reduces the blood [glucose] because it increases glycogen synthesis in the liver and muscles. Insulin increases the uptake of glucose through GLUT 4 (in adipocytes, heart and skeletal muscles). Insulin inhibits the gluconeogenesis from glycogenic amino acids in the liver.

Insulin promotes the storage of energy, the synthesis of glycogen, mRNA and proteins. Insulin thus reduces urea-genesis.

Insulin promotes lipogenesis in the fat stores; however, it inhibits lipolysis. It may be noted that the glycerol portion of the triglyceride molecule is a derivative of glucose.

Insulin increases the synthesis of cholesterol in the liver, in particular the rate of VLDL formation (Very Low Density Lipoprotein). The dangerous cholesterol fraction is LDL (Low Density Lipoprotein).

Insulin increases the GLUT 4 transfer of glucose and K^+ into the muscle cell interior.

Certain major tissues (kidney, brain, and intestine) are insensitive to the direct action of insulin.

3 c. Somatostatin

D-cells or d-cells are the source of somatostatin, a potent and multipotent hormone inhibitor. Somatostatin contains a disulphide bridge and 14 amino acid molecules. Somatostatin produced in the islets inhibits the local secretion of the other islet hormones, while glucagon stimulates the local release of insulin and somatostatin. Somatostatin is also produced in the hypothalamus, where it functions as the Growth Hormone Inhibiting Hormone (GHIH). Pancreatic somatostatin is released in response to high blood [glucose] and [alanine]. Somatostatin inhibits the secretion of the gastrointestinal tract (but not its motility) and functions as a synaptic transmitter in the CNS. Persons with somatostatin-producing tumours develop diabetes and gallstones.

GHIH is synthesized both in the hypothalamic-pituitary system and in the pancreatic islets. The D-cells of the pancreatic islets of Langerhans produce GHIH, which controls the function of the other islet cells by paracrine action. Somatostatin is a multipotent hormone inhibi
glucagon. Somatostatin blocks the gastrin secretion in the gastric antrum. Somatostatin inhibits the secretion of

digestive fluids, but increases gastrointestinal motility.

3 d. Pancreatic polypeptide

The cells responsible for pancreatic polypeptide (PP) secretion are particularly abundant in the head of the pancreas. PP contains 36 amino acid residues in a linear polypeptide. The plasma [PP] increase markedly after a protein rich meal, but it is not released by alanine infusion. The PP secretion is increased by exercise (with high plasma [alanine]), by fasting and by hyperglycaemia. The plasma [PP] is suppressed by glucose infusion. PP inhibits the exocrine pancreas and reduces the gallbladder contractions. This is an appropriate response during exercise and fasting, where any reduction in blood glucose would trigger a PP release.

Meals, rich in protein and fat, release pancreatic polypeptides (from the PP-cells).

Pancreatic polypeptide inhibits both enzyme secretions from the pancreas and the emptying of bile into the small intestine. This leads to a delay in the absorption of nutrients including glucose.

Patients with pancreatic islet cell neoplasm have elevated plasma [PP].

4. Pancreatic exocrine control

Endocrine glandular tissue is localised in the Langerhans islets that produces insulin in the b-cells, glucagon in the a-cells, somatostatin (GHIH) plus gastrin from the d- and G-cells, and pancreatic polypeptide (PP) from the P-cells (Fig. 27-3). Bombesin, galanin, and neuropeptides are present in pancreatic neurons and act as transmitters.

Stimulation of vagal fibres to the pancreas enhances the rate of enzyme secretion into the pancreatic juice. Stimulation of sympathetic fibres reduces bloodflow to the pancreas, and thus inhibits pancreatic secretion. Gastrin enhances enzyme secretion and insulin potentiates the effect, whereas somatostatin inhibits secretion from both acinar and ducts cells (Fig. 27-3: - all).

Fig. 27-3: Liberation of pancreatic islet hormones.

Pathophysiology

Diabetes mellitus (DM)

DM is a collective term for a multitude of metabolic disorders, where lack of insulin (type I diabetes) or insulin resistance (type II) dominates.

Insulin resistance is defined as insufficient sensitivity to insulin.

The diabetic condition is characterized by an abnormal glucose tolerance that is documented by the use of a glucose tolerance test.

Besides hyperglycaemia, the overall phenomena in the diabetic condition are *protein depletion* and *increased lipolysis* with deposition of lipids in the vascular walls of the brain, heart, kidneys, eyes and muscles.

All cases of DM where the cause is not fully explained are termed *primary* DM, whereas *secondary* DM is explainable and sometimes directly curable.

Secondary DM is caused by hypersecretion of one or more of the many catabolic hormones with hyperglycaemic effect (adrenaline, noradrenaline, glucagon, glucocorticoids and growth-hormone, HG) or by total destruction of the pancreas from pancreatitis or carcinoma. The hormone disorders are phaeochromocytoma, glucagonoma, Cushing's syndrome and acromegaly.

Most of the patients with primary DM can be classified into the two groups already presented above Insulin-Dependent DM (IDDM) or type I diabetes, and Non-Insulin-Dependent DM (NIDDM) or type II diabetes.

This paragraph deals with *1. IDDM*, *2. NIDDM*, *3. Insulin shock (hypoglycaemia)*, *4. Oral and intrveneous glucose tolerance tests*, *5. treatment of diabetes mellitus*, *and 6. Summary of the diabetic condition.*

1. IDDM or Type I diabetes

The presentation of IDDM is typically a young person with a few week history. This is a serious, life threatening metabolic disease, where continuation of life depends upon insulin treatment. The first treatment with insulin took

place in Canada (1922). Until then these patients died within half a year in ketoacidotic coma. Persons, often with hereditary predisposition, are suddenly attacked by autoimmune destruction of *all* b-cells in the pancreatic islets, which results in the complete absence of insulin. This autoimmune destruction occurs more often in populations, where breast feeding is unpopular, and protein-rich cow's milk is used generally.

Lack of insulin leads to extracellular hyperglycaemia, and increased lipolysis.

The classical triad is:

1. Polyuria (osmotic diuresis due to extracellular hyperglycaemia and glucosuria).
2. Polydipsia and thirst due to the loss of salt and water.
3. Weightless due to extracellular fluid volume (ECV) depletion and the breakdown of tissue stores (ie catabolic effect of insulin deficiency) with rapid wasting.

The intracellular lack of glucose activates glycogenolysis in the liver and muscle cells; lipolysis and muscular proteolysis is accelerated. The liberated free fatty acids (FFA) are converted to ketone bodies, whereby a metabolic acidosis (ketoacidosis) is produced.

A patient with a blood (glucose) above 25 mM loses consciousness and to such degrees that contact is impossible and reactions to pain are absent (coma).

Diabetic ketoacidosis is a condition of insulin deficiency causing increased hepatic ketogenesis. This condition of uncontrolled catabolism occurs in IDDM, only.

In a young person, the first sign of IDDM can be coma with diabetic ketoacidosis - a life-threatening condition, if the patient is alone. This is particularly likely, when the patient is under the stress of intercurrent illness such as infection with high fever. Also a recognized IDDM patient can be hit by ketoacidosis during intercurrent illness, where his insulin demand is increased, or the patient may take too little insulin because of lost appetite or for any other reason.

Insulin deficiency has two consequences. First of all, the hepatic glucose release accelerates, and secondly the uptake of glucose by muscle and fat cells in the periphery is reduced. Progressive hyperglycaemia causes osmotic diuresis with loss of salt and water. The abnormally low ECV is termed the dehydrate state, and with falling blood volume also renal bloodflow falls. Insulin deficiency also accelerates the lipolysis (Fig. 27-4).

Fig. 27-4: The development of ketoacidosis during insulin deficiency (FFA = free fatty acids).

Triglycerides are liberated from adipose tissue, and the concentration of free fatty acids (FFA) in the blood is elevated. FFA are broken down to fatty acyl carnitine within the liver cells, and this molecule is converted to acetyl CoA, which in turn reach the mitochondria, where ketone bodies (aceto-acetate, acetone, b-hydroxybutyrate) are formed (Fig. 27-4).

The breath of the patient smell by acetone, and there is ketosis in the urine. The concentration of ketone bodies in the blood passes *5 mM*, and when pH falls below 7, there is life-threatening or terminal coma. The condition is called an acute metabolic acidosis, characterized by a negative base excess. The patient tries to compensate the metabolic acidosis by hyperventilation, so-called Kussmaull-breathing.

Most patients with recent IDDM have circulating antibodies to islet cells, and tend to develop other organ-specific autoimmune disorders (Addison disease, Hashimotos thyroiditis and pernicious anaemia). Identical twins show a 40% concordance in developing IDDM, so life-style must also play a role. There is an association with HLA-DR4, if also HLA-B8 or HLA-DR3 is present.

2. NIDDM or Type II diabetes

This is a frequent type of DM in particular in populations with a sedentary life style and obesity. The incidence increases with age and development of obesity, which is reflected in the name *maturity-onset diabetes*. The prevalence is high in Afro-Caribbean and Asian population groups. The onset of NIDDM is sometimes triggered by Intercurrent illness or by pregnancy, but not by immunological reactions.

NIDDM is a complex of polygenic disorders. Certain families show an autosomal dominance. the genetic defects differ and many mutations are known. One is the gene on chromosome 7, which code for *glucokinase.* Identical twins

show almost absolute concordance in development of NIDDM. The much more frequent *type II diabetes* (maturity-onset) is the result of *insulin resistance* and b-cell defects. Type II diabetes also occurs in younger persons, especially in persons with a *high fat-low muscle mass*. A strong genetic element is always present, but *inactivity* and *stress* (an inactive life style with a low endurance capacity) seems to be involved in the development of type II diabetes. - Lack of exercise predisposes one to obesity, a condition that greatly decreases *insulin sensitivity* of the target cells (adipocytes, heart and skeletal muscle tissues). Reduced glucose combustion creates hyperglycaemia. The hyperglycaemia elicits insulin secretion from defective b-cells in some patients, resulting in raised serum [insulin]. Since insulin is present, the acute complications such as ketonaemia and metabolic acidaemia (often found with IDDM) are rare in these patients. The high serum [insulin] may further down-regulate the activity of their *insulin receptors.* The insulin secreted in NIDDM patients does not increase the uptake of glucose as in normal persons. Many NIDDM patients need much more insulin for a given test effects than IDDM patients and healthy people.

An inactive life style for years, with redundant food intake, seems to be involved in the development of NIDDM in persons with a genetic predisposition. Lack of regular physical activity with development of overweight, increases the incidence if NIDDM. The *impaired glucose tolerance* is demonstrated by a *glucose tolerance test*.

The insulin secretion is abnormal in patients with NIDDM, although they typically possess half of their *β-cell mass* at autopsy. The destroyed b-cells is filled with amyloid material (islet amyloid polypeptide, IAPP). IAPP is a possible antagonist to insulin, and explain some cases of *insulin resistance*.

Many older patients with NIDDM have no symptoms, but a routine examination reveals glucosuria or a raised blood [glucose]. Other patients are tired, have minor genital infections or sugar spots on their underwear. Some patients present with established *late-complications* such as retinopathy (blindness), nephropathy, arteriosclerotic disorders (cerebrovascular insults, myocardial infarction, intermittent claudication, gangrene), susceptibility to infection or neuropathy.

NIDDM can be caused, theoretically, by *1. β-cell defects* including genetic defects, resulting in abnormal insulin production, or by *2. target cell defects* including receptor failure. The possible defective sites in 1. and 2. have one common denominator. They are all *key proteins* (hormone, receptors and transporters). Muscular activity is required to stimulate the normal production of *key proteins*. NIDDM relates to inactive life style.

The basic problem is therefore possibly a genetically and activity dependent *defect in key protein production* in the cell interior. Actually, a genetic defect has just been demonstrated at certain steps of insulin action in a subset of patients of late-onset NIDDM.

3. Insulin shock (hypoglycaemia)

A high blood [insulin] will cause tissues to store away the available blood glucose rapidly, mainly through muscular GLUT 4, and stop simultaneously the production of new glucose.

A low blood glucose level elicits a large secretion of glucagon to the portal blood. Glucagon is the most important insulin-antagonist. Glucagon increases hepatic glucose production (enhancing glycogenolysis, gluconeogenesis and protein breakdown). Low glucose levels trigger glucagon production, even from denervated, pancreatic a- cells; hence, they must be glucose sensitive.

An increased catecholamine secretion from sympathetic nerve endings (NA) and Ad from the adrenal medulla (elicited from the hypothalamic glucostat via the sympathetic nervous system) helps within minutes to compensate, as catecholamines stimulate glycogenolysis, increase lipolysis and inhibit peripheral glucose uptake. Hours later, cortisol and GH also contribute. An appropriate rise of plasma [cortisol] in response to insulin-induced hyperglycaemia documents an intact CRH-ACTH-adrenal axis, and this is the most widely used stress test.

Adrenergic effects, such as trembling fingers, tachycardia, and muscular stiffness warn the hypoglycaemic patient. The glucose consumption by the heart and brain continues. The lack of glucose in the brain makes the patient uneasy at first; he is then inse
and denies with slurred speech to take glucose.

Blood [glucose] below 2.5 mM elicits hypoglycaemic shock with loss of consciousness (somnolence, sopor or coma), universal cramps and respiratory stop (Fig. 27-5).

Intravenous injection of glucose (50%) is the rational therapy for hypoglycaemic coma. The patients wake up almost immediately, and are then often in a hyperglycaemic state.

Fig. 27-5: Consequences of hypoglycaemic shock.

The b-cell defects are insufficiently described. Type 2 diabetics do not produce sufficient levels of ATP in the pancreatic b-cells to completely block the K^+ -ATP channels (Fig. 27-2). Thus, the b-cell does not hypopolarize adequately in response to hyperglycaemia. Therefore, the voltage dependent Ca^{2+} -channels is insufficiently activated, and intracellular [Ca^{2+}] do not increase enough to trigger the insulin exocytosis needed. Sulfonylurea compounds close the K^+-ATP-channels and thus help to treat type 2 diabetes.

4. Oral and intravenous glucose tolerance tests

The test is performed orally or intravenously.

Oral test. The patient drinks a glucose solution containing 75 g glucose within four minutes (WHO). The blood concentration in venous plasma ([glucose]) is followed over 3 hours by blood sampling.

Normal individuals start from a low fasting [glucose] such as 5.5 mM or less. The blood [glucose] peaks after one hour and *returns to normal within two hours* (less than 6.7 mM in Fig. 27-5). Persons with *impaired glucose tolerance* start with a fasting level less than 7.8 mM, and after 2 hours the level is 7.8-11.1 mM.

Fig. 27-5: Oral glucose tolerance curves for a normal person, a diabetic, a patient with hyperthyroidism and a patient with myxoedema.

The *diabetic patient* typically starts from a high fasting [glucose], such as values above 7.8 mM, and increase to a very high level. The blood [glucose] is not back to normal within two hours, but stays above 11.1 mM. This test pattern is the clinical WHO criterion of diabetes (Fig. 27-5).

A patient with hyperthyroidism (Graves disease or Morbus Basedowii) has a rapid intestinal absorption and a rapid combustion of glucose. The myxoedematous patient has a slow absorption and a slow combustion of glucose.

Fig. 27-6: Results of i.v. glucose tolerance tests from a normal person and from a diabetic.

Intravenous (i.v.) test. We inject 25 g glucose intravenously over a period of 4 minutes. We then measure the blood [glucose] every 10 min for at least an hour in order to determine the half-life from a semi-logarithmic plot (Fig. 27-6). The metabolic combustion rate for glucose is exponential, so it is easy to calculate the metabolic rate constant (k) expressed in percentages.

Note that the *metabolic rate constant k* here is the amount of glucose combusted divided by the total amount of glucose in a mainly extracellular distribution volume. The half-life ($T_{1/2}$) is equal to 0.693/k.

All glucose combustion rates *above 1.2% per min* are normal.

Fig. 27-7: Intravenous Glucose Tolerance Test

5. Treatment of diabetes mellitus

A normal person with three meals per day will have *three* peak concentrations of glucose and insulin in his blood. It is possible to obtain such a time profile in a diabetic person by the following strategy. Inject a fast-acting insulin three times a day just before meals and a slow-acting insulin at night. This is the *physiologic* principle. The aim of this procedure is to reduce the number of acute and chronic complications for diabetics.

All diabetics are recommended to eat healthy just like anyone else. When diet alone is insufficient to achieve a satisfactory blood glucose, a slim type II diabetic is treated with sulphonylurea compounds. The obese type II diabetic is treated with a biguanide called metformin. Patients, who have been in metabolic acidosis, are usually treated best with insulin.

6. Summary of the diabetic condition

A poorly controlled diabetic condition leads to extracellular hyperglycaemia, glucosuria, metabolic acidosis, polyuria

(osmotic diuresis), dehydration and polydipsia. The osmotic diuresis leads to the excretion of Na and water, which results in Na^+ and ECV depletion.

Intracellular lack of glucose activates glycogenolysis in the liver and muscles, and accelerates muscular proteolysis and lipolysis. This liberates free fatty acids, which are converted to ketone bodies.

A patient with hyperglycaemia above 25 mM loses consciousness to such a degree that contact is impossible (ie, coma).

The increased rate of cholesterol production increases the occurrence of atherosclerosis and of diabetic nephropathy. Albuminuria, hypertension and low GFR characterise diabetic nephropathy.

Self-Assessment

Multiple Choice Questions

I. Each of the following five statements have True/False options:

A. Peptides and protein hormones are lipophobic.

B. Infertility is a diagnosis used on a couple, which have been unable to conceive during one year of unprotected intercourse.

C. Chronic hypoadrenalism is also called Addison's disease.

D. Nephrogenic diabetes insipidus, is a condition where the renal cells are resistant to ADH.

E. Glycerol and lactate are substrates for hepatic gluconeogenesis.

II. Each of the following five statements have True/False options:

A. Receptors are frequently glycosylated, so one signal molecule linked to a receptor is always enough to elicit a response.

B. The incidence of atherosclerosis is correlated to the total [cholesterol], the [LDL], and the [LDL]/ [HDL] ratio in blood plasma.

C. Oestrogens, exercise, nicotinic acid and alcohol increase the plasma-HDL.

D. Omeprazole stimulates the luminal gastric proton-pump.

E. Prostaglandins are primarily paracrine hormones, which act through receptors linked to G-proteins.

Case History A

A 19 year old male, body weight 80 kg, was suddenly complaining of fever during his work and ordered home to bed. The patient was living alone. Fortunately a colleague visited him the next morning. He had to break the door down and found the patient unconscious. The patient arrived at the hospital in deep coma. A blood sample from the radial artery showed the following results: β-cell antibodies, P_{aCO_2} 27 mmHg, pH 7.21, actual bicarbonate 10.5 mM, O_2 saturation 0.96 and [glucose] 32 mM (5.75 g/L). The Base Excess is -15 mM in the extracellular fluid volume (see Fig. 17-12). The urine contained glucose and ketone bodies.

The patient's breathing was deep and fast, his heart rate was 115 beats/min, and his blood pressure was 90/55 mmHg. The mucous membranes of the mouth were dry and the tonsils were enlarged and infected. The rectal temperature was 39.9 Centigrade.

1. Explain the condition of the patient concerning thermo-balance, carbohydrate metabolism, acid-base balance

and fluid balance.

2. Describe a rational treatment of the four homeostatic disturbances mentioned in 1.

3. Following eight hours of treatment the blood pH was 7.41 and P_{aCO_2} 42 mmHg, but the patient is still hyperventilating. Explain why.

4. Following 24 hours of treatment all blood gas values were normal and the patient was resituated. Why did the patient stop to hyperventilate?

Case History B

A female of 49 years and with a height of 1.52 m is in hospital due to obesity and related problems. Her weight is 74 kg. She has developed skin mycoses and multiple boils. In the morning (fasting state) a blood sample shows a blood [glucose] of 7.4 mM, but glucose is not found in the urine.

At the hospital her total body water is measured following intravenous injection of radioactive water (5.5×10^7 Becquerel or Bq tritium water). Her bladder is emptied at the time of injection. Two hours later the bladder is drained for 95 ml of urine with a concentration of 1 598 400 Bq per l.

At this time the indicator is evenly distributed in the total body water with a concentration of 1 520 700 Bq per l. The amount of indicator lost in the urine is 2/3 of the total loss.

1. What is the principle for estimation of total body water?

2. Calculate the total body water in litres and in fraction of her body weight.

3. Is this a normal result?

4. The patient is obese, but is this a serious overweight?

5. Does she have symptoms and signs of complications?

Case History C

A 23-year old male was saved after 30 days in the ruins of a house following earth quake. There was no food but sufficient water. At the arrival to the hospital the patient was in syncope with frequent, deep respiration, and the expired air smelled of acetone. The skin was dirty with brown pigmentation. The cardiac rate was 85 bpm, and the arterial blood pressure was 11.3/7.3 kPa (85/55 mmHg).

The blood [glucose] was 2.2 mM, and the plasma [FFA] was increased. The serum concentrations of proteins and essential amino acids were reduced. The blood [haemoglobin] was 95 g l^{-1}. There was moderate antidiuresis with ketonuria with signs of water retention and a high nitrogen loss in the urine.

The patient was treated with parenteral administration of glucose, amino acids and electrolytes. Following the glucose intake, the blood [glucose] was increased to 10 mM, and glucosuria occurred. A glucose tolerance test was performed and resulted in a high blood [glucose] level that had not reached the normal level within 2 hours.

1. Describe the energetic events leading to survival.
2. Why did the patient smell of acetone?
3. What happened to the carbohydrate metabolism of the patient?
4. Explain the high nitrogen loss in the urine.

See answers

Highlights

- *Glucose is absorbed through the luminal membrane of the intestinal cells in glucose-Na^+ transporter proteins. The two substances pass through the basolateral membrane via separate routes: Glucose passes in a special glucose-transporter, and Na^+ is transferred by the Na^+-K^+-pump.*
- *Somatotropin - human growth hormone (GH) - is an insulin-antagonist, but together with insulin probably the most important anabolic hormone.*
- *Glucose sensitive neurons in the hypothalamus (the glucostatic centre) react to hypoglycaemia by releasing glucagon from the pancreatic α-cells and catecholamines from the adrenal medulla by action of the sympathetic system.*
- *Since the hypophysis hormones ACTH and GH are insulin-antagonists the net effect of the hypophysis, when not balanced by a normal pancreatic insulin secretion, is a reduced glucose tolerance.*
- *The endocrine pancreas or the pancreatic islets are synonyms for the produc*
  hormones: Glucagon, insulin, somatostatin, and pancreatic polypeptide (PP).
- *Insulin is synthesized as proinsulin, which is stored in granules close to the cell membrane of the β-cells of the pancreatic islets. When the secretory granules release proinsulin to the portal blood and later the extracellular fluid volume, connecting peptide (C-peptide) and two amino acids breaks off.*
- *A poorly controlled diabetic condition leads to extracellular hyperglycaemia, glucosuria, metabolic acidosis, polyuria (osmotic diuresis), dehydration and polydipsia. The osmotic diuresis leads to the excretion of Na^+ and water, which results in Na^+ and ECV depletion.*
- *Intracellular lack of glucose activates glycogenolysis in the liver and muscles, and accelerates muscular proteolysis and lipolysis. This liberates free fatty acids, which are converted to ketone bodies.*
- *A patient with hyperglycaemia above 25 mM loses consciousness to such a degree that contact is impossible (ie, coma).*
- *The increased rate of cholesterol production increases the occurrence of atherosclerosis and of diabetic nephropathy.*
- *Albuminuria, hypertension and low glomerular filtration rate characterise diabetic nephropathy.*

Further Reading

Almind, K., C. Bjørbæk, H. Vestergaard, T. Hansen, S. Echwald, and O. Pedersen. "Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus." *The Lancet* 342: 828-832, 1993.

Ashcroft, F.M. and S.J.H. Ashcroft. "Insulin." *IRL Press at Oxford Univ*. Press, Oxford 1992.

Banting, F.G. and C. H. Best. "Internal secretion of pancreas." *J. Lab. and Clin. Med.* 7: 251-326, 1922.

Flatt, P.R. "Nutrient regulation of insulin secretion." *Portland Press Ltd.*, London 1991.

Return to top

Return to Content

Chapter 21

Thermo-Regulation, Temperature And Radiation

This Chapter is written following discussions with my colleague, Leif Vanggaard, MD, Arctic Institute, Copenhagen.

Study ObjectivesF

- To *define* body core and body shell, heat balance, heat exchange (conduction, conversion, evaporation and radiation), hyperthermia, hypothermia, mean body temperature, heat capacity, and thermal steady state.
- To *describe* fever (pyrogens), benignant and malignant hyperthermia, heat exhaustion, heat syncope, heat stroke, sun stroke, and hypothermia.
- To *describe* radiation sickness.
- To *calculate* one thermal variable, when relevant variables are given.
- To *explain* the concepts heat exchange, thermogenesis by food and shivering, the human temperature control system and its function at different environmental temperatures.
- To *use* the above concepts in problem solving and case histories.

Principles

- ***Newton's law of cooling****: The dry heat loss is proportional to the temperature difference between the human body (shell) and the surroundings.*
- ***The total energy of a system*** *is conserved in an interaction, not the kinetic energy or the mass (Einstein). If the mass changes during an interaction, there is a resultant change in kinetic energy, so that the total energy remains constant. – Heat energy is proportional to molecular movement rates – "heat energy equals movement."*
- ***Stefan-Boltzmanns rule****: The higher the temperature of an object, the more it radiates. The energy radiated from an object is proportional to the fourth power of its Kelvin temperature. – The energy radiating from an object and received by the human body is proportional to the temperature difference between the object and the skin (see Eq. 21-4). This is because human life implies relatively small temperature gradients.*

Definitions

- **Body core** consists of the thermoregulated deeper parts of the body and the proximal extremity portions of warm-blooded animals including man.
- **Body shell** refers to those outer parts of the body (skin and subcutaneous tissue) that change temperature at cold exposure.
- **Conductance** changes of the shell are used as a measure of skin bloodflow.
- **Conductive heat loss** describes a direct transfer of heat energy by contact between two bodies of different temperature (eg, skin and objects).
- **Convective heat loss** is defined as the heat loss by contact between the surface (skin) and a moving medium (air or water).
- **Evaporative heat loss** is defined as the heat loss by evaporation from the body surface or lungs.

- **Fever** occurs when the *core temperature of the body* is raised above normal steady state levels. The body reacts as if it is too cold. Fever implies a disorder resulting in shivering combined with vasoconstriction, headache, dedolation, and general discomfort (eg, malaria).
- **Heat flow** is defined as energy exchanged due to a temperature difference. Heat flow is transmitted along a temperature gradient.
- **Heat capacity** is the amount of heat required to produce a temperature increase for a given amount of substance.
- **Heat energy balance** in a resting person is a condition, where the heat production is equal to the heat loss. Thus the body temperature is constant and the heat storage is zero (*thermal steady state*). Usually, there is no internal heat energy flux between body core and shell.
- **Hyperthermia** is an increase in core temperature above normal.
- **Hypothermia** refers to a clinical condition with a lowered core temperature (below 35 ^{o}C).
- **Mean body temperature** is defined according to Eq. 21-1 (see end of Chapter).
- **Non-shivering thermogenesis** is a rise in metabolism, which is not related to muscular activity (shivering or exercise).
- **Insensible perspiration** (leakage of the skin) is the small cutaneous evaporation loss, which is unrelated to sweat gland function.
- **Insulation** refers to resistance to heat transfer.
- **Radiative heat loss** is a transfer of heat energy between 2 separate objects at different temperature. Heat energy is transferred via electromagnetic waves (photons). This heat transfer does not require a medium, and the temperature of any intervening medium is immaterial.
- **Shell temperature** is the temperature of the outer parts of the body (measured on the skin surface) and related to cold environments.
- **Shivering** is a reflex myogenic response to cold with asynchronous or balanced muscle contractions performing no external work.
- **Specific heat capacity** is the relationship between heat energy exchanged per weight unit of a substance and the corresponding temperature change. The *specific heat capacity* of ***water*** is 4.18 and of the *human body* (blood and tissues) 3.49 kJ kg^{-1} ^{o}C^{-1}, respectively. The specific heat capacity of atmospheric air is 1.3 kJ (m^{3})$^{-1}$ ^{o}C^{-1}.
- **Temperature** is the measurement of heat energy content.

Essentials

This paragraph deals with

1. The temperatures of the body, 2. Body responses to cold, 3. Body responses to heat, 4. Emotional sweating, 5. Metabolic Rate and environmental temperature, 6. Temperature control, 7. The human thermo-control system, and 8. Thermoregulatory effectors.

1. The temperatures of the body

The human body consists of a peripheral shell and a central core (Fig. 21-1). The *heat content* (***H*** or *enthalpy*) of the

human body is reflected by its temperature. By definition a thermometer only measures the temperature of the thermometer, so its location is essential. The mean core temperature is 37 ^{o}C in healthy adults at rest, but small children have larger diurnal variations.

The *skin* is the *main heat exchanger* of the body. The skin temperature is determined by the core temperature and by the environment (temperature, humidity, air velocity). Thus the shell temperature is governed by the needs of the body to exchange heat energy.

Fig. 21-1: Heat transfers, body cores and shells temperatures of a naked person standing in cold and warm air, respectively.

The *shell temperature* is measured on the *skin surface* and at the *hands and feet* to approach the room temperature of 19^{o}C in a person standing in a cold room for hours (Fig. 21-1, left). The shell temperature is several degrees lower than the temperature in the central core. The limbs have both a longitudinal and a radial temperature gradient. The shell temperature and the size of the shell vary with the environmental temperature and the termal state of the person. A naked person, standing on a cold floor in 19^{o}C air has a small core and a thick shell compared to the same person in a warm environment (Fig. 21-1). The shell temperature of the skin and distal extremities is difficult to evaluate. The best estimate is measurement of the infrared heat radiation flux with a radiometer.

The *core temperature* is the rather constant temperature in the *deeper parts of the body* and in the proximal extremity portions (see the red stippled lines of Fig. 21-1). However, the core temperature may vary several Centigrades between different regions depending on the cellular activity. The brain has a radial temperature gradient between its deep and superficial parts. In a sense, the temperature of the mixed venous blood represents an essential core temperature.

The rectal temperature

A high core temperature is found to be constant in the rectum about 10-15 cm from the anus. When measuring the *rectal temperature* a standard depth of 5-10 cm is used clinically. The venous plexus around the rectum communicate with the cutaneous blood in the anal region. The rectal temperature falls when the feet are cold, because cold blood passes the rectum in the veins from the legs for the same reason. The rectal temperature rises during heavy work involving the legs.

Parents should be advised to measure the rectal temperature in disease suspect children. The rectal temperature is a reliable estimate of the core temperature in resting persons.

Sublingual (oral) or axillary temperatures are unreliable measures of the core temperature - often more than half a degree lower than the rectal temperature.

The cranial temperature (tympanic and nasal)

The main control of temperature is performed by the anterior hypothalamus, which has a high bloodflow. Within the cranium the hypothalamus lies over the Circle of Willis, which supplies it with blood, and close to the cavernous sinus which drains it. Hypothalamus elicits heat loss responses when stimulated by heat. The tympanic membrane and areas in the nasal cavity (the anterior ethmoidal region, part of the sphenoid sinus) are supplied with blood from the internal carotid artery just like the hypothalamus. These *cranial* locations then serve as a substitute for the measurement of the inaccessible hypothalamic temperature.

Intake of 250 g of ice releases an abrupt fall in the nasal temperature in a warm person, whereas the change in rectal temperature is smaller and delayed (Fig. 21-2). The cranial core temperature is more dynamic than the rectal.

Fig. 21-2: Intake of ice reduces the temperature in a warm person resting at 45^{o}C.

In sports and in surgical hypothermia dynamic measurements of core temperature are essential. The cranial temperature is often preferred. During forceful movements the thermistor may be displaced. In such situations an *oesophageal* location is applied at heart level. This is an approximative measure of the temperature of the mixed venous blood of the right heart located close to the thermistor.

The *mean body temperature* is defined according to Eq. 21-1. The storage of heat energy in the body can be calculated according to its heat capacity (3.49 kJ* kg^{-1}*^{o}C^{-1}), the body weight (kg) and the change in mean body temperature in the period (Eq. 21-2).

According to the first law of thermodynamics, the storage of heat energy equals the metabolic energy change minus the heat loss (Eq. 21-3). Quantification of thermodynamics in humans is possible using equations 21-1 to 21-7 (later in this chapter).

The body is in heat energy balance, when the storage is zero. However, the core temperature may change with internal fluxes of heat energy between core and shell without storage or loss of heat energy at a constant activity.

Venous blood draining active muscles and the liver is likely to be warmer than pulmonary venous blood, since this has undergone evaporative cooling in the alveoli. A patient with high fever can be in *thermal steady state*, with a high constant heat production, if both core- and shell-temperatures are constant, and no internal energy flux occurs.

Warm-blooded animals, *homeotherms* such as humans, can change their metabolism in order to keep their heat production equal to the heat loss. Such animals have a *temperature control system* and thereby maintain a *rather constant core temperature*. Warm-blooded animals live with the advantage of an unchanged cell activity and temperature in their core. However, the human core temperature falls during the oestrogen phase of the menstrual cycle and during sleep (circadian rhythm). The lowest temperature is between 18 at night and 6 o'clock in the morning (Fig. 21-3). The temperature cycle is part of the circadian periodicity. Our biological clock seems to be synchronised with the rotation of the globe. Also meals, light and temperature plays a role.

