

Virology

Principles and
Applications

2nd Edition



John Carter and Venetia Saunders

Contents

Preface to Second Edition

Preface to First Edition

Abbreviations Used in This Book

Greek Letters Used in This Book

Color Coding for Molecules

Chapter 1: Viruses and Their Importance

1.1 Viruses are Ubiquitous on Earth

1.2 Reasons for Studying Viruses

1.3 The Nature of Viruses

1.4 The Remainder of the Book

Learning Outcomes

Sources of Further Information

Chapter 2: Methods Used in Virology

2.1 Introduction to Methods Used in Virology

2.2 Cultivation of Viruses

2.3 Isolation of Viruses

2.4 Centrifugation

2.5 Structural Investigations of Cells and Virions

2.6 Electrophoretic Techniques

2.7 Detection of Viruses and Virus Components

2.8 Infectivity Assays

2.9 Virus Genetics

2.10 Investigation of Protein-Protein Interactions

Learning Outcomes

Sources of Further Information

Chapter 3: Virus Structure

3.1 Introduction to Virus Structure

3.2 Virus Genomes

3.3 Virus Proteins

3.4 Capsids

3.5 Virion Membranes

3.6 Occlusion Bodies

3.7 Other Virion Components

Learning Outcomes

Sources of Further Information

Chapter 4: Virus Transmission

4.1 Introduction to Virus Transmission

4.2 Transmission of Plant Viruses

4.3 Transmission of Vertebrate Viruses

4.4 Transmission of Invertebrate Viruses

4.5 Permissive Cells

Learning Outcomes

Sources of Further Information

Chapter 5: Attachment and Entry of Viruses into Cells

5.1 Overview of Virus Replication

5.2 Animal Viruses

5.3 Bacteriophages

Learning Outcomes

Sources of Further Information

Chapter 6: Transcription, Translation, and Transport

6.1 Introduction to Transcription, Translation, and Transport

6.2 Transcription of Virus Genomes

6.3 Transcription in Eukaryotes

6.4 Translation in Eukaryotes

6.5 Transport of Molecules in Eukaryotic Cells

6.6 Transcription and Translation in Bacteria

Learning Outcomes

Sources of Further Information

Chapter 7: Virus Genome Replication

7.1 Overview of Virus Genome Replication

7.2 Locations of Virus Genome Replication in Eukaryotic Cells

7.3 Initiation of Genome Replication

7.4 Polymerases

7.5 DNA Replication

7.6 Double-Stranded RNA Replication

7.7 Single-Stranded RNA Replication

7.8 Reverse Transcription

Learning Outcomes

Sources of Further Information

Chapter 8: Assembly and Exit of Virions from Cells

8.1 Introduction to Assembly and Exit of Virions from Cells

8.2 Nucleocapsid Assembly

8.3 Formation of Virion Membranes

8.4 Virion Exit From the Infected Cell

Learning Outcomes

Sources of Further Information

Chapter 9: Outcomes of Infection for the Host

9.1 Introduction to Outcomes of Infection for the Host

9.2 Factors Affecting Outcomes of Infection

9.3 Non-Productive Infections

9.4 Productive Infections

Learning Outcomes

Sources of Further Information

Chapter 10: Classification and Nomenclature of Viruses

10.1 History of Virus Classification and Nomenclature

10.2 Modern Virus Classification and Nomenclature

10.3 Baltimore Classification of Viruses

Learning Outcomes

Sources of Further Information

Chapter 11: Herpesviruses (and Other dsDNA Viruses)

11.1 Introduction to Herpesviruses

11.2 The Human Herpesviruses

11.3 The Herpesvirus Virion

11.4 HSV-1 Genome Organization

11.5 HSV-1 Replication

11.6 Latent Herpesvirus Infection

11.7 Other dsDNA Viruses

Learning Outcomes

Sources of Further Information

Chapter 12: Parvoviruses (and Other ssDNA Viruses)

12.1 Introduction to Parvoviruses

12.2 Examples of Parvoviruses

12.3 Parvovirus Virion

12.4 Parvovirus Replication

12.5 Other ssDNA Viruses

Learning Outcomes

Sources of Further Information

Chapter 13: Reoviruses (and Other dsRNA Viruses)

13.1 Introduction to Reoviruses

13.2 Rotavirus Virion

13.3 Rotavirus Replication

13.4 Other dsRNA Viruses

Learning Outcomes

Sources of Further Information

Chapter 14: Picornaviruses (and Other Plus-Strand RNA Viruses)

14.1 Introduction to Picornaviruses

14.2 Some Important Picornaviruses

14.3 Picornavirus Virion

14.4 Picornavirus Replication

14.5 Picornavirus Recombination

14.6 Picornavirus Experimental Systems

14.7 Other Plus-Strand RNA Viruses

Learning Outcomes

Sources of Further Information

Chapter 15: Rhabdoviruses (and Other Minus-Strand RNA Viruses)

15.1 Introduction to Rhabdoviruses

15.2 Some Important Rhabdoviruses

15.3 The Rhabdovirus Virion and Genome Organization

15.4 Rhabdovirus Replication

15.5 Other Minus-Strand RNA Viruses and Viruses with Ambisense Genomes

15.6 Reverse Genetics

Learning Outcomes

Sources of Further Information

Chapter 16: Influenza Virus

16.1 Introduction to Influenza Viruses

16.2 The Influenza Virion

16.3 Influenza A Virus Replication

16.4 Importance of Influenza Viruses

16.5 Reverse Genetics

Learning Outcomes

Sources of Further Information

Chapter 17: Retroviruses

17.1 Introduction to Retroviruses

17.2 Retrovirus Virion

17.3 Retrovirus Replication

17.4 Examples of Retroviruses

17.5 Retroviruses as Gene Vectors

17.6 Endogenous Retroviruses

Learning Outcomes

Sources of Further Information

Chapter 18: Human Immunodeficiency Viruses

18.1 Introduction to HIV

18.2 HIV Virion

18.3 HIV Genome

18.4 HIV-1 Replication

18.5 HIV-1 Variability

18.6 Progression of HIV Infection

18.7 Prevention of HIV Transmission

Learning Outcomes

Sources of Further Information

Chapter 19: Hepadnaviruses (and Other Reverse-Transcribing DNA Viruses)

19.1 Introduction to Hepadnaviruses

19.2 Importance of HBV

19.3 HBV Virion

19.4 Non-Infectious Particles

19.5 Soluble Virus Protein

19.6 HBV Genome

19.7 HBV Genetic Groups

19.8 HBV Replication

19.9 Prevention and Treatment of HBV Infection

19.10 Other Reverse-Transcribing DNA Viruses

Learning Outcomes

Sources of Further Information

Chapter 20: Bacterial Viruses

20.1 Introduction to Bacterial Viruses (Bacteriophages)

20.2 Single-Stranded RNA Phages

20.3 Double-Stranded RNA Phages

20.4 Single-Stranded DNA Phages

20.5 Double-Stranded DNA Phages

Learning Outcomes

Sources of Further Information

Chapter 21: Origins and Evolution of Viruses

21.1 Introduction to Origins and Evolution of Viruses

21.2 Origins of Viruses

21.3 Evolution of Viruses

Learning Outcomes

Sources of Further Information

Chapter 22: Emerging Viruses

22.1 Introduction to Emerging Viruses

22.2 Viruses in New Host Species

22.3 Viruses in New Areas

22.4 Viruses in New Host Species and in New Areas

22.5 New viruses

22.6 Recently Discovered Viruses

22.7 Re-Emerging Viruses

22.8 Virus Surveillance

22.9 Dealing with Outbreaks

Learning Outcomes

Sources of Further Information

Chapter 23: Viruses and Cancer

23.1 Introduction to Viruses and Cancer

23.2 Papillomavirus-Linked Cancers

23.3 Polyomavirus-Linked Cancers

23.4 Epstein-Barr Virus-Linked Cancers

23.5 Kaposi's Sarcoma

23.6 Adult T Cell Leukemia

23.7 Hepatocellular Carcinoma

23.8 Virus-Associated Cancers in Animals

23.9 Cell Lines Derived From Virus-Associated Cancers

23.10 How Do Viruses Cause Cancer?

23.11 Prevention of Virus-Induced Cancers

Learning Outcomes

Sources of Further Information

Chapter 24: Survival of Infectivity

24.1 Preservation of Virus Infectivity

24.2 Destruction of Virus Infectivity

24.3 Inactivation Targets in Virions

24.4 Inactivation Kinetics

24.5 Agents that Inactivate Virus Infectivity

Learning Outcomes

Sources of Further Information

Chapter 25: Virus Vaccines

25.1 Introduction to Virus Vaccines

25.2 Live Attenuated Virus Vaccines

25.3 Inactivated Virus Vaccines

25.4 Virion Subunit Vaccines

25.5 Live Recombinant Virus Vaccines

25.6 Mass-Production of Viruses for Vaccines

25.7 Virus-Like Particles

25.8 Synthetic Peptide Vaccines

25.9 DNA Vaccines

25.10 Storage and Transport of Vaccines

Learning Outcomes

Sources of Further Information

Chapter 26: Anti-viral Drugs

26.1 Introduction to Anti-viral Drugs

26.2 Development of Anti-viral Drugs

26.3 Examples of Anti-viral Drugs

26.4 Drug Resistance

26.5 Anti-viral Drug Research

Learning Outcomes

Sources of Further Information

Chapter 27: Prions

27.1 Introduction to Prions

27.2 Transmissible Spongiform Encephalopathies

27.3 The Nature of Prions

27.4 Prion Diseases

27.5 Prion Strains

27.6 Prion Transmission

27.7 The Protein-Only Hypothesis

Learning Outcomes

Sources of Further Information

VirologistsB5™ Vocabulary

Index



VIROLOGY

PRINCIPLES AND APPLICATIONS

John B. Carter
and
Venetia A. Saunders

[StormRG]
KAT | TPB | ET | h33t

WILEY

This edition first published 2013 by John Wiley & Sons Ltd

Copyright © 2013, 2007 John Wiley & Sons Ltd

Cover image: Glass Sculpture of HIV virus courtesy of Wellcome Images

All effort has been made to trace and acknowledge ownership of copyright. The publisher would be glad to hear from any copyright holders whom it has not been possible to contact.

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The rights of John Carter and Venetia Saunders to be identified as the authors of this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit www.wiley.com.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Carter, John B., 1944-

Virology : principles and applications / John Carter, Venetia Saunders. — 2nd ed. p.; cm.

Includes bibliographical references and index.

ISBN 978-1-119-99143-4 (cloth) — ISBN 978-1-119-99142-7 (pbk.)

I. Saunders, Venetia A., 1949- II. Title.

[DNLM: 1. Viruses. 2. Virus Diseases. QW 160]

616.9'101—dc23

2012041238

ISBN 978-1-1186-2979-6 (ebk)

ISBN 978-1-1186-2977-2 (ebk)

ISBN 978-1-1186-2976-5 (ebk)

ISBN 978-1-1199-9142-7 (pbk.)

ISBN 978-1-1199-9143-4 (hbk)

A catalogue record for this book is available from the British Library

Senior Commissioning Editor: Andrew McAleer

Assistant Editor: Katharine Earwaker

Marketing Managers: Fran Hunt and Jo Underwood

Project Editor: Juliet Booker

To Myra, Robert, Jon, and Mark

Preface to Second Edition

Virology continues to be an exciting subject and to develop at a rapid pace. The introduction of new laboratory techniques and the continued application of established techniques are producing a wealth of new information. There has been an explosion in the publication of virology papers reporting the discovery of new viruses and providing deeper insights into many facets of the subject. We have tried to reflect these developments in the second edition, which contains much new material, including an additional chapter (on influenza virus). In an attempt to improve clarity, sections of the text have been rewritten.

In the diagrams, many of which have been redrawn for inclusion in this edition, there is a standard color code to differentiate various types of nucleic acid and protein molecules. Please note that, in the interests of clarity, there have been some modifications to the color code used in the first edition. There is a key to the color code on page xxix. In addition, the book has a website (www.wiley.com/college/carter) with accompanying teaching and learning resources, including animations of virus replication cycles.

We are grateful to the many people who provided feedback on the first edition and to those who made suggestions for the second edition. We have incorporated many of the suggestions for new material, though space constraints prevented us from incorporating them all. Again, we would be grateful to receive feedback.

We thank the team at John Wiley & Sons Ltd for all their help with this new edition and to our families for their continuing support.

John B. Carter—

John.Carter147@ntlworld.com

Venetia A. Saunders—

—Authors are now retired; both were previously at School of Pharmacy & Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK.

Preface to First Edition

Virology is a fascinating and rapidly developing subject, and is worthy of study purely because viruses are interesting! Furthermore, virology is a branch of science that is of immense relevance to mankind for a host of reasons, not least of which are the threats to human health caused by viruses, such as HIV, hepatitis B virus, papillomaviruses, measles, and influenza viruses, to mention just a few. There is a continuing need for trained virologists and it is hoped that this book will play a small role in helping to fulfill that need. To a large extent the material in the book is based on virology taught at Liverpool John Moores University.

This is not a textbook of fundamental virology, medical virology, veterinary virology, plant virology or of bacteriophages, but a bit of each of these! The general pattern of the book is that principles of virology are covered earlier and applications are covered later. There is no strict demarcation between the two, however, so the reader may be made aware of important applications while principles are being introduced.

The first 10 chapters cover basic aspects of virology. A chapter on methods used in virology comes early in the book, but could be skimmed to gain an overview of its contents and thereafter used for reference. There is one chapter on each of the seven Baltimore classes, concentrating mainly on animal viruses. There is a chapter devoted entirely to HIV and an extended chapter on phages, reflecting the renewed interest in their biology and applications. After a chapter on origins and evolution of viruses, there follow five chapters covering various aspects of applied virology, including vaccines and antiviral drugs. The final chapter is on prions, which are not viruses but are often considered along with the viruses.

Each chapter starts with *At a glance*, a brief summary with the dual aim of giving a flavor of what is coming up and providing a revision aid. Each chapter ends with a list of learning outcomes and a guide to further reading in books and journals. The references are mainly from the twenty-first century, but there is a selection of important papers from the last century.

The book has a web site (www.wiley.com/go/carter) where you can find:

- many references additional to those in the book;
- links to the journal references (to the full text where this is freely available, otherwise to the abstract);
- links to virology web sites;
- self-assessment questions and answers for each chapter, to reinforce and extend concepts developed in the book.

A key feature of our book is a standard color code to differentiate various types of nucleic acid and protein molecules in the diagrams. The color code is explained in

the following pages. It is appreciated that color coding may be of limited value to individuals who have difficulty in differentiating colors, so we have also labeled many of the molecules.

A number of virus replication cycles are described and the reader should be aware that these are models based on evidence to date; the models may have to be modified in the light of future evidence. We present the virus replication cycles as fitting within a general framework of seven steps:

1. Attachment of a virion to a cell
2. Entry into the cell
3. Transcription of virus genes into mRNAs
4. Translation of virus mRNAs into virus proteins
5. Genome replication
6. Assembly of the virus proteins and genomes into virions
7. Exit of the virions from the cell.

We hope that this helps in appreciating how virus replication fits into a general pattern, and in comparing the replication cycles of different types of virus. For some groups of viruses the framework has to be modified, and we make clear when that is the case.

If you come across an unfamiliar term please consult the Virologists' B^o™ Vocabulary at the back of the book. This glossary includes not only virology-specific terms, but also a selection of terms from cell biology, molecular biology, immunology, and medicine.

A list of abbreviations that are used throughout this book appears on the following pages.

We wish to thank the many people who have made the production of this book possible. We thank all those who supplied images and those who gave permission for the use of their images; we are especially grateful to David Bhella, Tom Goddard, Kathryn Newton, and Jean-Yves Sgro. Thanks also to Robert Carter for assistance with images. We acknowledge the contributions of the many students who have acted as guinea pigs for our teaching materials and who have provided us with feedback. Grateful thanks also to those who reviewed material for the book and provided valuable feedback. We are sorry that we were unable to include all the topics suggested, but if we had done so the book would have run to several volumes! Many thanks to Rachael Ballard and all at John Wiley & Sons Ltd who helped the book come to fruition. Finally, thanks to our families for their support and for their patience during those many hours we spent ensconced in the study.

We hope you find the book useful and we would be interested to hear what you think of it. We have tried to ensure that there are no errors, but it is probable that some have slipped through; if you come across any errors please inform us.

John B. Carter

John.Carter147@ntlworld.com

Venetia A. Saunders

School of Pharmacy & Biomolecular Sciences, Liverpool John Moores University,
Byrom Street, Liverpool, L3 3AF, UK

Abbreviations Used in This Book

(+) DNA	plus strand (positive strand) DNA
(-) DNA	minus strand (negative strand) DNA
(+) RNA	plus strand (positive strand) RNA
(-) RNA	minus strand (negative strand) RNA
A	adenine
ADP	adenosine diphosphate
AIDS	acquired immune deficiency syndrome
AP-1	activator protein 1
ATP	adenosine triphosphate
b	base(s)
BL	Burkitt's lymphoma
bp	base pair(s)
BSE	bovine spongiform encephalitis
C	cytosine
C terminus	carboxy terminus
cccDNA	covalently closed circular DNA
CD	cluster of differentiation
cDNA	copy DNA
CJD	Creutzfeldt-Jakob disease
cos	cohesive end
CP	coat protein
CPE	cytopathic effect
DIP	defective interfering particle
DNA	deoxyribose nucleic acid
ds	double-stranded
DTR	direct terminal repeat

E	early
EBV	Epstein–Barr virus
EF	elongation factor
ELISA	enzyme-linked immunosorbent assay
ERV	endogenous retrovirus
E. coli	Escherichia coli
Ff	F-specific filamentous
G	guanine
GFP	green fluorescent protein
gp	(1) glycoprotein (as in HIV-1 gp120)
	(2) gene product (as in phage T4 gp5)
HA	hemagglutinin
HAART	highly active anti-retroviral therapy
HAV	hepatitis A virus
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HPV	human papillomavirus
HSV	herpes simplex virus
HTLV-1	human T-lymphotropic virus 1
IC50	50% inhibitory concentration
ICTV	International Committee on Taxonomy of Viruses
ICTVdB	International Committee on Taxonomy of Viruses database
IE	immediate early
IG	intergenic
IRES	internal ribosome entry site
ITR	inverted terminal repeat

kb	kilobase(s)
kbp	kilobase pair(s)
kD	kiloDalton(s)
KSHV	Kaposi's sarcoma-associated herpesvirus
LDI	long-distance interaction
LE	left end
LIN	lysis inhibition
LPS	lipopolysaccharide
LTR	long terminal repeat
Mbp	megabase pair(s)
MHC	major histocompatibility complex
MJ	Min Jou
m.o.i.	multiplicity of infection
MP	movement protein
mRNA	messenger RNA
N terminus	amino terminus
NA	neuraminidase
NF- κ B	nuclear factor kappa B
NK cell	natural killer cell
nm	nanometer(s) (10^{-9} meter)
NPC	nasopharyngeal carcinoma
NSP	non-structural protein
O	operator
ORF	open reading frame
ori	origin (replication)
P	promoter
PBS	primer binding site
PCR	polymerase chain reaction

pfu	plaque-forming unit
phage	bacteriophage
ppm	parts per million
PPT	polypurine tract
pRb	retinoblastoma protein
PS	packaging signal
RBS	ribosome binding site
RE	right end
RF	replicative form
RI	replicative intermediate
RNA	ribose nucleic acid
RNAi	RNA interference
RNase H	ribonuclease H
RNP	ribonucleoprotein
rRNA	ribosomal RNA
RT-PCR	reverse transcriptase-polymerase chain reaction
S	Svedberg unit
SARS	severe acute respiratory syndrome
SD	Shine-Dalgarno
SI	selectivity index
SIV	simian immunodeficiency virus
Sp1	stimulatory protein 1
ss	single-stranded
ssb	single-stranded binding
SV40	simian virus 40
T	thymine
T antigen	tumor antigen
TCID50	virus dose that infects 50% of tissue cultures

TK	thymidine kinase
tRNA	transfer RNA
TSE	transmissible spongiform encephalitis
U	uracil
UV	ultraviolet
vCJD	variant Creutzfeldt-Jakob disease
VP	virus protein
VPg	virus protein, genome linked
VSV	vesicular stomatitis virus

Greek letters used in this book

α	alpha
β	beta
γ	gamma
ϵ	epsilon
θ	theta
κ	kappa
λ	lambda
σ	sigma
ϕ	phi
ψ	psi

Color Coding for Molecules

With the aim of maximizing clarity of the diagrams the following standard color code is used to depict molecules.

Virus molecules

DNA

DNA

Where there are distinct (+) and (-) strands these are coded: (+) DNA
(-) DNA

RNA

primary transcripts and mRNA

(+) RNA (other than primary transcripts and mRNA)

(-) RNA

A cap on the 5' end of an RNA molecule is indicated by the symbol 

Protein

Some virus proteins are in this color
Some virus proteins are in this color

N.B. The colours of capsid images derived from X-ray crystallography and cryo-electron microscopy do not conform to the color coding in the drawings.

Cell molecules

DNA

mRNA

tRNA

protein

CHAPTER 1

Viruses and Their Importance

CHAPTER 1 AT A GLANCE

Photographs reproduced with permission of

1 World Health Organization.






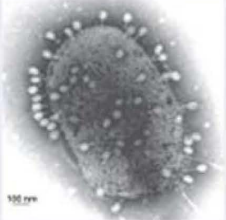
2 From Umeshappa et al. (2011) *Veterinary Immunology and Immunopathology*, 141, 230. Reproduced by permission of Elsevier and the authors.

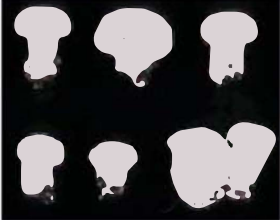
3 MacFarlane and Robinson (2004) Chapter 11, *Microbe-Vector Interactions in Vector-Borne Diseases*, 63rd Symposium of the Society for General Microbiology, Cambridge University Press. Reprinted with permission.

4 University of Warwick.

5 Cornell Integrated Microscopy Center.

Viruses infect:

- Humans

Smallpox¹
- Other vertebrates

Bluetongue virus-infected sheep²
- Invertebrates

Tipula sp. larvae (leatherjackets) infected with invertebrate iridescent virus 1
- Plants

Delayed emergence of potato caused by tobacco rattle virus infection³
- Fungi

Damaged potato (spraing) caused by tobacco rattle virus infection³
- Bacteria

Escherichia coli cell with phage T4 attached⁵


Mushroom virus X⁴

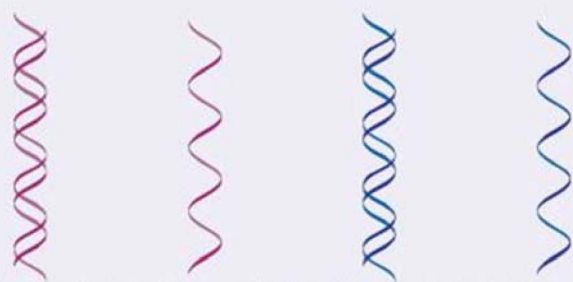
Some viruses are useful:

- Phage typing of bacteria
- Sources of enzymes
- Pesticides
- Anti-bacterial agents
- Anti-cancer agents
- Gene vectors

Viruses are parasites; they depend on cells for molecular building blocks, machinery, and energy.

Virus particles are small; dimensions range approx. from 20–400 nm.

A virus genome is composed of one of the following:



double-stranded DNA single-stranded DNA double-stranded RNA single-stranded RNA

1.1 VIRUSES ARE UBIQUITOUS ON EARTH

Viruses infect all cellular life forms: eukaryotes (vertebrate animals, invertebrate animals, plants, fungi) and prokaryotes (bacteria and archaea). The viruses that infect bacteria are often referred to as bacteriophages, or phages for short.

The presence of viruses is obvious in host organisms showing signs of disease. Many healthy organisms, however, are hosts of non-pathogenic virus infections, some of which are active, while some are quiescent. Furthermore, the genomes of many organisms contain remnants of ancient virus genomes that integrated into their host genomes long ago. As well being present within their hosts, viruses are also found in soil, air and water. Many aqueous environments contain very high concentrations of viruses that infect the organisms that live in those environments.

There is a strong correlation between how intensively a species is studied and the number of viruses found in that species. Our own species is the subject of most attention as we have a vested interest in learning about agents that affect our health. It is not surprising that there are more viruses known that infect mankind than any other species, and new human viruses continue to be found. The intestinal bacterium *Escherichia coli* has also been the subject of much study and many viruses have been found in this species. If other species received the same

amount of attention it is likely that many would be found to be hosts to similar numbers of viruses.

It is undoubtedly the case that the viruses that have been discovered represent only a tiny fraction of the viruses on the Earth. Most of the known plants, animals, fungi, bacteria and archaea have yet to be investigated for the presence of viruses, and new potential hosts for viruses continue to be discovered. Furthermore, the analysis of DNA from natural environments points to the existence of many bacterial species that have not yet been isolated in the laboratory; it is likely that these non-cultivable bacteria are also hosts to viruses.

1.2 REASONS FOR STUDYING VIRUSES

1.2.1 Some viruses cause disease

Viruses are important agents of many human diseases, ranging from the trivial (e.g. common colds) to the lethal (e.g. rabies), and viruses also play roles in the development of several types of cancer. As well as causing individuals to suffer, virus diseases can also affect the well-being of societies. Smallpox had a great impact in the past and AIDS is having a great impact today.

There is therefore a requirement to understand the nature of viruses, how they replicate and how they cause disease. This knowledge permits the development of effective means for prevention, diagnosis, and treatment of virus diseases through the production of vaccines, diagnostic reagents and anti-viral drugs. Vaccines, such as rotavirus and measles vaccines, have saved millions of lives and improved the quality of life for millions more. Smallpox has been eradicated as a result of vaccination. Anti-viral drugs, such as those used against HIV and herpes simplex virus, play major roles in the treatment of infectious disease. Medical applications therefore constitute major aspects of the science of virology.

Many viruses cause disease in domestic animals (such as cattle, sheep, dogs, poultry, fish, and bees) and in wild animals (such as red squirrels and seals). Vaccines are used to control some of these diseases, for example foot and mouth disease and bluetongue. Crop plants are also hosts to large numbers of viruses, such as rice yellow mottle virus and cucumber mosaic virus, a virus with a very wide host range. These, and other plant viruses, can cause devastating outbreaks of disease in crop plants, with significant impact on the quantity and quality of food produced.

Another area where viruses can cause economic damage is in those industries where the products result from bacterial fermentation. In the dairy industry phages can destroy the lactic acid bacteria used to produce cheese, yogurt, and other milk products, while other phages can destroy *Corynebacterium* species used in the industrial production of amino acids.

1.2.2 Some viruses are useful

Some viruses are studied because they have useful current or potential applications.

- Phage typing of bacteria. Some groups of bacteria, such as some *Salmonella* species, are classified into strains on the basis of the spectrum of phages to which they are susceptible. Identification of the phage types of bacterial isolates can provide useful epidemiological information during outbreaks of disease caused by these bacteria.
- Sources of enzymes. A number of enzymes used in molecular biology are virus enzymes. Examples include reverse transcriptases from retroviruses and RNA polymerases from phages.
- Pesticides. Some insect pests are controlled with baculoviruses, and myxoma virus has been used to control rabbits.
- Anti-bacterial agents. In the mid-twentieth century phages were used to treat some bacterial infections in humans. Interest waned with the discovery of antibiotics, but has been renewed with the emergence of antibiotic-resistant strains of bacteria.
- Anti-cancer agents. Genetically modified strains of viruses, such as herpes simplex virus and vaccinia virus, are being investigated for treatment of cancers. These strains have been modified so that they are able to infect and destroy specific tumor cells, but are unable to infect normal cells.
- Gene vectors for protein production. Viruses, such as certain baculoviruses and adenoviruses, are used as vectors to take genes into animal cells growing in culture. This technology is used to make cells produce useful proteins, such as vaccine components. Some genetically modified cells are used for mass production of proteins.
- Gene vectors for treatment of genetic diseases. Children with the genetic disease Severe Combined Immunodeficiency (baby in the bubble syndrome) have been successfully treated using retroviruses as gene vectors. The viruses introduced into the children's stem cells a non-mutated copy of the mutated gene (Section 17.5).

1.2.3 Virus studies have contributed to knowledge

Much basic knowledge of molecular biology, cell biology, and cancer has been derived from studies with viruses. Here are a few examples.

- A famous experiment carried by Alfred Hershey and Martha Chase, and published in 1952, used phage T2 and *E. coli* to provide strong evidence that genetic material is composed of DNA.
- The first enhancers to be characterized were in genes of simian virus 40 (SV40).

- The first transcription factor to be characterized was the transplantation (T) antigen of SV40.
- The first nuclear localization signal of a protein was identified in the T antigen of SV40.
- Introns were discovered during studies of adenovirus transcription.
- The role of the cap structure at the 5' end of eukaryotic messenger RNA (mRNA) was discovered during studies with vaccinia virus and a reovirus.
- The first internal ribosome entry site to be discovered was found in the RNA of poliovirus.
- The first RNA pseudoknot to be discovered was in the genome of turnip yellow mosaic virus.

1.3 THE NATURE OF VIRUSES

1.3.1 Viruses are small particles

Evidence for the existence of very small infectious agents was first provided in the late nineteenth century by two scientists working independently: Martinus Beijerinck in Holland and Dimitri Ivanovski in Russia. They made extracts from diseased plants, which we now know were infected with tobacco mosaic virus, and passed the extracts through fine filters. The filtrates contained an agent that was able to infect new plants, but no bacteria could be cultured from the filtrates. The agent remained infective through several transfers to new plants, eliminating the possibility of a toxin. Beijerinck called the agent a *contagium vivum fluidum* and the term has been in use ever since.

At around the same time, Friedrich Loeffler and Paul Frosch transmitted foot and mouth disease from animal to animal in inoculum that had been highly diluted. A few years later Walter Reed and James Carroll demonstrated that the causative agent of yellow fever is a filterable agent.