Ovulation releases a sharp rise in morning temperature. Progesterone effects seem to explain the higher temperature in the last phase of the menstrual cycle (Fig. 21-3).

Fig. 21-3: Variations of the core temperature during 24 hour (above), and variations related to phases of the menstrual cycle (below).

Cold-blooded animals (*poikilotherms*) live with a behavioural temperature rhythm, but have no autonomic temperature control. The core- and shell-temperatures vary with the environment and the cellular activity. Reptiles, premature and low weight-premature newborn babies are cold-blooded. These babies have no thermoregulation (see later). However, their capacity for heat production is 5-10 times as great per unit weight as that of adults.

Humans have a *warm-blooded* (homeothermic) *core* and a *cold-blooded* (poikilothermic) *shell* in a cold environment.

Persons exposed to general anaesthesia, alcohol, and certain drugs lose the autonomic thermoregulation. Cold-blooded animals must live with varying core and shell temperature, whereby the rate of their cellular activities varies with the surrounding temperature (Fig. 21-4).

Fig. 21-4: The body core temperature and the environmental body temperature for a warm-blooded animal (cat) and a cold-blooded animal (lizard).

a) Convection. The convective heat loss is calculated by Eq. 21-7. A healthy person in sports clothes experiences thermal comfort at three times the resting metabolic rate (3 MET), when the surrounding temperature is 20^{o}C, the humidity is 50% and the wind velocity is 0.5 m*s-1.

Diving (water has a high thermal conductivity) illustrates the importance of conduction andconvection in heat energy transfer.

The *dry diving suit* excludes water from contact with the skin and traps low-conductance air in insulating clothing worn inside the watertight sealing.

The *wet suit* traps water next to the skin but prevents its circulation. The water is warmed through contact with the skin, and the high insulation of the foam rubber wet diving suit, with its many pockets of trapped air, minimises the rate of heat energy loss to the surrounding water. Air is a poor heat conductor and thus a good insulator. During deep diving high pressures compress these air pockets and thus reduce the insulation properties of wet diving suits.

b) Radiation describes a transfer of energy between objects in the form of electromagnetic waves (photons). This includes ultraviolet and visible (sun light) radiation from the outside and from the body infrared or warm heat radiation.

Radiative heat transfer can be calculated for a naked person according to Eq. 21-4.

When the skin temperature (T_{skin}) is less than the temperature of the surrounding objects, heat is gained by radiation.

At wintertime, heat can be lost through a window glass by radiation from the body to the cold environment irrespective of the room temperature. This is because the skin temperature is higher than the outside temperature.

c) Conduction. Sitting on a cold stone is a typical example of conduction loss, just as standing on a cold floor (Fig. 21-1). – Conduction heat can also be gained, although it is really possible to walk on glowing coals with speed and a thick epidermal horn layer.

d) Evaporative heat loss- see sweat secretion below.

2. Body-responses to cold

Cutaneous vasoconstriction lowers skin temperature, and thereby reduces the *conductive-convective heat loss* that is determined by the temperature gradient from the skin surface to the environment. Cutaneous vasoconstriction directs the peripheral venous blood back to the body core through the deep veins and the commitant veins. These veins are located around the arteries with warm blood, so that the venous blood receives part of the heat energy from the arterial blood - so-called *counter current heat exchange* (Fig. 21-5). The vasoconstriction is so effective, that the bloodflow through the arterio-venous anastomoses in the fingers and toes can fall to below one percent of the flow at normal temperature. The cooling of the shell is immediate, and the size of the shell increases (Fig. 21-1). Obviously, the shell is large for a naked person in cold air. The resistance vessels of the hands may open periodically to nourish the tissues, but the high viscosity of the cold blood can endanger the tissue nutrition and result in trench foot.

The *arterio-venous shunts* of the hands and feet are closed, so the bloodflow to the limbs is a nutritive minimum.

The deep arteries and veins of the limbs lie in parallel, so the arterial bloodflow loses heat to the incoming venous blood partially surrounding the arteries (Fig. 21-5). This is a typical *counter-current heat exchange*. In a cold environment, where vasoconstriction and heat exchange produces cold extremities, the total insulation is increased at the expense of reduced neuromuscular efficiency.

Fig. 21-5: Counter-current exchange in a human arm conserving heat energy in a cold climate (left). Superficial venous cooling ribs eliminate heat energy in a warm climate (right).

In a *warm climate* the high bloodflow of the extremities ensures an optimal temperature of the deeper structures (eg, the neuromuscular system). The temperature of the arterial blood is maintained (Fig. 21-5, right) and the arterio-venous anastomoses are wide open conveying warm blood to the superficial veins. The superficial veins also act as **cooling ribs** and transfer large amounts of heat to the skin surface, where it is eliminated from the body by convection, conduction and evaporation (Fig. 21-5, right).

Shivering is a *reflex myogenic response* to cold with *asynchronous* or balanced muscle contractions elicited from the hypothalamus via cutaneous receptors. The activity in agonist and antagonist muscles balance, so there is no external work. Without outside work, all energy is liberated as metabolic heat energy. Heat production is also increased by thyroid gland activity and by release of catecholamines from the adrenal medulla.

External work, such as running, is helpful in maintaining body temperature when feeling cold. Cold increases the motivation for *warm-up exercises* and illustrates the voluntary, cortical (feedforward influence) on temperature homeostasis. The core temperature increases proportionally to the work intensity during prolonged steady state work (Fig. 21-6). The mean skin temperature falls with increasing work intensity at 20^{o}C, because the sweat evaporation cools the skin.

Fig. 21-6: Muscular and oesophageal temperature during steady state exercise. The levels of exercise range from zero to 100% of the maximum oxygen uptake.

The temperature in the active muscles determines the level of the rectal temperature. Following marathon rectal temperatures of more than 41^{o}C have been measured and heat strokes have occurred. A marathon is even more difficult to accomplish in warm, humid environments and strong sun may cause sunstroke (see later).

People may adapt to prolonged exposure to cold by increasing their basal metabolic rate up to 50% higher than normal. This *metabolic adaptation* is found in Inuits (Eskimos) and other people continuously subject to cold.

The environmental temperature, where we maintain our autonomic temperature control, is in the range of zero to 45^{o}C. Below and above this range we adapt to the environment by behaviour (adding or removing clothing, warm or cold

bath, sun or shadow). A core temperature above 44^{o}C starts protein denaturation in all cells and is incompatible with life. Below 32^{o}C humans lose consciousness and below 28^{o}C the frequency of malignant cardiac arrhythmia's is increasing, ending with ventricular fibrillation and death at a core temperature below 23 ^{o}C (Fig. 21-7).

Fig. 21-7: Environmental temperature variations and temperature control. Lack of vital signs in the clinic (respiration, heart rate, EEG) must not be taken as death. Treatment must be instituted until death signs are developed.

3. Body-responses to heat

***Sweat secretion*.** Three million sweat glands produce sweat at a rate of up to 2 litres *per hour* or more during exercise in extreme warm conditions. If not compensated by drinking, such high sweat rates lead to circulatory failure and shock. Sweat resembles a dilute ultrafiltrate of plasma. Healthy humans cannot maintain their body temperature, if the environmental air reaches body temperature and the air is saturated with water vapour. Primary sweat is secreted as an isosmotic fluid into the sweat duct, and subsequent NaCl reabsorption results in the final hypo-osmotic sweat. *Thermal sweating* is abolished by atropine, proving that the postganglionic fibres are cholinergic. Cholinergic drugs provoke sweating just as adrenergic agonists do. *Evaporation* of water on the body surface eliminates 2428-2436 J g^{-1} at mean shell temperatures of 30-32^{o}C. Evaporation of a large volume of sweat per time unit ($\dot{V}_{sweat}$) implies a substantial loss of heat according to Eq. 21-5.

Normally, the skin temperature falls with increasing work intensity, because the sweat evaporation cools the skin (Fig. 21-6). Danger occurs when the average skin temperature and the body core temperature converge towards the same value.

Condensation of water on the skin gains heat energy, which is stored in the body. This is what happens in a Sauna.

Vasodilatation of skin vessels in warm environments results in increased cardiac output. The *arterio-venous anastomoses* in the hands and feet are open, and the bloodflow can rise up to at least 10 folds. The shell is minimal, when a naked person is in warm air (Fig. 21-1, right). The *skin bloodflow*, mainly in the extremities, determines the amount of heat energy, which is carried from the body core to be lost on the surface. The heat energy is transported from the large body core to the skin by *convection in the blood*. A substantial part of the heat energy is lost through the superficial veins of the extremities acting as *cooling ribs* (Fig. 21-5). The blood of the superficial veins is thus arterialized, when the person is warm.

A piece of steak has the same composition as human skin but of course no blood flow and no sweat evaporation. Thus the steak will be cooked at an air temperature that humans can survive. A person can stay in a room with dry air at 128^{o}C for up to 10 min during which time the steak is partially cooked.

4. Emotional sweating

This is a *paradoxical response* in contrast to the *thermal sweating* of thermoregulation. Emotional stress elicits *vasoconstriction* in the hands and feet combined with *profuse sweat secretion* on the palmar and plantar skin surfaces.

5. Metabolic Rate and environmental temperature

The total heat loss consists of the evaporative heat loss and the dry heat loss (ie, the sum of convective, conductive and radiative loss). *Newton's law of cooling* states that the dry heat loss is proportional to the temperature difference between the human body (shell) and the surroundings.

Let us look at a healthy, lightly dressed sitting person in thermal balance. His heat loss is plotted as the ordinate and the environmental temperature as the abscissa (Fig. 21-8). The two types of heat loss are added in order to provide the total heat energy loss. At a room temperature of 37^{o}C there is a dry heat loss of zero, and below there is an increasing dry heat loss. Above 37^{o}C the heat loss turns into heat input. Obviously, the dry heat loss also depends upon conduction and convection of heat inside the body by contact and by the perfusion. At extremely low environmental temperature the dry heat loss becomes larger than the metabolic heat liberation and the body is cooled down.

The person is in *thermal steady state*, and the metabolic rate is almost constant in the *thermoneutral zone* between 20 and 30^{o} C (Fig. 21-8). The *law of metabolic reduction* reflects the tendency for heat production to match the rate of

heat loss. The thermoneutral zone, where minimal compensatory activity is required, is separated in the ***lower vasomotor*** and the *upper sudomotor control zone.*

In the lower, comfortable zone (20-26^{o} C) the *total heat dissipation* is maintained equal to the metabolic rate by cutaneous, vasomotor alterations. The small evaporation loss is termed leakage of the skin or *insensible perspiration*, which is unrelated to sweat gland function and rather constant at the basal metabolic rate.

Fig. 21-8: Metabolic rate and environmental temperature in a fasting dressed human at rest. The wet and the dry heat loss, as well as the metabolic heat and the basal metabolic rate (*BMR*) is measured in Watts.

In the *upper sudomotor zone* above 26^{o} C environmental temperature, the bloodflow through the skin rises, as does sweat secretion and evaporation. At 37^{o} C the rise in energy loss occurs via evaporation (Fig. 21-8).

When the environmental temperature falls the metabolic rate increases - first by increasing muscle tone and then by shivering. The *chemical* or *metabolic temperature control* is in the environmental region from 20^{o} C and below (Fig. 21-8), where shivering, decreased bloodflow through skin and non-myogenic heat production take place. Here, metabolism controls the core temperature by increasing metabolic rate with falling temperature in the environment. Above 20^{o} C, the *physical temperature control* takes over, as an autonomic capacity for alterations in heat loss. In this thermoneutral zone the body temperature is kept constant almost without either heat-producing mechanisms or sweat secretion. - The *thermal comfort* for light clothed, seated persons is about an air temperature of 26^{o}C when the humidity is 50%, 30^{o}C for nude persons, and about 36^{o}C sitting in water to the neck.

In the *metabolic zone*, the total heat loss rises with falling environmental temperature, but below 5 ^{o}C in the environment, the dry heat loss exceeds the metabolic rate, and the body is cooled down (Fig. 21-8). This is the *zone of hypothermia*, where cold death is inevitable without treatment.

The *zone of hyperthermia* begins at an environmental temperature of 37 ^{o}C, where humans soon reach the maximal capacity for evaporation and there is an unbalanced heat influx to the body ending in *heat death.*

The metabolism varies with the shell and the core temperature. These relations were elucidated by series of similar experiments. A person was placed in stirred water to the neck, where the water temperature thus defined the mean shell temperature. By pre-treatment in other baths an array of core-and shell- temperature combinations were obtained and measured simultaneously with the metabolic rate (*metabolic heat liberation*). The core temperature was reduced by intake of up to 2 l of crushed ice in water.

At a shell temperature of 20 ^{o}C, intake of ice water reduces the core temperature below 37.1 ^{o}C (here called the set point) and the metabolic rate increases by shivering (Fig. 21-9) with falling core temperature.

Fig. 21-9: The metabolic rate as a function of the core and the shell temperature.

Warmer skin makes the set point fall and the rise in metabolic rate per ^{o}C (of core temperature fall) is less steep. Core temperatures above 37.1 ^{o}C all have a low metabolic rate, regardless of the shell values. At a shell temperature of 30 ^{o}C the set point decreases from 37.1 to 36.7 ^{o}C. The hypothalamus functions as a *thermostat* using the rest of the body to stabilise its own temperature. With rising core temperature the metabolic rate is maintained low and the body tries to cool down. When the body core temperature falls below set point, the metabolic rate increases by shivering and the heat energy storage as well. The cutaneous cold receptors are maximally active at 20 ^{o}C and they are silent above 33 ^{o}C - they can only trigger shivering, when the core temperature is below set point. Shivering ceases immediately, when we take a warm shower and the skin is warmed. The hypothalamic, preoptic heat receptors inhibit shivering, and shivering is totally blocked above the set point.

Another series of experiments were directed towards heat loss. The heat loss from evaporated sweat was recorded during exercise at different shell temperatures.

Fig. 21-10: Evaporative heat loss as a function of the core and the mean skin temperature during different intensities of exercise.

Cold stimulates cutaneous cold receptors with connections to the hypothalamus (Fig. 21-12).

At mean skin temperatures of 33-39 ^{o}C, where the cutaneous cold receptors are silent, the person is unable to sweat before the core temperature is above 36.9 ^{o}C (the set point of Fig. 21-10). Below the set point, the evaporation is low (perspiratio insensibilis). With increasing muscle activity and core temperature the evaporative heat loss may rise towards 20 kJ each min.

Mean skin temperatures below 33 ^{o}C reduce the evaporative heat loss. The rising hypothalamic temperature releases the sweat secretion, but the local secretion is inhibited by the cutaneous cold receptors.

Heat has a direct effect on preoptic heat receptors in the hypothalamus. At increased core temperature, the preoptic heat receptors totally block shivering, although the hypothalamic centre also receives shivering signals from cutaneous cold receptors in cold surroundings with mean skin temperatures below 33 ^{o}C. The falling mean skin temperature can also act directly to lower cutaneous bloodflow and sweat secretion.

6. Temperature control

Humans have a rather *constant core temperature* although the metabolism and environmental temperature may vary considerably. This implies that *control* is exerted.

Fig. 21-11: Thermoregulation by dynamic gain and set point systems.

Information about the environmental temperature is provided by *peripheral thermo-sensors,* which are located in the skin, abdominal organs, and muscles. Internal or blood temperature is monitored by *central thermo-sensors* in the preoptic hypothalamus and the medulla.

A rise in *hypothalamic temperature* causes vasodilatation in the skin and reduces muscular tone. The person looses motivation for physical activity and reduces clothing. Then thermal sweat is observed, and after some time, reduced activity of the adrenal cortex and of the thyroid gland is also observed.

A *fall in hypothalamic temperature* by cooling of the shell and core releases cutaneous vasoconstriction together with increased muscular tone and shivering. There is a sympathetic activation with secretion of catecholamines, oxidation of fatty acids and glucose, and increased secretion of the thyroid and adrenal gland. The muscle tone is increased, and shivering is triggered reflexly as asynchronous muscle contractions without external work (movements), so all the metabolic energy is released as heat. The capacity for muscular thermogenesis by shivering is high. Up to five folds basal metabolic rate is observed, which corresponds to heavy industrial work.

Shivering may be suppressed voluntarily at the beginning. Transmission signals for shivering passes the rubrospinal pathways to a- and g-motor neurons of antagonist muscles.

Dynamic gain and set point control

A: *A dynamic gain system* responds *continuously to feedback signals* - regardless of the core temperature. With rising tissue temperature, the neural activity of heat sensors increases linearly, whereas the activity of cold sensors decreases (Fig. 21-11A).

This combined sensory input to the hypothalamus increases the core temperature and thus increases the activity of *heat loss effectors*, while inhibiting *heat production effectors* (Fig. 21-11 A). This determines the *reference signal*.

The dynamic gain system has a floating *reference signal* moving with the continuous heat loss and heat production (Fig. 21-11A).

B: *A set-point system* does not respond to a rise in temperature before a certain set point is reached. The set point is the core temperature at which neither heat loss mechanisms or heat production mechanisms are active.

When a thermal disorder reaches a certain *set-point* in the hypothalamus, signals passes to the effectors. The desired core temperature (t_{set} = *set point temperature*) is compared to the actual value (t_{core} in Fig. 21-11B). The caudal hypothalamus works as a thermostat. Error signals - a deviation from t_{set} - evoke responses that tend to restore core and hypothalamic temperature toward the set point. When the actual core and hypothalamic temperature rises above the desired set point such as 37 ^{o}C, effectors are turned on, and the compensatory heat

energy loss is almost linear (Fig. 21-11B). These compensatory mechanisms (vasodilatation, sweat, reduced muscle tone) do not turn off until the temperature drops to the set point (ie, an all-or-non response).

When the actual core and hypothalamic temperature is just below the set point, the compensatory mechanisms (vasoconstriction and shivering) are relatively inactive.

The hypothalamic set point change with the physiological conditions and is elevated in fever by *pyrogens* from microorganisms. The rise in metabolism is mainly accomplished by shivering.

7. The human thermo-control system

Human temperature control exhibits both *dynamic gain* and *set point* characteristics. The control system implies widespread cutaneous and deep sensors. Their afferents converge towards the hypothalamic integrator, which acts as a *thermostat.* The hypothalamus also contains thermosensors in the preoptic region, and inhibitory neurons perform *crossing inhibition* (Fig. 21-12). The *central heat drive* from the *preoptic hypothalamus* is maintained (Fig. 21-12). The stability of the core temperature is maintained by the *large heat capacity of the body mass,* and by the deep thermosensors, which are dominant.

Shivering is released from cutaneus cold sensors firing maximally at 20 ^{o}C. These cold sensors are silent above 33 ^{o}C. The cold shell (Fig. 21-12) activates deep cold sensors in the preoptic hypothalamus. This increases heat production by shivering. The preoptic thermostat simultaneously reduces heat loss by crossing inhibition.

Fig. 21-12: The hypothalamic thermostat and its connections.

Sweat secretion is released by preoptic warm sensors as soon as their temperature is 37 ^{o}C or above (t_{set}). Cutaneous cold sensors inhibit sweat secretion at shell temperatures below 33 ^{o}C, since they are silent above this temperature.

In conclusion, preoptic warm sensors show set point characteristics below the set point, and preoptic cold sensors show set point characteristics above the set point.

Apart from that, preoptic sensors show *dynamic gain*: With rising tissue temperature, the neural activity of heat sensors increases linearly, whereas the signal frequency of cold sensors increases with falling temperature (Fig. 21-11**A**).

Cutaneous sympathetic vasodilatation is probably also released by preoptic warm sensors above set point (Fig. 21-12). A fall in skin temperature below 33 ^{o}C will reduce skin bloodflow by crossing inhibition (Fig. 21-12).

Alcohol seems to off set the thermocontrol mainly by inhibition of the hypothalamic thermostat (Fig. 21-12).

Newborns, down to premature babies above 1000 g, possess certain thermoregulatory functions. The newborn can increase thermogenesis by a factor of three without shivering shows vasomotor reactions, sweat secretion, and reduces the surface area in cold air. However the baby has special problems: The surface-volume ratio is 3-fold higher than that of an adult. The baby has a thin shell due to the thin subcutaneous fat layer, so even a maximal vasoconstriction cannot limit heat loss with a capacity comparable to that of an adult. Newborns are specially equipped to perform non-shivering thermogenesis (chemical thermogenesis). Non-shivering thermogenesis is any rise in metabolism, which is not related to shivering. In babies this form of thermogenesis is particularly large in their brown adipose tissue. This tissue is abundant around vital organs, in neck and mediastinum, between scapulae and in the armpits. Brown adipose tissue cells contains multilocular droplets in the fat phase and many mitochondria. The tissue receives sympathetic innervation and is stimulated by catecholamines and thyroid hormones. Cooling increases the bloodflow and temperature of brown adipose tissue. Noradrenaline injections cause vasodilatation via b-receptors, and increase the metabolism to the same extent. The cause of the increased metabolism is an increased cell membrane permeability for Na^{+}- and K^{+}-ions, whereby the Na^{+}- K^{+}-pump (high ATP demand) is activated.

Fig. 21-13: Thermocontrol in the newborn.

Newborns are in thermal balance at a minimal metabolism only when the surrounding temperature is high (32-34 ^{o}C). With other words, newborns have an extremely high lower threshold for the thermoneutral zone, namely 32 ^{o}C. The threshold for maximal use of shivering is approximately 23 ^{o}C compared to a naked adult about 5 ^{o}C. In general, these

threshold values for newborn increase with falling body weight in premature. Premature below 1000 g has hardly any thermoneutral zone (Fig. 21-13). They are actually cold-blooded and their temperature control is maintained with a *couveuse*.

8. Thermoregulatory effectors

The *sympathetic* and the *somatomotor nervous system* participate in thermoregulation (Fig. 21-14).

Noradrenergic sympathetic neurons control the bloodflow through fingers, hands, ears, lips and nose. *Arterioles* contract and *arteriovenous anastomoses* close (thermal insulation) following an increase in sympathetic tone, and dilatate following a decrease in tone. When arterioles and arteriovenous anastomoses open, the bloodflow is markedly increased and thus the *convective heat loss* from the skin is increased.

Fig. 21-14: The thermoregulatory feedback system.

Cholinergic sympathetic fibres control sweat secretion. The vasodilatator *bradykinin* is liberated in the skin. Thus, profuse sweat secretion is always accompanied by vasodilatation.

Sympathetic activation releases *thyroid hormones* from the thyroid gland and *catecholamines* from the adrenal medulla. These hormones liberate fatty acids and glucose for combustion. A reduced sympathetic tone also reduces the activity of the adrenal and the thyroid gland.

The thermogenic response to cold also involves a *non-myogenic* or *non-shivering component* probably in adipocytes. Non-shivering heat production is controlled by the sympathetic nervous system via adrenergic b-receptors. The noradrenaline (NA in Fig. 21-14) released at the nerve terminals close to the adipocytes, stimulates the liberation of free fatty acids and their subsequent oxidation. *Non-myogenic heat production* includes a contribution from the brown fat of babies, but is insignificant in adults.

Shivering is induced by way of the motor system. The *central shivering pathway* passes from the hypothalamus to the motor neurons in the spinal cord. Shivering is abolished by blockade of the neuromuscular end plate with curare.

Thermoregulatory behaviour such as fanning and adding or removing clothing is effective in changing the thermal insulation. Several layers of clothing with trapped air act as a good insulator.

In healthy persons, heat energy is liberated from cellular metabolism and transferred to the environment through the skin by sweating and vasodilatation. Sweating occurs during exercise, and its evaporation is the most important mechanism in maintaining the core temperature as close to 37^{o} C as possible. The thermoregulatory centre in the hypothalamus controls all processes.

Fig. 21-15: Heat and cold adaptation

Cold adaptation is found among Australian aborigines and Inuits in Greenland. Inuits have relatively more sweat glands in the face and less on the body.

Aborigines can sleep naked on the ground even at low temperature. Inuits have a basal metabolic rate 50% higher than persons living in a temperate climate do. The threshold for shivering is shifted towards the left in cold-adapted persons, but they maintain normal function at the new set point. Very old people may show the same phenomenon, and live with a core temperature of 35 ^{o}C without shivering. Obviously, cold adaptation implies non-shivering thermogenesis, which is oeconomic metabolic heat liberation (Fig. 21-15).

Heat acclimatization is actually a *sweat gland adaptation*, and a 2-week process following arrival to a hot climate such as the tropics or a desert. Gradually sweat-evaporation is increased and the NaCl loss is reduced. The sweat secretion capacity may reach 4 l each hour with a thin sweat. The adaptation is caused by increased aldosterone secretion from the adrenal cortex. Aldosterone increases the reabsorption of NaCl and the secretion of K^{+} from the sweat, during its passage of the sweat gland tubules. The larger sweat loss the thirstier one feels. This is because the large sweat secretion reduces the time period for NaCl- reabsorption in the sweat gland tubules. The resulting high NaCl concentration in the plasma implies thirst, so heat adapted persons have to drink a lot. Thirst is an extremely late indicator of dehydration during work in a warm climate.

Tropic inhabitants are of course heat-adapted. They have an increased core temperature (set point) and their threshold

for sweat and vasodilatation is typically 0.5^{o}C higher than that of a person living in the *temperate zone* (Fig. 21-15).

Pathophysiology

This paragraph deals with *1. Heat cramps, 2. Heat exhaustion and heat stroke, 3. Malignant hyperthermia, 4. Hypothermia, 5. Frostbite, and 6. Fever and hyperthermia.*

A paragraph concerning nuclear energy radiation is given at the end.

1. Heat Cramps

Painful cramps in the leg muscles occur following exercise, when athletes run too fast in a hot climate. Heat cramps respond to salt and water replenishment in a normal diet, and the cramps are probably caused by hyponatraemia. During prolonged sweating, the runner is losing salt and water. If only the water loss is replenished, the result is water-induced hyponatraemia - a parallel to the classical *miner's cramps.*

2. Heat Exhaustion and Heat Stroke

When the water-salt balance is at risk in a hot climate, it is always a threat to the circulation. Profuse sweat secretion, in a subject who is not acclimatized, results in salt- and water depletion, with a daily loss of more than one mol of NaCl and more than 6 l of water. Within a period of one-hour strenuous working endurance athletes have lost up to 8 l of water.

The falling extracellular fluid volume and increasing body and brain temperature to above 40^{o}C elicit severe symptoms and signs. As the volume and salt depletion develops, the sweat production goes down in spite of extreme vasodilatation. The falling blood pressure stimulates the high-pressure baroreceptors resulting in a rising heart rate. The dehydration with imminent shock frequently results in cerebral, renal and hepatic failure. The low brain bloodflow through an overheated brain leads to fatigue (*heat exhaustion),* confusion and unconsciousness or syncope (*heat syncope*). The confusion may develop into a veritable *delirium* (an acute impairment of consciousness) with brain oedema.

Heat stroke - in the sun it is called *sunstroke* - is heat collapse that occurs suddenly, hereby creating a life-threatening condition. This heat collapse often occurs in warm, humid environments, when an unacclimatized subject exercises. The subject - without sweating (hypothalamic failure) -suddenly falls into coma, if not preceded by a short period of confusion and delirium. The condition is fatal, if not relieved by rapid cooling.

3. Malignant hyperthermia

Malignant hyperpyrexia is often caused by a *genetic defect* (autosomal dominant) in the sarcoplasmic reticulum of skeletal muscles. *General anaesthesia* (often halogen-substituted ethane) triggers an allergic reaction with sudden opening of Ca^{2+} - channels in the muscle cells. The following influx of Ca^{2+} elicits generalized and maintained muscle contraction (rigidity), which liberates enormous quantities of heat energy. This condition is life threatening and often results in sudden death during or just after anaesthesia.

4. Hypothermia

Hypothermia is a fall in core temperature to values below 35 ^{o}C. Hypothermic subjects lose consciousness, when the core temperature falls below 32 ^{o}C - a potentially lethal condition called *severe hypothermia*.

During anaesthesia and surgery the core temperature of the patient falls and stabilise around 34-35 ^{o}C after 4-5 hours. Such a *surgery hypothermia* increases the risk of cardiac complications, bleeding tendency and prolonged wound healing.

The condition is a prominent cause of death in climbers and skiers, as well as in persons being immersed in cold water or living in the Antarctis. The climbers and skiers are often exposed to a cold, wet and windy environment carrying insufficient clothing. As the neuromuscular function suffers, they can no longer move. General hypothermia develops and they die.

In mild hypothermia the subject can still take action to rewarm by exercise and clothing, but as consciousness is lost the core temperature falls further, because shivering is abolished. The patient feels cold to touch, is in a developing

coma, the circulation and respiration fall, as does the metabolic rate. The dissociation curve of oxyhaemoglobin is moved to the left, and the solubility of gasses increases as the blood temperature falls.

The heart is the target organ in hypothermia. Below 30 ^{o}C, spontaneous remission is practically impossible, and death ensues from ventricular fibrillation occurring around 28 ^{o}C.

Careful monitoring of all vital functions is required during rewarming, which is performed either passively (rewarming by the patients own metabolism) with insulation and space blankets or actively by a warm water bath, while monitoring the patient. A good strategy in treating hypothermic victims is the slogan: "*They are not dead until they are warm and dead*."

An arterial blood sample from a hypothermic patient is routinely analysed at 37 ^{o}C. The P_{aO2} and pH_a is falsely higher, the P_{aCO2} is falsely lower, than in the circulating blood of the hypothermic patient. This is because more CO_2 is bound also as carbamino-haemoglobin, and less O_2 is bound in the cold blood. *Base Excess* is defined at 37 ^{o}C and thus a true metabolic variable (see Chapter 17).

Artificially induced hypothermia is used in brain- and heart-surgery, where the usual thermocontrol is inactivated by general anaesthesia. The procedure becomes dangerous, when it elicits *ventricular fibrillation*.

5. Frostbite

Frostbite is a local cold injury due to the formation of extracellular ice crystals in the skin and other tissues. This leads to extracellular dehydration and *hyperosmosis*, whereby the cells lose water until they die. As the skin temperature falls *below + 7 ^{o}C* the subjects have lost their sensory functions, and thus do not recognise the developing frostbite.

The most effective treatment is *warming* of the still frozen area by immersion in 40-42 ^{o}C hot water following hospitalisation. The patient experiences severe pain (above *+ 7 ^{o}C*) and morphine must be administered. Tissue which is no longer frozen must be treated sterile.

6. Fever and hyperthermia

Fever occurs when the *core temperature of the body* is raised above normal steady state levels. The body reacts as if it is too cold, while the temperature rises up to the new higher set point. Fever attacks imply shivering combined with vasoconstriction, headache, dedolation pains, and general discomfort. Fever is the result of one of two phenomena: the set point may be set to a higher level, or the efficacy of the temperature control system may be impaired. Fever implies *hyperthermia;* however many cases of hyperthermia do not constitute fever. Fever results from the action of *endogenous pyrogens* on the *hypothalamic heat control centre* (they increase the set point for the core temperature via prostaglandins). Exogenous pyrogens from microbes cause these *endogenous polypeptides* to be released from the defence cells of the body (ie, the reticuloendothelial system, RES). Antipyretic drugs inhibit cyclo-oxygenase activity, hereby interfering with the synthesis of prostaglandins and thromboxanes.

Following an attack of fever, vasodilatation and sweat evaporation reduces the core temperature.

Physiologic hyperthermia is an increase in core temperature caused by extreme heat stressor exercise. During work the body temperature rises up to 39^{o}C without clinical consequences.Hereby, the heat loss capacity is exceeded. During hyperthermia the *heat loss effectors* are strained to the utmost. The high body temperatures of *exercise* activate cooling mechanisms and elicit sweat loss, which strive to return the core temperature to its normal level.In extreme hyperthermia the core temperature may rise to more than 41^{o} C (heat stroke). Irreversible protein denaturation occurs above 44^{o}C with brain oedema and destroyed thermoneurons in the hypothalamus. Clinically, the brain damage is shown with disorientation, lack of sweat secretion, delirium and *universal cramps* before death.

Ionizing Radiation Hazards

Ionizing radiation implies destruction of tissue molecules.

Dosages of absorbed radiation is measured in Joules per kg (1 J kg $^{-1}$ is known as a ***gray, Gy***). *One Gy* is equivalent to *100 rads*. Radioactivity is measured as the number of degradations per second in *bequerels* or *Bq*. One degradation per

s from radioactive material equals one Bq.

The absorbed dose of radiation is balanced for damaging ionizing tissue effect, since different types of radiation have different density of ionization. A dose equivalent termed sievert (Sv) causes a rather large damage. The annual background radiation is 2.5 milli-Sv (*2.5 mSv*). Non-penetrating radiation occurs from alpha- and beta particles. These particles are stopped by paper, but when they enter a tissue - such as the bone marrow – they stay there and spoil everything.

Penetrating radiation consists of either gamma rays (neutron) or X-rays. Survivors from nuclear power plant accidents, with whole body absorption greater than 100 rads, are threatened by acute and chronic radiation sickness.

Acute radiation sickness appears as vomiting and malaise following exposure to 1 Gy (100 rad) or more. Lymphocyte production is reduced immediately, soon followed by leucopenia and thrombocytopenia with bleeding. The villi of the gastrointestinal tract are destroyed, absorption of nutrients is impaired, and new villous cells are not produced. Diarrhoea, often with blood loss, results in dehydration and anaemia. The skin is red and blistering, and the hair is loosed. The immunodeficiency system is destroyed, and secondary infections have a high mortality - especially pulmonary infections. *Cerebral oedema* may kill the victim within hours when exposed to 35 Gy or more.

Contamination with radioactive iodine is treated by immediate intake of potassium iodide, which block the major part of the thyroid absorption. The treatment of seriously exposed radiation victims is supportive and frustrane.