Figure 1.1 gives some indication of the size of these agents, which are known as virus particles or virions. The virion of a herpesvirus, which is a fairly large virus, is about ten million times smaller than a large balloon, while the balloon is smaller than the Earth by the same factor. The virions of most viruses are too small to be seen with a light microscope and can be seen only with an electron microscope (Figure 1.2).

Figure 1.1 Comparative sizes of a herpesvirus particle, a balloon, and the Earth. A large balloon is about ten million times larger than a herpesvirus particle, while the Earth is larger than the balloon by the same factor.

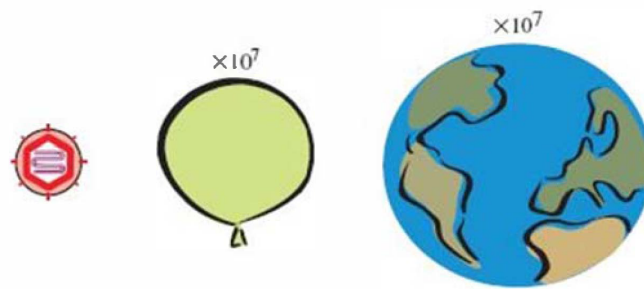


Figure 1.2 Transmission electron microscope. This is a microscope in which the image is formed by electrons transmitted through the specimen.

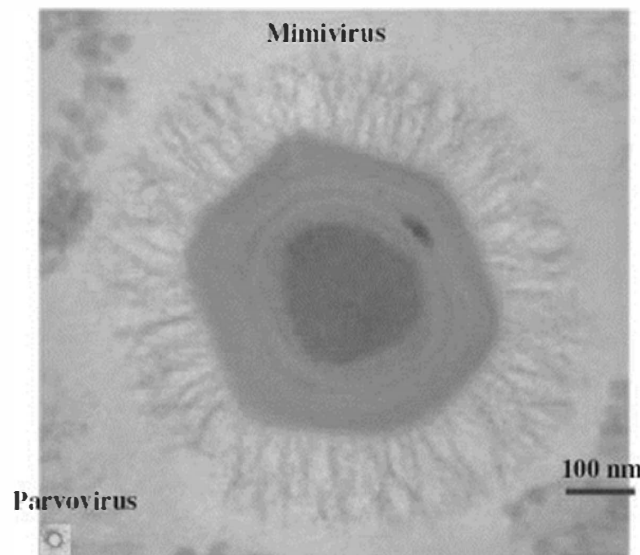
Source: Photograph courtesy of JEOL.



The units in which virions are normally measured are nanometers ($1 \text{ nm} = 10^{-9} \text{ m}$). Although virions are very small, their dimensions cover a large range. Amongst the smallest are parvoviruses (Figure 1.3), with diameters about 20 nm, while the megavirus and mimivirus (microbe-mimicking virus), isolated from amoebae, are amongst the largest.

Figure 1.3 Virions of mimivirus, one of the largest viruses, and a parvovirus, one of the smallest viruses.

Source: Electron micrograph of mimivirus from Claverie et al. (2009) *Journal of Invertebrate Pathology*, 101, 172. Reproduced by permission of Elsevier and the authors. Electron micrograph of parvovirus by permission of Professor M. Stewart McNulty and The Agri-Food and Biosciences Institute.



Virology is therefore concerned with very small particles, though often with very large numbers of those particles! A concentrated suspension of virions might contain 10^{12} virions/ml. A single virus-infected cell might produce 10^5 virions. A person infected with HIV might produce 10^{11} virions in a day.

Virions are not cells. They do not contain organelles, except for the virions of arenaviruses, which contain cell ribosomes that were packaged when the virions were assembled.

1.3.2 Viruses have genes

The virion contains the genome of the virus. Whereas the genomes of cells are composed of double-stranded DNA, there are four possibilities for a virus genome:

- double-stranded DNA (dsDNA);
- single-stranded DNA (ssDNA);
- double-stranded RNA (dsRNA);
- single-stranded RNA (ssRNA).

The genome is enclosed in a protein coat known as a capsid. The genome plus the capsid, plus other components in many cases, constitute the virion. A major function of the capsid is to protect the genome until it can be delivered into a cell in which it can replicate.

Generally, virus genomes are much smaller than cell genomes and the question arises as to how viruses encode all their requirements in a small genome. Viruses achieve this in a number of ways.

- Viruses use host cell proteins. The genomes of large viruses duplicate some functions of the host cell, but the small viruses rely very heavily on host cell functions. There is, however, one function that an RNA virus must encode, no matter how small its genome. That function is an RNA polymerase, because cells do not encode enzymes that can replicate virus RNA. A significant proportion of an RNA virus genome is taken up with the gene for an RNA polymerase.

- Viruses code efficiently. There may be overlapping genes and genes encoded within genes. The small genome of hepatitis B virus is a good example (see Section 19.6). In viruses most, or all, of the genome encodes proteins, in contrast to mammals where only about 1.5% of the genome encodes proteins.
- Many virus proteins are multifunctional. Some virus proteins have multiple enzyme activities, for example the rhabdovirus L protein replicates RNA, caps and polyadenylates mRNA, and phosphorylates another virus protein. Some virus proteins can bind to cell proteins with roles in immune responses, and hence interfere with immunity of the host.

1.3.3 Viruses are parasites

Viruses differ from cells in the way in which they multiply. A new cell is always formed directly from a pre-existing cell, but a new virion is never formed directly from a pre-existing virion. New virions are formed by a process of replication, which takes place inside a host cell and involves the synthesis of components followed by their assembly into virions.

Viruses are therefore parasites of cells, and are dependent on their hosts for most of their requirements, including:

- building-blocks such as amino acids and nucleosides;
- protein-synthesizing machinery (ribosomes);
- energy, in the form of adenosine triphosphate.

A virus modifies the intracellular environment of its host in order to enhance the efficiency of the replication process. It does this by producing proteins, and in some cases small RNAs, that interact with cell components. It has been demonstrated that the proteins of a virus can interact with hundreds of host proteins in an infected cell. The expression of thousands of host genes can be affected, with some genes up-regulated and some genes down-regulated.

Virus-induced modifications to a host cell might include the formation of new membranous structures or a reduced immune response. Some large phages of photosynthetic bacterial hosts encode proteins that enhance photosynthesis, thereby probably boosting the yields of virus from the cells.

A point has now been reached where the nature of viruses can be summarized in a concise definition (see the box).

VIRUS DEFINITION

A virus is a very small, non-cellular parasite of cells. Its genome, which is composed of either DNA or RNA, is enclosed in a protein coat.

1.3.4 Some viruses are dependent on other viruses

Some viruses, known as satellite viruses, are unable to replicate unless the host cell is infected with a second virus, referred to as a helper virus. The helper virus

provides one or more functions missing from the satellite virus, thereby enabling the latter to complete its replication cycle. Examples are given in Table 1.1

Table 1.1 Examples of satellite/helper viruses

Type of host	Satellite virus	Genome	Helper virus	Genome
Animal	Hepatitis delta virus	ssRNA	Hepatitis B virus	dsDNA
	Adeno-associated viruses	ssDNA	Adenovirus	dsDNA
Plant	Satellite tobacco necrosis virus	ssRNA	Tobacco necrosis virus	ssRNA
Ameba	Sputnik	dsDNA	Mimivirus	dsDNA
Bacterium	Enterobacteria phage P4	dsDNA	Enterobacteria phage P2	dsDNA

Viruses belong to biology because they possess genes, replicate, evolve, and are adapted to particular hosts, biotic habitats, and ecological niches. However, . . . they are nonliving infectious entities that can be said, at best, to lead a kind of borrowed life.

Marc van Regenmortel and Brian Mahy (2004)

It's life, Jim, but not as we know it!

Dr McCoy speaking to Captain Kirk of the Starship Enterprise, Star Trek

1.3.5 Are viruses living or nonliving?

There is an ongoing debate as to whether viruses are living or nonliving; the view taken depends on how life is defined. Viruses have genes and when they infect cells these genes are replicated, so in this sense viruses are living. They are, however, very different to cellular life forms, so Dr McCoy's stock phrase (see the box) on finding new life forms in the galaxy could be applied to viruses. When viruses are outside their host cells they exist as virions, which are inert, and could be described as nonliving, but viable bacterial spores are inert and are not considered to be nonliving. You might form your own view as to whether viruses are living or nonliving as you progress through this book.

When Beijerinck selected the word *virus* he chose the Latin word for poison. This term has now been in use for over a century and virology has developed into a huge subject. More recently, the term virus has acquired further meanings. Computers are threatened by infection with viruses that can be found in the wild once they have been released by their authors. These viruses are specific for certain file types. Infected files may be put on several web sites and a virus epidemic may ensue. Another use of the term virus is exemplified in John Humphrys' book *Lost For Words*, in which he talks about the deadly virus of management-speak infecting language. All the italicized terms in this paragraph are also used in the context of the viruses that are the subject of this book.

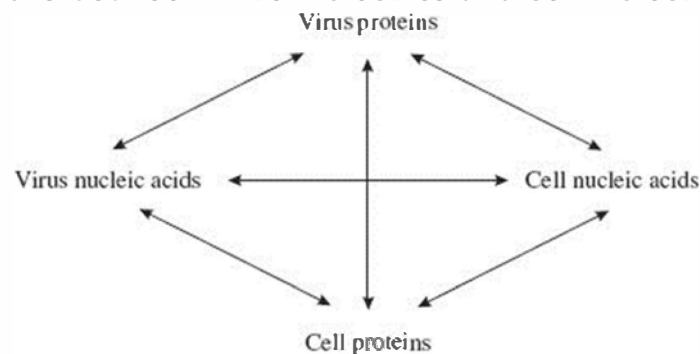
1.4 THE REMAINDER OF THE BOOK

Having outlined the nature of viruses and why they are important, the remainder of the book will examine many aspects of fundamental and applied virology. The early

chapters cover principles, such as the structure of virions, virus replication, and the classification of viruses. There are then ten chapters devoted to reviews of particular groups of viruses, where both principles and applications of virology are covered. Towards the end of the book we consider specific applications of virology, including viral vaccines and anti-viral drugs. The final chapter is devoted to prions, which are not viruses!

It is important to point out that much of virology is concerned with characteristics of the proteins and nucleic acids of viruses, and with interactions between these molecules and the proteins and nucleic acids of cells (Figure 1.4). Most of these interactions rely on specific binding between the molecules. We shall also be discussing cellular structures, and processes such as transcription, translation, and DNA replication. A good background in molecular biology and cell biology is therefore essential; some useful sources of information for plugging any gaps can be found under Sources of further information.

Figure 1.4 Interactions between virus molecules and cell molecules.



LEARNING OUTCOMES

After review of this chapter you should be able to:

- discuss reasons for studying viruses;
- explain how viruses differ from cellular organisms;
- define the term *virus*.

SOURCES OF FURTHER INFORMATION

Cell biology and molecular biology books

Alberts, B. et al. (2008) *Molecular Biology of the Cell*, 5th edition, Garland

Bolsover, S. R. et al. (2011) *Cell Biology: A Short Course*, 3rd edition, Wiley-Blackwell

Karp, G. (2010) *Cell Biology*, 6th edition, Wiley

Lodish, H. F. et al. (2008) *Molecular Cell Biology*, 6th edition, Freeman

Pollard, T. D. and Earnshaw, W. C. (2008) *Cell Biology*, 2nd edition, Saunders

Watson, J. D. et al. (2007) *Molecular Biology of the Gene*, 6th edition, Pearson

Weaver, R. F. (2011) *Molecular Biology*, 5th edition, McGraw-Hill

Historical paper

Hershey, A. D. and Chase, M. (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *Journal of General Physiology*, 36, 39-56

Recent papers

Breitbart, M. and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends In Microbiology*, 13, 278-284

Enquist, L. W. (2009) Virology in the 21st century. *Journal of Virology*, 83, 5296-5308

Hunter, W. et al. (2010) Large-scale field application of RNAi technology reducing Israeli acute paralysis virus disease in honey bees (*Apis mellifera*, Hymenoptera: Apidae). *PLoS Pathogens*, 6 (12): e1001160

Kutateladze, M. and Adamia, R. (2010) Bacteriophages as potential new therapeutics to replace or supplement antibiotics. *Trends in Biotechnology*, 28, 591-595

Roossinck, M. J. (2011) The good viruses: viral mutualistic symbioses. *Nature Reviews Microbiology*, 9, 99-108

Scholthof, K.-B. G. et al. (2011) Top 10 plant viruses in molecular plant pathology. *Molecular Plant Pathology*, 12, 938-954

van Regenmortel, M. H. V. and Mahy, B. W. J. (2004) Emerging issues in virus taxonomy. *Emerging Infectious Diseases*, 10 (1): <http://www.cdc.gov/ncidod/eid/vol10no1/03-0279.htm>

Young, L. S. et al. (2006) Viral gene therapy strategies: from basic science to clinical application. *Journal of Pathology*, 208, 299-318

CHAPTER 2

Methods Used in Virology

CHAPTER 2 AT A GLANCE

1 Plaques formed by influenza virus (large plaques) and canine adenovirus-1 (small plaques) in MDCK cells. From Murata et al. (2011) *Vaccine*, 29, 3155. Reproduced by permission of Elsevier and the authors.

2 Plaques formed by phage MS2 in *Escherichia coli* cells.

3 Separation of reovirus full virions and empty virions in a cesium chloride gradient. From Mendez et al. (2000) *Journal of Virological Methods*, 90, 59. Reproduced by permission of Elsevier and the authors.

4 Fluorescence microscopy of cells expressing a cytomegalovirus protein. From Liu et al. (2012) *Virus Research*, 166, 31. Reproduced by permission of Elsevier.

5 An endosome (labeled with a red fluorophore) containing rabies virus protein (labeled with green fluorescent protein) in the cytoplasm of an infected cell. From Finke and Conzelmann (2005) *Virus Research*, 111, 120. Reproduced by permission of Elsevier Limited and the authors.

6 Negatively stained particles of an orbivirus. The bar represents 50 nm. From Attoui et al. (2005) *Journal of General Virology*, 86, 3409. Reproduced by permission of the Society for General Microbiology and the authors.

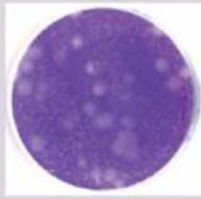
7 Reconstructed images from cryo-electron microscopy: measles virus nucleocapsid (left), echovirus type 12 bound to a fragment of its cell receptor (right). Courtesy of Dr David Bhella (MRC Virology Unit, Glasgow). Reinterpretations of data in Bhella et al. (2004) *Journal of Molecular Biology*, 340, 319 (by permission of Elsevier Limited) and Bhella et al. (2004) *Journal of Biological Chemistry*, 279, 8325 (by permission of The American Society for Biochemistry and Molecular Biology).

8 Bluetongue virus proteins. From Belhouchet et al. (2011) *PLoS ONE*, 6 (10): e25697.

9 Orbivirus RNA segments separated by electrophoresis through an agarose gel. From Attoui et al. (2005) *Journal of General Virology*, 86, 3409. Reproduced by permission of the Society for General Microbiology and the authors.