Chronic radiation sickness or *late radiation damage* implies an increased rate of mutagenesis, which includes a high frequency of leucaemia, cancer of the brain, the thyroid and the salivary glands, infertility and *cataract* (ie, an eye disease where the vision is blurred by an opaque lens). The radiation liberates large amounts of highly reactive ions in the cells. The reactive ions rupture DNA strands and cause mutations with production of cancer cells. Cancer cells multiply exponentially and their energy demand approach the total nutritive energy available, causing malnutrition and death.

The nuclear power plant accident in *Tjernobyl*, Ukraine, had many victims from radiation exposure. The nearby town, Pripatja, previously with 40,000 residents is now abandoned.

The frequency of thyroid cancer among children in Ukraine is more than 100 times the expected. Most of the children and young persons exposed have developed leucaemia or cancer. Their airway epithelium and respiration is seriously affected, and they have an unusual high frequency of cerebral haemorrhage.

Equations

- The mean body temperature (T_{body}) is calculated with the following equation:

$$\text{Eq. 21-1: } T_{body} = (0.7 \times T_{core} + 0.3 \times T_{shell}),$$

where the mean shell temperature (T_{shell}) is estimated from a series of representative skin temperatures. The obvious assumption in this equation is that 70% of the body weight is core (T_{core}), and the balance is shell, but the size of the shell varies with the environmental temperature.

- The storage of heat energy in the body is calculated as follows:

$$\text{Eq. 21-2: } \Delta H_{STORE} = 3.49 * \text{BODY WEIGHT} * (T_2 - T_1)$$

with the unit kJ per hour.

- According to the first law of thermodynamics, the following relation is valid:

$$\text{Eq. 21-3: } \Delta H_{STORE} = \Delta H_{METABOLIC} - (\Delta H_{RADIATION} + \Delta H_{CON} + \Delta H_{EVAPORATION})$$

where ΔH_{CON} is the change in heat loss by convection and conduction.

- Radiative heat loss from a warm object (T_{obj}) can be calculated for a naked person with a known skin temperature (T_{skin}):

$$\text{Eq. 21-4: } \Delta H_{RADIATION} = 0.5 \times A \times (T_{skin} - T_{obj})$$

where 0.5 is kJ min^{-1} m^{-2} K^{-1}, A is the area of the human body (radiating or receiving radiation), and T_{obj} is the temperature of the object exchanging energy. This equation is an approximation of Stefan-Boltzmanns rule.

- Evaporation of water on the body surface eliminates 2430 J g^{-1}. Evaporation of large volume rates of sweat ($V°_{sweat}$) implies a substantial loss of evaporative energy (J/min) according to the equation:

$$\text{Eq. 21-5: } \Delta H_{EVAPORATION} \text{ (J min}^{-1}) = 2430 \text{ (J g}^{-1}) \times V°_{sweat} \text{ (g min}^{-1}).$$

- Changes in heat conduction through the shell ($conductance_{shell}$) reflect changes in skin bloodflow. The conductance can be calculated by the formula:

$$\text{Eq. 21-6: } Conductance_{shell} = \text{Metabolic Rate}/(T_{core} - T_{shell}).$$

A large temperature difference implies an effective isolation, whereas a small difference implies a low isolation capacity.

- The heat loss of a naked person resting in quiet air by convection (including a minor part by conduction without movement) is given by the following equation:

$$\text{Eq. 21-7: } H_{CON} = 0.5 *(T_{shell} - T_{air}) \text{ in kJ per min.}$$

This equation is an approximation of Newton's law of cooling.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. Humans can survive at a temperature that would cook a piece of steak, because the steak cannot dissipate core heat energy.

B. Heat conductance of air at high pressure exceeds that at low pressure, whereby more heat energy is lost from the divers body by conduction through air inside a diving bell deep under water than at one atmosphere of pressure.

C. Temperature homeostasis is present when heat energy production equals heat loss.

D. Basal metabolic rate (BMR) is lower before than after a meal.

E. The neutral environmental temperature defines the level, where the resting metabolic rate is minimal.

II. Each of the following five statements have True/False options.

A. Water-induced hyponatraemia is called miners cramps.

B. Substantial influx of Ca^{2+} to the cells is probably involved in some cases of malignant hyperthermia.

C. Hypothermia is a fall in core temperature below 32^{o}C.

D. One degradation per second from radioactive material equals one bequerels (1 Bq), which is also equal to one curie.

E. Acute radiation reduces leucocyte and thrombocyte production.

Case History A

A male with a body weight of 70 kg stops his malaria prophylaxis with primaquine when leaving the endemic area. Three weeks later, a sudden attack of fever increases his core temperature from 37 to 40^{o}C within 30 min. The heat capacity of the human body is 3.47 kJ/kg and per ^{o}C. The metabolic rate of the subject increased substantially during the rise in temperature. Following the cold stage with uncontrollable shivering, the patient develops a delirious condition with severe headache. Two hours later the patient develops a profuse sweating. Partially evaporation of the water in 32 ml sweat/min (25% evaporates) occurs from the body surface (eliminating 2,436 J/g), during which body temperature drops.

Calculate the extra heat energy stored in his body after 30 min.

1. *Calculate the smallest possible metabolic rate in the 30-min period.*

2. *What causes the extra heat energy stored in the body?*

3. *Calculate the reduction in body temperature by ingestion of one L of ice water, when the fever is at its highest level.*

4. *Calculate the time it takes to lose the accumulated heat energy by evaporation.*

Case History B

A male sedentary person, weight 70 kg, has a daily food intake of 400 g carbohydrate (17.5 kJ/g), 100 g fat (39 kJ/g), and 100 g protein (17 kJ/g). There is a metabolic water formation of 32 mg/kJ. The man excretes 2100 ml of water (i.e., 1200 ml in the urine, 100 ml in faeces, and 800 ml through lungs and skin).

1. *Calculate the metabolic rate (in kJ/day or MJ/day).*

2. *The man is in water balance by a water intake of 1,700 ml as a total. Explain this apparent imbalance.*

Case History C

A 79-year old male is found apparently dead in the snow following a winter storm, where all traffic was arrested by snow. His muscles are stiff, and the heart rate is not palpable. The tendon reflexes are depressed, and the pupillary and other brainstem reflexes are lost.

The body is placed in a chapel at the hospital until the funeral. The next day the personnel are disturbed by noises from the chapel. Obviously, the man is alive.

1. *What has awakened the man?*

2. *Suggest a likely core temperature, at the time where the man was admitted to the hospital.*

Case History D

A 20-year old person, with a body surface area of 1.8 m^{2}, is at rest with a metabolic rate (MR, or heat energy production and transfer) of 80 Watts. His rectal temperature is 37^{o}C and his mean skin temperature is 33 ^{o}C. -

Suddenly, a malaria attack develops with a 7-fold rise in MR, and the patient reach a temperature plateau of 40 , with a mean skin temperature of 34 ^{o}C.

The conductance of the shell (C_{shell}) is a measure of skin bloodflow. This can be calculated by the formula: $C_{shell} = MR/(T_{core} - T_{shell})$.

1 *Calculate the conductance of the shell at rest.*

2. *Calculate the conductance of the shell at the fever plateau.*

3. *How does the body accomplish this increase?*

Try to solve the problems before looking up the answers..

Highlights

- ***The core temperature*** *is 37 ^{o}C in healthy adults at rest, but small children have larger diurnal variations.*
- ***Parents*** *should be advised to measure the rectal temperature in disease suspect children. The rectal temperature is a reliable estimate of the core temperature in resting persons.*
- ***A constant core temperature*** *does not guarantee heat energy balance, because the energy content of the shell may change with the surroundings.*
- ***The human core temperature*** *falls during the oestrogen phase of the menstrual cycle and during sleep. The lowest temperature is between 18 at night and 6 o'clock in the morning. Progesterone effects seem to explain the higher temperature in the last phase of the menstrual cycle*
- ***The temperature cycle*** *is part of the circadian periodicity. Our biological clock seems to be synchronised with the rotation of the globe. Also meals, light and temperature plays a role.*
- ***The thermal comfort point*** *for light-clothed, seated persons is about 26^{o}C when the humidity is 50%, 30^{o}C for nude persons, and about 36^{o}C sitting in water to the neck.*
- ***Vasodilatation*** *in warm environments permits increased skin bloodflow as the cardiac output also begins to rise. The arterio-venous anastomoses are dilatated, and the bloodflow can rise to at least 10 folds in hands and feet.*
- *In the* ***metabolic zone****, the total heat loss rises with falling temperature, but below 5 ^{o}C the dry heat loss exceeds the metabolic rate, and the body is cooled down. This is the zone of hypothermia, where cold death is inevitable without therapy.*
- ***Sweat secretion****. Three million sweat glands produce sweat at a rate of up to 2 litres per hour or more during extreme conditions. Such high sweat rates lead to circulatory failure and shock. Sweat resembles a dilute ultrafiltrate of plasma.*
- ***The zone of hyperthermia*** *begins at an environmental temperature of 37 ^{o}C in humid air, where we soon reach the maximal capacity for sweat secretion, and there is a heat influx to the body ending in heat death.*
- ***The hypothalamus*** *functions as a thermostat using the rest of the body to stabilise its own temperature.*
- ***A dynamic gain system*** *responds continuously to feedback signals, regardless of the size of the core temperature. With rising tissue temperature, the neural activity of heat sensors increases linearly, whereas the activity of cold*

sensors decreases

- ***A set-point system*** *does not respond to a rise in temperature before a certain set point is reached. The set point is the core temperature at which neither heat loss mechanisms or heat production mechanisms are active.*

- ***Human temperature control*** *exhibits both dynamic gain and set point characteristics. The control system implies widespread cutaneous and deep sensors. Their afferents converge towards the hypothalamic integrator and thermostat.*

- ***The thermogenic response*** *to cold also involves a non-myogenic or non-shivering component probably in adipocytes. Non-shivering heat production is controlled by the sympathetic nervous system via adrenergic β-receptors.*

- ***Non-shivering thermogenesis*** *is any rise in metabolism, which is not related to muscular activity. In babies this form of thermogenesis is particularly large in their brown adipose tissue. The brown adipose tissue of babies is abundant around vital organs, in neck and mediastinum, between scapulae and in the armpits.*

- ***Newborns*** *are in thermal balance at a minimal metabolism only when the surrounding temperature is high (32-34 ^{o}C).*

- ***Thermoregulatory behaviour*** *such as fanning and adding or removing clothing is effective in changing the thermal insulation. Several layers of clothing with trapped air are good insulators.*

- ***Heat acclimatization*** *is actually a sweat gland adaptation, and a 2-week process following arrival to a hot climate. Gradually sweat-evaporation is increased and the NaCl loss is reduced. The adaptation is caused by increased aldosterone secretion from the adrenal cortex.*

- ***Hypovolaemia*** *with low brain bloodflow through an overheated brain leads to fatigue (heat exhaustion), confusion and unconsciousness or syncope (heat syncope). The confusion may develop into a veritable delirium with brain oedema.*

- ***Heat stroke*** *- in the sun called sunstroke - is heat collapse with high brain temperature that occurs suddenly, hereby creating a life-threatening condition.*

- ***Hypothermia*** *is a fall in core temperature to values below 35 ^{o}C. Hypothermic subjects lose consciousness, when the core temperature falls below 32 ^{o}C - a potentially lethal condition called severe hypothermia.*

- ***Artificially induced hypothermia*** *is used in brain- and heart-surgery, where the usual thermocontrol is inactivated by general anaesthesia.*

- ***Penetrating radiation*** *consists of either gamma rays (neutrons) or X-rays. Survivors from nuclear power plant accidents with whole body absorption greater than 100 rads are threatened by acute and chronic radiation sickness.*

- ***Acute radiation sickness*** *appears as vomiting and malaise following exposure to 1 Gy (100 rad) or more. Lymphocyte production is reduced immediately, soon followed by leucopenia and thrombocytopenia with bleeding. The villi of the gastrointestinal tract are destroyed, absorption of nutrients is impaired, and new villous cells are not produced. Diarrhoea, often with blood loss, results in dehydration. The skin is red and blistering, and the hair is loosed.*

- ***Chronic radiation sickness*** *or late radiation damage implies an increased rate of mutagenesis, which includes a high frequency of leucaemia, cancer of the brain, the thyroid and the salivary glands, infertility and cataract.*

Further Reading

Scientific American. Monthly journal published by Scientific American Inc., 415 Madison Avenue, N.Y., USA.

Benzinger, T. H. "The human thermostat." *Sci. Am.* 204: 134-147, 1961.

Hong, S. K. *News in Physiol. Sci.* 2: 79, 1987.

Wasserman, D.H. and A.D. Cherrington. "Hepatic fuel metabolism during muscular work: Role and regulation." *Am. J. Physiol.* 260: E811, 1991.

Return to top

Return to content

Chapter 32

Immunology And Immune System Disorders

Study Objectives

- To *define* the following concepts and disorders: AIDS, allergy, autoimmune reactions, HIV, interleukins, neutropenia, the reticulo-endothelial system (RES), serum disease, and vaccination.
- To *describe* anaphylactic reactions, the role of natural killer cells, of soluble lysozymes, active and passive immunisation.
- To *explain* the mechanisms involved in allergy, acquired and congenital immunity, haemopoiesis, phagocytotic killing, myasthenia gravis, ulcerative colitis and Crohns disease, rheumatoid arthritis, insulin-dependent diabetes mellitus, pernicious anaemia, and Graves hyperthyroidism.
- To *use* the above concepts in problem solving and case histories.

Principles

- *The immune system defends the body against invading agents, participates in autoimmune and hypersensitivity disorders, and determines transplant tissue reactions.*
- *The ability to recognise foreign antigens allows destruction and removal of invading organisms by various effector mechanisms.*
- *Inappropriate immune reactions against self-antigens or host cells result in autoimmune disorders.*
- *Overt responses to an antigen result in hypersensitivity disorders.*

Definitions

- **Acquired Immuno-Deficiency Syndrome** (AIDS) results from infection with *Human Immuno-Deficiency Virus* (HIV). This is a substantial T-lymphocyte defect. The HIV is bound to receptors on $CD4^+$ T-lymphocytes, but also to monocytes and macrophages.
- **Autoimmune disorders** occur when the immune system kills the body's own cells or *self-antigens*, because the system fails to recognise them. Autoantibodies are produced and immune complexes are deposited in the tissues. Here, complement and neutrophils are activated and accumulated. The neutrophils release proteolytic enzymes and toxic cytokines.
- **Delayed allergic reactions** occur when highly resistant allergens (mycobacterium such as TB, fungi, contact allergens) get into contact with T-lymphocytes, and do not involve production of antibodies. Repeated contact with antigen *activates* helper T- and killer T-cells.
- **Interleukins** are toxic lymphokins and monokines from lymphocytes and monocytes.
- **Natural killer cells** destroy tumour cells and virus-infected cells. They are unspecific, non-phagocytotic lymphocytes that are activated by *interferon* produced by the affected cell. Interferon induces a high degree of resistance in the affected cell.
- **Neutropenia** are too few neutrophils in the peripheral blood.
- **The reticuloendothelial system** (RES) is also called the *mononuclear phagocytotic system* (MPS). Lymphoid

organs belonging to the RES are the bone marrow, the liver, the spleen, the lymph nodes, the microglia of the brain, the thymus, tonsils, as well as MALT, BALT, GALT, and SALT (see below).

- **Serum disease** results from passive immunisation with too many antibodies. Antigen-antibody complexes accumulate in the blood. They precipitate in the slow bloodstream of the capillaries.

- **Soluble lysozymes** are enzymes in plasma, lymph, extracellular fluid, saliva, gastric fluid and other secretions, which can destroy the bacterial wall.

- **Vaccination** is iatrogenous immunity. At the first vaccination with a certain antigen, some plasma cells transform to memory B-cells that remain in the RES. At the second vaccination, the memory B-cells evoke an exaggerated antibody production that rapidly deactivates the antigens.

Essentials

This paragraph deals with *1. The immune system*, *2. Congenital immunity*, *3. Acquired immunity*, *and* *4.Vaccination.*

1. The immune system

Burnet received the Nobel Prize in 1960 for his *clonal selection theory*. Pluripotent stem cells differentiate into millions of different B-lymphocytes, T-lymphocytes, erythrocytes, polymorphonuclear leucocytes, monocytes, macrophages and mast cells.

Lymphocytes are the most important cells involved in immune responses. Without exposure to all the antigens each of the B-lymphocytes have inherited the ability to divide into a clone of plasma cells. The first contact with the specific antigen starts the clone production. The clone of plasma cells produces the specific immunoglobulins. This understanding of the immuno-reaction made transplantation possible.

Thomas made the first *transplantation* of kidneys in 1956, and received the Nobel Prize in 1990 for his contribution to science and therapy. The second successful transplantation was the transplantation of bone marrow to treat *leukaemia* (ie, uncontrolled proliferation of *impotent* leucocytes).

Overactivity in the immune system causes allergic and autoimmune disorders, whereas underactivity results in immunodeficiency.

The *immune system* is a complex of cells and humoral factors controlled by the hypothalamo-hypophysary axis in concert with the adrenal and probably other endocrine glands (Fig. 32-1). The major organs of the *reticuloendothelial system*, RES (bone marrow, lymph nodes, spleen, and thymus) receive sympathetic efferents - in particular the T-lymphocyte regions (see below). Hereby, the CNS (via the hypothalamus) modulates the intensity of immunoreactions. Endotoxins from the normal bacterial flora of the intestine constantly enter the blood. Ordinarily they are inactivated by phagocytic activity of the RES mainly in the liver. Macrophages not only inactivate endotoxins; they also release hydroxylase, proteases, certain coagulation factors and arachidonic acid derivatives (ie, prostaglandins, thromboxanes, leucotrienes and monokines). *Monokines* are control proteins that modulate metabolism (Chapter 20), temperature control (Chapter 21), hormone secretion (Chapter 26>30), and the immune defence systems. The important role of RES in haemorrhagic and endotoxic shock is described in Chapter 12.

Fig. 32-1: Control of the immune system by the hypothalamo-pituitary axis during an antigen attack.

The *immune system* protects us against disease. The system confers congenital (inborn) and acquired immunity. Both subsystems activate soluble humoral factors as well as fixed and mobile cells.

Congenital immunity involves T-lymphocytes, which are derived from the thymus, whereas acquired immunity involves B-lymphocytes and the production of antibodies.

The *bone marrow* is the site of haemopoiesis, since all blood cells are derived from the *pluripotent stem cell* or haemocytoblast (Fig. 32-1). This is a primitive cell type, which can divide rapidly and differentiate into *committed*

stem cells. The committed stem cells are colony- forming in that they are committed to produce large quantities of *erythrocytes, granulocytes* (neutrophils, eosinophils and basophils), *monocytes-macrophages, megacaryocytes-blood platelets*, and *B- & T-lymphocytes* (Fig. 32-1) depending upon various growth inducers or cytokines.

Interleukines are *toxic lymphokines* and *monokines* from lymphocytes and monocytes. They inhibit the hypothalamic production and release of *corticotropin-releasing- hormone* (CRH) just as *cortisol* - thus reducing the immuno-response. Cortisol also inhibits lymphocyte and monocyte production from stem cells (Fig. 32-1). Normally, CRH stimulates the synthesis and release of *adrenocorticotropic hormone* (ACTH). Stimulation of the sympathetic nervous system by immunological stress releases *adrenaline* from the adrenal medulla. Adrenaline probably stimulates most of the blood cell formation from stem cells (erythro- granulo- lympho- monocyto- thrombo- poiesis) via a_2-adrenergic receptors.

2. Congenital immunity

The inborn immune defence system is *unspecific* and responsible for *immediate responses* to infection (bacteria, fungi, parasites, and viruses) and other pathogens (from tumours or other sources).

The inborn system is immediately activated with all the elements of congenital immunity: *Phagocytes* (neutrophils and macrophages), *cytotoxic eosinophils*, histamine-containing *basophils* and *mast cells*, and *essential* proteins (complement, acute phase proteins, heat shock proteins).

Phagocytes comprise a large number of neutrophils, which are released from the bone marrow during acute infection. *Neutrophilic granulocytes* have an extremely short life cycle, namely 24 hours. They are leucocytes formed in the bone marrow. The production of neutrophils is increased by the action of *granulocyte-colony stimulating factor* (G-CSF) and *granulocyte-macrophage-colony stimulating factor* (GM-CSF). During severe long-lasting infections the bone marrow is exhausted and too few neutrophils are released to the blood (ie, *neutropenia*). Neutrophils are important in the defence against microorganisms.

Fig. 32-2: Congenital immune defence against bacteria.

Granulocytes can leave the blood by moving between endothelial cells to reach the interstitial space of different tissues.

2.1. Steps of microbial destruction

Bacterial invasion. When bacteria invade the body, macrophages release the complement cascade, and B-lymphocytes release immunoglobulins (Fig. 32-2).

Neutrophilic chemotaxis. Complement cascade products and leucotriene B_4 (see later) are released from cells in the infection area. These molecules attract neutrophils from the blood into the infected tissue by so-called *chemotaxis* (ie, attraction of cells by foreign chemical substances). The neutrophils pass the endothelial wall by *diapedesis* (ie, they squeeze through the capillary wall - see Fig. 32-2). Neutrophils surround the microbe with their *pseudopodia* and engorge them. Neutrophils are large enough to phagocytize bacteria and fungi, but they cannot phagocytize larger organisms such as parasites.

Phagolysosomes. The pseudopodia form a membrane bound vesicle around the microbe, and the vesicle is then released as a free-floating *phagosome*. Inside the neutrophil, the phagosome fuses with *neutrophil granules* to form *phagolysosomes*, where the killing occurs. Phagocytes get hungry from *opsonization* of the pathogen surface with complement or with specific immunoglobulins such as IgM and IgG.

Microbial perforation. The complement released from many macrophages also fights its own battle. Besides being bound to immunoglobulins, complement is also bound to the surface of bacteria, whereby they get leaky.

Microbial breakdown. Phagocytotic killing occurs in the *phagolysosomes*. The method of execution is by a *respiratory burst* or by *gas*. Oxygen is reduced to reactive oxygen metabolites by an NADPH oxidase. These reactive metabolites are hydrogen peroxide and oxygen radicals. Many toxic proteins or enzymes (lipases, proteases) take part in the

destruction. Immuno-stimulated macrophages produce nitrite and nitrate, and their killer activity is related to the unstable gas, *nitric oxide* (NO). NO is produced in large quantities by the macrophages, kills microbes and cancerous cells. NO is synthesized from one of the guadino nitrogens of L-arginine by the enzyme *nitric oxide synthase*. Several synthases have been purified and cloned. The enzymes represent a new family that contains a haeme moiety.

Soluble lysozymes are enzymes in plasma, lymph, extracellular fluid, saliva, gastric fluid and other secretions, which can destroy the bacterial wall (Fig. 32-2).

Specific antibodies and complement *cascade* substances ease the execution of microbes. Neutrophils carry receptors for immunoglobulins and complement on their surfaces, which increase the binding force between the cell and the microbe, and simultaneously transduct signal molecules to increase the enzymatic killing activity. This is a typical co-operation between congenital and acquired immunity. The capillaries in the area dilate and get leak for proteins. This is why the site of invasion gets hot, red, swells, and pains (Latin: calor, rubor, tumour, and dolor -the classical signs of inflammation).

2.2. Cytotoxic eosinophilia

Eosinophils contain granules with substances, which become cytotoxic, when they are released on the surface of parasites. Thus, eosinophils are mainly involved in reactions against parasitic infections. Eosinophils are not phagocytic, but they intoxicate nematodes and other parasites and bacteria. The cytotoxic substances are *major basic protein*, which kill helminths, *eosinophil cationic protein* (an extremely efficient killer of parasites and potent neurotoxins) and *eosinophil peroxidase* (kills bacteria, helminths and tumour cells). Eosinophils are involved in hypersensitivity reactions.

2.3. Histamine containing cells

Circulating basophils and mast cells residing in the tissues are morphologically similar with granules that contain histamine and other vasoactive amines. These histamine containing cells are involved in hypersensitivity reactions (see IgE-mediated allergy). The binding of IgE to the cells stimulate the release of histamine, but also of prostaglandins, leucotrienes and cytokines. These substances cause immediate (Type I) hypersensitivity. The T- mast cell contains trypsin and cytoplasmic IgE, and the TC- mast cell contains both trypsin and chymotrypsin.

2.4. Natural killer cells

Such cells destroy tumour cells and virus-infected cells. They are unspecific, non-phagocytotic lymphocytes that are activated by *interferon* produced by the affected cell. Interferon induces a high degree of resistance in the affected cell.

2.5. Essential proteins

Complement extirpates microbes and immune complexes. The complement system includes several serum glycoproteins that are activated in a *cascade* similar to the *coagulation cascade* (Chapter 8). Complement activation destruct and removes microbes, immune complexes and tumour cells, recruits cells and proteins to infection sites by chemotaxis, and modulates the *B-cell immune response*.

Acute phase proteins (C-reactive protein, complement complex, fibrinogen, haptoglobulin, caeruloplasmin, a_1-antitrypsin) are produced in response to infection and inflammation (ie trauma, necrosis, tumours etc). The disease activity is measured in blood serum as *C-reactive protein*.

Heat shock proteins preserve the protein structure of cells during infections. They resemble antigens and are involved in immunity and autoimmunity.

3. Acquired immunity

Antigen stimulation of inactivated lymphocytes results in development of *humoral*- or *cell-mediated* immune responses. Humoral responses involve antibodies from B-lymphocytes activated to large antigen-producing plasma cells. Also macrophages and T-helper cells are required. Cell-mediated responses require cells that produce antigens and cytokines to T-helper cells.

Some pathogens can prevent phagocytosis or suppress the formation of lysosomes or kill the neutrophils. When attacked by such pathogens we must rely on acquired or specific immunity. This is produced by rearrangements of germ-line DNA in B- and T-lymphocytes. Hereby, specific *antibodies* and specific antigen-binding T-cell *receptors* are created. Tonegawa has shown how a rearrangement of DNA in only a few genes can produce millions of different antibodies in an individual. This is enough to cover all antigens encountered.

In foetal life, cells from the bone marrow pass through the gastrointestinal lymph nodes. Here the inactive cells become immunologically *active B-lymphocytes*. The cells re-enter the blood and migrate to the foetal spleen, liver and other lymph nodes. When an antigen binds to receptors on these cells, the lymphocytes divide, and from now on the whole clone of plasma cells can produce the specific antibody.

3.1. The reticuloendothelial system (RES)

This system is also called the *mononuclear phagocytotic system* (MPS). Lymphoid organs belonging to the RES are the bone marrow, the liver, the spleen, the lymph nodes, the microglia of the brain, the thymus, tonsils, as well as MALT, BALT, GALT, and SALT (see below). These organs contain macrophages originating from monocytes. Inactivated and circulating macrophages are called monocytes, but when they nitigate into extravascular tissues they are known as macrophages. Macrophages contain lysosomes filled with various catabolic enzymes. The macrophage membrane contains receptors for binding complement components and immunoglobulins. Macrophages destroy other phagocytized organisms or molecules by production of free radicals and digestive enzymes. Tumour necrosis factor (TNF) is produced by macrophages stimulated by bacterial cell wall components. TNF turns a tumour into haemorrhagic necrosis. Recombinant TNF is available in the form of TNF-a and TNF-b (lymphotoxin).

The cell content of the RES organs covers fixed and locally wandering macrophages as well as B-lymphocytes, which produce the antibodies after antigen exposure, and are now called plasma cells. B-lymphocytes comprise 25% of all lymphocytes. The remaining lymphocytes (75%) are T-lymphocytes, which are undergoing a maturation process in the thymus. T-lymphocytes possess distinct cell surface antigens. RES receive sympathetic efferents. Hereby, the hypothalamus can modulate the intensity of immunoreactions. This is what is termed the *psycho-immune coupling process* (Fig. 32-1).

The *spleen* is the largest lymphoid organ in the body containing both B- and T-lymphocytes. The *lymph nodes* are distributed all over the body. The *thymus* contains cells that originate from the bone marrow. The lymphocytes derived from the thymus are called T-lymphocytes. The immature T-lymphocytes are matured to $CD4^+$ and $CD8^+$ by thymic hormones. The thymus also deletes cells that are reactive to the body's own tissues (*clonal deletion*). MALT, BALT, GALT, and SALT are lymphoid tissues found in the intestinal mucosa (mucosa-associated lymphoid tissue = MALT), in the wall of the main bronchi (bronchus-associated lymphoid tissue = BALT), in the gut (gut-associated lymphoid tissue = GALT), and in the skin (skin-associated lymphoid tissue = SALT). *Tonsils* combat airborne antigens by the help of antigen producing B-lymphocytes.

Fig. 32-3: Formation of sensitised lymphocytes, lymphokines and antibodies. B-lymphocytes are involved in acquired, humoral immunity, and T-lymphocytes in congenital, cellular immunity.

3.2 T-lymphocytes

acquire their immune competence in the thymus (Fig. 32-3). They are divided into *helper T-cells* and *killer T-cells*. Helper T-cells carry CD4 protein on their surface and produce *lymfokines* (interferon and interleukin-2 and -4). Helper T-cells help the killer T-cells to proliferate, to destruct antigen and to reinforce antibody production. Some external antigen molecules are processed in macrophages before they bind to the lymphocytes (Fig. 32-3). Helper T-cells activate *resting B-lymphocytes*, so they differentiate to *plasma cells* or to *active B-lymphocytes*. Some new cells develop to plasma cells and remain in the lymph nodes.

Interleukin-2 is a peptide of 133 amino acid moieties. This substance stimulates the production of lymphokine-activated killer cells that destroy tumour cells without affecting normal cells. Interleukin-3 stimulates the primitive stem cell.

Interferon is called so, because they interfere with viral RNA and protein synthesis; interferon probably induces enzymes that destroy viral RNA and other proteins. Human can produce at least 3 types of interferon (a, b, g); they are glycoproteins with a molecular weight of 20-25 kD.

With viral or RNA stimulation, a-interferon is synthesized in macrophages (Fig. 32-4) and b-interferon in fibroblasts and macrophages. The g-interferon (no sequence homology to the two other forms) is produced in antigen-stimulated T-lymphocytes. The g-interferon stimulates the antigen production in macrophages and B-lymphocytes.

Recombinant interferon is commercially available. Interferon is used in the treatment of severe attacks of condylomata acuminata, chronic hepatitis B or C, and certain types of sarcomata.

Fig. 32-4: Production of interferon in macrophages following stimulation with RNA or virus.

T-lymphocytes constitute the majority of blood lymphocytes. The lymphocytes proliferate at first contact between antigen and T-lymphocytes. Some new cells bind the antigen in an antigen-antibody reaction and destroy the antigen. *Killer T-cells* is the proper name for these cells, but the destruction of antigen requires the co-operation of *helper T-cells*. Helper T-cells stimulate the proliferation and differentiation of killer T-cells to increase their number. A subgroup of *effector T-cells* can suppress antibody formation by *B-lymphocytes* and inhibit other effector T-cells. the so-called *suppressor T-cells*. Congenital immunity is a delayed form of immunity. The response reaches a peak after 2 days. Delayed immunity reaction encompasses the rejection of transplants, contact allergies and defence reactions against certain viruses and fungi. The T-cell number is deficient in AIDS victims (*Acquired Immune Deficiency Syndrome*).

The T-lymphocytes recognise self-antigens, known by the body's own cells and *non-self antigens* from cancer cells, foreign cells (transplantates) and foreign molecules (external antigens). This recognition ability is acquired early in life, when lymphoid stem cells migrate into the thymus, where a few are modified into *memory T-cells* and released to the blood.

At the first contact between antigen and T-lymphocyte the cell proliferates. Some of the new lymphocytes are killer T-cells. They bind the antigen in an antigen-antibody complex and destroy the antigen. Killer T-cells carry *CD8 protein* on their surface and kill other cells suffering from cancer or virus infection. Some of the killer T- cells are actually *suppressor T-cells*, because they can suppress antibody formation by the B-lymphocytes and inhibit other effector T-cells. Hereby they can down-regulate the immune response when necessary to prevent autoimmune responses.

3.3. B-lymphocytes

produce *specific antibodies* or immunoglobulins to antigens (Fig. 32-3). The immunoglobulins form part of the gamma-globulin fraction in plasma. The B-lymphocytes multiply by *clonal expansion*. A specific antigen is bound and recognized to the B cell through special surface immunoglobulins. B cells also contain CD19 and CD20 proteins. Cytokines activate the B-lymphocyte, so it divides and the resulting cells differentiate to enormous plasma cells with an overwhelming surplus of protein-producing endoplasmic reticulum (ER). This is why plasma cells produce large amounts of antibodies and release them into the blood as Y-shaped molecules (Fig. 32-3). The plasma cells have a short life cycle, and die when they have fulfilled their defence mission. Hereby, the B-lymphocyte population is reduced to its normal size apart from a few cells remaining as *memory cells*. The antibodies are also called *immunoglobulins* (Ig). They are specific serum glycoproteins. Each antibody is Y-shaped and consists of heavy and light polypeptide chains. The heavy chains with complement receptors provide the *constant domain* of the Ig molecule, which is the same in all antibodies (Fig. 32-5). The light chain region constitutes the *variable domain*, which is functionally important. Antibodies deactivate antigens by forming a complex, which causes agglutination and precipitation, by masking the active sites of the antigens, or by activating the complement cascade. A single Ig with its antigen activates a complement cascade with mobilisation of up to 10^9 new complement molecules carrying lots of enzymes that rapidly lyse the antigen-carrying microbe.