10 Analysis of varicella-zoster virus transcription using microarrays. From Kennedy et al. (2005) *Journal of General Virology*, 86, 2673. Reproduced by permission of the Society for General Microbiology and the authors.

Virus Isolation and Culture



Animal virus plaques in a cell culture¹



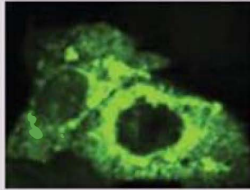
Phage plaques in a lawn of bacterial cells²

Density Gradient Centrifugation



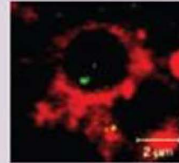
Separation of virus particles in a density gradient³

Fluorescence Microscopy

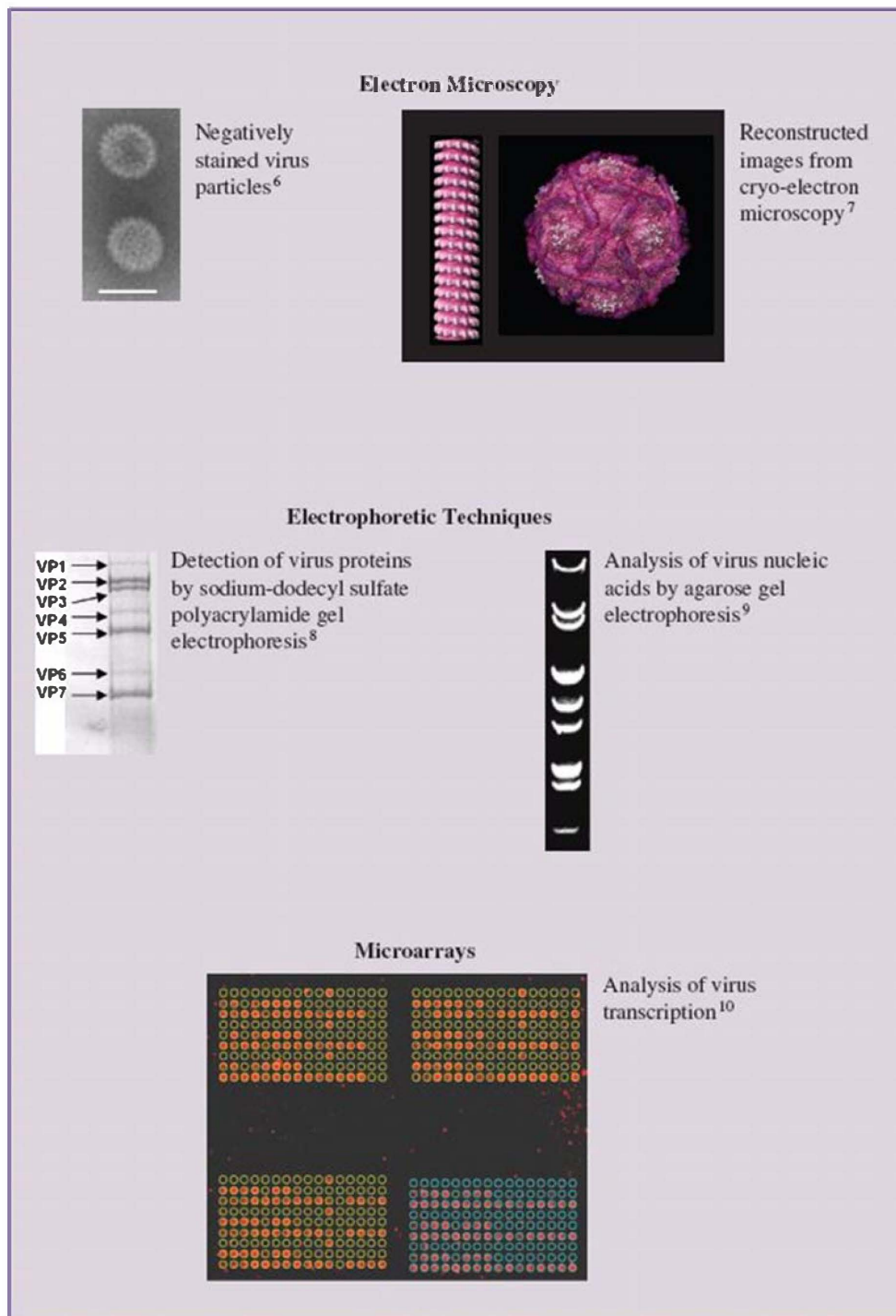


Cells transfected with a plasmid encoding a virus protein and a fluorescent protein⁴

Confocal Microscopy



An endosome (labeled red) containing virus protein (labeled green) in an infected cell⁵



2.1 INTRODUCTION TO METHODS USED IN VIROLOGY

Methods used in virology are introduced early in this book in order to provide an appreciation of the nature of the techniques that have been used to achieve our current level of knowledge and understanding of viruses. Virology is a huge subject and uses a wide range of methods. Many of the techniques of molecular biology and cell biology are used, and constraints on space permit us to mention only some of them. Much of the focus of this chapter is on methods that are unique to

virology. Many of these methods are used not only in virus research but also in the diagnosis of virus diseases of humans, animals, and plants.

Initially this chapter could be skimmed to gain an overview of its contents and thereafter used for reference. Details of the methods outlined here, and of other methods important in virology, fill many volumes, some of which are listed at the end of the chapter.

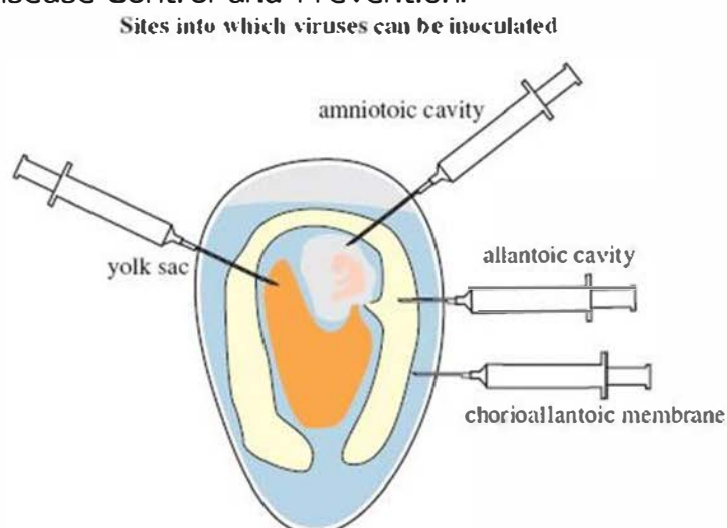
2.2 CULTIVATION OF VIRUSES

Virologists need to be able to produce the objects of their study, so a wide range of procedures has been developed for cultivating viruses. Virus cultivation is also referred to as propagation or growth, all terms borrowed from horticulture! A few techniques have been developed for the cultivation of viruses in cell-free systems, but in the vast majority of cases it is necessary to supply the virus with appropriate cells in which it can replicate.

Phages are supplied with bacterial cultures and plant viruses may be propagated using specially cultivated plants, tissue cultures or protoplasts (plant cells from which the cell wall has been removed). Animal viruses may be propagated in whole organisms, such as mice, which may be genetically modified. Transgenic mice are used in studies of hepatitis B and hepatitis C viruses. Some vertebrate viruses are propagated in eggs containing chick embryos (Figure 2.1), while some insect viruses are propagated in insect larvae. For the most part, however, animal viruses are grown in cultured animal cells.

Figure 2.1 Cultivation of viruses in eggs containing chick embryos.

Source: Photograph courtesy of Taronna Maines and Greg Knobloch, Centers for Disease Control and Prevention.



Inoculation of an egg

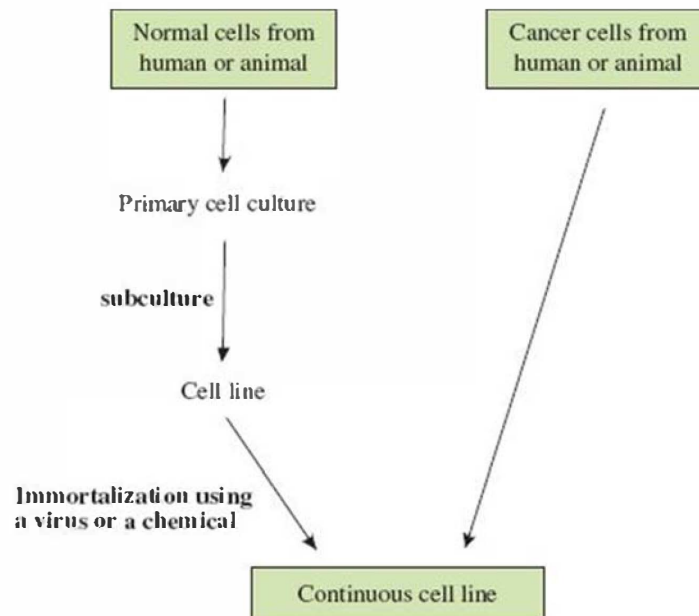


2.2.1 Animal cell culture

Animal cell culture techniques are well developed and most of the cells used are from continuous cell lines derived from humans and other animal species. Continuous cell lines consist of cells that have been immortalized, either in the laboratory or in the body (Figure 2.2); they can be subcultured indefinitely. The

HeLa cell line is a widely used continuous cell line that was initiated in the middle of the twentieth century from cells taken from a cervical carcinoma.

Figure 2.2 Derivation of continuous cell lines of human and animal cells. Most types of cell taken from the body do not grow well in culture. If cells from a primary culture can be subcultured they are growing as a cell line. They can be subcultured only a finite number of times unless they are immortalized, in which case they can be subcultured indefinitely as a continuous cell line. Cancer cells are already immortalized, and continuous cell lines may be established from these without further treatment.



Sometimes it is difficult to find a cell line in which a virus can replicate. For many years no suitable system could be found to support the replication of hepatitis C virus, but eventually the obstacles were overcome and this virus can now be cultured in a human hepatoma cell line and in primary human hepatocytes.

Cells are cultured in media that provide nutrients. Most media are supplemented with animal serum, which contains substances that promote the growth of many cell lines. Other important roles for the medium are the maintenance of optimum osmotic pressure and pH for the cells. Viruses can be cultivated in cells growing on the surface of a variety of plastic vessels (Figure 2.3) with the cells bathed in the growth medium. Most cells grow on a plastic or glass surface as a single layer of cells, known as a monolayer. Alternatively the cells can be suspended in the medium, which is stirred to keep them in suspension.

Figure 2.3 Cell culture flasks, dishes, and plates.

Source: Photographs of TPP cell culture products courtesy of MIDSCI.



Contamination with bacteria and fungi can cause major problems in cell culture work; in order to minimize these problems work is normally done in a sterile cabinet (Figure 2.4) and most media contain antibiotics. Many cell types require a relatively high concentration of carbon dioxide, which can be supplied in a special incubator.

Figure 2.4 Cell culture work. Precautions to avoid contamination include working in a sterile cabinet and wearing gloves and mask.

Source: Courtesy of Sanofi Pasteur ВТ "Copyright.



2.3 ISOLATION OF VIRUSES

Many viruses can be isolated as a result of their ability to form discrete visible zones (plaques) in layers of host cells. If a confluent layer of cells is inoculated with an amount of virus to infect a small proportion of the cells, then plaques may form

where areas of cells are killed or altered by the virus infection. Each plaque is formed when infection spreads radially from an infected cell to surrounding cells. Plaques can be formed by many animal viruses if the cell monolayers are overlaid with agarose gel to maintain the progeny virus in a discrete zone (Figure 2.5). Plaques can also be formed by phages in lawns of bacterial growth (Figure 2.6). Figure 2.5 Method for production of plaques by animal viruses.

Source: Photograph of influenza virus plaques from Ma et al. (2011) *Virology*, 410, 1. Reproduced by permission of Elsevier.

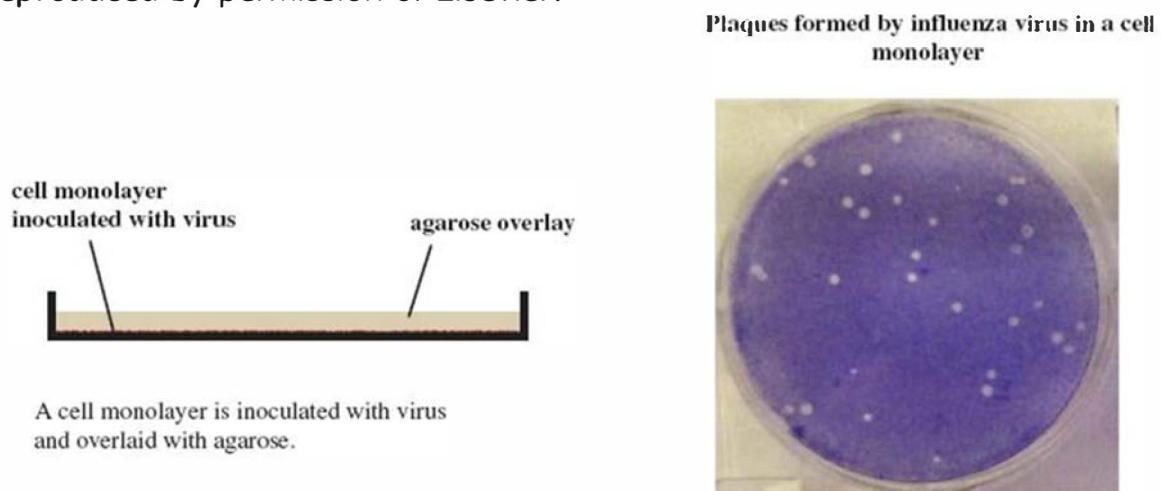


Figure 2.6 Plaques formed by a phage in a bacterial lawn. The control plate on the left was inoculated with only the bacterial host. The plate on the right was inoculated with phage and bacterial host.



It is generally assumed that a plaque is the result of the infection of a cell by a single virion. If this is the case then all virus produced from virus in the plaque should be a clone, in other words it should be genetically identical. This clone can be referred to as an isolate, and if it is distinct from all other isolates it can be referred to as a strain. This is analogous to the derivation of a bacterial strain from a colony on an agar plate.

There is a possibility that a plaque might be derived from two or more virions so, to increase the probability that a genetically pure strain of virus has been obtained, material from a plaque can be inoculated onto further monolayers and virus can be derived from an individual plaque. The virus is said to have been plaque purified.

When a virus is first isolated it may replicate poorly in cells in the laboratory, but after it has gone through a number of replication cycles it may replicate more efficiently. Each time the virus is subcultured (to borrow a term from

bacteriology) it is said to have been passaged. After a number of passages the virus may be genetically different to the original wild strain, in which case it is now a laboratory strain.

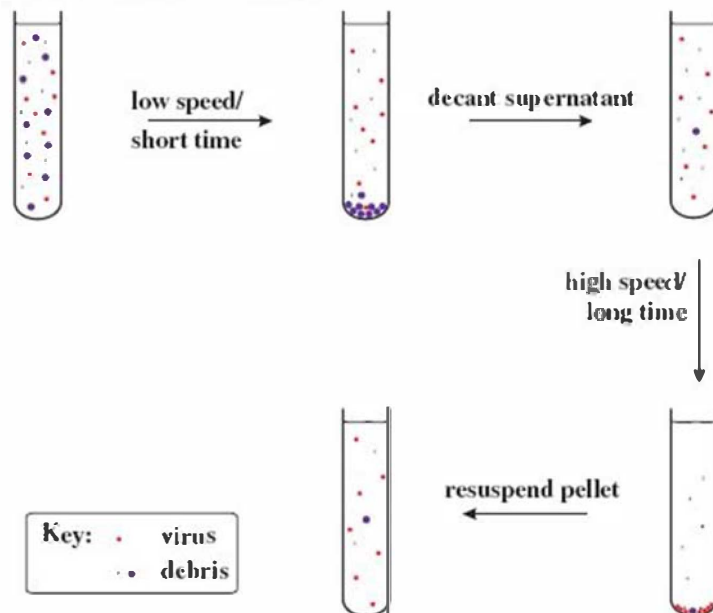
2.4 CENTRIFUGATION

After a virus has been propagated it is usually necessary to remove host cell debris and other contaminants before the virus particles can be used for laboratory studies, for incorporation into a vaccine, or for some other purpose. Many virus purification procedures involve centrifugation; partial purification can be achieved by differential centrifugation and a higher degree of purity can be achieved by some form of density gradient centrifugation.

2.4.1 Differential centrifugation

Differential centrifugation involves alternating low-speed centrifugation, after which most of the virus is still in the supernatant, with high-speed centrifugation, after which the virus is in the pellet (Figure 2.7).

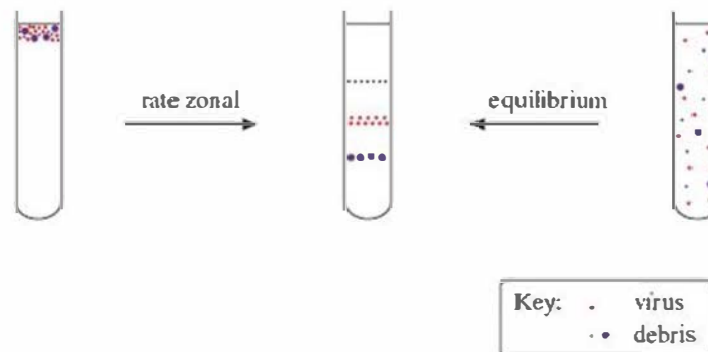
Figure 2.7 Partial purification of virions by differential centrifugation. A crude preparation of virus containing host debris is subjected to low-speed/short-time centrifugation (e.g. 10 000 g/20 minutes) followed by high-speed/long-time centrifugation (e.g. 100 000 g/2 hours). This cycle can be repeated to obtain a higher degree of purity. The final pellet containing partly purified virus is resuspended in a small volume of fluid.



2.4.2 Density gradient centrifugation

Density gradient centrifugation involves centrifuging particles (such as virions) or molecules (such as nucleic acids) in a solution of increasing concentration, and therefore density. The solutes used have high solubility; sucrose is commonly used. There are two major categories of density gradient centrifugation: rate zonal and equilibrium (isopycnic) centrifugation (Figure 2.8).

Figure 2.8 Purification of virions by density gradient centrifugation. A partly purified preparation of virus is further purified in a density gradient. Rate zonal centrifugation involves layering the preparation on top of a pre-formed gradient. Equilibrium centrifugation can often be done starting with a suspension of the impure virus in a solution of the gradient material; the gradient is formed during centrifugation.



In rate zonal centrifugation a particle moves through the gradient at a rate determined by its sedimentation coefficient, a value that depends principally on its size. Homogeneous particles, such as identical virions, should move as a sharp band that can be harvested after the band has moved part-way through the gradient.

In equilibrium centrifugation a concentration of solute is selected to ensure that the density at the bottom of the gradient is greater than that of the particles/molecules to be purified. A particle/molecule suspended in the gradient moves to a point where the gradient density is the same as its own density. This technique enables the determination of the buoyant densities of nucleic acids and of virions. Buoyant densities of virions determined in gradients of cesium chloride are used as criteria in the characterization of viruses.

2.5 STRUCTURAL INVESTIGATIONS OF CELLS AND VIRIONS

2.5.1 Light microscopy

The sizes of most virions are beyond the limits of resolution of light microscopes, but light microscopy has useful applications in detecting and studying virus-induced changes in cells, for example observing cytopathic effects (Section 2.7. 2). Fluorescence microscopy is used to detect virus products labeled with fluorophores (molecules that absorb light at certain wavelengths then emit light at longer wavelengths; Section 2.7. 3).

Confocal microscopy is proving to be valuable in virology. The principle of this technique is the use of a pinhole to exclude light from out-of-focus regions of the specimen. Most confocal microscopes scan the specimen with a laser, producing exceptionally clear images of thick specimens and of fluorescing specimens. Furthermore, optical slices of a specimen can be collected and used to

create a three-dimensional representation. The techniques can be used with live cells and can be applied to investigations of protein trafficking, with the virus or cell protein under investigation carrying a suitable label, e.g. green fluorescent protein (a jellyfish protein).

2.5.2 Electron microscopy

Many investigations of the structure of virions or of virus-infected cells involve observations at large magnifications using transmission electron microscopes. The specimen may be a preparation of virions, a virion component, virus-infected cells or an ultra-thin section of a virus-infected cell. In order to observe detail in the specimen it is either negatively stained, or cooled to a very low temperature, or both.

Negative staining techniques generate contrast by using heavy metal-containing compounds, such as potassium phosphotungstate and ammonium molybdate. In electron micrographs of virions the stains appear as dark areas around the virions, allowing the overall virion shape and size to be determined. Further structural detail may be apparent if the stains penetrate any crevices on the virion surface or any hollows within the virion. Negative staining techniques have generated many high quality electron micrographs, but the traditional techniques have limitations, including structural distortions resulting from drying.

Cryo-electron microscopy techniques are more recent. For these techniques wet specimens are rapidly cooled to a temperature below -160°C , freezing the water as a glass-like material. The images are recorded while the specimen is frozen. They require computer processing in order to extract maximum detail, and data from multiple images are processed to reconstruct three-dimensional images of virus particles. This may involve averaging many identical copies or combining images into three-dimensional density maps (tomograms).

2.5.3 X-ray crystallography

X-ray crystallography is another technique that is revealing detailed information about the three-dimensional structures of virions (and DNA, proteins, and DNA-protein complexes). This technique requires the production of a crystal of the virions or molecules under study. The crystal is placed in a beam of X-rays, which are diffracted by repeating arrangements of molecules/atoms in the crystal. Analysis of the diffraction pattern allows the relative positions of these molecules/atoms to be determined.

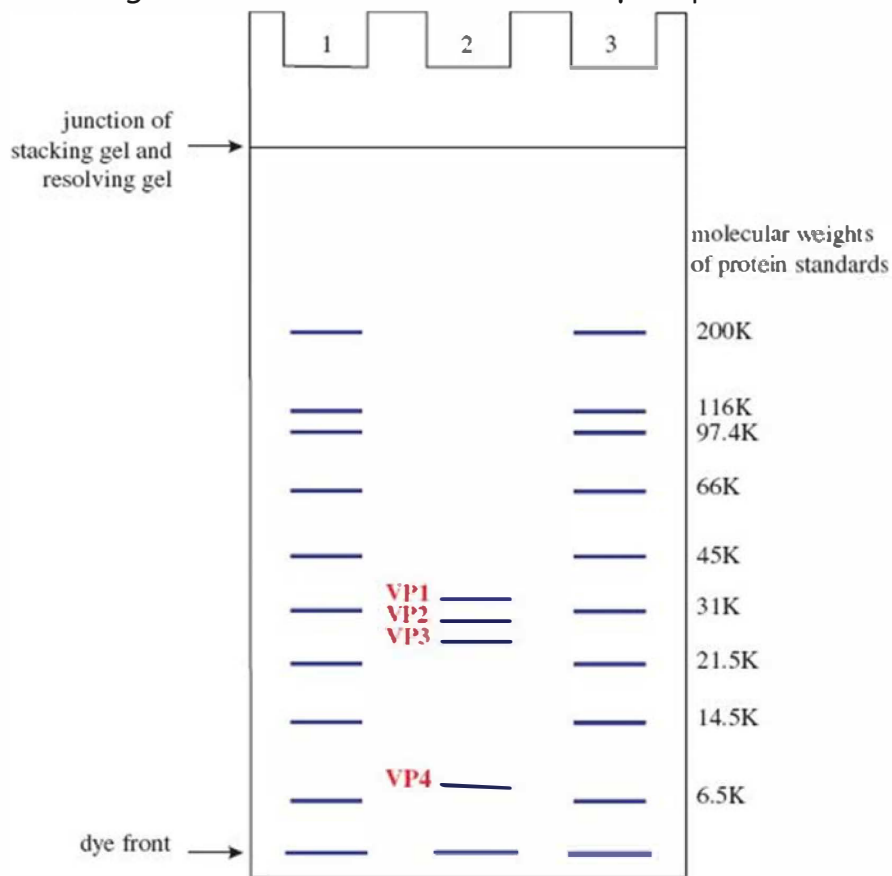
Other techniques that are providing useful information about the structure of viruses are nuclear magnetic resonance and atomic force microscopy.

2.6 ELECTROPHORETIC TECHNIQUES

Mixtures of proteins or nucleic acids can be separated by electrophoresis in a gel composed of agarose or polyacrylamide. In most electrophoretic techniques, each

protein or nucleic acid forms a band in the gel. Electrophoresis can be performed under conditions where the rate of movement through the gel depends on molecular weight. The molecular weights of the protein or nucleic acid molecules can be estimated by comparing the positions of the bands with positions of bands formed by molecules of known molecular weight electrophoresed in the same gel. The technique for estimating molecular weights of proteins is polyacrylamide gel electrophoresis in the presence of the detergent sodium dodecyl sulfate (SDS-PAGE; Figure 2.9).

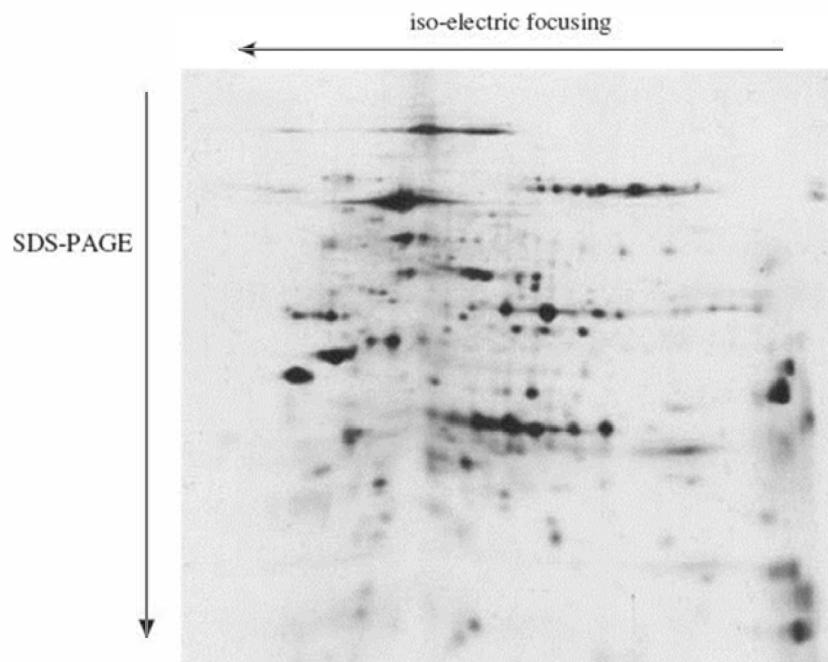
Figure 2.9 Diagram illustrating separation of proteins and estimation of their molecular weights using SDS-PAGE. Lanes 1 and 3 contain standard proteins of known molecular weight. Lane 2 contains the four capsid proteins of a picornavirus.



Complex mixtures of virus proteins can be separated by two-dimensional electrophoresis, where the proteins are first separated by their iso-electric points (iso-electric focusing) and then in a second dimension by their molecular weights (SDS-PAGE) (Figure 2.10).

Figure 2.10 Separation of mimivirus proteins by two-dimensional electrophoresis. The proteins were first separated by iso-electric focusing, and then by SDS-PAGE.

Source: Raoult et al. (2007) *Clinical Infectious Diseases*, 45, 95, with the permission of Oxford University Press and the authors.



The patterns of nucleic acids and proteins after electrophoretic separation may be immobilized by transfer (blotting) onto a membrane. If the molecules are DNA the technique is known as Southern blotting, named after Edwin Southern; if the molecules are RNA the technique is known as northern blotting, and if the molecules are protein the technique is known as western blotting!

2.7 DETECTION OF VIRUSES AND VIRUS COMPONENTS

A wide range of techniques has been developed for the detection of viruses and virus components and many of them are used in laboratories involved with diagnosis of virus diseases. The techniques can be arranged in four categories: detection of (1) virions, (2) virus infectivity, (3) virus antigens, and (4) virus nucleic acids.

2.7.1 Detection of virions

Specimens can be negatively stained (Section 2.5.2) and examined in an electron microscope for the presence of virions. Limitations to this approach are the high costs of the equipment and limited sensitivity; the minimum detectable concentration of virions is about 10^6 /ml. An example of an application of this technique is the examination of feces from a patient with gastroenteritis for the presence of rotavirus particles.

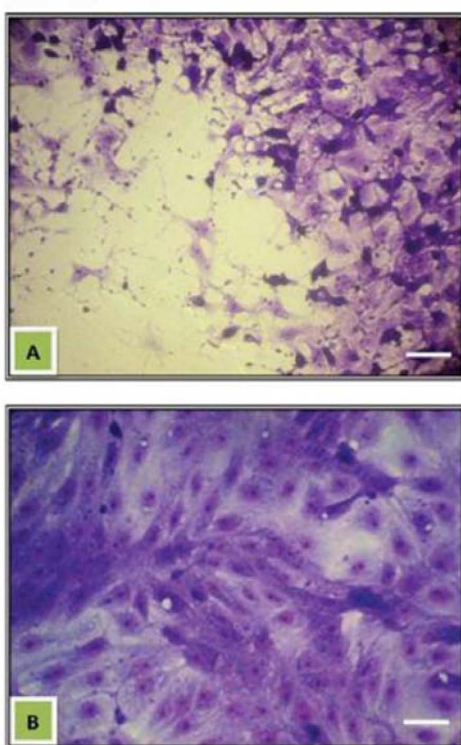
2.7.2 Detection of infectivity using cell cultures

Not all virions have the ability to replicate in host cells. Those virions that do have this ability are said to be *infective*, and the term *infectivity* is used to denote the capacity of a virus to replicate. Virions may be non-infective because they lack part of the genome or because they have been damaged.