The most abundant is *IgG*, which has a high antigen affinity and is the antibody of the secondary response to protein antigens (viruses and tetanus toxin). IgG can cross the placental barrier and protect the newborn for a couple of

months.

IgM is confined to the blood, because it is a pentameric molecule (5 IgM molecules joined together). IgM cannot cross the placental barrier, and is responsible for the primary immune response.

IgA_1 predominates in serum, whereas IgA_1 and IgA_2 are present in equal amounts in secretions such as saliva, gastric juice, pancreatic and intestinal juice. IgA protects mucosal surfaces in the gut, respiratory and urinary tracts, by preventing the attachment of poliovirus, enterovirus, bacteria, and enterotoxin.

The concentration of *IgD* in serum is high in disorders with B-lymphocyte activation such as *AIDS* and *Hodgkin's disease*.

IgE is mainly bound to basophils and mast cells, and involved in the pathogenesis of *allergic* and *nematode diseases*.

Fig. 32-5: An immunoglobulin (Ig) or antibody molecule with two antigen molecules attached (left). The immunoglobulins IgA and IgM are build up of two or more immunoglobulins moieties connected with disulphide bonds (right).

4. Vaccination

This is iatrogenous immunity. At the first vaccination with a certain antigen, some plasma cells transform to memory B-cells that remain in the RES. At the second vaccination, the memory B-cells evoke an exaggerated antibody production that rapidly deactivates the antigens.

Vaccine from death microbes is used for bacterial diseases such as diphtheria and typhoid fever. Other vaccines are derived from toxins that are deactivated without losing their antigen specificity (tetanus, botulism). Vaccines against viral disease have passed through a series of other organisms, until a mutant originates without pathogenic activity but with intact antigen specificity (measles, polio, smallpox, and yellow fever).

Tumour-antigen vaccines are under development in order to stimulate an immune reaction against tumour cells.

Pathophysiology

This paragraph deals with *1.Congenital and acquired immuno-deficiencies due to underactivity of the immune system, and 2. Allergic and autoimmune diseases caused by overactivity of the immune system.*

1. Congenital and acquired immuno-deficiencies

The 5 major types of *congenital immuno-deficiency* are:

1. Chronic granulomatous disease, which is a congenital defect of the neutrophil killing mechanism.
2. Complement cascade deficiencies.
3. B-lymphocyte defects with antibody deficiency.
4. Absent thymus with T-lymphocyte deficiency.
5. Combined immunodeficiency with severe defects of both B- and T- lymphocyte function.

Acquired immunodeficiency (AID)

Iatrogenic AID is caused by splenectomy, chemotherapy or iatrogenically induced malnutrition. Glucocorticoids suppress immune responses and the movement of lymphocytes from the blood to the tissues.

Autoimmune suppression by infection is *Acquired Immuno-Deficiency Syndrome*. AIDS results from infection with *Human Immuno-deficiency Virus* (HIV). This is an overwhelming T-lymphocyte defect. The HIV is bound to receptors on $CD4^+$ T-lymphocytes, but also to monocytes and macrophages. HIV and AIDS are described in Chapter 33.

2. Allergy & autoimmunity

Allergic and autoimmune disorders are typically *hypersensitivity reactions*.

Allergic reactions are either *immediate anaphylactic (type I) or delayed (type II). Autoimmune conditions are also called type III hypersensitivity reactions.*

2.1.Immediate anaphylactic reactions (type 1)

occur immediately when the allergen sensitises B-lymphocytes with allergen specific IgE antibodies (ie, within minutes). At the next exposure to the same allergen, plasma cells recognise the allergen and release large amounts of IgE (Fig. 32-6). The allergen-IgE complex is bound to IgE-receptors on the surface of the mast cells and basophils. Hereby, histamine, serotonin (a vasoconstrictor), lymphokines, *platelet activating factor* (PAF) and *toxic eichosanoids* (prostaglandins and leucotrienes) are released (Fig. 32-6). PAF activates both thrombocytes and phagocytes. Histamine is a powerful vasodilatator and bronchoconstrictor. *Leucotrienes* were previously called *slow reacting substances for anaphylaxis* (SRS-A). Leucotrienes are strong bronchoconstrictors causing bronchial asthma attacks. Cells participating in inflammatory and immune responses are suppressed by prostaglandin E_2 (PGE_2). When PGE_2) stimulates adenylcyclase in activated neutrophils, they release less leucotriene (LTB_4) than else; the same mechanism is probably functioning in other cells.

Atopic allergy is genetic and characterized by large amounts of IgE in the blood. An antigen molecule with multiple binding sites can bind to many IgE molecules on the mast cell and release large amounts of anaphylactic substances (Fig. 32-6). Three typical disorders are the following:

a. The so-called *anaphylactic shock* is often fatal shortly after the histamine release. Adrenaline injected intravenously or intracardially may save the victim.

b. *Urticaria* and *hay fever* are anaphylactic reactions in the skin and the nasal mucosa, respectively. The histamine released causes vasodilatation, with increased capillary pressure and ultrafiltration causing oedema and red colour.

c. *Bronchial asthma* in an atopic person is a bronchiolar anaphylactic response. Leucotrienes from the bronchiolar mast cells cause bronchoconstriction, mucosal infiltration with inflammatory cells, mucosal oedema and hypersecretion. Leucotrienes are blocked by a variety of bronchodilatators.

Fig. 32-6: Anaphylactic reaction as it occurs in mast cells and basophils.

2.2. Type II:

Delayed allergic reactions occur when highly resistant allergens (mycobacterium such as TB, fungi, contact allergens) get into contact with T-lymphocytes, and do not involve production of antibodies. Repeated contact with antigen *activates* helper T- and killer T-cells. These cells diffuse into the skin, lungs or other organs and there release a cellular immune response.

2.3 Type III:

Autoimmune disorders occur when the immune system kills the body's own cells or *self-antigens*, because the system fails to recognise them. Autoantibodies are produced and immune complexes are deposited in the tissues. Here complement and neutrophils are activated and accumulated. The neutrophils release proteolytic enzymes and toxic cytokines. They may kill cells in A) a single organ system or B) affect multiple systems.

2.3.A) Single organ disorders:

Insulin-dependent diabetes mellitus (IDDM). Newly presenting patients possess islet-cell antibodies that have destroyed all the insulin producing b-cells of the pancreatic islets (Chapter 27). There is an inherited increased susceptibility to IDDM.

Pernicious anaemia. Typically parietal cell antibodies are found in the blood. They may kill the entire parietal cell population leading to atrophy of the mucosa. The atrophic gastric mucosa fails to produce hydrochloric acid and

intrinsic factor for vitamin B_{12} (Chapter 8). Also intrinsic factor antibodies have been found, either able to block the complex binding between intrinsic factor and vitamin B_{12}, or capable of blocking the binding of the intrinsic factor and vitamin B_{12} -complex to its ileal receptors.

Graves's or Basedow's disease is hyperthyroidism combined with eye signs (*exophtalmus*). Normally, the thyroid stimulating hormone (TSH) increases the thyroid hormone production after its binding to the *thyroid TSH receptors*. Bacterial infection in a genetically susceptible person may be the cause of autoimmune production of the TSH-receptor antibodies. These IgG -antibodies behave exactly like TSH itself, and thus stimulate the thyroid hormone production (Chapter 28). Retroorbital swelling and damage of the extraocular muscles cause the thyroid eye disease from specific antibodies.

Atrophic hypothyroidism is caused by microsomal autoantibodies in the thyroid. The hypersensitivity reactions result in thyroid atrophy and hypothyroidism (Chapter 28).

Hashimoto's thyroiditis is caused by other microsomal autoantibodies. The inflammatory reaction produces goitre (struma) with or without hypothyroidism.

Primary hypo-adrenalism (Addison's disease). The entire adrenal cortex is destroyed by organ-specific autoantibodies, and its steroid hormone production (sex hormones, glucocorticoids, and mineralocorticoids) is lost (Chapter 30). Autoimmune disease is the cause of destruction in 4 of 5 patients. Other causes are TB, cancer, infections, and AIDS.

Myasthenia gravis. This is a rare disease caused by autoantibodies against the *acetylcholine receptors* of the neuromuscular endplate (Chapter 2). Complement complexes and IgG molecules are deposited at the endplates. Many patients have thymic hyperplasia.

Crohns disease and ulcerative colitis

These two disorders may be different manifestations of a single disease, *non-specific inflammatory bowel disease.*

Ulcerative colitis is a disorder located to the *colon* only, whereas *Crohns disease* can involve any part of the gastrointestinal tract.

Fig. 32-7: Inflammatory bowel disease (Crohns disease).

Transmissible agents are suspected but not found. Both antigen specific, cellular and autoimmune responses have been postulated. The immune-hypersensitivity is evident from activation of T-lymphocytes, neutrophils, eosinophils, basophils and mast cells. A variety of cytokines, Eicosanoids, oxygen radicals and NO are produced. Dysfunction of the terminal ileum has serious consequences, because bile salts are not reabsorbed and vitamin B_{12} is not absorbed.

The cause of inflammatory bowel disease is unknown, but some features are common to the two conditions:

1. *Autoimmune disease* in the gut lumen and wall.
2. *Cytotoxic T lymphocytes* sensitised by antigens destroy the mucosa or the whole gut wall.
3. *Genetic predisposition* (concordance in monozygotic twins, familial occurrence),
4. Infective agents are *suspected* but not identified.
5. Uveitis, episcleritis, arthritis, anchylosing spondylitis, erythema nodosum occur in both types of bowel disease.
6. The clinical picture includes abdominal pain and diarrhoea often with blood and mucus. *Steatorrhoea* typically points to ileal involvement, whereas frequent *bloody diarrhoeas* indicate colic involvement as in ulcerative colitis. There is an increased risk of colorectal carcinoma in ulcerative colitis.

Crohns disease is a chronic infection or inflammation of the gut with a particular prevalence for the *terminal ileum*, but it can be located all the way along the tract. The patients with terminal ileitis suffer from malabsorption of bile salts and vitamin B12.

Ulcerative colitis is always confined to the colon and it is a mucosal inflammation with haemorrhage and rectal bleeding.

Special hospital units accomplish different diagnosis and management.

Gluten-sensitive enteropathy or *coeliac disease* (sprue) describes a condition where the duodenal and jejunal mucosa is more or less destroyed by hypersensitivity towards gluten. Gluten is found in barley, rye, wheat, and oats. Gluten consists of 4 *gliadin peptides* out of which α*-gliadin* destroys the jejunal mucosa. There is a sequence homology between a-gliadin and *adenovirus 12*.

Nontropical sprue. In allergic persons *gluten* causes an immunological reaction with desquamation of the luminal part of the intestinal mucosa - in particular most of the microvilli. The marked fall in area available for absorption causes malabsorption. In severe cases the malabsorption involves fats causing steatorrhoe, Ca^{2+} causing osteomalacia, vitamin K causing bleeding disturbances, vitamin B_{12} causing pernicious anaemia, and folic acid causing folic acid deficiency.

The inheritance is unknown, but this is a disease occurring in atopic families, and an immunogenetic mechanism is likely - in particular after infection with *adenovirus 12*. Symptoms and signs include stomatitis with ulcers, diarrhoea, steatorrhoea, abdominal pain, osteomalacia and malnutrition with oedema. The desquamation of the mucosa sometimes descends along the villous region in the small intestine and thus includes the terminal ileum.

A gluten-free diet is necessary for successful treatment.

Tropical sprue is found in tropical areas, and probably caused by gastrointestinal infection, although the bacterial diagnosis is seldomly confirmed. Tropical sprue is often curable with antibacterial agents.

2.3.B) Multiple system disorders:

Rheumatoid arthritis (symmetrical inflammatory polyarthritis with progressive joint damage and lung, cardiac, renal and many other organ manifestations) is an autoimmune disease. The synovial fluid contains IgG, lymphokines and immune complexes. The cause is probably a persistent external antigen, which is not removed.

Connective tissue diseases are *Systemic Lupus Erythematosus* (SLE = Arthralgia, rash, cerebral and renal dysfunction), systemic sclerosis (sclerosis of the skin and oesophagus, organ fibrosis, Raynaud's phenomenon), polymyositis (necrosis of muscle fibres with proximal weakness), dermatomyositis (polymyositis with rash) and other rare disorders. The aetiology is unknown, but many of these conditions have circulating autoantibodies and immune complex deposition.

Several autoimmune disorders show a strong association.

Self-Assessment

Multiple Choice Questions

Each of the following five statements have True/False options:

A. The pluripotent stem cell is not identical with the haemocytoblast.

B. Interferon induces a high degree of resistance in the affected cell.

C. The hypothalamus can modulate the intensity of immunoreactions through sympathetic efferents to the RES.

D. Hyperthyroidism is always combined with exophtalmus.

E. The terminal ileum does not have essential functions.

Case History

*A female, 39 years of age, works as a secretary. For 6 months she has been working with **a** new copy machine. Over the last three months she has developed a blistering rash on both hands and the rash is now turned into deep scaling of the skin. There is a mild scaling at the site of both her earrings. The patient has used a bland cream without effect. The dermatologist examines her with patch tests, where solutions of common allergens, chemicals, metals including nickel are placed on her back and occluded with dressings. The next day (24 hours later) the dressings are removed and the exposed areas examined. The examination is repeated 48 hours after the beginning of the exposure. There is a severe rash at the site of the colour powder used in the copy machine, and a mild rash at the nickel spot.*

1. What is the diagnosis?

2. Which therapeutic strategy is recommendable?

Try to solve the problems before looking up the answers.

Highlights

- *The immune system is a complex of cells and humoral factors controlled by the hypothalamo-hypophysary axis in concert with the adrenal and probably other endocrine glands.*
- *The inborn immune defence system is unspecific and responsible for immediate responses to infection (bacteria, fungi, parasites, and viruses) and other pathogens (from tumours or other sources).*
- *Congenital immunity involves T-lymphocytes, which are derived from the thymus, whereas acquired immunity involves B-lymphocytes and the production of antibodies.*
- *The bone marrow is the site of haemopoiesis, since all blood cells are derived from the pluripotent stem cell or haemocytoblast, which can divide rapidly and differentiate into committed stem cells.*
- *The committed stem cells are colony- forming in that they are committed to produce large quantities of erythrocytes, granulocytes (neutrophils, eosinophils and basophils), monocytes-macrophages, megacaryocytes-blood platelets, and B- & T-lymphocytes depending upon various growth factors or cytokines.*
- *Anaphylactic shock is often fatal shortly after the histamine release. Adrenaline injected intravenously or intracardially may save the victim.*
- *Urticaria and hay fever are anaphylactic reactions in the skin and the nasal mucosa, respectively. The histamine released causes vasodilatation, with increased capillary pressure and ultrafiltration causing oedema and red colour.*
- *Bronchial asthma in an atopic person is a bronchiolar anaphylactic response. Leucotrienes from the bronchiolar mast cells cause bronchoconstriction, mucosal infiltration with inflammatory cells, mucosal oedema and hypersecretion. Leucotrienes are blocked by a variety of bronchodilatators.*
- *Iatrogenic AID is caused by splenectomy, chemotherapy or iatrogenically induced malnutrition.*
- *Autoimmune suppression by infection is the Acquired Immuno-Deficiency Syndrome (AIDS) resulting from infection with Human Immuno-Deficiency Virus (HIV). This is a substantial T-lymphocyte defect.*
- *Immuno-stimulated macrophages produce nitrite and nitrate, and their killer activity is related to the unstable gas, nitric oxide (NO). NO, produced in large quantities by the macrophages, kills microbes and cancerous cells.*
- *Phagocytotic killing occurs in the phagolysosomes. The method of execution is by a respiratory burst or by gas.*
- *Myasthenia gravis. This is a rare disease caused by autoantibodies against the acetylcholine receptors of the*

neuromuscular endplate. Complement complexes and IgG molecules are deposited at the endplates. Many patients have thymic hyperplasia.

- *lcerative colitis and Crohns disease (transmural enteric ulceration with a particular affinity for the terminal ileal cells) are of unknown aetiology, but they may represent two aspects of the same disease. Both antigen specific, cellular and autoimmune responses have been postulated.*

- *Rheumatoid arthritis (symmetrical inflammatory polyarthritis with progressive joint damage and lung, cardiac, renal and many other organ manifestations) is an autoimmune disease. The synovial fluid contains IgG, lymphokines and immune complexes. The cause is probably a persistent external antigen, which is not removed.*

- *Insulin-dependent diabetes mellitus (IDDM). Newly presenting patients possess islet-cell antibodies that have destroyed all the insulin producing β-cells of the pancreatic islets. There is an inherited increased susceptibility to IDDM.*

- *Pernicious anaemia. Typically parietal cell antibodies are found in the blood. They may kill the entire parietal cell population leading to atrophy of the mucosa. The atrophic gastric mucosa fails to produce hydrochloric acid and intrinsic factor for vitamin B_{12}.*

- *Graves's or Basedow's disease is hyperthyroidism combined with eye signs (Exophtalmus). Normally, the thyroid stimulating hormone (TSH) increases the thyroid hormone production after its binding to the thyroid TSH receptors. Bacterial infection in a genetically susceptible person may be the cause of autoimmune production of the TSH-receptor antibodies. These IgG -antibodies behave exactly like TSH itself, and thus stimulate the thyroid hormone production.*

Further Reading

*Scandinavian Journal of Immunology***.** Monthly journal published by Blackwell Science Ltd., Osney Mead Oxford OX2 OEL, UK.

*Scientific American***.** Monthly journal published by Scientific American Inc., 415 Madison Avenue, N.Y., USA.

Mims C, Playfair J, Wakelin D, and R Williams. *Medical Microbiology*. Mosby, London, 1998.

Return to top

Return to content

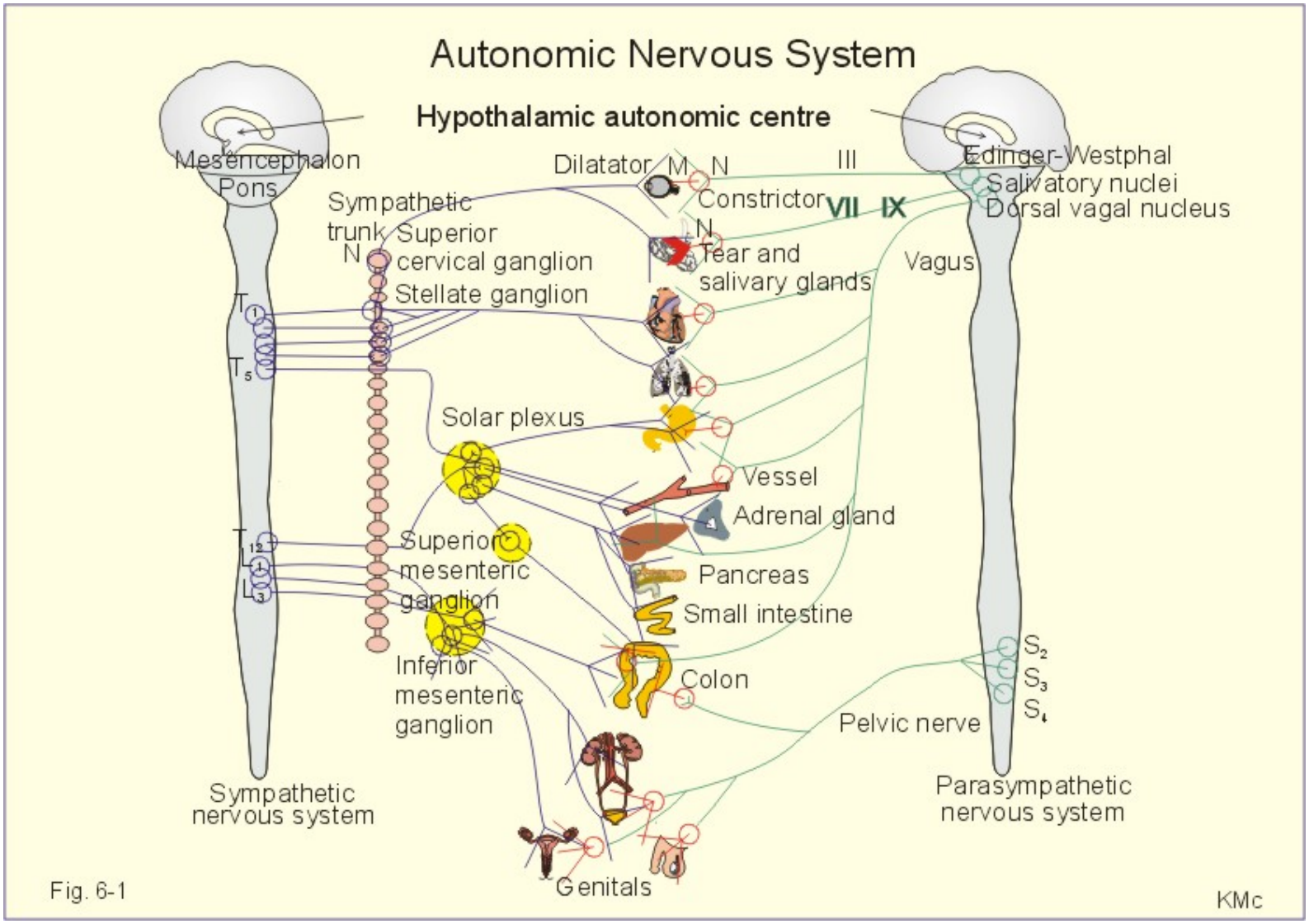
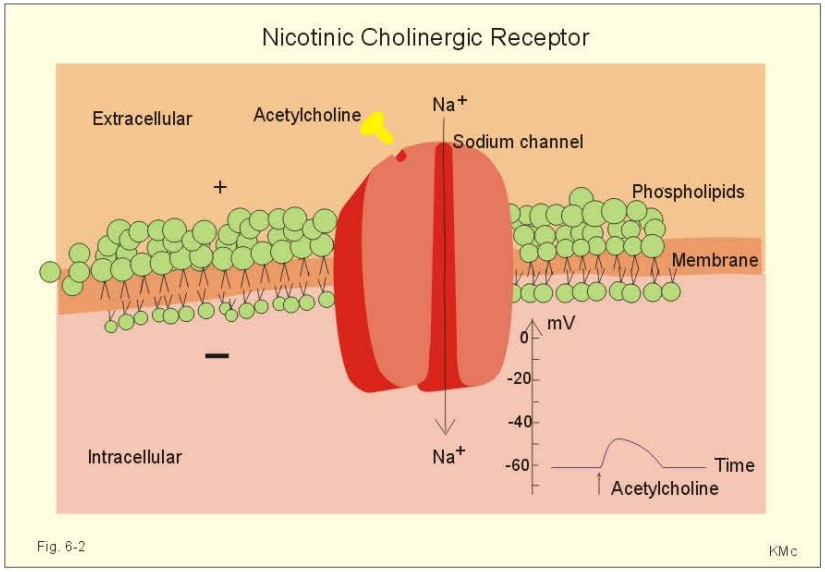
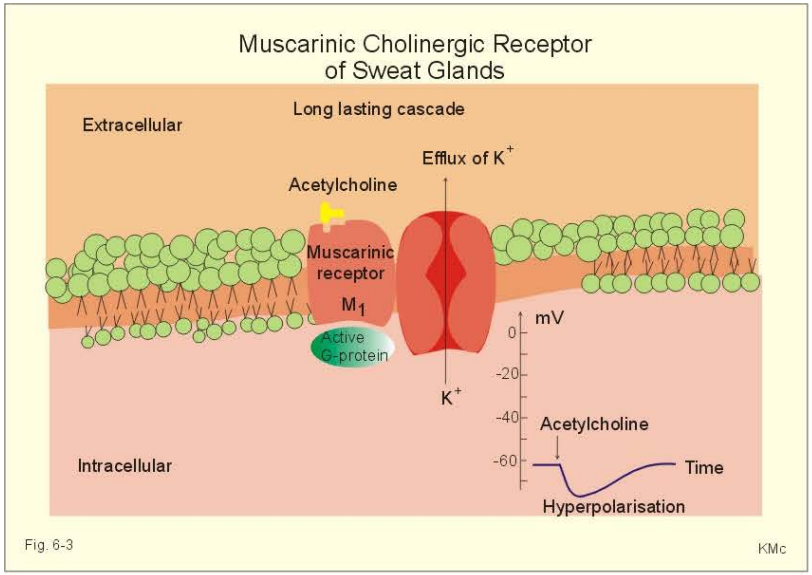

Autonomic Nervous System
Hypothalamic autonomic centre
Mesencephalon
Pons
Sympathetic trunk
N
Superior cervical ganglion
Stellate ganglion
T1
T5
T12
L1
L3
Solar plexus
Superior mesenteric ganglion
Inferior mesenteric ganglion
Sympathetic nervous system
Dilatator M
N
III
Constrictor
N
VII
IX
Tear and salivary glands
Vessel
Adrenal gland
Pancreas
Small intestine
Colon
Genitals
Edinger-Westphal
Salivatory nuclei
Dorsal vagal nucleus
Vagus
S2
S3
S4
Pelvic nerve
Parasympathetic nervous system

Fig. 6-1

KMc


Nicotinic Cholinergic Receptor
Extracellular
Acetylcholine
Na+
Sodium channel
+
Phospholipids
Membrane
mV
0
-20
-40
-60
−
Intracellular
Na+
Time
Acetylcholine
Fig. 6-2
KMc


Muscarinic Cholinergic Receptor
of Sweat Glands
Long lasting cascade
Extracellular
Efflux of K+
Acetylcholine
Muscarinic
receptor
M1
Active
G-protein
K+
mV
0
-20
-40
-60
Acetylcholine
Time
Hyperpolarisation
Intracellular
Fig. 6-3
KMc

Sympathetic Ganglion Synapse With Sif-cell

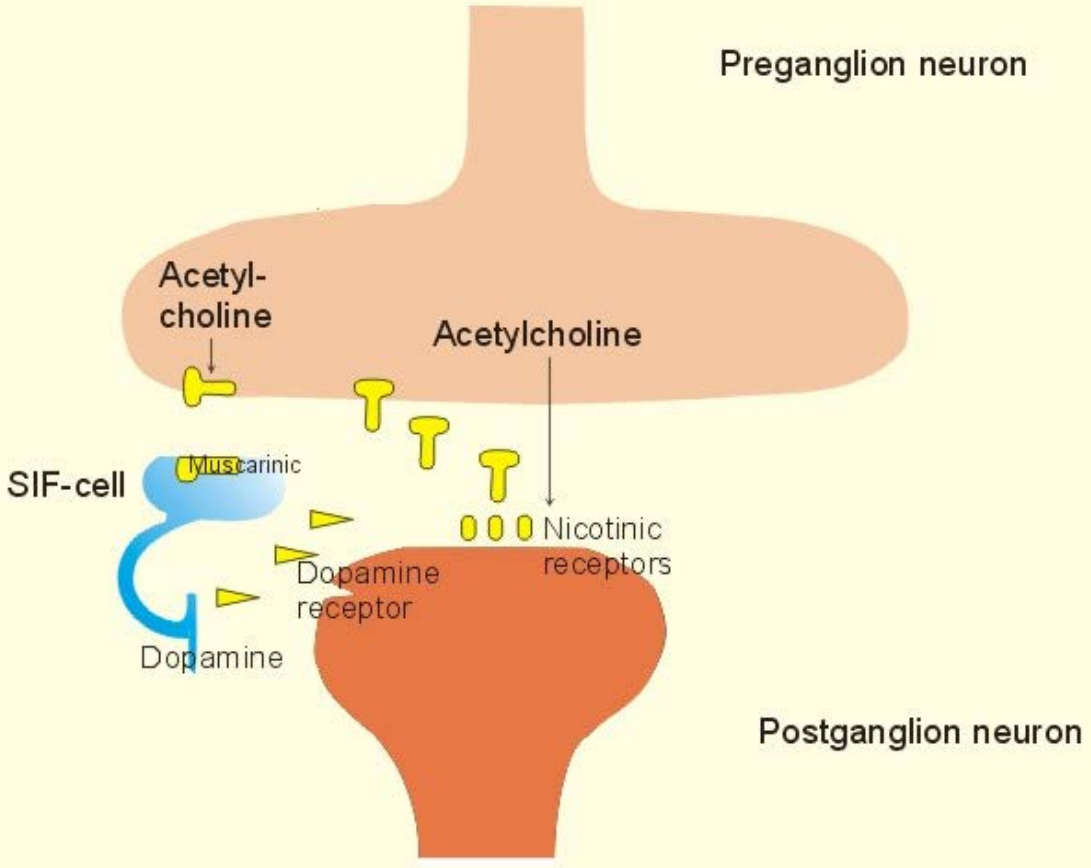


Fig. 6-5

KMc

Chapter 6. Answers

Multiple Choice Questions

I. Answers **A, B, D** and **E** are true statements, whereas **C** is false.

II. Answers **B, D** and **E** are true statements, whereas **A** and **C** are false.

III. Answers **B, C** and **E** are true statements, whereas **A** and **D** are false.

Case History

- *1. The **vagal nerve is traumatised**.*
- *2. This is a **bilateral lesion** of the vagal nerves – a bilateral vagal paresis.*

Return to chapter 6

Return to content

Chapter 7.

Multiple Choice Questions

I. Answers **A, B, C** and **E** are true statements, whereas **D** is not.

II Answers **A, D** and **E** are true statements, whereas **B** and **C** are not.

III. Answers **A, C, D** and E are true, whereas **B** is false.

Case History A

1. Degeneration of dopaminergic neurons in the substantia nigra results in severe lack of dopamine in the striatum. Lack of dopamine overactivates the GABA pathways to the motor thalamus, which activates the motor cortex neurons. The descending motor command probably increases the discharge of both alpha- and gamma-motor neurons in the spinal cord.

2. Parkinson's disease.

3. The GABA pathways overactivate thalamus due to the lack of dopamine in the striatum. Hereby, the motor cortex neurons are activated, and they increase the discharge of static gamma- motor neurons in the spinal cord.

Case History B

- *1. The event started as a partial or focal seizure, with the focus in the face and hand area of the right motor cortex. Then the overactivity spread to become generalized - a so-called grand mal seizure. The spread in a particular pattern is called Jacksonian march.*
- *2. The EEG probably shows high-voltage-high-frequency discharge over the entire cortex during the grand mal seizure. Following the seizure the EEG is possibly normal, as many epileptic patients have normal EEG activity between attacks.*
- *3. Tonic convulsions of the entire body and loss of consciousness characterise general epileptic seizures (eg. grand mal). Most of the brain is involved in a neuronal circuit activation. The hyperactive nerve cells release K^+ and glutamate/aspartate. Epileptic seizures are either initiated or propagated through NMDA-receptors on the neurons. NMDA-receptors contain ion channels that are also voltage-gated. NMDA-receptors are also called excitatory amino acid receptors (EAA-receptors), because both amino acids are excitatory. Lack of inhibitory neurotransmitters, such as GABA, may also be involved in the neuronal overactivity.*
- *4. The blood glucose was normal (5 mM), so hypoglycaemia was excluded.*

Case History C

***1**. Alzheimers disease is a form of presenile dementia (occurring before the age of 70), where the loss of mental powers is rapid. - Simple dementia is insidious but also caused by cortical atrophia.*

***2.** The age of the patient is compatible with presenile dementia, and the rapid loss of mental powers in a case of cortical atrophia is characteristic of Alzheimer's disease.*

***3.** The prognosis is poor for both disorders, and the Alzheimer patient may die within a few years of opportunistic infections.*

***4.** Most cases of Alzheimer's disease occur sporadically, but Alzheimer families exist. The gene defect causing familial Alzheimer disease is located on chromosome 21, close to the pro-A4 gene.*

***5**. At autopsy argentophilic plaques filled with amyloid protein A4 are found in the hippocampus, basal ganglia,*

thalamus and the cortex. The weight of the brain is 925 g.

Return to chapter 7

Return to content

Ion Channels in Neuronal Membrane

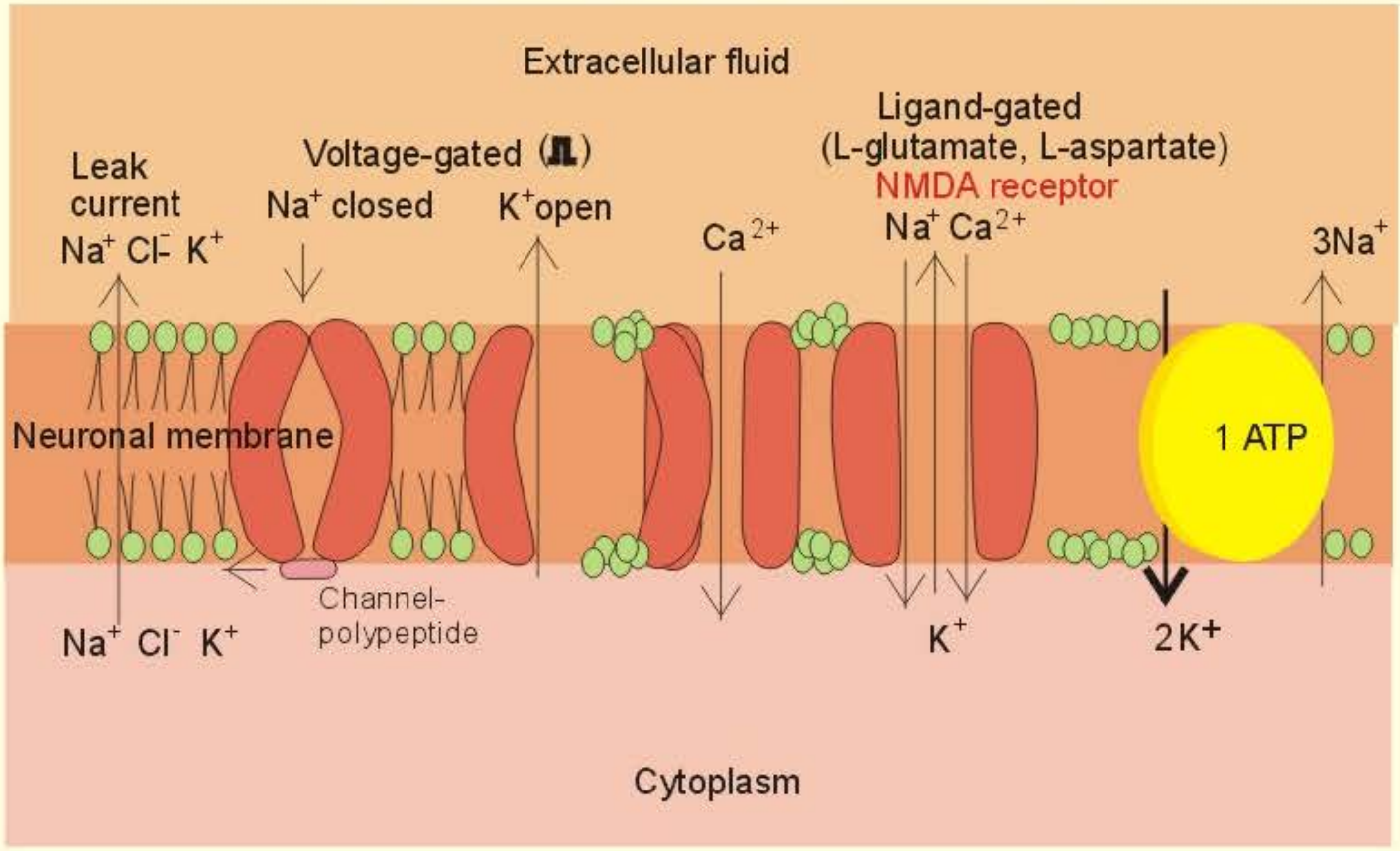


Fig. 7-1

KMc

Chapter 22.

Gastrointestinal Function And Disorders

Study Objectives

- To *define* concepts such as achlorhydria, enterogastrones, haematemesis, incretins, macrolide, malabsorption, melaena, migrating motor complex, paracrine secretion, peptide hormone families, peptic ulcer disease, peristalsis, segmentation, slow waves, and spike potentials.

- To *describe* the extrinsic and intrinsic enteric nervous system including neurotransmitters and gastrointestinal hormones, cholesterol and lipid metabolism,

- To *explain* gastrointestinal motility, gastrointestinal secretion (saliva, gastric juice, pancreatic juice, bile), digestion and intestinal absorption of nutrients, vitamins, water and iron. To *explain* the pathophysiology of common gastrointestinal disorders including malabsorption of carbohydrate, amino acids and fat, osmotic and secretory diarrhoea, and iron deficiency.

- To *use* the above concepts in problem solving and case histories.

Principles

- ***The central autonomic nervous system*** *(hypothalamus and brain stem) mediates its influence on the gastrointestinal function through the intrinsic, enteric nervous system (the so-called "little brain").*

- ***Cannons law of the gut****: The peristalsis of the small intestine always proceeds in the oral- aboral direction.*

Definitions

- **Achlorhydria** refers to absence of HCl production in the stomach

- **Defaecation** is a reflex act involving colon, rectum, anal sphincters and many striated muscles (diaphragm, abdominal and pelvic muscles). The motor pathway is the pelvic nerves. Defaecation implies a temporal release of anal continence brought about by a reflex. The coordinating centre is in the sacral spinal cord.

- **Enterogastrones** are enterogastric inhibitory hormones liberated from the duodenal mucosa by acid chyme (ie, cholecystokinin: CCK, gastric inhibitory peptide: GIP, secretin, somatostatin, neurotensin and vasoactive intestinal peptide: VIP).

- **Haematemesis** is defined as vomiting of whole blood or blood clots.

- **Incretins** are hormones, which increase *insulin secretion* from the b-cells of the pancreatic islets much earlier and to a greater extent, than when the blood glucose concentration is elevated by intravenous infusion (GIP, glicentin, glucagon-like peptides-1 and -2).

- **Intrinsic, enteric nervous system** refers to the large number of neuronal connections in the gut wall, in particular the *submucosal Meissner plexus,* which regulates the digestive glands, and the *myenteric Auerbach plexus,* primarily connected with gut motility.

- **Macrolides** are antibiotics, which bind to and prevent translocation on bacterial ribosomes.

- **Malabsorption** describes the condition resulting from inefficient absorption of nutrients by the gastrointestinal tract.

- **Melaena** is defined as passage of dark tarry stools (coal-black, shiny, sticky, and foul smelling).

- **Migrating motor complex** refers to a gastric sequence of events, where contractions occur each 90 min during fasting. There is a quiet period (I) followed by a period of irregular contraction (II), and culminated with a *peristaltic rush* (III) accompanied by increased gastric, pancreatic and biliary secretion.

- **NANC neurons** are non-adrenergic, non-cholinergic postganglionic neurons, which liberate gastrin-releasing peptide (GRP) to the gastrin producing G-cells.

- **Nitric oxide** (NO) is a possible neurotransmitter between the preganglionic and the NANC postganglionic neurons.

- **Paracrine secretion** is the release of signal molecules to neighbour cells.

- **Peptide hormone families** are groups of hormones that exhibit *sequence homology:* They possess a common amino acid sequence, such as the *gastrin family,* which has sequence homology in their terminal penta-peptide. *Peptide hormones* have *autocrine* and *paracrine* functions in the gastrointestinal tract.

- **Peristalsis** is a propagating contraction of successive sections of circular smooth muscle preceded by a dilatation. The dilatated intestinal wall is drawn over its content in this reflex mechanism, which transports the content aborally and is called the *law of the gut.*

- **Segmentation** divides the small intestine into many segments by localised circular smooth muscle contractions. Segmentation mixes the intestinal content and propagate it at a slow rate, which allows sufficient time for digestion and absorption.

- **Slow waves** (*basic electrical rhythm*) are slow gastrointestinal depolarisation's occurring at a frequency of 3-18 per min. The slow waves change the resting membrane potential of smooth muscles from -50 to -40 mV.

- **Spike potentials** are *periodic fast waves of depolarisation* that most

often follow a slow wave, and then always initiate gastric contractions (elicited by a rise in cytosolic [Ca^{2+}]).

- **Vaso-active intestinal peptide** (VIP) is a vasodilatator in line with adenosine, ATP, NO. The increased bloodflow increases intestinal secretion.

Essentials

This paragraph deals with *1. The autonomic and enteric nervous system*, *2. The cephalic, gastric and intestinal digestive phase*, *3. Mastication and swallowing*, *4. Gastric and intestinal motility*, *5. Vomiting*, *6. Colonic motility and defecation*, *7. Gastrointestinal hormones*, *8. Saliva*, *9. Gastric secretion*, *and 10. Intestinal digestion and absorption*.

1. The autonomic and the enteric nervous system

The digestive system is innervated with nerve fibres of both the sympathetic and parasympathetic divisions, although the *parasympathetic control* dominates (Fig. 22-1). Movements of the gastrointestinal tract are brought about by smooth muscle activity. There is an outer longitudinal layer, an inner circular layer, and a submucosal muscle layer (muscularis mucosae) with both circular and longitudinal fibres that moves the villi of the mucosa. The inner surface is lined with mucosal epithelium (Fig. 22-1). The outer muscle layer is covered by the serosa, which is continuous with the mesentery containing blood vessels, lymph vessels and nerve fibres.

The main CNS centres regulating digestive functions are located in the brain stem, where the sensory taste fibres from gustatory, tactile and olfactory receptors terminate on the cell bodies of the motor vagal and salivary nuclei. Many afferent, sensory fibres in the vagus nerve inform the *central autonomic system* about the condition of the gut and its content. The higher cortical and olfactory centres influence these *brain stem motor centres* and their parasympathetic outflow.

The parasympathetic system *increases* digestive activity (secretion and motility), and the sympathetic system has a net *inhibitory* effect. The generally inhibitory digestive effects of the *sympathetic nervous system* are caused indirectly by vasoconstriction, which reduces bloodflow in the digestive tract.

The vagus nerve innervates the gastrointestinal tract down to the transverse colon and contains both efferent and afferent fibres. The last part of the gastrointestinal tract receives parasympathetic innervation from the pelvic nerves.

The efferent parasympathetic fibres enhance digestive activities by stimulating local neurons of the *intrinsic, enteric nervous system* located in the gut wall (Fig. 22-1).

Fig. 22-1: The autonomic innervation of the gastrointestinal system and the structure of the enteric wall. – A sensory neuron to the CNS is shown to the left.

The intrinsic, enteric nervous system consists of two sets of nerve plexi. The *submucosal Meissner plexus* mainly regulates the digestive glands, whereas the *myenteric Auerbach plexus*, located within the muscle layers, is primarily connected with gut motility (Fig. 22-1). The nerve plexi contain

local sensory and motor neurons as well as interneurons for communication. Motor neurons in the myenteric plexus release acetylcholine and Substance P. Acetylcholine contracts smooth muscle cells, when bound to muscarinic receptors. Inhibitory motor neurons release *vasoactive intestinal peptide* (VIP) and *nitric oxide* (NO). These molecules relax smooth muscle cells.

Sensory neurons are connected to mucosal chemoreceptors, which detect different chemical substances in the gut lumen, and to stretch receptors, which respond to the tension in the gut wall, caused by the food and chyme. The short effector neurons increase digestive gland secretion and induce smooth muscle contraction. The large number of neuronal connections constitutes the intrinsic, enteric nervous system, mediating brain influence on digestive functions. The enteric nervous system is also called the *little brain*.

2. The cephalic, gastric and intestinal digestive phase

The secretion related to a meal occurs in three phases (Box 22-1).

2a. The cephalic phase is elicited even before food arrives to the stomach. The thought, smell, sight, or taste of food signals to the limbic system (including the hypothalamus) that elicits an unconditioned reflex secretion with intensity dependent upon the appetite.

Box 22-1: The secretion related to a meal from salivary, gastric and exocrine pancreatic glands.	
Cephalic phase	
	Unconditioned reflexes secrete saliva, gastric and pancreatic juice
	Conditioned reflexes (the thought of food) also.
Gastric phase (distension of the stomach)	
	Vagal reflexes - cholinergic, muscarinic receptors
	Intrinsic peptidergic neurons (VIP, GRP)
	Histamine
	Somatostatin (multipotent inhibitor)
	Gastrin
Intestinal phase	
	Gastrin from duodenal G-cells increases gastric secretion
	Secretin (from S-cells) and bulbogastrone inhibit gastrin-stimulated acid secretion
	Cholecystokinin (CCK) and gastric inhibitory peptide (GIP) inhibit gastrin release from G-cells, and acid secretion by the parietal cells
	All these entero-gastric *inhibitory* hormones are called *enterogastrones*

2b. The gastric phase

is brought about when food enters and distends the stomach. Distension stimulates stretch receptors and peptide sensitive chemoreceptors. They provide afferent signals for both long, central vago-vagal reflex loops as well as local, enteric reflexes. Signals in these fibres reach cholinergic, muscarinic receptors on the basolateral membrane of the parietal cells.

Distension of the body of the stomach can release gastrin from the antral mucosa by vagal reflexes. Most of the daily gastric secretion of 1.5 l is accounted for by the gastric phase.

2c. The intestinal phase is elicited by duodenal and jejunal mechanisms that both stimulate and inhibit gastric acid secretion. Gastric secretion and motility are at first increased to promote further digestion and emptying. This fills the duodenum with acidic and fatty chyme. Acid chyme reaching the duodenum with peptides and amino acids releases gastrin from duodenal G-cells, which increases gastric secretion. Normally, the inhibitory intestinal mechanisms dominate, when the pH of the chyme is low. Acid chyme in the duodenum causes release of secretin (from S-cells) and of bulbogastrone (Box 22-1).

3. Mastication and swallowing

The process of chewing or *mastication* requires co-ordination of the chewing muscles, the cheeks, the palate and the tongue. Chewing is normally a reflex action. The forces involved in grinding and cutting the food are enormous, and sufficient to fragment cellulose membranes. Finally, the food is mixed with saliva and formed into a bolus. The bolus is pushed back into the pharynx, when the tongue is pressed against the hard palate.

Fig. 22-2: Swallowing of a food bolus in three steps (OES stands for the upper Oesophageal sphincter).

The gastrointestinal tract moves ingested materials and secretions from the mouth to the anus. These movements, as well as nonpropulsive contractions, are called *motility*.

Gastrointestinal sphincters possess adrenergic a_1-receptors. Stimulation of these receptors results in contraction.

Swallowing (*deglutition*) begins as a voluntary process by which the tongue pushes a portion of the food back against the soft palate (Fig. 22-2). Elevation of the soft palate closes the nasopharynx, and the food enters the pharynx, the larynx is elevated closing the epiglottis and respiration stops. The upper pharyngeal constrictor contracts, initiating sequential contractions of the other pharyngeal constrictors. These contraction waves are involuntary and push the food towards the oesophagus. Peristalsis in the oesophagus is started as the pharyngeal wave passes through the upper oesophageal sphincter (Fig. 22-2). When the propulsive wave reaches the lower oesophageal sphincter (LES), the relaxed muscle wall preceding the bolus momentarily relaxes the LES, and the food passes the cardia to enter the stomach. Vagal stimulation relaxes both sphincters (see achalasia, below).

The upper third of the oesophagus is composed of striated muscle, the middle third contains mixed smooth and striated muscle, and the lower third contains only smooth muscle.

Swallowing is controlled by brainstem neurons. They form a swallowing centre (Fig. 22-2). The vagus nerve contains both somatic motor neurons (originate in the nucleus ambiguus) that form motor endplates on striated muscle fibres, and visceral, preganglionic motor neurons (from the dorsal motor vagal nucleus to the myenteric plexus). The swallowing reflex coordinate motor signals from both oesophageal striated and smooth muscles as well as signals to the upper and lower oesophageal sphincters.

Sympathetic stimulation contracts the LES mediated by noradrenaline acting on a-receptors. When a swallow is initiated via touch receptors in the pharynx, or when the lower oesophagus is distended by a bolus, it will relax the LES by reflexes in inhibitory vagal fibres joining the enteric nervous system. VIP and NO act as transmitters.

4. Gastric and intestinal motility

In the stomach, digestion continues (*salivary amylase*) and the stomach regulates emptying of its content into the duodenum. The fundus has a *high compliance*, so food can accumulate without much increase in gastric pressure. Vagal fibres releasing VIP to inhibitory neurons of the myenteric plexus mediate this receptive relaxation. The body of the stomach mixes and grinds the food with gastric juice - also by *retropulsion* (backward or oral movement) - and then propels the content toward the antrum and pyloric region for regulated emptying. The distal stomach reduces solids to a fluid consistently composed of particles less than 2 mm. Here is a *forceful peristalsis* (ie, propagating contractions), so the pyloric sphincter opens and the chyme is ejected into the duodenum (Fig. 22-3).

Fig. 22-3: Intestinal smooth muscle potentials (left) and contractions (right).

Along the greater curvature of the stomach is a region of rapid spontaneous depolarization, which is called the *gastric pacemaker* establishing the maximum rate of gastric contractions. The gastric smooth muscle wall generates two types of electrical activity. *Slow waves* (basic electrical rhythm) are *slow depolarisation's* occurring at a frequency of three in the stomach, up to 18 in the duodenum and 8 per min in the terminal ileum. The slow waves are oscillations of the resting membrane potential (Fig. 22-3). Voltage-gated (potential sensitive) Ca^{2+}-channels open at a certain threshold of depolarization, causing a Ca^{2+}-influx to the smooth muscle cell resulting in the so-called spikes and contractions. *Spikes* are *periodic fast waves of depolarisation* that always initiate gastric contractions, elicited by the rise in cytosolic [Ca^{2+}]. These contractions last up till 3 s, because the Ca^{2+} -channels open slowly and remain open longer than the Na^{+} -channels. Spikes are elicited by vagal signals, by acetylcholine (muscarinic receptors), by stretch, by myenteric signals and by gastrin (Fig. 22-3). Adrenaline and noradrenaline relax smooth muscle by hyperpolarization through *α-adrenergic receptors.* Relaxation occurs when intracellular Ca^{2+} is returned to the extracellular fluid and to the endoplasmic reticulum.

The small intestine is about 8 m long and commonly divided into three segments: the duodenum, jejunum and ileum. The intestinal contents must be moved in a manner that brings them into contact with the mucosa of the intestine, and propels the contents along this tubular organ. Several pacemaker regions in the small intestine control the slow waves. The

pacemaker rate is highest in the duodenum (about 18 each minute), and decreases down to 8 waves each min in the terminal ileum.

During fasting, a migrating sequence of events called the *migrating motor complex* occurs each 80-90 min. The complex consists of an 80-90 min long quiet period (I) followed by a period of irregular propulsive contractions (II), culminating in a *peristaltic rush* (III) to begin in the stomach, accompanied by increased gastric, pancreatic and biliary secretion. The migrating motor complex is the "*intestinal housekeeper*", which cleanses the digestive tract of non-absorbable substances, and provides an effective emptying of the tract all the way.

During the fed state, *segmentation* serves to mix chyme with enzyme-containing digestive fluid, and brings the mixture into contact with the mucosal surface for absorption. Segmentation divides the small intestinal content into many segments by localised circular smooth muscle contractions with only a small propulsive effect (Fig. 22-3).

Propulsive motility is accomplished by *peristalsis.* Peristalsis is a propagating contraction of successive sections of circular smooth muscle preceded by a dilatation (Fig. 22-3). The dilatated intestinal wall is drawn over its content in this reflex mechanism, which has been called the *law of the gut*. Peristaltic contractions usually travel along a small length of the small intestine, except for the *peristaltic rush* related to the migrating motor complex.

The *ileocoecal sphincter* prevents retrograde flow of colonic matter. The sphincter regulates emptying of ileum five hours after a meal. The emptying of ileum is stimulated by *gastrin*, possibly via the *gastro-ileal reflex*, but a distended colon inhibits the emptying. The gastro-ileal reflex is an increased motility of the terminal ileum caused by elevated gastric activity. On the other hand, distension of the terminal ileum decreases gastric motility. The ileocoecal sphincter is normally passed by *one litre* of faecal matters per day.

5. Vomiting

The feeling of nausea, and an array of sympathetic and parasympathetic responses initiate vomiting or emesis. *Sympathetic responses* include sweating, pallor, increased respiration and heart rate and dilatation of pupils. *Parasympathetic responses* include profuse salivation, pronounced motility of the oesophagus, stomach, and duodenum, relaxation of the oesophageal sphincters. Duodenal contents can be forced into the stomach by anti-peristalsis (Fig. 22-4). During the expulsion of gastric contents, the person *takes a deep breath*, the pylorus is *closed*, the glottis is *closed* so respiration stops, and the stomach is *squeezed* between the diaphragm and the abdominal muscles, causing rapid emptying (Fig. 22-4). Vomiting is co-ordinated by the *vomiting centre* in the medulla.

Fig. 22-4: Vomiting co-ordinated by the vomiting centre.

Vomiting is stimulated in certain areas of the brain (hypothalamus) and the cerebellum through sensory stimuli or injury. Vomiting is also provoked by certain labyrinthine signals, and from the chemoreceptive *trigger zone* located on the floor of the 4th ventricle close to *area postrema*.

During deep anaesthesia the vomiting and swallowing mechanisms are *paralysed*. Any patient must abstain from food and water for at least six

hours before deep anaesthesia is administered. Otherwise, the patient may vomit into the pharynx, and suck his own vomit into the trachea. Over the years, many patients have choked to death due to this mechanism. The survivors develop *aspiration pneumonia*. Such events are clearly malpractice.

The swallowing mechanism is also cut-off by injury of the 5th, 9th, or 10th cranial nerve, by *poliomyelitis*, by *myasthenia gravis* and by *botulism* (Chapter 33).

An acute loss of H^+ from the extracellular fluid (ECF) by vomiting creates a *metabolic alkalosis* (high pH with high Base Excess, see Chapter 17).

6. Colonic motility and defaecation

Colonic transit is measured in *days*. Mixing occurs in the ascending colon, because peristalsis is followed by anti-peristalsis. *Slow waves* of contraction move the content in the oral direction to delay propulsion and increase absorption of water and electrolytes. Colonic segmentation is a mixing of the content by regular segments called *haustrae*. Prominent haustration along the length of the colon is characteristic for the X-ray image of the normal colon. The colon provides an optimal environment for bacterial growth. *Peristaltic rushes* in the colon occur several times per day. They often start in the transverse colon as a *tight ring*, continuing as a *long contraction wave*. Gastro-colic and duodeno-colic reflexes assisted by gastrin and by cholecystokinin (CCK) promote peristaltic rushes**.**

Defaecation is a complex act involving both voluntary and reflex actions in colon, rectum, anal sphincters and many striated muscles (diaphragm, abdominal and pelvic muscles). Defaecation is a temporal release of anal continence brought about by a reflex. The rectum is usually empty, and its wall has a rich sensory supply. Distension of the recto-sigmoid region with faecal matter releases awareness of the urge to defaecate, an *intrinsic defaecation reflex*, and a *strong, spinal reflex*. There is a reflex contraction of the descending colon and the recto-sigmoideum.

The smooth internal *anal sphincter muscle* maintains a tonic contraction during continence, due to its sympathetic fibres from the *lumbar medulla* (through hypogastric nerves and the inferior mesenteric ganglion). The muscle relaxes due to its parasympathetic, cholinergic fibres in the pelvic splancnic nerves (S_2-S_4). The strong spinal reflex produces relaxation of the smooth muscles of the internal anal sphincter (Fig. 22-5) and contraction of the striated muscles of the external anal sphincter (innervated by somatic fibres in the pudendal nerve) inhibiting the reflex and causing *receptive relaxation*. This is the last decision - before defaecation.

Fig. 22-5: Defaecation reflexes.

The *levator ani muscle* contributes to the closure of anus, because contractions increase the angle between the rectum and the anus.

Destruction of the lower sacral medulla (the *defaecation centre*) destroys the spinal reflex and thus the normal defecation. Higher spinal lesions destroy the voluntary control, whereas the defaecation reflexes persist. An acceptable status is obtainable in paraplegics by mechanical release of the reflex (manual expansion of the external sphincter) once daily following a meal.

7. Gastrointestinal hormones

Gastrointestinal hormones are peptides secreted by the gastrointestinal mucosa, and controlling all gastrointestinal functions together with other hormones and transmitters. As an example insulin works together with acetylcholine and parasympathomimetics to *stimulate* secretion and motility, whereas catecholamines, sympatomimetics and parasympatolytics, such as atropine, *inhibit* gastrointestinal secretion and motility.

Peptide hormone families are groups of regulatory peptides that exhibit *sequence homology* (ie, they possess a common amino acid sequence). The *gastrin-family* and the *secretin-glucagon family* are the most important.

7a. The gastrin family

consists of gastrin and cholecystokinin (CCK) in three different forms (CCK-8, CCK-22, and CCK-33). Gastrin and CCK release *pancreatic glucagon* from the islet cells. There are two major forms of gastrin in the plasma, *normal gastrin* or *G-17* and *big gastrin* or *G-34*. They are 17 and 34 amino acid polypeptides, respectively. Gastrin is produced by G-cells of the gastric antrum and duodenum. The duodenal Brunner glands secrete half of the G-34.

Gastrin is the strongest stimulator of gastric acid secretion. Gastrin also imposes *tropic* (growth-stimulating) actions on the parietal cells, the mucosa of the small and large intestine and possibly the pancreas. Gastrin stimulates the *pepsin secretion* from peptic cells, and the *glucagon secretion* from the a-cells of the pancreatic islets.

Gastrin is derived from *parietal* or *oxyntic* cells in the stomach. When stimulating gastric acidity, gastrin relaxes the gastric muscles, thus retarding the passage of chyme into the duodenum.

Feeding induces the secretion of gastrin to the interstitial fluid and then to the blood. Neural signals pass through the vagal nerve to the gastrin-secreting *G-cells* of the gastric antrum and duodenum (Fig. 22-6). The afferent input begins with the smell and taste of food, and is reinforced by vago-vagal reflexes elicited by oesophageal and gastric distension. *Digested protein* (polypeptides and amino acids) act directly on G-cells.

Fig. 22-6: Gastric HCl secretion following feeding. GRP: Gastrin Releasing Peptide. NANC: Non-adrenergic, Non-cholinergic postganglionic neurons.

Vagal, cholinergic preganglionic fibres transfer signals to the G-cells via *non-adrenergic, non-cholinergic* (NANC) postganglionic neurons. These enteric neurons liberate *gastrin-releasing peptide* (GRP) to the gastrin producing G-cells. The gastrin released reaches the parietal cells through the blood and increases the HCl secretion. GRP thus releases gastrin and hereby stimulates the secretion of gastric acid. - GRP consists of 27 amino acid moieties and is also released from neurons in the brain.

An indirect vagal route to the G-cells is via *postganglionic cholinergic enteric neurons* to *somatostatin cells* that are located close to the G-cells (Fig. 22-6). When these enteric neurons release acetylcholine, the response of the somatostatin cells is inhibition of somatostatin release. Somatostatin inhibits G-cell secretion by paracrine action. The result of both vagal inputs to the G-cells is *gastrin release* (Fig. 22-6). An elevated [H^+] in the

duodenal lumen inhibits gastrin release.

Cholecystokinin, CCK, according to its function and structure, belongs to the *gastrin family*. Cholecystokinin empties the gall bladder as the name implies, and stimulates pancreatic secretion of an enzyme rich juice. However, CCK has a higher affinity for receptors stimulating gallbladder contraction and pancreatic enzyme secretion. CCK has a maximal effect only in the presence of secretin (potentiation) and normal vagal influence.

Both gastrin and CCK release glucagon from the a-cells of the pancreatic islets.

CCK is cleaved from *pre-pro*-CCK in the duodenum, upper jejunum (I-cells) and in the *brain*. CCK molecules consist of a group of peptides. CCK-8, CCK-22 and CCK 33 are the dominant forms in the blood.

The most important stimulus for CCK liberation is amino acids and fatty acids, which reach the duodenal mucosa. Bile is ejected into the duodenum, where fat is emulgated to ease its absorption. CCK also acts as an *enterogastrone* - an intestinal hormone that *inhibits* gastric activity and emptying. This leaves more time for the bile to emulgate fat.

7 b. The secretin-glucagon family

Secretin exhibits sequence homology with pancreatic glucagon, vasoactive intestinal peptide (VIP), growth hormone-releasing hormone (GHRH) and gastric inhibitory polypeptide (GIP). A family of five genes code for these five hormones.

Secretin is secreted by S-cells in the mucosa of the upper small intestine, when acid chyme (pH below 4.5) arrives to the first part of the duodenum. Fatty acids from fat digestion also contribute to secretin release.

Secretin stimulates the secretion of bicarbonate and water by pancreatic duct cells, and of bicarbonate-rich aqueous bile. Secretin potentiates the action of CCK including an *enterogastrone effect* (gastric inhibiting effect). Secretin antagonises gastrin - and potentiates CCK. Secretin is an enterogastrone that is released by H^+ to stimulate pancreatic juice secretion.

Gastric inhibitory polypeptide (GIP or Glucose-dependent Insulin releasing peptide) works as the two names imply: GIP inhibits the gastric mucosa and releases insulin from the a-cells of the pancreatic islets.

Glucagon is actually two different molecules: *Intestinal* glucagon (*glicentin*) and *pancreatic* glucagon. Both are hepatic insulin-antagonists. Glucagon stimulate glycogenolysis, gluconeogenesis (urea genesis-glycogenic amino acids), and ketogenesis.

The function of other peptide hormones is given in Box 22-2.

Box 22-2: Effects of some gastrointestinal hormones and transmitters.

Duokrinin stimulates duodenal secretion.

Endogenous (*enkephalins*) and exogenous opiates inhibit ganglionary transmission.

Enterokrinin stimulates secretion in the small intestine.

Gastrin releasing peptide (GRP) and **bombesin** release gastrin from G-cells.

Glicentin (intestinal glucagon) stimulates insulin secretion as other incretins.

Motilin stimulates gastrointestinal motility.

Neuropeptide Y and neurotensin stimulate neurotransmission.

Nitric oxide (NO) is a possible neurotransmitter between the preganglionic and the NANC postganglionic neurons.

Pancreatic Polypeptide (PP) from the PP-cells inhibits pancreatic and biliary secretion, which delay the absorption of nutrients. PP is released by meals.

Pancreotonin: Inhibits the pancreatic exocrine secretion.

Somatostatin (Growth hormone-inhibiting hormone, GHIH; 14 amino acid moieties) is a strong, universal inhibitor - both blood-born and paracrine.

Substance P (11 amino acid residues) stimulates smooth muscle contraction and thus the gastrointestinal motility.

Vasoactive intestinal peptide (VIP; 28 amino acid residues; vessel wall and brain neurons) is a vasodilatator in line with adenosine, ATP, and NO. The increased bloodflow increases intestinal secretion. VIP is also involved in penile erection and in bronchiolar dilatation.

Villikrinin: Stimulates the rhythmic movement of villi in the intestine.

Traditionally, the important peptides are also divided into two *functional* groups: *Enterogastrones inhibit* gastric motility and secretion. When gastric acid, fats, and hyperosmolar solutions have entered and distended the duodenum, GIP and other enterogastrones (somatostatin, CCK, and secretin) are released and suppress gastric acid secretion and motility of the stomach

Incretins stimulate insulin secretion. Incretins are liberated to the blood as gastric chyme enters the duodenum - and before the glucose of the chyme can be absorbed. Incretins increase *insulin secretion* from the b-cells of the pancreatic islets much earlier and to a greater extent, than when the blood [glucose] is elevated by intravenous infusion. Incretins are GIP, glicentin, and glucagon-like peptides: GLP-1 and -2.

8. Saliva

Saliva is a watery solution of electrolytes (bicarbonate and K^+) and organic substances, which is a mixture of secretions from three pairs of glands. The

parotid is the largest and serous (watery saliva), the *sublingual* is mucous (viscous, containing mucin), and the *submandibular* salivary gland is build of mucous acini surrounded by serous *half moons*. The *primary* saliva is produced in the acini, but *secondary* processes in the salivary ducts (secretion and reabsorption) are involved in the final saliva production. Salivary glands have a high bloodflow and produce up to one l of saliva daily. The maximal secretion rate is one ml of saliva per g salivary tissue per min (ie, 60 times that of pancreas).

Salivary *mucin* (a glycoprotein) and *water* lubricate food, dissolve particles, and salivary enzymes initiate digestion. Ptyalin or a-amylase cleaves a-1-4 glycoside bindings in starch. *Salivary buffers* maintain the pH-optimum (6.8) of amylase during the first period in the stomach. The saliva dilutes injurious agents.

Saliva cleans the mouth and pharynx (prevents *caries*), and ease swallowing. Salivary *lysozyme* lyses bacterial cell walls. The salivary *epidermal growth factor* promotes the healing of wounds. Animals instinctively lick their wounds. Saliva contains immuno-defensive secretory globulin A (IgA), amino acids, urea, and blood-type antigens in secreting persons. Saliva may inactivate human immunoactive virus (HIV). The most common infection of the salivary glands is *acute parotitis* caused by the mumps virus.

The virus causing infectious mononucleosis is probably transferred with saliva by "deep kissing". *Infectious mononucleosis* is a disease characterised by lympadenopathy, lympho-cytosis and duration longer than an ordinary tonsillitis. The condition is dangerous, because spontaneous rupture of the spleen occurs.

Salivary secretion is controlled by the autonomic nervous system, and minimally influenced by hormones. Unconditioned reflexes (taste-, olfactory- and mechano-receptors) control salivation as well as *conditioned reflexes* (the thought of food). These signals reach the *brain stem salivary centres*, which activate the parasympathetic nerves to the salivary glands. The *primary* salivary secretion into the acini resembles an ultrafiltrate of plasma, but the final saliva is hypotonic.

Parasympathetic, cholinergic fibres, originating in the *salivary nuclei* of the brain stem, synapse with postganglionic neurons close to the secretory cells. These neurons transmit signals to the *cholinergic, muscarinic receptors* (Fig. 22-7). Parasympathetic activity can release maximal salivary secretion and bloodflow resulting in a amylase-rich saliva with mucin (glycoproteins). *Atropine* blocks the muscarinic, cholinergic receptors (during anaesthesia where the mouth becomes dry). The rise in bloodflow is atropine-resistant and caused by the vasodilatating VIP, which is released from peptidergic nerve terminals that also contain acetylcholine. *β_1-adrenergic agonists* and VIP elevate cAMP in the acinar cells, an effect potentiating the secretory effect of acetylcholine. The vascular smooth muscle relaxation by VIP is probably also mediated via cAMP.

Fig. 22-7: Salivary enzymes, ions and mucin production from two acinar cells. Solid and dashed arrows indicate active and passive transport, respectively. Circles are carrier molecules, whereas tubes symbolise transport channels. - To the left is shown receptors and second messengers.

1. Neural or humoral (acetylcholine) stimulation of cholinergic, muscarinic receptors on the basolateral membrane of acinar cells leads to a rise in intracellular [Ca^{2+}].