To determine whether a sample or a specimen contains infective virus it can be inoculated into a culture of cells, or a host organism, known to support the replication of the virus suspected of being present. After incubation of an inoculated cell culture at an appropriate temperature it can be examined by light microscopy for characteristic changes in the appearance of the cells resulting from virus-induced damage. A change of this type is known as a cytopathic effect (CPE); an example of CPE induced by bluetongue virus can be seen in Figure 2.11

Figure 2.11 Cytopathic effect caused by replication of bluetongue virus in Vero (monkey kidney) cells. (A) bluetongue virus-infected cells. (B) non-infected Vero cells. Scale bar: 90 Ojm.

Source: Ali Al Ahmad et al. (2011) *Theriogenology*, 76, 126. Reproduced by permission of Elsevier and the authors.



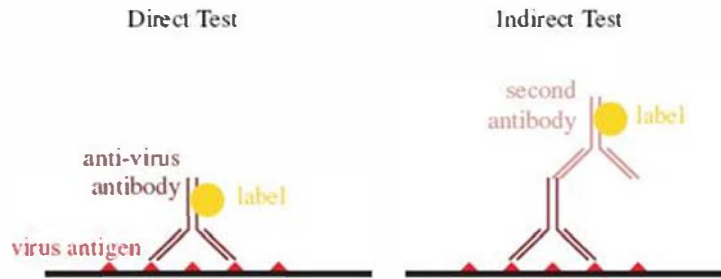
The quantity of infective virus in a specimen or a preparation can be determined (Section 2.8).

2.7.3 Detection of virus antigens

Virus antigens can be detected using virus-specific antisera or monoclonal antibodies. In most techniques positive results are indicated by detecting the presence of a label, which may be attached either to the anti-virus antibody (direct tests) or to a second antibody (indirect tests) (Figure 2.12). The anti-virus antibody is produced by injecting virus antigen into one animal species and the second antibody is produced by injecting immunoglobulin from the first animal species into a second animal species.

Figure 2.12 Principles of tests to detect virus antigens. The specimen is treated with anti-virus antibody. In direct tests the anti-virus antibody is labeled. In indirect tests

a labeled second antibody detects any anti-virus antibody that has bound to antigen.



Antibodies can have many types of label attached and the labels can be detected using a range of methods. Some types of label and some methods for detecting them are listed in Table 2.1.

Table 2.1 Molecules used to label antibodies (and nucleic acids) and techniques used to detect them

Label	Detection technique
Enzyme	Enzyme-linked immunosorbent assay
Fluorescent	Fluorescence microscopy
	Fluorimetry
Gold	Electron microscopy
Radioactive	Autoradiography

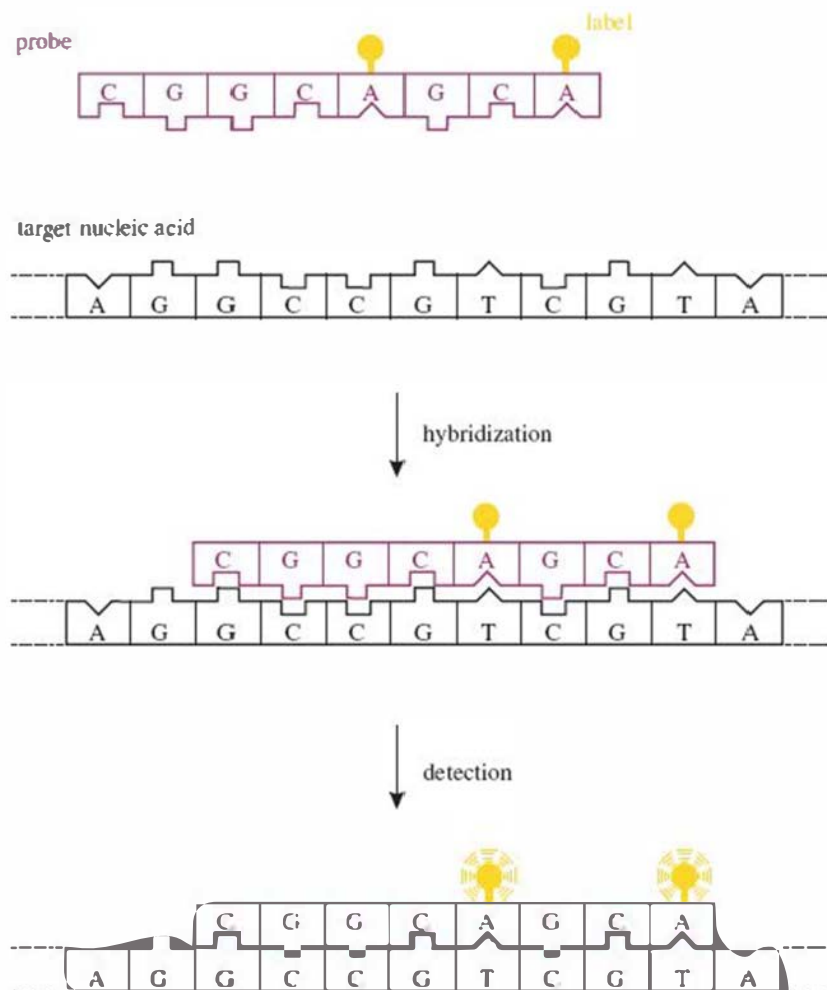
2.7.4 Detection of virus nucleic acids

Techniques for the detection of virus genomes and other virus-specific nucleic acids are widely used in virus research and in diagnostic virology. Some viruses were discovered, not by detecting virions or infective virus, but by detecting virus nucleic acids in infected hosts. Examples of viruses discovered in this way include Kaposi's sarcoma-associated herpesvirus (Section 11.2.7), Merkel cell polyomavirus (Section 23.3), and torque teno virus.

2.7.4.a Hybridization

Virus genomes or virus messenger RNAs (mRNAs) may be detected using sequence-specific DNA probes carrying appropriate labels (Figure 2.13). Some of the labels that are used for antibody detection can be used to label the probes (Table 2.1).

Figure 2.13 Detection of a specific target nucleic acid (DNA or RNA) using a labeled DNA probe.



Hybridization may take place on the surface of a membrane after Southern blotting (DNA) or northern blotting (RNA) (Section 2.6). Thin sections of tissue may be probed for the presence of specific nucleic acids, in which case the technique is known as in situ hybridization.

The principles of nucleic acid hybridization are also applied in DNA microarrays. A DNA microarray consists of a substrate, such as a glass slide, with hundreds, thousands, or even millions of small spots of DNA attached. Each spot contains a specific DNA probe and specific DNA or RNA molecules in a sample are detected by hybridization to specific spots.

2.7.4.b Polymerase chain reaction (PCR)

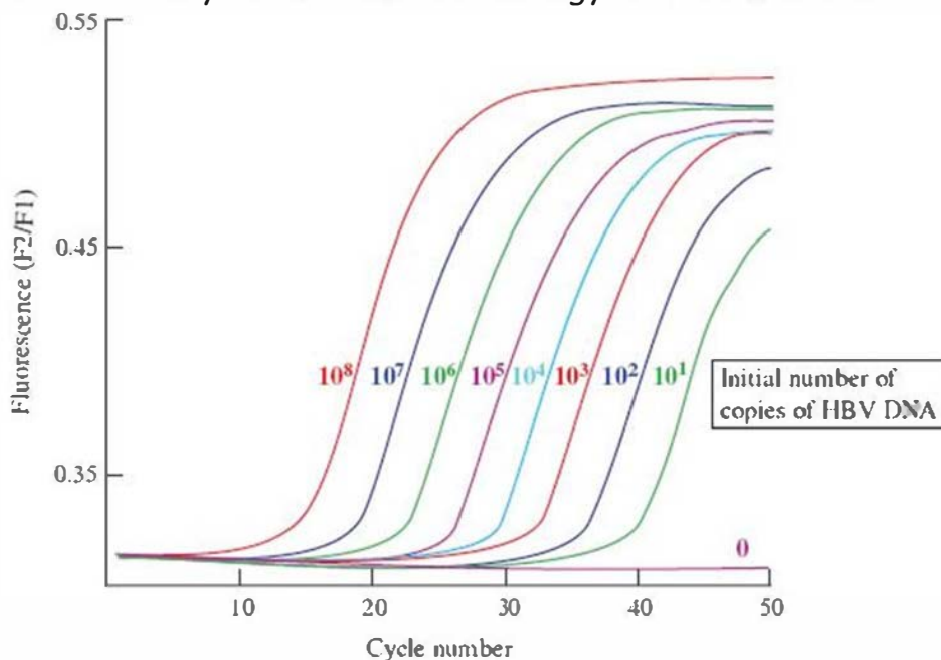
When a sample is likely to contain a low number of copies of a virus nucleic acid, the probability of detection can be increased by amplification. DNA can be amplified using a PCR, while RNA can be copied to DNA using a reverse transcriptase (RT) and then amplified (RT-PCR). The procedures require oligonucleotide primers specific to viral sequences. An amplified product can be detected by electrophoresis in an agarose gel, followed by transfer to a nitrocellulose membrane, which is incubated with a labeled probe.

There are also PCR techniques available for determining the number of copies of a specific nucleic acid in a sample. Real-time PCR is commonly used for this purpose. In this technique the increase in DNA concentration during the PCR is monitored

using fluorescent labels; the larger the initial copy number of DNA, the sooner a significant increase in fluorescence is observed. The PCR cycle at which the fluorescent signal passes a defined threshold is determined and this gives an estimate of the starting copy number. An example is given in Figure 2.14.

Figure 2.14 Quantification of hepatitis B virus (HBV) DNA by real-time PCR. When there are 10^8 copies of HBV DNA in the sample, fluorescence starts to increase at an early cycle number. The increase starts at progressively later cycles with decreasing numbers of HBV DNA copies.

Source: Data from Ho et al. (2003) *Journal of Medical Microbiology*, 52, 397; by permission of the Society for General Microbiology and the authors.



2.8 INFECTIVITY ASSAYS

An infectivity assay measures the titer (the concentration) of infective virus in a specimen or a preparation. Samples are inoculated into suitable hosts, in which a response can be observed if infective virus is present. Suitable hosts might be animals, plants or cultures of bacterial, plant or animal cells. Infectivity assays fall into two classes: quantitative and quantal.

2.8.1 Quantitative assays

Quantitative assays are those in which each host response can be any one of a series of values, such as number of plaques (Section 2.3). A plaque assay can be carried out with any virus that can form plaques, giving an estimate of the concentration of infective virus in plaque-forming units (pfu).

The assay is done using cells grown in Petri dishes, or other appropriate containers, under conditions that will allow the formation of plaques. Cells are inoculated with standard volumes of virus dilutions. After incubation, dishes that received high dilutions of virus may have no plaques or very few plaques, while dishes that received low dilutions of virus may have very large numbers of plaques or all the

cells may have lysed. Dishes are selected that have plaque numbers within a certain range, e.g. 30-300. The plaques are counted and the concentration of virus in the sample (pfu/ml) is calculated.

2.8.2 Quantal assays

In a quantal assay each inoculated subject either responds or it does not; for example, an inoculated cell culture either develops a CPE or it does not; an inoculated animal either dies or it remains healthy. The aim of the assay is to find the virus dose that produces a response in 50% of inoculated subjects. In cell culture assays this dose is known as the TCID₅₀ (the dose that infects 50% of inoculated tissue cultures). In animal assays the dose is known as the ID₅₀ (the dose that infects 50% of inoculated animals), or where the virus infection kills the animal the dose is known as the LD₅₀ (the dose that is lethal for 50% of inoculated animals).

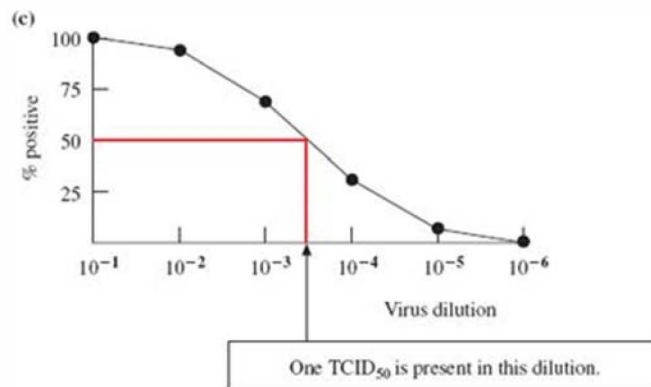
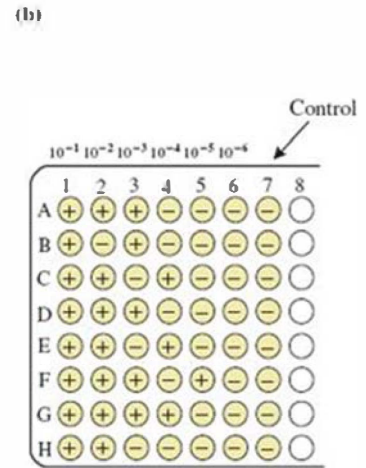
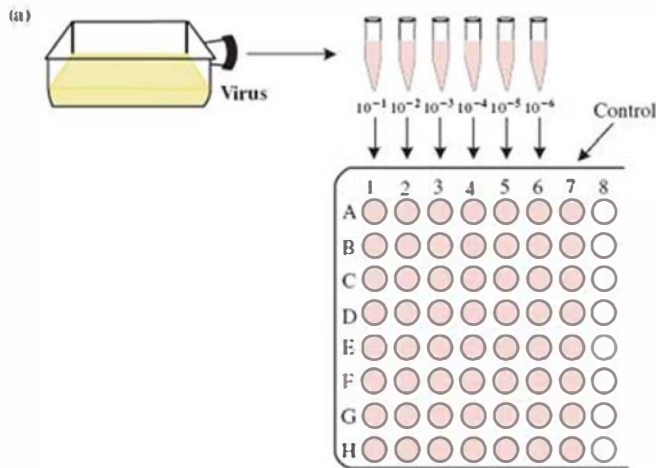
The outcome of a TCID₅₀ determination can be used to estimate a virus titer in pfu, or vice versa, using the formula

$$1 \text{ TCID}_{50} = 0.7 \text{ pfu}$$

This formula applies only if the method used to visualize plaques does not alter the yield of virus.

An example of a TCID₅₀ determination is depicted in Figure 2.15.

Figure 2.15 Example of a TCID₅₀ assay. (a) Tenfold dilutions of the virus were inoculated into cell cultures grown in the wells of a plate. Each well received 1 ml virus suspension, except for the control wells, which each received 1 ml diluent. (b) After incubation the cell culture in each well was scored + or - for CPE. (c) The results of the assay.



It is clear that the dilution of virus that contains one TCID₅₀ lies between 10^{-3} and 10^{-4} . Several methods have been developed for estimating this end point; one of these, the Reed-Muench method, is shown in the box.

THE REED-MUENCH METHOD FOR ESTIMATING TCID₅₀

Virus dilution	Infections per number inoculated	Observed values		Cumulative values ¹		Infection ratio ²	% infection ³
		Positive	Negative	Positive	Negative		
10 ⁻¹	8/8	8	0	24	0	24/24	100
10 ⁻²	7/8	7	1	16	1	16/17	94
10 ⁻³	5/8	5	3	9	4	9/13	69
10 ⁻⁴	3/8	3	5	4	9	4/13	31
10 ⁻⁵	1/8	1	7	1	16	1/17	6
10 ⁻⁶	0/8	0	8	0	24	0/24	0

¹The cumulative values are derived by adding up the observed values in the direction of the arrows.