2. This rise triggers luminal Cl^- - and basolateral K^+ -channels. Hereby, K^+ is transferred to ISF and Cl^- to the acinar lumen in a balanced relationship (Fig. 22-7). Therefore, Cl^- flows down its electrochemical potential gradient into the lumen of the acinus. K^+ flows down its gradient to the ISF through activated channels. These ion flows create a negative electric field in the lumen.

3. The initial fall in intracellular [K^+] increases the driving force of the electroneutral Na^+-K^+-$2Cl^-$ co-transporter to transport two Cl^- into the cell together with Na^+ and K^+. Thus the electrochemical potential of Cl^- and K^+ is greater in the cell, than in the interstitial fluid (ISF) and in the saliva.

4. The negative field provides an electric force that drives a passive Na^+ flux into the acinar lumen through leaky tight junctions. Osmotic water transport through leaky junctions and trans-cellularly through water channels in the cell membranes follow the NaCl flux into the lumen. The trans-cellular Cl^- transport is coupled to the paracellular Na^+ transport. The net result is an isosmotic NaCl transport produced by a secondary active Cl^-- secretion.

5. The basolateral membranes of acinar cells contain a Na^+-K^+-pump that provides the energy for the primary salivary secretion (Fig. 22-7). The rise in intracellular [Na^+] from 2., activates the Na^+-K^+-pump, whereby [Na^+] is kept almost constant. Ouabain inhibits salivary secretion, because it blocks the pump.

Sympathetic nerve signals, and circulating catecholamines via b-adrenergic receptors, inhibit the bloodflow and the secretion of serous saliva (b_1-receptors in Fig. 22-7). A small, transient, mucous secretion with a high [K^+] and [bicarbonate], and a low [Na^+] is produced, because of the low secretion rate. Noradrenaline (NA) stimulates both a_1-adrenergic and b_1-adrenergic receptors. Binding of NA or b-adrenergic agonists elevates intracellular cAMP, which correlates with a small increase in primary salivary secretion. This explains why the mouth becomes dry during events, where the sympathetic system dominates (anxiety, excitement etc).

The salivary ducts are almost watertight. Therefore, the final salivary flow is dependent upon the primary salivary secretion rate in the acini.

The duct systems, in particular the small-striated ducts with a substantial O_2 consumption reabsorb large amounts of Na^+ and Cl^-, whereas bicarbonate and K^+ are secreted. Saliva becomes more and more hypotonic at low secretion rates, because the Na^+ and Cl^- reabsorption dominate.

1. The reabsorption of Na^+ and the secretion of K^+ are processes stimulated by the mineralo-corticoid, aldosterone. Aldosterone stimulates Na^+-influx through the luminal Na^+-H^+-exchanger (Fig.

22-8). Na^+ enters the cell in exchange with H^+. The resulting intracellular rise in [Na^+] activates the basolateral Na^+-K^+-pump. Thus, Na^+ is reabsorbed trans-cellularly from the salivary duct. The pump maintains the electrochemical potential gradients of Na^+ and K^+.

2. The Cl^- follows passively, and is partly exchanged with bicarbonate along the duct system through a luminal Cl^- - bicarbonate exchanger (Fig. 22-8). The secretion of bicarbonate is so great that its concentration in the final saliva exceeds that in plasma.

3. At the basolateral membrane Cl^- leaves the cell via an electrogenic Cl^- channel, while Na^+ is pumped out.

4. K^+, taken up by the Na^+-K^+-pump, leaves the cell through K^+-channels in the basolateral membrane, recycling K^+ to balance the Cl^- efflux.

5. Some of the K^+ leaves the cell by luminal H^+-K^+-exchange. At low secretion rates the H^+-K^+-exchanger (antiport) in the luminal membrane transfers sufficient K^+ for the [K^+] in the final saliva to exceed the concentration in plasma. The net result is K^+-secretion from blood to the duct lumen.

The final salivary [Na^+] and [Cl^-] increase with increasing salivary secretion rate, because the high flow provides less time for reabsorption in the duct system. Bicarbonate may be secreted even without Cl^--reabsorption. At low salivary secretion rates the final saliva becomes hypotonic down toward half of the osmolarity of plasma.

Fig. 22-8: Secretion from salivary duct cells.

The aldosterone effects described above (increased Na^+ reabsorption and increased K^+ secretion) are similar to those in the distal, renal tubules and in the sweat glands.

9.Gastric secretion

The stomach is divided into three main regions: the fundus, corpus and pyloric antrum. The gastric mucosa is highly invaginated and is mainly composed of gastric glands, with mucous neck cells, parietal cells secreting HCl, and peptic (chief) cells secreting pepsinogen. The parietal cells also secrete the peptide intrinsic factor, which is necessary for absorption of vitamin-B_{12}. G-cells in the mucosa produce the hormone gastrin (Fig. 22-6). The gastric secretions include hydrochloric acid (HCl), pepsin and basic mucus, which contains mucin (glycoproteins) and salts.

Efferent signals from the dorsal motor nuclei of the vagi stimulate gastric motility and HCl production. Acetylcholine is released from the short postganglionic vagal fibres and directly stimulates parietal cells to secrete HCl. The parietal cells contain muscarinic receptors on the basolateral membrane. Vagal fibres work together with intrinsic, peptidergic neurons containing vasoactive intestinal peptide (VIP) and gastrin releasing peptide (GRP). VIP controls the bloodflow of the gastric mucosa; GRP releases the

important gastrin from the antral G cells and the peptic cells secrete pepsinogen.

The secretion related to a meal occurs in three phases (cephalic, gastric and intestinal).

The gastric juice is hyperosmotic (325 mOsmol/l), contains 10 mM of K^+ and is low in Na^+ at moderate and high secretion rates; the $[H^+]$ is 170 mM and the $[Cl^-]$ is 180 mM. Gastric juice has an approximate pH of 1, forming a million-fold gradient of H^+ across the gastric mucosa to the blood. The HCl activates pepsinogen, maintains the optimal pH for pepsin activity and denatures proteins and microbes.

The peptic cells, located in the base of the gastric gland, produce pepsinogen. Pepsinogen is stored in granules of the peptic cell. Pepsinogen secretion is stimulated by cholinergic, muscarinic substances and by b-adrenergic agents, but peptic cells have no histamine receptors. Exocytosis releases pepsinogen into the gastric juice, where it is cleaved into pepsin, if HCl is present. Pepsin is the major hydrolytic enzyme in the stomach, but it is only active in the acidic gastric juice.

Fig. 22-9: Secretion of parietal and non- parietal cell juice.

Adult humans produce up to two l of gastric juice daily. The gastric juice is produced from two different sources: The *parietal cell juice* with 170 mM [HCl], 10 mM $[K^+]$, and a low $[Na^+]$. A juice with an ionic composition similar to that of plasma is produced from other cells - the *non parietal juice*. Each of the two secretion products has almost a constant composition.

Increased secretion of gastric juice means increased secretion of parietal cell juice. This explains why the [HCl] increases more and more in the mixed product, whereas $[Na^+]$ falls with increasing secretion rate.

Fatty chyme entering the duodenum delays gastric emptying by negative feedback through duodenal reflexes and by the release of gut inhibiting hormones (so-called enterogastrones: somatostatin, VIP, gastric inhibitory peptide, GIP, neurotensin and secretin). These inhibitors not only inhibit gastric motility; they also inhibit the gastrin release from the antral G cells, and also the HCl production from the parietal cells. Mucus contains mucin (glycoproteins) and electrolytes with bicarbonate that protect the gastric mucosa from adversive effects.

Stimulation of the parietal cells with acetylcholine, histamine and gastrin has two consequences for their content of second messengers (Fig. 22-10, right). The cellular $[Ca^{2+}]$ and [cAMP] is elevated.

Fig. 22-10: HCl secretion from parietal cell in the stomach (left). Secretory receptors on the parietal cell are also shown (right).

1. These second messengers activate luminal Cl^-- and K^+-channels. Cl^- and K^+ pass into the lumen, whereby their cellular concentrations decrease (Fig. 22-10 left). The luminal $[K^+]$ activates the K^+-H^+-pump. In addition, more pumps are inserted into the luminal membrane from cellular tubulo-vesicles.

2. The fall in cellular $[Cl^-]$, and a rise -see below - in cellular

[bicarbonate], stimulates the basolateral Cl -bicarbonate exchanger, whereby the cellular [bicarbonate] is reduced. The fall in cellular [H^+] and [bicarbonate] stimulates formation of H^+ and bicarbonate, under the influence of carbo-anhydrase (*). The H^+ and bicarbonate are derived from metabolic carbon dioxide from the blood. Bicarbonate diffuses from the interstitial fluid space (ISF) into the blood. Every time the gastric juice receives one H^+, the blood will receive one HCO_3^-. This explains why the pH of the gastric venous blood increases after a meal - the alkaline tide.

3. Cellular [H^+] is a substrate for the luminal gastric proton pump (the K^+-H^+-pump), already activated by K^+. The net result is H^+-secretion to the lumen in a balanced relationship to Cl^--secretion. The surface of the gastric mucosa is always electrically negative with respect to the serosa. H^+ moves against a large concentration gradient into the gastric lumen. The intracellular [H^+] of the parietal cells is 10^{-7} mol/l, so with a [H^+] of 10^{-1} mol/l in the gastric juice, a million-fold concentration gradient is present across the luminal membrane. Accordingly, energy is required for the transport of both ions. The HCl secretion requires ATP.

4. The cellular concentration of cations is maintained by the basolateral Na^+-K^+-pump.

   The parietal cells contain more mitochondrial mass per volume unit than any other cells in the body, indicating a rich oxidative metabolism.

Histamine, acetylcholine and gastrin stimulate acid secretion. We have two types of histamine receptors in the human body: H_1 receptors (blocked by diphenhydramine) and H_2 receptors. Only H_2 receptors are located on the parietal cells.

1. The H_2 receptors make histamine a potent stimulant of HCl secretion. When histamine is bound to the H_2 receptor it activates adenylcyclase, an enzyme generating cAMP from ATP. This increase in intracellular [cAMP] is specific for histamine. The cAMP binds to and activates cAMP-dependent protein kinase (consisting of a regulatory and an active catalytic subunit). The cAMP binding releases the active catalytic subunit, which phosphorylate a variety of target proteins.

   H_2 receptor antagonists (cimetidine and ranitidine) prevent histamine from binding to the H_2 receptors of the basolateral membrane of the parietal cells, which reduces acid secretion. Synthetic analogues of prostaglandin E can inhibit both the cAMP and the Ca^{2+} release mechanisms, thus promoting ulcer healing (see later).

2. Acetylcholine (ACh) is released by vagal stimulation that leads to a stimulation of acid secretion. This secretion is inhibited by atropine. Thus the parietal cells contain muscarinic, cholinergic receptors (M_3).

3. Gastrin is the most potent stimulant of acid secretion in humans. Gastrin receptors were previously supposed not to be present on human parietal cells. Gastrin from G-cells was thought to release histamine from the granules of the mast cells in the gastric glands (Fig. 22-10). This is probably not the case. A direct gastrin effect on human gastrin receptors occurs, and an additional indirect effect via histamine increases the HCl secretion markedly (H_2 receptors). However, the three-receptor hypothesis is still under debate.

   Gastrin and acetylcholine release inositol-triphosphate (IP_3), which is produced with diacylglycerol (DAG) by a membrane phospholipase. The target system for IP_3 is a Ca^{2+}-channel protein located in the endoplasmic reticulum. Ca^{2+} is released from the reticulum, and Ca^{2+} also enters the cell through the basolateral membrane.

   Combined stimulation of all three receptors results in maximal gastric secretion (potentiation).

10. Intestinal digestion and absorption

Almost all of the dietary nutrients, water and electrolytes that enter the upper small intestine are absorbed. The small intestine, with its epithelial folds, villi, and microvilli, has an internal surface area of 200 m^2.

10a. Carbohydrates

Carbohydrates are the most important energy-containing components of the diet. The energetic value of most carbohydrates is 17.5 kJ per g, so that a daily diet of 400 g carbohydrates covers 7 000 kJ, which is 56% of the usable energy in a diet of 12 500 kJ daily. The formation of metabolic water on a mixed diet is 0.032 g of water per J.

Fig. 22-11: Absorption of carbohydrates by the enterocyte.

The common sources of digestible carbohydrates are starches (amylose), table sugar, fruits and milk. Plant and animal starch (amylopectin and glycogen) are branched molecules of glucose monomers. Indigestible carbohydrates are present in vegetables, fruits and grains (cellulose, hemicellulose, pectin) and in legumes (raffinose). Indigestible carbohydrates are also referred to as dietary fibres.

Digestion of starches to simple hexoses occurs in two phases: The luminal phase begins in the mouth with the action of salivary amylase (ptyalin), but most of this phase occurs in the upper small intestine as pancreatic a-amylase reach the chyme. The starch polymer is reduced to maltose, maltriose and a-limit dextran or dextrins (Fig. 22-11). The three substrates are pushed through the intestine and are now ready for the brush-border phase. Some of the substrate molecules get into contact with the brush-borders of the absorbing mucosal cell via the unstirred water layer. Enterocytes carry disaccharidases and trisaccharidases (oligosaccharidases) on their surface that cleave these substrates to glucose, G.

Milk sugar (lactose) and cane sugar (sucrose) only require a brush-border phase of digestion, since they are disaccharides. Sucrose is reduced to glucose and fructose (G-F), and lactose to glucose and galactose (G-Ga) by

the action of disaccharidases (sucrase and lactase).

Glucose in the intestinal lumen is absorbed by active transport.

1. The mechanism of active glucose transport is a carrier-mediated, Na^+ - glucose cotransport. As the luminal [glucose] falls below the fasting blood [glucose], active glucose transport becomes essential and sequesters all remaining luminal glucose into the blood. Glucose and Na^+ bind to apical membrane transport proteins (a glucose-transporter, GLUT 5). The two substances are deposited in the cytoplasm, because of conformational changes in GLUT 5, whereby the affinity of GLUT 5 for glucose-Na^+ changes from high to low. Glucose accumulates inside the cell to a level that exceeds blood [glucose].
2. Glucose therefore diffuses down its concentration gradient, through a specific uniport carrier in the basolateral membrane, out into the interstitial space and into the blood (Fig. 22-11). The basolateral uniport carrier for glucose is highly specific (glucose only), and does not depend upon Na^+. Galactose is also actively transported by the luminal glucose carrier system, and is a competitive inhibitor of glucose transport. Phlorrhizin blocks the glucose absorption, when its glucose moiety binds to the transporter instead of glucose.
3. Cytoplasmic Na^+ is actively pumped out through the basolateral membrane by the Na^+-K^+-pump. The low intracellular [Na^+] creates the Na^+ gradient and energises the transport of hexoses over the luminal enterocyte membrane.
4. Fructose has no effect on the absorption of glucose and galactose. Fructose is not actively transported by the Enterocytes, but is absorbed by a carrier-mediated, facilitated diffusion system, where energy is not required.

10 b. Proteins

The typical Western diet contains *100 g* of protein, which is equivalent to an energy input of 1700 kJ daily, although an adult needs only less than one g pr kg of body weight. This luxury combustion is an inappropriate use of global resources. Moreover, a high protein intake implies a long-term risk of uric acid accumulation from purine degradation (Chapter 20). Meats, fish, eggs, and diary products are high in proteins and expensive. Vegetable proteins are not as expensive as animal proteins.

Residents of areas with carbohydrate dominated nutrition and protein hunger develops diseases of protein deficiency, such as *Kwashiorkor* (Chapter 20).

Digestion of dietary proteins begins in the stomach, with the action of the gastric enzyme pepsin (pH optimum is 1), which cleaves proteins to proteoses, peptones and polypeptides. Pepsin is produced from pepsinogen in the presence of HCl. Pepsinogen is secreted by the gastric chief cells. The digestion is continued in the intestine by proteolytic enzymes of the pancreas. Enteropeptidase converts trypsinogen to trypsin. Trypsin acts auto-catalytically to activate trypsinogen, and also convert chymo-trypsinogen, pro-carboxy-peptidases A/B, and pro-elastase to their active

form. When the chyme is pushed into the duodenum, the pancreatic juice neutralises the chyme and the activity of pepsin is stopped. The proteolysis in the small intestine plays the major role, because the digestion and absorption of dietary protein is not impaired by total absence of pepsin.

Cytosolic peptidases from the enterocytes and brush border peptidases from the brush borders of the villous cells then cleave the small peptides into single amino acids (Enteropeptidase, amino-polypeptidase and di-peptidases). The end products of protein digestion by pancreatic proteases and brush border peptidases are di- and tri-peptides and amino acids. The cytosolic peptidases are abundant and particularly active against di- and tri-peptides.

Hydrolytic digestive products such as tripeptides, dipeptides and amino acids can be absorbed intact across the intestinal mucosa and into the blood. Two transport routes are dominant:

1. A peptide transporter, with high affinity for di- and tri-peptides, is absorbing the small peptides (Fig. 22-12). The system is stereospecific and prefers peptides of physiologic L-amino acids. This peptide transport across the brush border membrane is a secondary active process powered by the electrochemical potential difference of Na^+ across the membrane. The total amount of each amino acid that enters the enterocytes in the form of small peptides is considerably greater than the amount that enters as single amino acids.

2. The absorption of single amino acids from the intestinal lumen is an active process that involves a Na^+-dependent, carrier-mediated cotransport system similar to that for glucose. Competitive inhibition, saturation kinetics, Na+ dependency, and expenditure of metabolic energy in this case also characterise active transport.

   Selective carrier systems appear to be present for certain groups of amino acids: neutral, acidic, imino and basic groups. The neutral brush border (NBB) system transports most of the neutral amino acids. The imino acid system handles proline and hydroxyproline.

Fig. 22-12: Absorption of peptides and single amino acids by the enterocyte.

Basic amino acids and phenylalanine are absorbed primarily through facilitated diffusion from the gut lumen to the blood.

The basolateral membrane is more permeable to amino acids than is the brush border membrane. Therefore diffusion is more important for the basolateral transport, especially for amino acids with hydrophobic side chains.

The amino acids are carried in the blood to the liver via the portal vein.

Half of the amino acids absorbed in the intestine are from the diet, the remaining part is from digestive secretions and from desquamated mucosal cells.

Only 1 % of the dietary protein is excreted in the faeces, the remaining faecal protein is derived from micro-organisms and desquamated cells.

The reabsorption of amino acids (and glucose) in the renal tubules bares

many similarities to the active absorption mechanism in the intestine.

A rare genetic disease involves defective intestinal absorption of neutral amino acids and a similar defective renal reabsorption. This condition is called Hartnups disease, which is caused by defects in the NBB transport system of the brush border coated epithelial cells of the jejunum and the proximal renal tubules.

10 c. Lipids

The typical Western diet contains *100 g* of lipids (3900 kJ) daily. Most of the dietary lipids consumed are triglycerides (only 2-4% is made up of phospholipids, cholesterol, cholesterol esters etc). Lipids would comprise just above 30% (ie, 100 g = 3900 kJ) of a standard diet of 12 500 kJ daily. An optimal diet should contain only 20% lipids, such as the lipids of fish oil and olive oil.

Absorption of excess lipids results in accumulation (obesity). The consequences of long term obesity are described in relation to diabetes mellitus in Chapter 27.

Essential dietary fatty acids are poly-unsaturated and cannot be synthesized in the body (linoleic acid, linolenic acid and arachidonic acid).

Dietary triglycerides are broken down into simpler molecules, to facilite absorption. A small fraction of the triglycerides is digested in the mouth and stomach by salivary, lingual lipase.

Most dietary triglycerides (TG) are digested in the small intestine. However, two problems must be solved before digestion can occur. Triglycerides are insoluble in water, and the chyme in the intestine is an emulsion of large fat particles in water. All the lipase proteins by contrast are water-soluble. It follows that, triglycerides must be dissolved in the aqueous phase before they can be digested.

The lipolytic activity requires the emulsifying action of bile salts in order to dissolve triglycerides in water. Pancreatic lipase binds to the surface of the small emulsion particles.

1. Simple bile micelles are aggregates of bile salt monomers that form spherical structures with a diameter of 5 nm, and the micelles have a negative charge. Following a meal, bile micelles are formed above a certain concentration of bile salts, called the critical micellar concentration. The lipophilic, hydrophobic, apolar end of the bile acids faces inward creating a hydrophobic core (Fig. 22-13). The hydrophilic polar end of the bile salts (hydroxyl-, carboxyl- and amino- groups) points outward, so that they are mixed with the polar water molecules. The simple lipids must pass a diffusion barrier - an unstirred water layer, which is the water layer immediately adjacent to the mucosa, where the intestinal flow rate is essentially zero. This water layer contains the water-soluble lipases and cholesterol esterases.

Fig. 22-13: Absorption of lipids by the enterocyte (2-MG is 2-monoglyceride).

2. Mixed micelles. Simple lipid molecules (cholesterol, phospholipids, fatty acids, 2-monoglycerides or 2-MG, fat-soluble vitamins and lyso-lecithin) diffuse into the lipophilic core of the simple bile

micelles and form a mixed micelle (Fig. 22-13). A solution of micelles is water-clear and stable.

The mixed micelles carry the major part of all the lipids that are absorbed by the intestinal microvilli. When the lipids of the mixed micelle have diffused into the enterocyte, emulsifying more hydrolysed lipids recycles the empty bile micelle. Neither bile salt micelles nor bile salt molecules diffuse into the enterocyte (Fig. 22-13).

The fatty acids with a short chain (up to 12 C-atoms) are more hydrophilic than the rest. They can diffuse directly to the portal blood as fatty acids. Once fatty acids enter the enterocyte, they are primarily activated to acetyl coenzyme A by a process that requires ATP and acetyl coenzyme A synthetase. Acetyl coenzyme A enters one of two pathways: the 2-MG and the a-glycerol phosphate pathways. Both bring about the resynthesis of triglycerides (TG) in the enterocyte.

In the enterocyte the lipids are reformed to triglycerides, cholesterol, phospholipids etc. The reformed triglycerides, cholesterol, phospholipids, fatty acids, esters and fat-soluble vitamins reach the endoplasmic reticulum, where they are packed in another lipid-carrying particle: the chylomicron.

The centre of the chylomicron is a cholesterol ester (E in Fig. 22-13). Chylomicrons are packed into vesicles in the Golgi-system. These vesicles reach the basolateral membrane, and their contents pass through this membrane by exocytosis. Thus the chylomicrons reach the lymphatic channel of the villus (the central lacteal). The lymph delivers the chylomicrons to the blood through the thoracic duct. Plasma is milky (lipaemic) following a fatty meal.

All of the dietary lipid is normally absorbed in the intestine. Faecal fat derives from bacterial lipids and lipids of desquamated mucosal cells. - Disorders such as gallstones, pancreatitis, Crohn's disease, and liver disease can lead to fat malabsorption (steatorrhoea or fat-diarrhoea).

Lipids are mainly absorbed through the enterocyte and transported by the lymph, which reaches the blood via the thoracic duct. Lipids thus reach the liver through the hepatic artery, with the exception of short-chain fatty acids that enter the portal blood directly. Other nutrients are absorbed directly to the blood and reach the liver through the portal vein.

Fat-soluble vitamins, such as vitamin A, D and K, are absorbed in the chylomicrons along with lipid nutrients (Fig. 22-13). In contrast, the water-soluble vitamins, such as vitamin- B and -C, cross the mucosa by diffusion and by association to specific membrane transporter proteins. Vitamin B_{12} (cyanocobalamine) is the largest of the vitamins, and its absorption in the terminal ileum utilises a specific transport mucoprotein called intrinsic factor.

10.d Fluids and electrolytes

The intestinal content is isosmolar with plasma, and the water is absorbed from the lumen to the blood by passive osmosis. The membranes of the intestinal mucosal cells and even the tight junctions are highly permeable to water. Hereby, active transport of Na^+ and Cl^- from the lumen to the small interstitial space builds up a forceful osmotic gradient, drawing water the same way by a passive process. In the small interstitial space water creates a

hydrostatic overpressure. Since the capillary and lymph endothelial membranes are no barriers for Na^+, Cl^- and water, a bulk flow of fluid from the interstitial space passes into the blood- and lymph vessels. The intestinal mucosa possesses elevations called villi, and pitted areas called crypts. The villous cells have a typical brush border responsible for net absorption of ions and water, whereas the crypt cells contain secretory mechanisms causing net secretion.

The villous cells absorb Na^+ through the luminal brush border membrane by three mechanisms:

1. An inward diffusion gradient through a Na^+-channel,
2. A Na^+-H^+-exchange, and
3. A Na^+ -solute coupled cotransport (the solute being glucose, galactose, bile salts, water-soluble vitamins and amino acids).

Fig. 22-14: Ion transport processes in jejunal enterocyte.

Ad 1.: The $[Na^+]$ is kept low (14 mM) in the cell, whereas $[Na^+]$ is 140 mM in the intestinal lumen. This concentration gradient work together with an electrical gradient, since the cytosol of the cell is -40 mV with the intestinal content as a reference (Fig. 22-14). Thus Na^+ can easily pass the luminal brush border membrane passively. The intestinal mucosa has ion permeable tight junctions - it is leaky. This paracellular transport is so great that the net absorption of Na^+ and Cl^- through the cells only amounts to 10% of the total transport through the mucosa.

Ad 2.: The transport of Na^+ into the enterocyte (Fig. 22-14) is through a co-exchange protein (Na^+/H^+). Part of the energy released by Na^+ moving down its gradient is used to extrude H^+ into the intestinal lumen. Here H^+ reacts with bicarbonate from bile and pancreatic juice to produce CO_2 and water, thus reducing the pH of the intestinal fluid.

Ad 3.: Na^+ -solute coupled cotransport.

The basolateral membrane of the enterocyte contains a Na^+-K^+-pump, which maintains the inward directed Na^+-gradient. The pump is energised by the hydrolysis of ATP, which provides the driving force for Na^+ entry. Thus an active process pumps Na+ out in the small interstitial space and K^+ is pumped into the cell. The basolateral membrane also contains many K^+-channel proteins, so K^+ will leak back to the interstitial space almost as soon as it has entered the cell. The K^+ is absorbed by diffusion - a daily net total of 80 mmol.

A Na^+-K^+-2 Cl^- co-transporter located on the basolateral membrane (Fig. 22-15) maintains the Cl^- gradient, with an elevated intracellular $[Cl^-]$. This transporter drags Cl^- from the interstitial fluid (ISF).

-

Fig. 22-15: Net Cl -secretion by crypt cells of the small intestine.

The transporter system uses the electrochemical Na^+ gradient to transport K^+ and Cl^- into the cell (Fig. 22-15). The crypt cells hereby can secrete Cl^- through the luminal membrane via an electrogenic channel. The Cl^- secretion produces a net luminal electronegativity, which drags Na^+ across the tight junctions resulting in net secretion (Fig. 22-15). Water (about 2 l daily) is secreted by passive osmosis.

A dramatic rise in Cl^- and water secretion - caused by gut inflammation with cholera - can lead to secretory diarrhoea.

Fluid absorption in the colon is determined by the absorption of NaCl. The Na^+ transport involves 1. Electrogenic Na^+ transfers via Na^+ channels, and 2. Na^+-co-exchange as in the small intestine (Fig. 22-14). Both transport processes are driven by the Na^+ gradient maintained by the basolateral Na^+-K^+-pump. (The Na^+-solute coupled co-transporter is not present in the human colon). The colonic Na^+-K^+- pump is more sensitive to aldosterone than that in the small intestine. Aldosterone is a steroid hormone. Steroids bind directly to cytosolic receptors and do not need second messengers. The colonic Na^+-K^+-pump activity accumulates K^+ in the enterocyte, and this gradient drives the K^+ secretion across the luminal K^+ channel. The Cl^- absorption is accomplished by diffusion along a Cl^--gradient, and by a luminal Cl^--bicarbonate exchanger producing bicarbonate secretion. We have a bicarbonate-chloride-shift just as in the red cells. Since electrolyte absorption exceeds secretion, there is a net water absorption in the healthy colon (1-1.5 l daily and with a colonic salvage capacity of 4 500 ml).

Nutrient malabsorption of the small intestine increases the fluid volume delivered to the colon and can provide an osmotic effect in the colon with diarrhoea. Up till 4 600 ml of fluid normally passes the ileocoecal valve without causing diarrhoea.

In conditions such as cholera, the excess fluid from the ileum exceeds the colonic salvage, leading to life-threatening diarrhoea. The cholera toxin can enhance the Cl^--secretion drastically and cause secretory diarrhoea with large quantities of Cl^- and water.

In inflammatory diseases of the colon, the colonic salvage capacity is markedly reduced, resulting in colonic diarrhoea.

10.e Iron absorption

Two-third of the iron content of the body (3-4 g) is stored in the haeme group of haemoglobin. The ability to transport O_2 depends on the presence of haeme. Haeme gives the red cell its characteristic red colour. Only haemoglobin with iron in the ferrous state binds O_2, whereas the dark red methaemoglobin with the iron in ferric state cannot bind O_2. Soluble ferritin forms an intracellular store (25% of total). Essential, but minor amounts of iron, is bound in myoglobin and in the electron-transporting enzymes of the mitochondria in all respiring cells. Haemosiderin is an insoluble degradation product of ferritin that aggregates into cytoplasmic granules. Haemosiderin is a normal microscopic finding in the spleen, bone marrow and the

Kupffers cells of the liver.

1. Ascorbate in the food reduces Fe^{3+} to Fe^{2+}, and forms a soluble complex with iron, thereby effectively promoting the iron absorption. We normally ingest about 20 mg iron daily, and less than 1 mg is absorbed in healthy adults, because iron form insoluble salts and complexes in the gastrointestinal secretions.

2. Iron is transported from the lumen of the upper jejunum, across the mucosa, and into the plasma by an iron-binding protein called gut transferrin.

3. Receptor proteins in the brush border membrane bind the transferrin-iron complex, and the complex is taken up into the cell by receptor-mediated endocytosis (Fig. 22-16).

Fig. 22-16: Iron absorption through an enterocyte.

4. There is a free pool of iron in the cytosol. Iron exists in one of two states in the cytosol: The ferrous state (Fe^{2+}) or the ferric state (Fe^{3+}). The Fe^{2+} ions, after absorption into the mucosal cell, are oxidised to Fe^{3+} (Fig. 22-16).

5. When intracellular iron is available in excess, it is bound to apoferritin, an ubiquitous iron-binding protein, and stored within the mucosal cells as ferritin. The synthesis of apoferritin is stimulated by iron. This translational mechanism protects against excessive absorption.

6. At the basolateral membrane the Fe^{3+} are reduced to Fe^{2+} and passes from the interstitial space to the blood. Here Fe^{2+} are again oxidised to Fe^{3+} and binds to plasma transferrin. Cellular iron stores are mobilised by autophagocytosis of enterocyte ferritin, when body stores of iron are deficient.

Normally, serum-iron is 12-36 mM, which is about one-third of the total iron-binding capacity in the plasma of adults. This means that one-third of the circulating plasma transferrin is saturated with iron.

In iron deficiency the serum-iron is falling, whereas the iron binding capacity increases. The red cell count, haematocrit and the haemoglobin concentration fall in continued deficiency, as does the concentration of iron containing cellular enzymes. Latent (or untreated) iron deficiency anaemia is found in 25-33% of all fertile females.

Increase of the total iron content takes place by enhanced intestinal iron absorption or by blood transfusions.

Ferritin is further saturated with iron to form Haemosiderin in the liver and elsewhere, when abnormal amounts are ingested over months. Extreme accumulation of excess iron in cells throughout the body (heart, lungs, pancreas, kidneys, glands and skin) finally damages vital organs and is called *haemochromatosis*.

When blood-containing products are ingested, proteolytic enzymes release the haeme groups from the haemoglobin in the intestinal lumen. Haem is absorbed by facilitated transport. Approximately 20% of the haem iron ingested are absorbed. Blood containing products are effective in iron

deficiency anaemia.

Pathophysiology

The following is a short description of classical gastrointestinal disorders, such as:

1. Achalasia, 2. Gastro-oesophageal reflux, 3. Gastritis, 4. Peptic ulcer disease, 5. Gastric tumours, 6. Gastrointestinal bleeding, 7. Coeliac disease, 8. Crohns disease and ulcerative colitis, 9. Diarrhoea, 10. Acute abdomen, 11. Colon irritabile, diverticulosis and constipation, 12. Megacolon, 13. Colonic cancer, 14. Dry mouth, and 15. Carbohydrate malabsorption.

1. Achalasia

Achalasia is a disease characterised by *lack of peristalsis* in oesophagus and relaxation failure of the lower oesophageal sphincter (LOS or american LES) in response to swallowing (Fig. 14-2). Vomiting and weight loss is major symptoms. There is no *receptive relaxation*, because the myenteric plexus does not work. The aetiology is unknown.

There is *absence of ganglion cells* in the myenteric plexus of the oesophageal wall and the LOS. The peptidergic neurons in the LOS normally secrete VIP (Vasoactive Intestinal Peptide), which relaxes the LOS, but these neurons are lost in achalasia.

The food gets stuck because of the lack of peristalsis, the oesophagus dilates and the patient regurgitates. Intermittent dysphagia during meals is typical. Many patients leave the table, provoke vomiting and are relieved. Vomiting is a classical vagal reflex phenomenon relaxing LOS.

Fig. 22-17: Oesophageal disorders

The diagnosis is confirmed by chest X-ray in particular following a barium swallow, and oesophagoscopy is necessary to exclude malignancy in the region.

A pneumatic bag is placed in the LOS opening and pressurised until LOS is sufficiently dilatated. Surgical division of the LOS muscle is performed by laparoscopy.