²The infection ratio is the number of positives for the cumulative value out of the total for the cumulative value.

³The % infection is the infection ratio converted to a percentage.

It had already been determined that the dilution of virus that contains one TCID₅₀ lies between 10⁻³ and 10⁻⁴, so the end point can be expressed as 10^{-(3+x)}, where x is the value to be estimated.

$$\begin{aligned}
 x &= \log_{10} \text{ dilution factor} \left(\frac{\% \text{ infection at next dilution above } 50\% - 50}{\% \text{ infection at next dilution above } 50\% - \% \text{ infection at next dilution below } 50\%} \right) \\
 &= 1 \left(\frac{69 - 50}{69 - 31} \right) \\
 &= 0.5
 \end{aligned}$$

End point = 10^{-(3+0.5)} = 10^{-3.5}.

i.e. 1 ml of a 10^{-3.5} dilution contains one TCID₅₀ of virus.

i.e. 1 ml of a 1/3200 dilution contains one TCID₅₀ of virus.

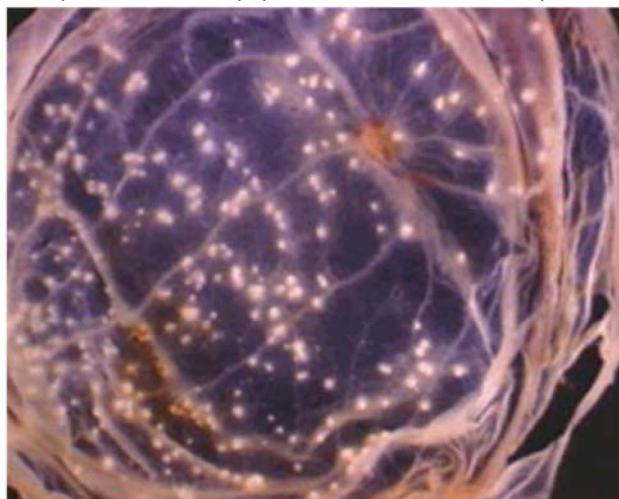
(3.2 is the antilogarithm of 0.5.)

The concentration of virus in the undiluted suspension is 3.2 × 10³ TCID₅₀/ml.

A similar approach can be used to assay the infectivity of any virus that causes the formation of discrete lesions. A number of animal viruses, including smallpox virus and other poxviruses, can cause the formation of pocks after inoculation onto the chorioallantoic membrane of a chick embryo (Figures 2.1 and 2.16). Similarly, many plant viruses cause the formation of lesions after they have been rubbed on the surface of leaves, though the titration is not directly equivalent to a plaque assay as the number of sites on a leaf through which a virus can enter is limited.

Figure 2.16 Virus-induced pocks on the chorioallantoic membrane of a chick embryo.

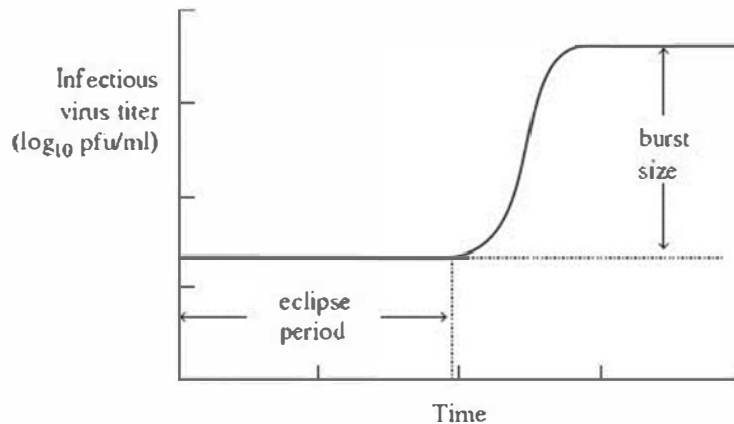
Source: Buxton and Fraser (1977) *Animal Microbiology*, Volume 2, Blackwell Scientific Publications. Reproduced by permission of the publisher.



2.8.3 One-step growth curve

A common type of experiment that involves assaying virus infectivity is a single-burst experiment, which provides data for a one-step growth curve (Figure 2.17). The first one-step growth curves produced were for phages, but they have since been produced for other viruses.

Figure 2.17 One-step growth curve. The eclipse period is the period during which no intracellular infectious virus can be recovered; infectious nucleic acid might be recoverable in some cases. The burst size is the average yield of infectious virus per cell.



For a period (the eclipse period), the titer of infective virus remains constant. During the eclipse period the virus is replicating, but no virus has been released from infected cells, so each infected cell gives rise to one plaque. As infected cells begin to die and release virus the titer begins to rise, and when all the infected cells have lysed the titer levels off.

The procedure for deriving a one-step growth curve is now outlined.

- A suspension of virions is mixed with a suspension of host cells. It is necessary to ensure that, as far as possible, all the cells are infected simultaneously, so the number of virions must greatly exceed the number of cells. The ratio of virions to cells is known as the multiplicity of infection (m.o.i.). In a one-step growth experiment an m.o.i. between 5 and 10 is commonly used.
- Adsorption of virions to cells is allowed to proceed for a suitable period (e.g. 2 minutes), and then adsorption is stopped by, for example, greatly diluting the mixture and/or by adding anti-virus antiserum.
- A sample is taken from the suspension at intervals until lysis is complete. For phages the experiment is usually complete in less than one hour. For animal viruses, however, the timescale is measured in hours or days, reflecting the much slower growth rates of eukaryotic cells and the longer times taken for replication of animal viruses.
- A plaque assay is carried out on each sample.
- The one-step growth curve is obtained by plotting the logarithm of the virus titer against time (Figure 2.17).

This type of experiment provides valuable information about the replication cycle of a virus in a particular host cell system. The average yield of infectious virus per cell (the burst size) can be calculated from the formula

$$\text{burst size} = \frac{\text{final virus titer}}{\text{initial virus titer}}$$

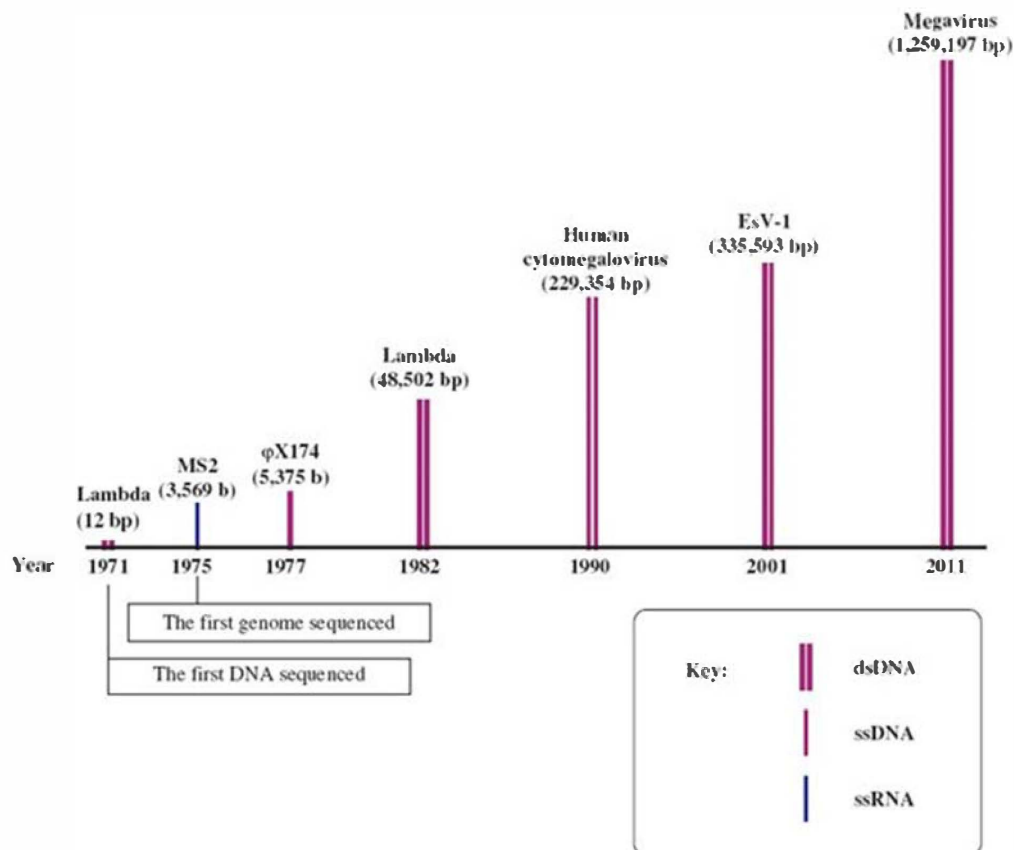
It should be noted that the burst size is an average of what may be a significant variable as there is usually much variation in burst size between cells. One factor that may affect the burst size is the physiological state of the cells. In one study with a phage, a burst size of 170 was obtained when growing bacteria were used, while a value of 20 was obtained with resting bacteria.

2.9 VIRUS GENETICS

2.9.1 Genome sequencing

A pioneering technique for determining the sequence of bases in a DNA molecule was the dideoxy chain termination method developed by Fred Sanger and his colleagues at Cambridge. The sequence of the genome of phage Π X174 DNA was determined by this group six years after the first determination of a DNA sequence, which itself was from a virus: 12 base pairs (bp) of phage lambda DNA (Figure 2.18). Newer sequencing technologies enable DNA sequences to be read in a fraction of the time taken by the Sanger method.

Figure 2.18 Time line depicting some landmarks in viral genome sequencing. Lambda, MS2, and Π X174 are phages. The host of EsV-1 is the marine filamentous brown alga *Ectocarpus siliculosus*. The host of megavirus is an amoeba. The first DNA ever sequenced was 12 bp of the genome of phage lambda. Eleven years later the complete genome sequence of this virus was published. Not to scale.



Computer programs, such as Artemis and BLAST, are essential to analyze the sequences; much useful information can be yielded.

- Open reading frames (ORFs) can be found. The ORFs can be translated into the amino acid sequences of the virus proteins, and these may allow characteristics and functions of the proteins to be deduced. Hydrophobic sequences may indicate membrane association; the sequence Gly β T β Asp β T β Asp may indicate RNA-dependent RNA polymerase activity.
- Sequences that regulate gene expression, e.g. promoters and enhancers, can be identified.
- During an outbreak of virus disease, such as foot and mouth disease or influenza, sequences of virus isolates can be compared. This can provide useful epidemiological information such as the source of the virus responsible and whether more than one strain is responsible for the outbreak.
- Phylogenetic trees can be constructed (Section 10.2.1).

2.9.2 Genome manipulation

The wide range of techniques available for the manipulation of nucleic acids can be applied to virus genomes. These techniques include the isolation of specific fragments of genomes using restriction endonucleases, the cloning of fragments in bacterial plasmids, and the introduction of site-specific mutations into virus genomes. The natural processes of recombination and reassortment (see Chapter 21) that produce new combinations of virus genes can be harnessed to produce new viral genotypes in the laboratory.

2.9.3 Investigation of gene function and expression

The function of a gene may be deduced if its expression is blocked, and many studies of gene function involve the creation of a virus with a mutated gene. Many of the mutants are unable to replicate under conditions in which the wild-type virus replicates, and are known as conditional lethal mutants. Most studies have used temperature-sensitive mutants, which are unable to replicate at temperatures at which the wild-type virus replicates (non-permissive temperatures), but are able to replicate at other temperatures (permissive temperatures).

Techniques for introducing mutations into DNA are well established and have been applied to the study of DNA viruses for some time. More recently, techniques for manipulating the genomes of RNA viruses have been developed. These techniques are referred to as reverse genetics, and they involve the reverse transcription of the RNA genome to DNA, mutation of the DNA, then transcription back to RNA. Reverse genetics are discussed further in Chapters 14 and 15.

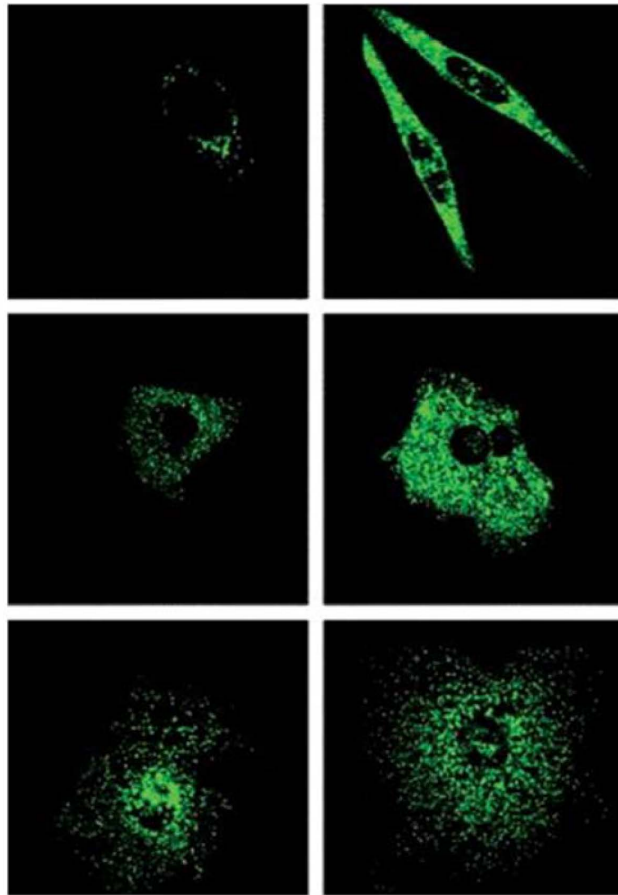
Another way in which the expression of a gene can be blocked is by RNA silencing, or RNA interference (RNAi). This approach exploits a cell defense mechanism (described in Section 9.2.3) that can destroy virus mRNAs after induction by specific dsRNA molecules. Short sequences of dsRNA can be used to inhibit the expression of virus genes, and hence to investigate the functions of those genes.

Yet another approach to the investigation of gene function involves the introduction of the gene into cells in which it is transiently expressed, and then monitoring any changes in the cells.

As well as the production of virus mutants there are other ways in which viruses can be genetically modified in order to aid their study. A sequence encoding a tag, such as green fluorescent protein, can be appended to the gene for a virus protein. When the gene is expressed the virus protein plus the tag is synthesized, enabling the distribution of the protein in infected cells to be monitored (Figure 2.19).

Figure 2.19 Cells expressing a hepatitis C virus protein tagged with green fluorescent protein. The distribution of the fluorescence indicates the distribution of the labeled virus protein.

Source: Realdon et al. (2004) *Journal of Hepatology*, 40, 77. Reproduced by permission of Elsevier.