American trypanosomiasis (*Chagas´ disease* in Latin America) produces achalasia by microbial destruction of the ganglion cells.

2. Gastro-oesophageal reflux disease

Gastroesophageal reflux with oesophagitis is caused by incomplete closure of the LOS. Gastric contents with acid reaction then reflux into the oesophagus causing inflammation, erosion and bleeding.

This disorder is also called *reflux oesophagitis*. It results from regurgitation of gastric contents (with HCl and pepsin) into the lower oesophagus causing long lasting damage of its mucosa. The wall becomes hyperaemic, and white patches are seen on the epithelium (leucoplakias). The dysphagia most often presents as *heartburn*. As dysphagia progress it is likely that an oesophageal stricture is developing. If the squamous epithelium of the lower oesophagus is replaced by columnar epithelium, as a response to long lasting injury, there is an increased risk of transformation of the epithelium into an adenocarcinoma.

The most important barrier to the reflux is the LOS. Normally, LOS contracts as soon as the food has passed into the stomach, and the oesophagus is cleared by secondary peristalsis.

Gastro-oesophageal reflux disease is usually treated with H_2-receptor antagonists, who inhibit the gastric acid production, or with *proton pump inhibitors*, which inhibit the gastric proton pump and thus effectively reduce gastric acidity. Major complications such as *strictures* usually need surgery.

3. Gastritis

Gastritis occurs as at least two typical manifestations: Acute, erosive gastritis and chronic, non-erosive gastritis.

Focal inflammatory lesions of the mucosa characterise acute gastritis. Sometimes the erosions extend into the deeper layers of the wall (beyond the lamina propria) to form acute ulcers (Fig. 22-18). Acute gastritis is produced by alcohol, drugs (corticosteroids, ASA and NSAIDs) or infections with Helicobacter pylori or virus. After severe stress the gastritis may develop into a life-threatening condition with stress ulcers and haemorrhage. The stress conditions are severe burns, trauma, shock, and sepsis.

Chronic gastritis is a long-lasting inflammation of the gastric wall. The superficial layers are infiltrated with lymphocytes and plasma cells. Atrophia develops with loss of both parietal and chief cells. Helicobacter pylori are the chief cause of chronic gastritis in the antrum. The loss of parietal cells leads to *achlorhydria* (absent HCl production), and to deficiency of *intrinsic* factor.

Autoimmune gastritis is a pangastritis, where autoantibodies to parietal cells can be demonstrated in the blood. Vitamin B_{12} is not absorbed in the ileum in the absence of intrinsic factor, so the result is pernicious anaemia (Chapter 8).

Fig. 22-18: Peptic ulcers extend beyond the lamina propria, whereas erosions are superficial.

4. Peptic ulcer disease

Peptic ulcer disease is a mucusal ulcer in an acid- producing zone in the distal stomach or the proximal duodenum.

The normal stomach produces enough mucus and alkaline juice to protect the gastric and duodenal mucosa against HCl. The mucine molecules swell and form a non-stirred layer covering the mucosa. In duodenum the pancreatic bicarbonate creates a pH of 7.5 at the luminal membrane of the mucosa.

Epidemiological occurrence can be explained on the prevalence of *Helicobacter pylori* infection of the stomach and the colonisation of the upper gastrointestinal tract with this bacteria. Helicobacter pylori infection destroys the protective system, and at the same time provokes excess acid secretion.

The patient, whose pain complains typically occur a few hours following a meal or awaken the patient at night, points out Epigastric pains.

Bleeding from ulcers can be *fatal*. Upper gastrointestinal tract bleeding

implies a significant loss of blood into the lumen of the foregut. Haematemesis and melaena demonstrate such a bleeding. *Haematemesis* is defined as vomiting of whole blood or blood clots. *Melaena* is defined as passage of dark tarry stools (coal-black, shiny, sticky, and foul smelling).

Risk factors for peptic ulcer disease are *drugs* (ASA, NSAIDs and corticoids), *hyperparathyroidism* (the high Ca^{2+} level stimulates gastric acid secretion), and *gastrin-producing tumours* of the pancreas (Zollinger-Ellisons syndrome). Other contributing factors are increased *pepsinogen* from the chief cells, increased parietal cell mass, reduced somatostatin secretion from the antral D cells, and damage of the mucosa. Acetylsalicylic acid and other non-steroid anti-inflammatory drugs deplete the gastric mucosa for prostaglandins, which leads to mucosal damage. Strong alcoholic beverages also damage the gastric mucosal barrier and stimulate acid secretion. Caffein stimulates gastric acid secretion.

Genetic factors must be considered, since persons who do not secrete blood group 0 antigen into the saliva and gastric juice, have an increased risk of developing duodenal ulcers.

The diagnosis is confirmed with endoscopy and biopsy or with double-contrast barium technique.

The following five therapeutic strategies are used in the treatment of peptic ulcer disease:

1. Eradication of Helicobacter pylori with *antibiotics* is the treatment of choice for most cases of peptic ulcer disease, since it seems to cure the patient. *Clarithromycin* is a macrolide that binds to and prevents translocation on Helicobacter pylori- ribosomes, which is an effective basic therapy of peptic ulcers.
2. Inhibition of the *gastric proton pum***p** in the luminal membrane of the parietal cells. Omeprazole is a *proton pump inhibitor*, which relieves symptoms and cure most duodenal ulcers within four weeks - often in combination with antibiotics. Omeprazole and similar antagonists to the gastric proton pump are especially effective in treatment of *persistent* HCl-secretion caused by the Zollinger-Ellison syndrome.
3. Histamine acts through H_2 receptors on the basolateral membrane of the parietal cells. The second messengers for histamine is cAMP. All other cells contain H_1 receptors. Accordingly, *H_2 receptor antagonists* (cimetidine, ranitidine, famotidine, and nizatidine) inhibit acid secretion because they fit the H_2 receptors specifically. The H_2 receptor antagonists prevent histamine from binding to the H_2 receptors on the basolateral membrane of the parietal cells.
4. *Prostaglandin E_1 analogues,* such as misoprostol, inhibits gastric acid secretion by unspecific inhibition of the second messenger, cAMP, in the parietal cell and elsewhere. Prostaglandin E_1 analogues hereby promote ulcer healing.
5. *Surgical management* is rarely used unless complications occur. Highly selective vagotomy, in which only the nerve fibres to the parietal cells were cut was previously used, but this is not an alternative to chemical vagotomy (procedure 2., 3., 4.).

All treatment procedures, which work by inhibition of gastric acid secretion, have a common drawback. To the extent that gastric acid secretion is reduced there is no inhibition of the *gastrin release* from the antral G cells. Accordingly, the blood [gastrin] increases, and during treatment of the patients this concentration is constantly increased. The high gastrin level counteracts the expected effect on the acid production. Since gastrin is a trophical hormone for the gastric mucosa, long-term treatment with acid suppression might result in *mucosal hypertrophy* with a further rise in acid production and in cellular modifications. These complications are probably related to the rather high ulcer recurrence rate of most treatment procedures. Obviously, the only rational strategy is to eliminate the cause of the peptic ulcer disease.

5. Gastrointestinal tumours

The *leiomyoma* is the most frequent benign gastric tumour. This is a tumour of smooth muscle cells. Leiomyoma are usually discovered at autopsies or by chance, as they do not produce symptoms except when they ulcerate and bleed.

Carcinoma of the stomach is frequently located in the antrum and is almost always adenocarcinoma.

Risk factors for gastric cancer are Helicobacter pylori colonisation with chronic gastritis, atrophia and metaplasia. Dietary factors include spiced, salted or smoked food (with benzpyren). Nitrosamines are probably carcinogenic in man, and they are produced in food and water with a high nitrate content.

One third of the general population have blood group A, but 50% of all patients with gastric cancer belong to blood group A.

Enterochromaffin cells of the intestinal wall form carcinoid tumours. The tumour secretes serotonin, bradykinin, histamine, tachykinins and prostaglandins.

Somatostatin is an almost universal hormone-inhibitor. A somatostatin analogue, octreotide, inhibits the secretion of many gut hormones including those outlined above. Often the typical signs of carcinoid tumour, facial flushing and diarrhoea are totally alleviated with octreotide treatment.

6. Gastrointestinal bleeding

Acute gastrointestinal bleeding occurs in the form of haematemesis or dramatic vomiting of blood.

A bleeding peptic ulcer causes most cases. Less frequent is bleeding oesophageal varicose veins, and gastric carcinoma.

The danger is bleeding shock, with tachycardia, falling blood pressure and pallor in a cold sweating patient. Urgent and adequate blood transfusion is life saving.

Ulcers, infections, tumours, polyps, and varicose veins throughout the gastrointestinal tract cause chronic gastrointestinal bleeding. These patients present with iron deficiency anaemia (Chapter 8).

The patients are first examined with gastroscopy, often followed by Colonoscopy or enteroscopy.

7. Coeliac disease

Gluten-sensitive enteropathy or coeliac disease (*sprue*) describes a condition where the duodenal and jejunal mucosa is more or less destroyed by hypersensitivity towards gluten (see Chapter 32).

8. Crohns disease and ulcerative colitis

These two disorders may be different manifestations of a single disease, *non-specific inflammatory bowel disease* (see Chapter 32).

9. Diarrhoea

This term is usually used for an *increased stool frequency* and implies a larger than normal stool weight (Fig. 22-19).

One pathophysiological differentiation of diarrhoea is the following:

1. Zollinger-Ellisons syndrome with tremendous gastric secretion can cause diarrhoea.

Fig. 22-19: Diarrhoea of different origin.

2. Bacterial or Secretory diarrhoea is caused by increased Cl and reduced Na^+ - reabsorption. Enterotoxins from bacteria on the microvillus surface affect the toxin receptors, which increases the cAMP level in the cell. This in turn activates the chloride- channel and inhibits the NaCl reabsorption process.
3. Inflammatory diarrhoea is caused by mucosal destruction with outflow of fluid and blood such as in ulcerative colitis.
4. Osmotic active substances in the gut lumen cause osmotic diarrhoea. These substances are normal nutrients in case of malabsorption, or non-absorbable substances taken for some reason or other.
5. Diarrhoea following ileal resection. Bile acids are normally reabsorbed in the terminal ileum. Following ileal resection the bile acids enter the colon. Bile acids are toxic to the colonic mucosa and stimulate colonic secretion of large volumes,

10. The acute abdomen

Acute appendicitis is the dominant cause of *acute abdomen*. Mechanical obstruction of the orifice of the appendix by a faecolith is demonstrated in less than half the operated cases. Secretions dilatate the obstructed appendix, until the mucosa ulcerates and the wall is invaded by intestinal bacteria. In many cases only generalized inflammation is found, and in 10% of all removed appendices, the microscopy is normal.

The patient typically experiences periumbilical or diffuse pain, which moves towards the right iliac fossa within hours. The patient is subfebrile and there is nausea and vomiting. The examinator finds a tender right iliac fossa with defence musculaire (guarding), showing local peritonitis. Rectal exploration often reveals tenderness to the right.

Perforation of an inflamed appendix can cause several severe complications: Periappendiceal or hepatic abscesses, fistulae, generalized peritonitis, and septicaemia with septic shock.

Appendectomy is performed as early as possible by open surgery or by

laparoscopy.

A history of more than 48 hours of abdominal pain, with a solid mass in the right fossa iliaca indicates disaster. Perforation is most likely present with formation of a periappendiceal abscess. Here, the patient is preferably treated with antibiotics for some days, and appendectomy is delayed (French: a fraud) until the danger of generalized spread to the peritoneal cavity is minimal.

Acute peritonitis is frequently caused by perforation and presented as a sudden, severe abdominal pain. High fever develops rapidly with nausea, vomiting and paralytic ileus. As the bacterial infection spreads to affect the peritoneum in general, the condition becomes serious and septic shock may develop.

Spontaneous peritonitis with ascites in adults is caused by hepatic, alcoholic cirrhosis with portal hypertension (see Chapter 23).

11. Colon irritabile (irritable bowel), diverticulosis and constipation

These are disorders of *slow* colonic motility.

The patient with *irritabile bowel syndrome* complains of abdominal pain (diffuse or localised to the left iliac fossa), which is relieved by defecation or flatulence. There are often frequent small-volume stools, but the patient feels that the emptying is incomplete. The abdomen is distended. This is a condition with painful spasms causing constipation alternating with mucous diarrhoea. The condition is related to stress and sedentary life style, and is relieved by daily exercise.

Fig. 22-20: Frequent colonic disorders

Diverticulosis or *diverticular disease* is a condition with herniation of the mucosa through the muscular layers of the colon, caused by increased intraluminal pressure. The diverticules are recognized following a barium enema, and if they are inflamed the condition is called *diverticulitis*. Persons with disturbed stool-habits are likely to develop increased intraluminal pressure during defaecation, and they may develop hernias at weak spots in the gut wall. The incidence is high in inactive persons and low in vegetarians or in persons with a high dietary fibre content.

Mild clinical cases can be treated with light daily exercise such as walking in a hilly environment. Emergency cases may need surgery.

Constipation is frequently caused by a low fibre intake in sedentary persons. They often exhibit *irregular* defaecation habits, and irrational use of laxatives. Such habits suppress the natural reflexes.

The condition is improved by a high-fibre diet or by daily walking. Suppositories may be necessary, but long-term use of laxatives is contraindicated.

12. Megacolon

Megacolon covers several disorders, where the colon is dilatated.

Congenital megacolon or *Hirschprungs disease* is colonic dilatation resulting from congenital absence of ganglion cells in the myenteric plexus at the region, where the colon passes into rectum. Migration of cells from the neural crest is disturbed.

The cause is mutation of a gene localised on chromosome 10. The a-ganglionic segment is permanently contracted and stenotic, so the intestinal content is accumulated proximal to stenosis. The markedly distended colon gives rise to the term *megacolon*. Large amounts of faecal matters accumulate, because peristalsis and mass movements are impossible.

The diagnosis is confirmed by a transmural rectal biopsy showing absent ganglion cells. Surgical removal of the segment cures most of the young patients.

Fig. 22-21: Hirschprungs disease with a-ganglionosis and megacolon.

Acquired Megacolon usually occurs in adults with Parkinsonism, diabetic neuropathy, Chagas disease (Chapter 33) or any other disorder that affect the innervation of the smooth muscles.

13. Colonic cancer

Colonic cancer is related to slow passage of faecal material with *carcinogens* through the colon. *Carcinogens* are chemicals, whose end-products bind to DNA and damage it.

Sedentary persons have a high frequency (morbidity) of *constipation* and a high mortality of *colon cancer*, but not of rectal cancer. The colon cancer is clearly related to an *inactive life style*, and regular exercise reduces morbidity and mortality.

Prolonged accumulation of faecal content with carcinogens in the colon increases the exposure time of the mucosa and may be of importance. *High-fibre diet* and *daily walking* reduce the exposure time. There is a firm correlation between colonic cancer and the activity level of persons in industrial societies. The same is true for groups of persons living on a low fibre diet with a high content of meat and animal fat.

Usually the recto-sigmoid area is involved, a location where the faecal content is moved to and fro for varying periods (Fig. 22-22).

Patients with chronic gastrointestinal bleeding usually present with iron deficiency anaemia. Measurements for faecal occult blood are easy to perform and of value as a mass population screening for large bowel malignancy.

Fig. 22-22: Colon cancer in the ascending colon (polypoid) and in the sigmoid (constricting cancer). – A rectal cancer tumour is shown in the upper rectum.

A correlation between *rectal cancer* and exposure time for carcinogens is not to be expected, because the faecal content passes this part of the tract without delay. A correlation has also been disproved in large population groups.

14.Dry mouth

Patients with a rare autoimmune disorder (the *Sjögren syndrome*) suffer from dry mouth (*xerostomia*), dry eyes (*xerophtalmia*) and rheumatoid arthritis.

In patients lacking functional salivary glands, xerostomia, infections of the buccal mucosa, and dental caries are prevalent.

In most cases of xerostomia the condition is therapy-resistant and

unexplained. Some cases are caused by dehydration or by antidepressants.

15. Carbohydrate malabsorption

The most common chronic disorder in humans is lactose malabsorption or *hypolactasia* (lactose-induced diarrhoea or lactose intolerance), which is due to a *genetically deficiency of lactase* in the brush-border of the duodeno-jejunal enterocytes (see Chapter 31).

Self-Assessment

Multiple Choice Questions

I. Each of the following statements has True/False options:

A. The receptive relaxation response of the stomach decreases Gastroesophageal reflux.

B. The intrinsic innervation of the digestive, secretory epithelium responds to parasympathetic input with decreased secretion.

C. The sympathetic nerve fibres to the gut act presynaptically to inhibit acetylcholine release in the myenteric ganglia and activate a-receptors. Hereby, sphincter muscles are contracted, blood vessels are constricted, and secretion is inhibited.

D. Relaxation of the lower oesophageal sphincter is not caused by increased vagal inhibitory fibre discharge.

E. Oesophageal reflex activity is controlled by primary peristalsis that are co-ordinated by a swallowing centre in the solitary tract nucleus, vagal nuclei, and reticular formation. Local distension stimulates the secondary peristalsis.

II. Each of the following statements has False/True options:

A. Gastrin originates in the antral and duodenal mucosa, where it is released from G-cells.

B. Secretin is a hormone that is released from the duodenum in response to HCl.

C. Pancreozymin (CCK) contracts the sphincter of Oddi.

D. GIP stimulates insulin secretion.

E. GRP is involved in vagal gastrin secretion.

III. The following five statements have True/False options.

A. The major source of cholesterol is food intake.

B. A sweat test resulting in a Na^+ -concentration above 60 mM in the sweat, is strongly indicative of cystic fibrosis.

C. G-cells in the pancreatic islets produce large amounts of a certain

hormone, but they are named after their G-protein systems, which amplify a signal, by production of second messengers.

D. Glucagon stimulate glycogenolysis, gluconeogenesis, ureagenesis and ketogenesis.

E. VIP controls the bloodflow of the gastric mucosa, and GRP releases gastrin from the antral G-cells.

IV. Each of the following five statements have False/True options:

A: The basic electrical rhythm is an electrical event that always causes contractions in the digestive system.

B: The basic electrical rhythm determines the maximal rate of peristaltic contractions.

C: Slow waves in the colon cannot result in anti-peristalsis.

D: The major role of the human colon is to reabsorb water and electrolytes.

E: The only entirely voluntary motor process of the motility patterns in the digestive tract is chewing.

V. Each of the following five statements have False/True options:

A. Hot and acidic liquids are buffered by saliva in the mouth, and the salivary epidermal growth factor promotes the healing of wounds.

B. The parotid secretion is watery and serves to solubilize food, so it can be tasted.

C. Salivary buffers maintain the activity of amylase during the first period in the stomach.

D. Saliva has bactericide effects due to lysozyme.

E. AIDS is transferred via saliva.

VI. Each of the following five statements have False/True options:

A. Acetylcholine, gastrin and histamine stimulate gastric acid secretion.

B. H_2 blockers bind to histamine receptors at the basolateral membrane.

C. The parietal cells increase their O_2 consumption, acid secretion, intracellular [cAMP] and [Ca^{2+}], when stimulated by histamine.

D. The H^+-K^+-ATPase is responsible for gastric acid secretion.

E. Gastrin and acetylcholine does not release IP_3.

Case History A

A resting male patient, age 54 years, body weights 76 kg, is suspected of

Zollinger-Ellison syndrome and examined in the morning after fasting overnight. The throat is sprayed with lignocaine and a gastroscope is introduced into the pharynx under direct vision and passed down the oesophagus into the stomach and duodenum. No ulcers, tumours or bleeding is found. A biopsy of the mucosa shows an overgrowth of parietal cells. A sample of gastric juice is aspirated. Following stimulation by an injection of pentagastrin, gastric juice is aspirated via a nasogastric tube for one hour. The hydrogen ion concentration in the aspirate is 150 mM, and the volume is 350 ml.

Due to lung complications the blood gasses of the patient are measured in the morning (P_{aCO2} 40 mmHg, pH_a 7.40 , Base Excess zero, actual [bicarbonate] 24 mM) and just after completion of the aspiration (P_{aCO2} 40 mmHg, pHa 7.48, Base Excess 7 mM, actual

[Bicarbonate] 30 mM). The next morning blood gases were normalised.

1. *Calculate the gastric acid secretion rate of the patient, and compare the result with a normal value of 30 mmol per hour.*

2. *Describe the acid-base status of the patient just following the aspiration.*

3. *Explain the normalisation of the acid-base status the following morning.*

4. *Suggest a better diagnostic tool for the Zollinger-Ellison syndrome.*

Case History B

A nervous, smoking male, age 36 years, is admitted to hospital with severe hunger Epigastric pain reduced by eating, acid hiccups, diarrhoea, and steatorrhoea (ie, fatty stools). He has a stressful work, and over the last months he has frequently used drugs containing acetyl salicylic acid for headache, and used whisky on the rocks. Radiological examination of the stomach and duodenum suggests the presence of an ulcer in the duodenal bulb. This is confirmed by endoscopy. Gastric juice is removed by aspiration. The basal rate of HCl secretion is found to be 5 times normal. Histological examination of the gastric mucosa reveals a higher density of parietal cells and gastric glands than normal, but no hyperplasia of antral G cells.

The serum [gastrin] of the patient is 10 times higher than normal, and does not increase following a test meal.

One dose of the proton pump blocker, Omeprazole, reduces the HCl secretion rate of the patient to normal for 24 hours.

1. *Present a likely explanation for the development of the patient's duodenal ulcer.*

2. *Why does the patient have elevated serum [gastrin]?*

3. *Explain why a test meal did not induce a rise in serum [gastrin]?*

4. *Explain the mechanism for the patient's steatorrhoea and diarrhoea?*

5. *Why is one dose of omeprazole effective for such a length of time?*

6. *Transportation of one mol of H^+ from the cytosol of the parietal cell to the gastric lumen costs at least an oxidation of 30 mmol of glucose. Calculate the free energy necessary for the active transport of one mol of H^+.*

Case History C

A 35-year old male computer expert visits his general practitioner complaining of exhaustion. For weeks he has suffered from constipation, fatty stools and abdominal pain. He is loosing weight and gets out of breath when he is stair climbing.The patient looks pale and emaciated. Palpation of the abdomen reveals a soft mass in the right iliac fossa. Haematological tests show that the blood haemoglobin is 5.2 mM, the red cell count is $(3.1*10^{12})$ l^{-1} and the mean cell volume is 68 fl. Endoscopy with a duodenal biopsy show a normal mucosa with long intact villi. Colonoscopy shows patchy reddening of the mucosa and biopsies show granulomas in the lamina propria. A barium examination reveals narrowing of the terminal ileum.

1. *What is the haematological diagnosis?*

2. *What is wrong with the intestine of the patient?*

3. *What is the therapy of this condition?*

4. *Describe the complications of this chronic condition.*

5. *Describe two disorders which may mimic the condition of this patient.*

Try to solve the problems before looking up the answers.

Highlights

- *Epithelial and glandular cells of the gastrointestinal tract produce important digestive secretions that contain electrolytes, enzymes and hormones. The control of gastrointestinal secretion is effected ny neurons and by hormones.*

- *Saliva is a hypotonic fluid with high bicarbonate and potassium concentrations, and an α-amylase that cleaves α-1-4-glycoside bindings in starch.*

- *Saliva cleans the mouth and pharynx (prevents caries), and ease swallowing. Salivary lysozyme lyses bacterial cell walls. The salivary epidermal growth factor promotes the healing of wounds.*

- *Swallowing is a reflex controlled by brainstem neurons forming a swallowing centre.*

- *The swallowing and vomiting mechanisms are blocked by deep anaesthesia and by injury of the 5.th, 9.th or 10.th cranial nerve.*

- *Gastric motility mixes food with gastric juice and subdivides solids to form a fluid composed of small particles.*
- *Gastric glands and mucosa secrete gastrin (G-cells), HCl (parietal cells), pepsinogen (peptic cells), and mucus (mucous neck cells). Mucus and bicarbonate protect the gastric mucosa from adverse HCl effects.*
- *Segmentation mixes the content of the small intestine.*
- *The migrating motor complex is the "intestinal housekeeper", which cleanses the gastrointestinal tract.*
- *Vagal, cholinergic preganglionic fibres transfer signals to the gastrin-producing G-cells in the mucosa via non-adrenergic, noncholinergic (NANC) postganglionic neurons. These enteric neurons liberate gastrin-releasing peptide (GRP) to the G-cells.*
- *The ileocoecal sphincter prevents retrograde flow of colonic matter. The sphincter regulates emptying of ileum some five hours after a meal. The emptying of ileum is stimulated by gastrin, possibly via the gastroileal reflex, but a distended colon inhibits the emptying. The ileocoecal sphincter is normally passed by one litre of faecal matters daily.*
- *In the ascending colon, peristalsis is followed by antiperistalsis, which allow time for absorption of water and electrolytes.*
- *Gluten-sensitive enteropathy or coeliac disease describes a condition where the duodenal and jejunal mucosa is more or less destroyed by hypersensitivity to wards gluten. Gluten is found in barley, rye, wheat and oats.*
- *Acetylsalicylic acid and other non-steroid anti-inflammatory drugs deplete the gastric mucosa for prostaglandins, which leads to mucosal damage. Strong alcoholic beverages also damage the gastric mucosal barrier and stimulate acid secretion. Caffein stimulates gastric acid secretion.*
- *Peptic ulcer disease is a mucosal ulcer in an acid-producing zone in the distal stomach or the proximal duodenum.*
- *All treatment procedures, which work by inhibition of gastric acid secretion in peptic ulcer disease, have a common drawback. To the extent that gastric acid secretion is reduced there is no inhibition of the gastrin release from the antral G cells. Accordingly, the blood [gastrin] increases, and during treatment of the patients this concentration is constantly increased. The high gastrin level counteracts the expected effect on the acid production.*
- *Eradication of Helicobacter pylori with antibiotics is the treatment of choice for most cases of peptic ulcer disease, since it seems to cure the patient. Clarithromycin is a macrolide that binds to and prevents translocation on Helicobacter pylori- ribosomes, which is an effective basic therapy of peptic ulcers.*

- *Inhibition of the gastric proton pump in the luminal membrane of the parietal cells. Omeprazole is a proton pump inhibitor, which relieves symptoms and cure most duodenal ulcers within four weeks - often in combination with antibiotics. Omeprazole and similar antagonists to the gastric proton pump are especially effective in treatment of persistent HCl-secretion caused by the Zollinger-Ellison syndrome.*

- *Histamine acts through H_2 receptors on the basolateral membrane of the parietal cells. The second messengers for histamine is cAMP. H_2 receptor antagonists (cimetidine, ranitidine, famotidine, and nizatidine) inhibit acid secretion because they fit the H_2 receptors specifically. The H_2 receptor antagonists prevent histamine from binding to the H_2 receptors.*

- *Crohns disease is a chronic infection or inflammation of the gut with a particular prevalence for the terminal ileum, but it can be located all the way along the tract.*

- *Ulcerative colitis is always confined to the colon. Ulcerative colitis is a mucosal inflammation with haemorrhage and rectal bleeding.*

Further Reading

Calver, A., J. Collier and P. Vallance. "Nitric oxide and cardiovascular control." *Experimental Physiology* 78: 303-326, 1993.

Furness, J.B. et al. "Roles of peptides in the enteric nervous system." *Trends Neurosci* 15:66, 1992.

Return to top

Return to content

Chapter 26

The Hypothalamo-Pituitary System

Study Objectives

- To *define* autocrine, endocrine, neurocrine and paracrine transmission, endocrine feedback, endocytosis, exocytosis, hormones, hormone receptors, membrane receptors, neurohormones, neurotransmitters, phagocytosis, transcription, and tropic hormones.
- To *describe* the structural relations between the brain, hypothalamus and hypophysis. To describe secretion mechanisms, second messengers, hormonal sensitivity, melanotropin secretion and function, and stimulation-secretion coupling. To describe the structure and function of hypothalamo-hypophyseal hormones, of tropic hormones from the adenohypophysis and of pro-opio-melano-corticotropin. To describe acromegaly, Cushings syndrome, gigantism, dwarf growth, panhypopituitarism, and hyperpituitarism.
- To *explain* the secretion and function of growth hormone, somatomedins and somatostatin. To explain the secretion and function of gonadotropins, prolactin, relaxin, oxytocin, vasopressin. To explain hormonal function tests.
- To *use* the above concepts in problem solving and case histories

Principles

- *The endocrine and nervous systems co-ordinate the functions of the other organ systems as regulators of the function of the whole body.*
- *The endocrine system exerts its influence through blood-borne substances (hormones) produced in glands without secretory ducts (endocrine glands).*
- *The endocrine glands comprise the hypothalamo-hypophyseal axis, which regulates the function of the thyroid, parathyroid, adrenal, and reproductive glands. Other important hormones are the growth factors, cytokines and gastrointestinal hormones.*

Definitions

- **Autocrine transmission** refers to liberation and diffusion of signal molecules inside a cell to control functions in the cell of origin.
- **Calmodulin** is a specific binding protein for Ca^{2+} inside cells. Different calcium-calmodulin complexes activate or inhibit the activities of *calcium-dependent enzymes*.
- **Calsequestrin** is a specific binding protein for Ca^{2+} inside the sarcoplasmic reticulum of muscle cells. Calsequestrin is a buffer for cytosolic Ca^{2+}-concentration.
- **Catecholamines** are substances consisting of catechol (an aromatic structure with two hydroxyl groups) linked to an amine. The important catecholamines in humans are adrenaline, noradrenaline and dopamine.
- **Cushing's disease** is *hypercorticism* (increased glucocorticoid production) caused by a pituitary basophilic adenoma.
- **Cushing's syndrome** refers to the consequences of increased plasma glucocorticoid concentration from any source.
- **Cytokines** are secreted polypeptides that affect the functions of other cells.

- **Domain** is a segment of a protein molecule with a functional role independent of the rest.
- **Endocrine feedback** is a system, whereby the first hormone, liberated to the blood stream, controls the secretion and liberation of the second. The second hormone acts by feedback and modulates the secretion of the first.
- **Endocrine transmission** refers to transport of hormones along the blood stream to a distant target organ.
- **Endocytosis** or *pinocytosis* refers to transport of molecules or material into the cell in vesicles of cell membrane. In some cases the coating is made by a surface protein called **clathrin**. Endocytosis requires metabolic energy (ATP).
- **Exocytosis** is a process whereby the contents of intracellular vesicles (hormones, transmitters) are released to the external environment.
- **Feedback systems** which are negative contain at least one step of inhibition. The total effect is to minimise any external change introduced to the system. Almost all hormone systems maintain homeostasis by negative feedback.
- **Feedback systems** which are positive are systems, where an external change leads to increased secretion of hormone 1, which also leads to a secondary rise in hormone 2's concentration. This is an auto-accelerating phenomenon and a rarity.
- **Hormones** are *messenger* or *signal* molecules. Classical hormones are conveyed by the blood (endocrine substances) and their target cells are equipped with *receptors* that recognise each hormone. Hormone molecules form a large *signal family* together with neurotransmitters, and local diffusive (autocrine- paracrine) substances.
- **Hormone receptors** are proteins, to which hormones bind, they are present in cell *membranes*, *cytoplasm* and *nucleus*, and serve two functions. Firstly, they are required for selectivity. Secondly, they are connected to an effector mechanism in the cell. In response to hormone binding, the receptor conformation is changed, and this activates a specific enzyme system that serves as an amplifier.
- **Membrane receptors** are surface glycoproteins (just like immunoglobulins), that bind the water-soluble hormones (catecholamines and peptides). Some receptors have an amino acid sequence similar to a sequence within the hormone.
- **Neurocrine** (neurosecretory) **transmission** refers to transport of a *neurohormone* first from the cell body of a neuron along its axon, and then with the blood to its target cells.
- **Neurotransmitters** are signal molecules functioning in axonal transfer or between neurons.
- **Phagocytosis** refers to transport into cells of bacteria and large foreign bodies. The cells a leucocytes and cells of the reticuloendothelial system that can destroy noxious substances.
- **Paracrine transmission** is a release and diffusion of signal molecules with regulatory action on neighbour cells.
- **Radio-immuno-assays** (RIA) refers to any method for detecting or quantitating antigens or antibodies utilising radiolabeled reactants. RIA utilises the competitive binding between a hormone and its induced antibody. RIA can be used to detect small quantities (high sensitivity), even in complex mixtures (high specificity).
- **Transcription** (copying) is defined in Chapter 31 together with other genetic concepts.
- **Tropic hormones** regulate the growth and hormone secretion from other cells. The five classical hormones from the adenohypophysis are tropic hormones.

Essentials

This paragraph deals with *1.Hormones in general, 2. Hormone receptors, 3. Monoamines, amino acids and peptides, 4. Radio-immuno-assays, 5. Hormone function tests, 6. Clinical application of hormones, 7. The hypothalamo-hypophyseal system, 8. The adenohypophysis, 9. The neurohypophysis.*

1. Hormones in general

The scientists of the past used extirpation, substitution and transplantation to obtain the classical part of our present knowledge on endocrinology (ie, the discipline covering the internal secretion of signal molecules to the blood). Removal of the pancreas produced diabetes in animals. Removal of the pituitary gland, followed a few days later by removal of the pancreas, produced no symptoms of diabetes. This is because the adenohypophysis produces a hormone that is antagonistic to the effect of insulin in glucose metabolism. The hormone is human growth hormone. Houssay received a Nobel Prize for this work in 1947.

Hormones are *messenger* or *signal* molecules. Classical endocrine hormones are secreted into the blood and transported to their distant target cells, which are equipped with *receptors* that recognise each hormone. The hormones co-ordinate the activities of different cells in order to maintain homeostasis and to secure growth and reproduction. Hormone molecules form a large *signal family* together with neurotransmitters, autocrine and paracrine acting substances.

Paracrine and *autocrine* signal molecules are secreted and diffuse into the interstitial fluid surrounding the cells and their actions are restricted either to nearby cells (paracrine) or to the cell of origin (autocrine).

Neurotransmitters (acetylcholine, adenosine, amines, amino acids, ATP, peptides) exert a type of paracrine action, since they are released in the synaptic region.

All reactions in the cell linking stimulation and secretion together are termed *stimulation-secretion couplings*. Stimulation-secretion coupling involves depolarization of the cell membrane or opening of Ca^{2+} -channels, so that Ca^{2+} can diffuse into the cell and combine with its Ca^{2+}-binding proteins. A rise in intracellular concentration [Ca^{2+}] is necessary for exocytosis.

Elimination of hormones takes place by metabolic processes such as the inactivation of peptide hormones by proteolytic enzymes, or the transformation of hormones in the liver. Hormones are also eliminated by excretion in the urine or bile. In the liver hormones are coupled to glucuronic acid or sulphate, but these hormones are in part reabsorbed in the *entero-hepatic- circuit*.

Protein binding protects small hormone molecules (such as the thyroid hormone) from elimination. Protein binding also eases the transportation of the lipid-soluble steroids, and main
so the concentration of free hormone is maintained.

Hormones can be divided into three chemical categories:

Peptides and proteins include neuropeptides, pituitary and gastrointestinal hormones.

Steroids consist of adrenal and gonadal steroids and vitamin D, which is converted to a hormone. Steroids are lipid soluble (lipophilic).

Monoamines (modified amino acids) comprise catecholamines, histamine, serotonin, and melatonin. Catecholamines (dopamine, noradrenaline and adrenaline) are derived from tyrosine - and serotonin/melatonin from tryptophan - by a series of enzymatic conversions. Monoamines and amino acid hormones are water soluble just as peptides. Thyroid hormones are iodinated derivatives of tyrosine, and thyroid hormones are lipophilic.

The water-soluble hormones are packed in the Golgi complex in secretory granules that migrate to the cell surface.

Exocytosis of the granule contents to the interstitial fluid (ISF) and diffusion through fenestrae to the capillary blood is a common method. The secretory cells are first stimulated by chemical or electrical signals.

Synthesis of protein or peptide hormones takes place as outlined in Chapter 31. Transcription of the hormone gene results in a specific mRNA determining the synthesis of a single hormone. However, a single gene may dictate the synthesis of different peptides in different cells. As the signal protein is cut off, the prohormone is formed and

transported to the Golgi apparatus and stored in granules. The hormone specific amino acid sequence is contained in the prohormone.

An *endocrine feedback system* is a system whereby the first hormone controls the secretion and liberation of the second. The second hormone acts by feedback to modulate the secretion of the first.

A *negative feedback system* contains at least one step of inhibition. The total effect is to minimise any external change introduced to the system. Almost all hormone systems maintain homeostasis by negative feedback.

A *positive feedback system* exagerates any primary change initiated. - This is an auto-accelerating phenomenon and a rarity.

The most important example in humans is the *steep* rise in blood [oestradiol] in the middle of the menstrual cycle. High [oestradiol], when maintained for longer than 35 hours, stimulates by positive feedback, the luteinizing hormone (LH) and follicle stimulating hormone (FSH) secretion from the adenohypophysis, which further stimulate oestradiol secretion etc.

By contrast, moderate plasma [oestradiol] levels, which are present during the other parts of the cycle, provide negative instead of positive feedback. *Long feedback systems* act on the hypothalamo-pituitary system from remote target organs.

Short feedback systems use a short distance feedback, such as the influence of the hypophysis back to the hypothalamus. *Auto-feedback* refers to the action of a liberated hormone that was secreted on the cell from where it came thereby modulating its own secretion.

2. Hormone receptors

These are proteins, to which hormones bind. They are present in cell membranes, cytoplasm and nucleus, and serve two functions. Firstly, they are required for selectivity. Secondly, they are connected to an effector mechanism in the cell (Fig. 26-1).

In response to hormone binding the receptor conformation is changed, and this activates a specific enzyme system that serves as an amplifier. In the cytosol, multiple second messengers have evolved to serve such purposes, whereas in the nucleus, the hormone-receptor complex binds to DNA and regulates gene expression (Fig. 26-1). The effector domain of the membrane receptor is directly coupled to the regulatory portion of the effector enzymes (such as adenylcyclase = adenylate cyclase). These effector enzymes control ion fluxes, membrane transport systems, the production of cyclic nucleotides, and the breakdown of phospholipids. Inactive kinases are activated by the use of ATP.

This phosphorylation is critical for for transformation of information and for cell viability (synthesis, transport and metabolism of vital molecules). Many hormones initiate a series of reactions when bound to membrane receptors. One family of coupling molecules, called G-proteins, links some of the receptors to nearby effector molecules (see Chapter 1). Other receptors make use of another system.

Fig. 26-1: Target cells activation by hormones acting at Membrane, Cytoplasmic, and Nuclear receptors.

Steroids and thyroid hormones are lipophilic and therefore pass easily through the cell membrane by diffusion. Steroids bind to specific cytosol-receptor proteins that are then translocated into the cell nucleus where they reversibly bind to DNA (Fig. 26-1). Some unbound receptor proteins may even exist in the nucleus. The binding of the steroid-receptor complex to the specific gene modulates mRNA transcription.

Tri-iodo-thyronine (T_3) binds to nuclear receptor proteins, which then attaches to a thyroid response unit in the gene in a manner similar to that of steroid receptors. The result is increased mRNA formation (Fig. 26-1).

Steroids and thyroid hormones frequently work in conjunction with each other (potentiate amplification of gene expression).

Cell membrane and intracellular receptors can change their affinity and number. A specific ligand for a receptor is able to modulate the total number of this receptor. In
drug) often reduces the number of receptors (down-regulation), and other hormones recruit their own receptors at low concentrations (up-regulation). Maximal effects of hormones are generally observed at receptor occupancy of less than 50%.

The myoepithelial cells (myometrium and breast) contain oxytocin receptors. Their number is up regulated by estrogens and down regulated by progesterone. The cardiac muscle contains nor-adrenergic receptors (b_1). Both affinity and number of receptors is increased by thyroid hormone stimulation (T_3/T_4).

Internalisation is the transport of hormone-receptor complex into the cell by an endocytotic vesicle. This is a means of terminating the action of the hormone. After destruction of the hormone by lysosomes, the receptor returns to the surface and is reused.

3. Monoamines, amino acids and peptides

Such water-soluble hormones (first messengers) bind to hormone receptors on the lipid-rich plasma membrane. Peptide hormone and catecholamine receptors are membrane receptors with a binding domain located extracellularly and an effector domain intracellularly (Fig. 26-1).

The second messengers involved are cyclic adenosine monophosphate (cAMP), cyclic guanosine monophosphate (cGMP), inositol trisphosphate (IP_3), Ca^{2+}, diacylglycerol (DAG) etc. The Ca^{2+}-ion is an important second messenger.

The Ca^{2+}-influx to the cytosol is controlled by hormone receptor binding, neural stimuli or modified by other second messengers.

Sutherland discovered cAMP and demonstrated its role as a second messenger in mediating body functions (Nobel Prize 1971).

Increased activity of the sympathetic nervous system including release of adrenaline triggers fight-or-flight reactions. In the heart, adrenaline molecules diffuse to the myocardial cells, where they bind to membrane b-receptors. A stimulatory signal is hereby transmitted to an associated enzyme called adenylcyclase. This enzyme catalyses the conversion of ATP to cAMP. The importance of cAMP is that it activates protein kinase A, which, among many other functions, phosphorylates the Ca^{2+}-channel protein. This activation is correlated with an increase in the magnitude of the Ca^{2+}-influx, the force of contraction, and the heart rate.

The parasympathetic system counteracts the sympathetic by slowing the heart rate and decreasing the force of contraction. Acetylcholine is bound to another set of specific membrane receptors located on the heart cell membrane. Acetylcholine reduces the Ca^{2+}-influx that was increased by adrenaline.

Most hormones have a blood concentration of approximately 10^{-10} mol per l. One molecule bound to a cell receptor releases 10 000 times more cAMP in the cell. Hence, cAMP works as an amplifier of the hormone signal.

Phosphodiesterase (PDE) destroys cAMP. PDE enhances hydrolysis of cAMP to the inactive 5'- AMP by a highly exergonic process.

Inhibitors of the PDE (theophylline and caffeine) act synergistically with hormones that use cAMP as a second messenger.

cAMP stimulates catabolic processes such as lipolysis, glycogenolysis (glucagon), gluconeogenesis, and ketogenesis. The cAMP also stimulates amylase liberation in the saliva by the parotid gland, the HCl secretion by the parietal cells, the insulin release by the b-cells in pancreas, and the increased ion permeability of many cell membranes.

When the glucose concentration increases in the arterial blood and close to the b-cells of the pancreatic islets of Langerhans, it triggers an increase in Ca^{2+} -influx to the cell.

The initial surge in insulin secretion is caused by calmodulin-dependent *protein kinases*.

The high cytosolic [Ca^{2+}] activates the membrane phospholipase A_2 and C. Phospholipase A2 releases arachidonic acid (AA) which stimulates insulin secretion. Phospholipase C catalyses the formation of IP_3 and DAG. The IP_3 releases more Ca^{2+} from the endoplasmic reticulum, and DAG activates protein kinase C.

The decrease in insulin secretion after the initial surge and its subsequent increase can be explained by the action of protein kinase C.

Initially, the active protein kinase C stimulates the Ca^{2+}-pump in the plasma membrane, reduces cytosolic [Ca^{2+}] and thus reduces the initial calmodulin-dependent insulin secretion. Later, protein kinase C stimulates the formation of cAMP and amplifies the induction of calmodulin-dependent protein kinase thereby causing a gradual increase in insulin secretion. Prolonged glucose stimulation probably leads to down-regulation of protein kinase C. An abnormally prolonged glucose stimulation may render b-cells glucose blind and thus spoil their function.

Insulin secretion is not only stimulated by glucose, but also potentated by acetylcholine via phospholipase C and by glucagon via activation of adenylcyclase. b-Agonists stimulate b-receptors on the glucagon producing a-cells, whereas a-agonists inhibit insulin secretion via a_2-receptors on the b-cells. Acetylcholine and glucagon react by activating protein kinase C and cAMP dependent protein kinase A, respectively. Both mechanisms potentate the Ca^{2+}-triggered insulin secretion.

Transcription in the cell nucleus produces a precursor messenger RNA molecule complementary to part of a DNA. The precursor is processed into messenger RNA and transported through the nuclear membrane into the cytoplasm. Messenger RNA carries the genetic information in triplet codons (Chapter 31). Messenger RNA binds to ribosomes and transfer RNA molecules synthesise peptides (ribosomal translation). Translation produces big precursor molecules (pre-pro-hormones). Precursors have a signal peptide that contains processing information to ensure that the protein enters the rough endoplasmic reticulum. Here enzymes split the precursor into a signal molecule and a prohormone. Finally, peptide hormones undergo post-translational processing (for eg, thyroid stimulating hormone, TSH, and gonadotropins are glycosylated; insulin forms a zinc-complex). The hormones reach the Golgi complex, where they are packed into secretory granules that migrate to the cell surface.

Roger Guillemin synthesized brain peptides that regulate the pituitary secretion in vitro. He received the Nobel Prize in 1977.

4. Radio-immuno-assays (RIA)

RIA refers to any method for detecting or quantitating antigens or antibodies utilising radiolabeled reactants. RIA is used to detect very small quantities of antigens or antibodies, even in complex mixtures.

First a specific antibody is produced towards the antigen (eg, hormone).

In one version of RIA for antigen detection, the antigen is radiolabeled and reacted with a limited amount of specific antibody. The complex containing bound antigen is then separated from free antigen. Unlabeled antigen in a test sample is used to compete with the binding of radiolabeled antigen. The test antigen is quantitated from the extent of inhibition obtained with standards containing defined amounts of the same antigen.

Rosalyn S. Yalow and Saul Berson developed the RIA method. Rosalyn Yalow received the Nobel Prize in 1977.

Recent variations of the RIA technique include immuno-radiometric, chemi-luminescent, and enzyme-linked radioimmuno-sorbent assays. - In the radioreceptor assay a hormone receptor is substituted for the antigen-antibody in RIA.

5. Hormone function tests

The following tests are clinical tools in the diagnosis of hormone disorders:

5.1. The hormone concentration in the blood is commonly used. It can be measured by taking advantage of the new methods described above.

5.2. The secretion flux of T_3 and T_4 from the thyroid gland.

5.3. The metabolic rate or the absorption rate of ^{131}I (radioactive iodine) in the thyroid gland. The physical half-life of ^{131}I is 8 days or 192 hours. The elimination rate constant (k) of a substance is the amount eliminated per unit time divided by the total amount present in the distribution volume, assuming exponential elimination. The variable k is easy to calculate:

$T_{½}$ = ln 2/k = 0.693/k. The value of k for iodine is 0.693/192 or 0.0036 $hours^{-1}$.

5.4. The elimination rate:

Abnormal amounts of catecholamines or VMA (Vanillyl mandelic acid) in a 24-hour urine suggest the presence of a catecholamine-producing tumour (phaeochromocytoma).

5.5. Stimulation test:

Stimulation with ACTH (Adrenocorticotropic hormone) without a substantial rise in plasma -[cortisol] suggests primary, adrenocortical atrophia.

5.6. Suppression test:

Dexamethasone (a cortisol synergist) is administered to a Cushing suspect patient in the evening. The next morning a measurement of plasma [cortisol] shows suppres
primary, adrenocortical hyperfunc
production by negative feedback.

Hypothalamic/pituitary Cushing is never suppressed by cortisol.

5.7. The glucose tolerance test:

A load of glucose normally triggers an increased rate of insulin production.

6. Clinical applications of hormones

Distribution of oestrogens and progesterone in *contraceptives* (P pills) is world-wide. Oestrogens are widely used to relieve postmenopausal discomfort. Now some females with osteoporosis are treated experimentally with calcitonin, because calcitonin inhibits osteoclastic bone resorption.

Insulin is a lifesaver for diabetics, and it is produced and distributed as pure human insulin.

In the affluent areas of the world many women deliver their babies following an oxytocin infusion.

Oestrogens and gonadotropins are used in treatment of sterility and menstrual disturbances.

Huggins received the Nobel Prize in 1966 for the introduction of a new form of cancer therapy in which sex hormones are used to retard their growth. He used androgens for breast cancer and oestrogens for prostate cancer.

7. The hypothalamo-hypophyseal system.

The human pituitary gland consists essentially of two parts both controlled by the hypothalamus.

The *glandular* part is the adenohypophysis or anterior lobe, and the *neural* part is the neurohypophysis or posterior lobe.

The adenohypophysis develops ectodermally from the primitive mouth cavity (Rathkes pouch). Blood-borne signal molecules from the hypothalamus regulate the cells of the adenohypophysis.

The neurohypophysis develops from the neuro-ectoderm in the floor of the third ventricle. The two parts combine to form one body called the adeno-neuro-hypophysis that weighs about 0.5 g.

The infundibular process of the neurohypophysis receives blood from the inferior hypophyseal artery, whose capillary plexus drains into the adenohypophysis (Fig. 26-2). The upper stalk and the median eminence is supplied with blood by the carotid artery to the superior hypophyseal artery, whose primary capillary plexus ends in long portal veins carrying blood to the highly permeable secondary capillary plexus of the adenohypophysis. From this plexus, blood drains into the dural sinus. The adenohypophysis lies outside the blood-brain barrier, and does not receive arterial blood directly. A third capillary plexus, between the neurohypophysis and the median eminence of the hypothalamus,

allows *short loop feedback* from the hypophysis to the hypothalamus.

The hypothalamus and the hypophysis connects in the following ways (Fig. 26-2):

1. The neurosecretory axons pass from the cell bodies in the supraoptic and paraventricular nuclei of the hypothalamus to the neurohypophysis. The neurosecretory granules are here stored in the terminals of these axons. The granules are released by exocytosis upon stimulation. The peptides from the granules then enter the capillary plexus of the inferior hypophyseal artery

2. The hypophysiotropic zone in the median eminence of the hypothalamus is connected to the adenohypophysis. Both releasing and inhibiting peptides are synthesized in hypothalamic neuron bodies and transported to the median eminence in granules via axonal transport. At the median eminence the inhibitory and releasing signal molecules are discharged to the capillary plexus of the superior hypophyseal artery. From here they follow the blood through the long portal veins to reach the specific cells in the adenohypophysis. Here, the releasing and inhibiting hormones modulate the output of tropic hormones.

Fig. 26-2: The hypothalamo-pituitary axis. The 5 classical tropic hormones are corticotropin, gonadotropins (FSH, LH), somatotropin, thyrotropin and mammotropin.

Neurosecretory neurons (which have nuclei in the hypothalamus and axons that lead to the median eminence and to the posterior lobe of the hypophysis) and peptidergic neurons (spread in the nervous system and gut) produce and liberate peptides in much the same way.

The secretory granules travel through the axons of the neurosecretory neurons that form the supraoptico-hypophyseal tract, with high velocity (more than 100 mm each hour). This tract runs through the pituitary stalk and end in the neurohypophysis. The transfer is known as axoplasmic transport. The neuro-hy
the capillary blood. During transport the pro-hormone splits into its subunits. Oxytocin and vasopression are then released.

The nerve endings of this tract in the neurohypophysis are the storage area for these two neurosecretory hormones; secretion to the blood takes place through fenestrated capillaries. The secretion granules release their content by regulated exocytosis. Exocytosis is triggered when the neurosecretory neuron is depolarised and an action potential is transferred to the terminals.

Even a small rise in the osmolarity of plasma stimulates osmoreceptors, located close to the neurosecretory cells in the hypothalamus.

The osmoreceptors stimulate both production and secretion of vasopressin (ADH) in the neurosecretory cells. The plasma [ADH] will then rise from the basal level that is 2 pmol per l. The normal secretion flux is 10^{-13}mol ADH per kg body weight per min, and the biological half-life in human plasma is 18 min. Some females increase their plasma [ADH] in the pre-menstrual phase.

The neuroregulatory peptides are endogenous opiates (endorphins and enkephalins), b-lipotropin, neurotensin, substance P, VIP or vasoactive intestinal peptide etc. Many of these peptides are cut off from a big mother molecule: pro-opio-melanocortin (POMC). These peptides may exhibit a permissive effect on other hormones (ACTH, growth hormone) related to behaviour and autonomic responses. During exercise (a Cooper test which lasts 12 min) the plasma [b-endorphin] and [ACTH] increases by 200-300% from the normal resting averages of 1.7 and 2.2 pmol per l, respectively. There is an increase in plasma [ACTH] and its accompanying neuroregulatory peptides during prolonged stress like exhaustive exercise and chronic disease. The endogenous opiates affect stress-adapting behaviour, such as the euphoria observed in the chronically ill.

The pituitary gland normally has a mass of approximately 500 mg, but it increases during pregnancy and decreases with aging.

8. The adenohypophysis

The *five* hormones from the adenohypophysis are tropic hormones - they regulate the growth and hormone secretion of target cells (Fig. 26-2). They include:

1. Thyrotropin or thyroid-stimulating hormone (TSH), which is produced in thyrotropic cells,
2. Gonadotropins (FSH and LH) from gonadotropic cells,
3. Corticotropin (ACTH) from corticotropic cells,
4. Somatotropin (human growth hormone, HGH or GH) from somatotropic cells and
5. Prolactin or mammotropin produced in mammotropic cells.

The five tropic peptide hormones have the following molecular characteristics:

1. One group contains glycoproteins with two peptide chains.

   There is a special, biologically active b-chains for each of the three hormones TSH, FSH and LH, although the inactive a-chain is the same for all of them.

2. Somato-mammotropins are single-chain peptides containing 200 amino acids of almost the same sequence.

   HGH and prolactin (PRL) are probably simple gene duplicates from the same prohormone molecule.

3. POMC peptides are neuroregulatory hormones: ACTH, endogenous opiates, b-endorphin, b-lipoprotein, a-MSH and b-MSH (MSH abbreviates melanocytic stimulating hormone).

   Histamine plays an important role in pituitary hormone secretion. Histamine stimulates the secretion of ACTH, b-endorphin, a-MSH, and PRL. Histamine participates in the release of these hormones during prolonged stress and possibly in the suckling- and oestrogen-induced PRL-release.

   The release of growth hormone (GH) and TSH are predominantly inhibited by histamine. GH is the main stimulator of body growth in humans (Chapter 30).

   Histamine increases the secretion of LH in females - mediated by GnRH (gonadotropin releasing hormone). Histamine probably affects the cell bodies in the supraoptic and paraventricular nuclei, stimulating the formation of arginine vasopressin and oxytocin.

9. The neurohypophysis

The neurohypophysis secretes two hormones: vasopressin and oxytocin.

Vasopressin or antidiuretic hormone (ADH) is a vasopressor with a strong antidiuretic effect as the names imply. Vasopressin is normally synthesized as a big pre-prohormone in the ribosomes of neurons in the supraoptic and paraventricular nuclei of the hypothalamus. The pre-prohormone consists of a signal peptide, ADH, neurophysin and a glycopeptide. First, the signal peptide is cut off, and then the precursors are packed in secretion granules in the Golgi complex. The secretion granules travel by axoplasmic transport through the axon of the neurosecretory neurons that form the supraoptic-hypophyseal tract, and then are stored in its terminals in the neurohypophysis. These terminals are located close to the fenestrated capillaries. The smallest rise in the osmolarity of plasma stimulates osmoreceptors located close to the neurosecretory cells of the hypothalamus. The osmoreceptors stimulate both production and release of ADH. Vasopressin is a nonapeptide with a molecular weight of 1084 Da.

ADH has the following effects:

1. ADH eases the renal reabsorption of water in the cortical collecting ducts (and not in the outer medulla but in the inner medulla) - leading to antidiuresis.
2. ADH probably stimulates the active solute reabsorption (NaCl) in the thick ascending limb of the renal Henle

loop. Thus, ADH helps maintain the concentration gradient in the kidney.

3. Vasopressin is a universal vasoconstrictor. Vasopressin reduces the small, medullary bloodflow through vasa recta along the Henle loop.

   ADH acts on the basolateral membrane of the cells, and the result is a rise of [cAMP] in the cytosol. The cAMP diffuses to the luminal side, where it causes vesicular structures to develop and fuse with the luminal membrane. Hereby, the membrane receives a large number of water channels, so the membrane becomes highly water permeable. Water diffuses through the cell to the basolateral membrane and into the interstitial fluid.

Oxytocin

Stimulation of *tactile receptors in the mammary nipple* causes the neurosecretory neurons to release oxytocin through a neuroendocrine reflex.

The latency between the stimulus and milk ejection is due mainly to the transport of oxytocin in the blood from the neurohypophysis to the milk ducts (20-30 s). Oxytocin stimulates the myoepithelial cells in the milk ducts of the lactating breast so that milk is ejected to the baby.

Oxytocin also stimulates the myoepithelial (myometrial) cells of the uterus satisfying the woman sexually during breast-feeding. Oxytocin can perhaps start labour.

Pathophysiology

- This paragraph deals with the five tropical hormones from the adenohypophysis and vasopressin secreted from the neurohypophysis: *1.Pituitary TSH-secreting tumours*, *2. Polycystic ovarian syndrome*, *3. Basophilic pituitary adenoma*, *5. Prolactinomas*, *6. Diabetes insipidus*, *7. Syndrome of inappropriate ADH secretion*, *8. Panhypopituitarism*.

1. Pituitary TSH-secreting tumours

A pituitary TSH-secreting tumour is an extremely rare cause of thyrotoxicosis. Thyrotoxicosis is dealt with in Chapter 28.

2. Polycystic ovarian syndrome

Chaotic LHRH secretion from the hypothalamus to highly sensitive gonadotropic cells in the pituitary increases the LH-level in the blood plasma. Actually, LHRH induces its own receptors. The gonadotropin level is so high in the follicular phase that androgens in excess are produced from the theca cells. These females produce immature or atretic follicles occurring as multiple cysts in enlarged ovaries.

A maintained LH-level can be produced by other causes. The excessive gonadotropin and androgen secretion causes irregular bleedings, subfertility, acne and hirsutism.

3. Basophilic pituitary adenoma

Basophilic pituitary adenoma is the cause of classical *Cushing's disease*. The excessive ACTH secretion induces adrenocortical hypersecretion of cortisol. The hyper-cortisolaemia causes the many symptoms and signs found in *Cushing's syndrome* (Chapter 30).

4. Somatotropic pituitary adenoma

Somatotropic pituitary adenomas produce large amounts of growth hormone leading to *gigantismus* in childhood and to *acromegaly* in adults. These cases of hyperpituitarism are dealt with in Chapter 30. In rare cases the cause is excessive GHRH secretion from the hypothalamus. Some acromegalics also produce excess prolactin in hypertrophic mammotrophs.

Pituitary adenoma cells with TRH receptors also secrete excess GH. TRH is used as a diagnostic test. Other pituitary adenoma cells have somatostatin receptors. Somatostatin and somatostatin agonists inhibit GH secretion, and make the

adenomas shrink

5. Prolactinomas

Prolactinomas are microlactinomas (less than 1 cm), which cause anovulatory, irregular bleedings, abnormal milk production (galactorrhoea), and subfertility. The constantly elevated plasma prolactin inhibits the LH-secretion necessary for ovulation. – Dopamine, dopamine agonists and somatostatin analogues inhibit prolactin secretion and can make the prolactinomas shrink.

6. Diabetes insipidus

The true form of *diabetes insipidus* is caused by deficiency of vasopressin (ADH deficiency). There are two types of diabetes insipidus. The primary or idiopathic type, which is due to a genetic defect that blocks the hormone production, and the secondary type, where the hypothalamo-hypophysary system is damaged by disease or surgery.

The renal tubule cells are rarely insensitive to ADH, and this infrequent condition is called *renal* diabetes insipidus or *nephrogenic* diabetes insipidus. This is a sex-linked recessive disorder or it is acquired from renal disorders or hypercalcaemia.

The symptoms and signs are mainly due to the large diuresis (polyuria), nocturia, and a tremendous thirst (polydipsia). A total lack of ADH can result in a diuresis of 25 l daily.

ADH eases the renal reabsorption of water in the cortical collecting ducts and in the inner medulla via a cAMP mechanism which increases the number of water channels in the luminal membrane. ADH stimulates NaCl reabsorption in the thick ascending limb of Henle, and vasopressin is a universal vasoconstrictor.

Synthetic vasopressin is given intra-nasally as a spray up to 3 times daily.

7. Syndrome of inappropriate ADH secretion

ADH producing tumours in the hypophysis or in the lungs causes the *syndrome of inappropriate ADH secretion*. Water retention, concentrated urine, hyposmolar plasma, and muscle cramps characterise this syndrome.

8. Panhypopituitarism

is typically due to total destruction (lesions or tumour invasion) of all hormones in the hypothalamo-hypophysary system. Lack of GH and somatomedins result in a dwarf without normal sex development (lack of LH and FSH). This dwarf has also hypothyroidism (lack of TSH), and a Cushing-like syndrome (hyper secretion).

The differentiation of somatotrophs, mammotrophs and thyrotrophs is dependent upon a *protein transcription factor* (Pit-1). Mutation of the *Pit-1 gene* leads to hypoplasia of the adenohypophysis and to insufficient production og GH, prolactin and TSH with hypopituitarism.

Self-Assessment

Multiple Choice Questions

The following five statements have True/False options:

A. High-pressure liquid chromatography is a sensitive analysis used for many hormones and biochemical key molecules.

B. he affinity and number of specific membrane receptors on a given cell is constant.

C. Monoamines and amino acid hormones are water-soluble just as peptides.

D. cAMP inhibits catabolic processes such as lipolysis, glycogenolysis, gluconeogenesis, and ketogenesis.

E. Hyperpituitarism is often caused by microadenomata, which typically cause dwarf growth.

Case History A

The adenohypophysis of a 23 year old woman contains approximately 300 μg of TSH with a molecular weight of 31 000. TSH has a half-life ($T_{½}$) in plasma of 55 min and a concentration of 100 pmol per l of plasma. The haematocrit of the patient is 0.5. The woman secretes TSH to her total blood volume (TBV), which is 4 l.

1. Develop an equation for the calculation of her TSH secretion (J mol/hour). The rate constant k can be used.
2. Calculate the secretion of TSH from her adenohypophysis
3. What fraction of her total TSH store is secreted per 24 hours?

Case History B

A woman (24 years of age; height: 1.70 m; weight: 60 kg) is in hospital due to a tremendous thirst, and she drinks large amounts of water. Since she is producing 10 or more litres of urine each day, the doctors suspect the diagnosis to be diabetes insipidus. The vasopressin concentration in plasma (measured by a RIA method) is 10 fmol per l. Her secretion of vasopressin is only 5% of the normal flux of 10^{-13} mol per min per kg body weight. The normal plasma [vasopressin] is 2 pmol per l as a mean. The extracellular volume (ECV) is 20% of her body weight.

Vasopressin is injected intravenously at several occasions. A dose of 3 μg vasopressin is the minimum necessary to normalise her diuresis for 4 hours. Before the injection her diuresis is 6 ml of urine per min, but within 25 min her urination is constantly around 0.5 ml/min.

1. Calculate the secretion of vasopressin (in mg/hour) from the neurohypophysis of a normal 60-kg person and of this patient.
2. Calculate the distribution volume for vasopressin, which is 20% higher than ECV.
3. Assume the 3 mg vasopressin injected to be distributed evenly immediately after the intravenous injection. Calculate the rise in vasopressin concentration in the distribution volume.
4. Estimate the relation between this concentration and that of a healthy individual.
5. Does this ratio have implications for the interpretation of her special type of diabetes insipidus?
6. Is it dangerous to lose 10 litres of urine per day?

Try to solve these problems before looking up the answers .

Highlights

- *The hypothalamo-pituitary system controls the function of the adrenal, thyroid, and reproductive glands, as well as regulating growth, lactation, milk secretion and water excretion.*
- *Protein and peptide hormone synthesis starts with processing of a primary gene transcript (code) called a prohormone. The processing includes proteolysis, glycosylation and phosphorylation.*
- *Catecholamines, peptide and protein hormones are stored in secretory granules and discharged by exocytosis.*
- *Catecholamines, peptide and protein hormones are water-soluble and cannot pass the cell membrane. They act on the surface of target cells via membrane receptors. The hormone-receptor-complex activates second messengers in the cell (cAMP, Ca^{2+}, DAG, IP_3) via stimulatory or inhibitory G-proteins.*
- *Thyroid and steroid hormones are lipid-soluble and act through specific nuclear receptors. The hormone-receptor-*

complex modulates elements in DNA molecules in order to change the expression of target genes.

- *Peptide hormones produced in the cell bodies of hypothalamic neurons pass down their axons inside secretory granules to be stored in the terminals of the neurohypophysis.*
- *Releasing and inhibiting peptides from the hypothalamus are released in pulses in the adenohypophysis and act via second messengers. They modulate transcription, translation and secretion of tropic hormones.*
- *Vasopressin or antidiuretic hormone is a vasopressor with a strong antidiuretic effect as the names imply. Vasopressin (ADH) is normally synthesized as a big pre-prohormone in the ribosomes of neurons in the supraoptic and paraventricular nuclei of the hypothalamus. The pre-prohormone consists of a signal peptide, ADH, neurophysin and a glycopeptide.*
- *POMC peptides are neuroregulatory hormones: ACTH, endogenous opiates, β-endorphin, β-lipoprotein, α-MSH and β-MSH.*
- *Stimulation of tactile receptors in the mammary nipple causes the neurosecretory neurons to release oxytocin through a neuroendocrine reflex. Oxytocin stimulates the myoepithelial cells in the milk ducts of the lactating breast. Oxytocin also stimulates the myoepithelial (myometrial) cells of the uterus. Oxytocin can perhaps start labour.*
- *The true form of diabetes insipidus is caused by deficiency of vasopressin (ADH deficiency). There are two types of diabetes insipidus. The primary or idiopathic type, which is due to a genetic defect that blocks the hormone production, and the secondary type, where the hypothalamo-hypophysary system is damaged by disease or surgery.*
- *ADH producing tumours in the hypophysis or in the lungs causes the Syndrome of inappropriate ADH secretion. Water retention, concentrated urine, hyposmolar plasma, and muscle cramps characterise this syndrome.*
- *Panhypopituitarism is due to total destruction (lesions or tumour invasion) of all hormones in the hypothalamo-hypophysary system. Lack of GH and somatomedins result in a dwarf without normal sex development (lack of LH and FSH). This dwarf has also hypothyroidism (lack of TSH), and a Cushing-like hyper secretion).*
- *Hyperpituitarism is often caused by prolactin producing microadenomata, which cause abnormal milk production. This leads to disturbance of the menstrual cycle and infertility. Other pituitary adenomas produce large amounts of GH leading to gigantismus in childhood and to acromegaly in adults.*

Further Reading

Nature. Weekly journal published by Macmillan Magazines Ltd, Porters South, 4 Crinan Street, London N1 9XW, UK.

Cell. Bi-weekly journal published by Cell Press, 1050 Massachusetts Av., Cambridge Massachusetts 02138, USA.

Yalow, R. S. “Radioimmunoassay: Historical aspects and general considerations”, in *Handbook of Experiment. Pharmacol*., 1987.

Return to top

Return to Content

nt