The expression of virus genes in infected cells can be monitored with the aid of DNA microarrays (Section 2.7.4.a). Microarray technology allows investigators to monitor the expression of hundreds or thousands of genes, and is thus ideal for studies of viruses with large genomes such as the herpesviruses and the large DNA phages. Virus mRNAs are detected by copying RNA from infected cells to DNA using a reverse transcriptase, amplifying the DNA with a PCR, labeling it with a fluorophore, and adding it to a microarray. The probes that bind DNA from the sample are detected by scanning with a laser at a wavelength that excites the fluorophore.

DNA microarrays can also be used to investigate the expression of host cell genes that are relevant to viruses. Cell surface molecules used by viruses as receptors have been identified using microarrays.

2.10 INVESTIGATION OF PROTEIN-PROTEIN INTERACTIONS

During the replication of a virus there are multiple interactions between different virus proteins and between virus proteins and cell proteins. Several techniques are available to help unravel these interactions by determining whether two proteins bind. One technique commonly used in virology is the yeast two-hybrid system, where the two proteins of interest are fused to separate domains of a transcription factor. If the two proteins bind, transcription factor activity is restored, initiating

transcription of a reporter gene. Pull-down assays are also used to investigate whether two proteins interact.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- outline methods for:
 - cultivation of viruses;
 - purification of viruses;
 - detection of viruses and their components;
 - assay of virus infectivity;
 - investigation of virus gene function;
- assess the value of virus genome sequencing.

SOURCES OF FURTHER INFORMATION

Books

Brown, T. (2010) *Gene Cloning and DNA Analysis: An Introduction*, 6th edition, Wiley-Blackwell

Coleman, W. B. and Tsongalis, G. J. (2006) *Molecular Diagnostics for the Clinical Laboratorian*, 2nd edition, Humana

Davis, J. editor, (2011) *Animal Cell Culture: Essential Methods*, Wiley-Blackwell

McPherson M. J. and Møller, S. G. (2006) *PCR*, 2nd edition, Taylor & Francis

Schena, M., editor (2008) *DNA Microarrays*, Scion

Historical paper

Ellis, E. L. and Delbruck, M. (1939) The growth of bacteriophage. *Journal of General Physiology*, 22, 365–384

Recent papers

Beerenwinkel, N. and Zagordi, O. (2011) Ultra-deep sequencing for the analysis of viral populations. *Current Opinion in Virology*, 1, 413–418

De Carlo, S. and Harris, J. R. (2011) Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron*, 42, 117–131

Friedel, C. C. and Haas, J. (2011) Virus–host interactomes and global models of virus-infected cells. *Trends in Microbiology*, 19, 501–508

Grigorieff, N. and Harrison, S. C. (2011) Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Current Opinion in Structural Biology*, 21, 265–273

Kuznetsov, Y. G. and McPherson, A. (2011) Atomic force microscopy in imaging of viruses and virus-infected cells. *Microbiology and Molecular Biology Reviews*, 75, 268–285

Leland, D. S. and Ginocchio, C. C. (2007) Role of cell culture for virus detection in the age of technology. *Clinical Microbiology Reviews*, 20, 49–78

Miller, M. B. and Tang, Y.-W. (2009) Basic concepts of microarrays and potential applications in clinical microbiology. *Clinical Microbiology Reviews*, 22, 611-633





Ratcliff, R. M. et al. (2007) Molecular diagnosis of medical viruses. *Current Issues in Molecular Biology*, 9, 87-102

Sivaraman, D. et al. (2011) Detecting RNA viruses in living mammalian cells by fluorescence microscopy. Trends in Biotechnology, 29, 307-313

CHAPTER 3

Virus Structure

CHAPTER 3 AT A GLANCE

Main types of virion structure		Genomes			
		dsDNA	ssDNA	dsRNA	ssRNA
Icosahedral, naked		✓	✓	✓	✓
Icosahedral, enveloped		✓		✓	✓
Helical, naked		✓	✓		✓
Helical, enveloped					✓

3.1 INTRODUCTION TO VIRUS STRUCTURE

Outside their host cells, viruses survive as virus particles, also known as virions. The virion is a gene delivery system; it contains the virus genome, and its functions are to protect the genome and to aid its entry into a host cell, where it can be replicated and packaged into new virions. The genome is packaged in a protein structure known as a capsid.

Many viruses also have a lipid component, generally present at the surface of the virion, forming an envelope which also contains proteins that play a role in aiding entry into host cells. A few viruses form protective protein occlusion bodies around their virions. Before looking at these virus structures we shall consider characteristics of the nucleic acid and protein molecules that are the main components of virions.

3.2 VIRUS GENOMES

A virion contains the genome of a virus in the form of one or more molecules of nucleic acid. For any one virus the genome is composed of either RNA or DNA. If a new virus is isolated, one way to determine whether it is an RNA virus or a DNA virus is to test its susceptibility to a ribonuclease and a deoxyribonuclease. The virus nucleic acid will be susceptible to degradation by only one of these enzymes.

Each nucleic acid molecule is either single-stranded (ss) or double-stranded (ds), giving four categories of virus genome: dsDNA, ssDNA, dsRNA, and ssRNA. The dsDNA viruses encode their genes in the same kind of molecule as animals, plants, bacteria, and other cellular organisms, while the other three types of genome are unique to viruses. It is interesting to note that most fungal viruses have dsRNA genomes, most plant viruses have ssRNA genomes, and most prokaryotic viruses have dsDNA genomes. The reasons for these distributions presumably concern diverse origins of the viruses in these very different host types.

A further categorization of a virus nucleic acid can be made on the basis of whether the molecule is linear, with free 5' and 3' ends, or circular, as a result of the strand(s) being covalently closed. Examples of each category are given in Figure 3.1. In this figure, and indeed throughout the book, molecules of DNA and RNA are color coded.








Figure 3.1 Linear and circular viral genomes.

ss: single-stranded

ds: double-stranded

Dark blue and light blue depict (+) RNA and (-) RNA respectively; these terms are explained in Section 6.2.

There are no viruses known with circular dsRNA genomes.

DNA genomes		Examples
	ss, linear	Parvoviruses
	ds, linear	Poxxviruses
	ss, circular	Phage φX174
	ds, circular	Baculoviruses
RNA genomes		
	ss, linear	Tobacco mosaic virus
	ds, linear	Reoviruses
	ss, circular	Hepatitis delta virus

It should be noted that some linear molecules may be in a circular conformation as a result of base pairing between complementary sequences at their ends (see Figure 3.7). This applies, for example, to the DNA in hepadnavirus virions and to the RNA in influenza virions.

Figure 3.7 Terminal repeats in virus genomes.

1 DTR: direct terminal repeat.

ITR: inverted terminal repeat.






X and x represent complementary sequences.

Y and y represent complementary sequences.


ssRNA (+) has the same sequence as the virus mRNA.

ssRNA (вЂ“) has the sequence complementary to the virus mRNA.

The RNAs of ssRNA viruses with ITRs can circularize; a вЂ“panhandleвЂ“ is formed by base pairing between the complementary sequences at the termini.

Nucleic acid	Type of repeat	Examples
	DTR	Some herpesviruses T phages
	ITR	Adenoviruses Tectiviruses (phages)
	ITR	Some parvoviruses
	DTR	Retroviruses
	ITR	Influenza viruses Bunyaviruses





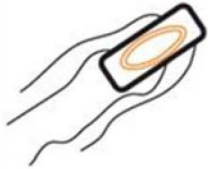
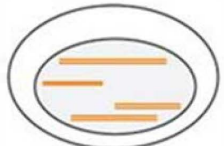
↓



3.2.1 Genome size

Virus genomes span a large range of sizes. The smallest virus known is tobacco necrosis satellite virus, with a ssRNA genome of 1239 bases, while at the other end of the scale there are viruses with dsDNA genomes of over 1000 kilobase pairs (kbp). The largest virus genomes, such as that of the mimivirus, are larger than the smallest genomes of cellular organisms, such as some mycoplasmas (Figure 3.2).

Figure 3.2 Genome sizes of some DNA viruses and cells. Not to scale.

			Approximate genome size (kbp DNA)
Viruses		Hepatitis B virus	3.2
		Phage λ	48.5
		Mimivirus	1200
Cells		<i>Mycoplasma genitalium</i>	580
		<i>Escherichia coli</i>	4700
		Human	3200000

The maximum size of a virus genome is subject to constraints, which vary with the genome category. As the constraints are less severe for dsDNA all of the large virus genomes are composed of dsDNA. The genomes of some dsDNA viruses have been found to encode small RNAs (transfer RNAs and micro-RNAs), as well as proteins. The largest RNA genomes known are those of some coronaviruses, which are 33 kb of ssRNA.

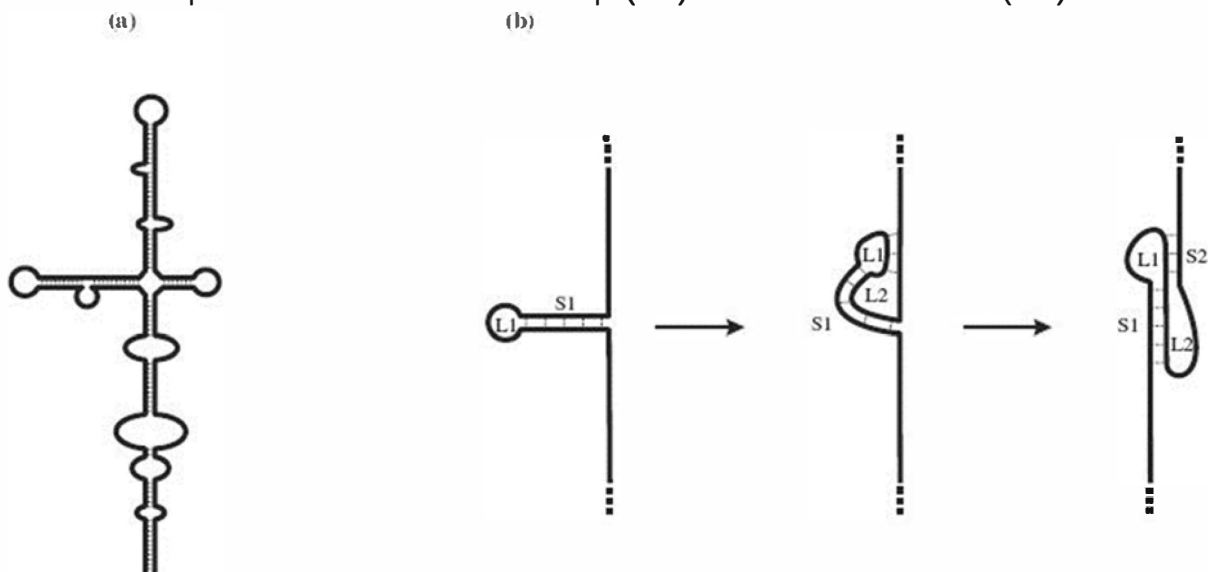
Viruses with small genomes have evolved economical ways of using their limited coding capacities. Some viruses, such as hepatitis B virus (Section 19.6), use every nucleotide for protein coding and encode proteins in overlapping reading frames. Another way that small viruses have circumvented the limitation of a small genome is by evolving proteins that can perform two or more functions (Section 3.3).

3.2.2 Secondary and tertiary structure

As well as encoding the virus proteins, and in some cases RNAs, to be synthesized in the infected cell, the virus genome carries additional information, such as signals for the control of gene expression. Some of this information is contained within the nucleotide sequences, while for the single-stranded genomes some of it is contained within structures formed by intramolecular base pairing.

In ssDNA, complementary sequences may base pair through G₃T₂C and A₃T₂T hydrogen bonding; in ssRNA weaker G₃T₂U bonds may form in addition to G₃T₂C and A₃T₂U base pairing. Intramolecular base pairing results in regions of secondary structure with stem-loops and bulges (Figure 3.3(a)). In some ss-RNAs, intramolecular base pairing results in structures known as pseudoknots, the simplest form of which is depicted in Figure 3.3(b). Some pseudoknots have enzyme activity, while others play a role in ribosomal frameshifting (Section 6.4.2).

Figure 3.3 Secondary structures resulting from intramolecular base-pairing in single-stranded nucleic acids. (a) Stem-loops and bulges in ssRNA and ssDNA. (b) Formation of a pseudoknot in ssRNA. A pseudoknot is formed when a sequence in a loop (L1) at the end of a stem (S1) base-pairs with a complementary sequence outside the loop. This forms a second loop (L2) and a second stem (S2).



Regions of secondary structure in single-stranded nucleic acids are folded into tertiary structures with specific shapes, many of which are important in molecular interactions during virus replication. For an example see Figure 14.5, which depicts the 5' end of poliovirus RNA, where there is an internal ribosome entry site to which cell proteins bind to initiate translation. Another example can be seen in Figure 18.3, which depicts the 5' end of HIV-1 RNA, where there are a number of regions with specific functions.

3.2.3 Modifications at the ends of virus genomes

It is interesting to note that the genomes of some DNA viruses and many RNA viruses are modified at one or both ends (Figures 3.4 and 3.5). Some genomes have a covalently linked protein at the 5' end. In at least some viruses this is a vestige of a primer that was used for initiation of genome synthesis (Section 7.3.1).

Figure 3.4 DNA virus genomes with one or both ends modified. The 5' end of some DNAs is covalently linked to a protein. One of the hepatitis B virus DNA strands (the (+) strand) is linked to a short sequence of RNA with a methylated nucleotide cap.

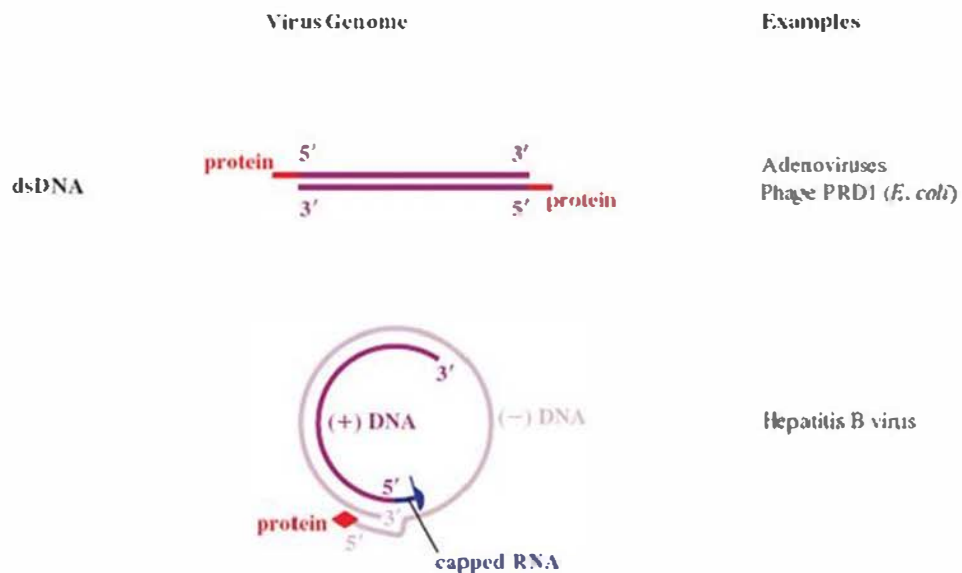
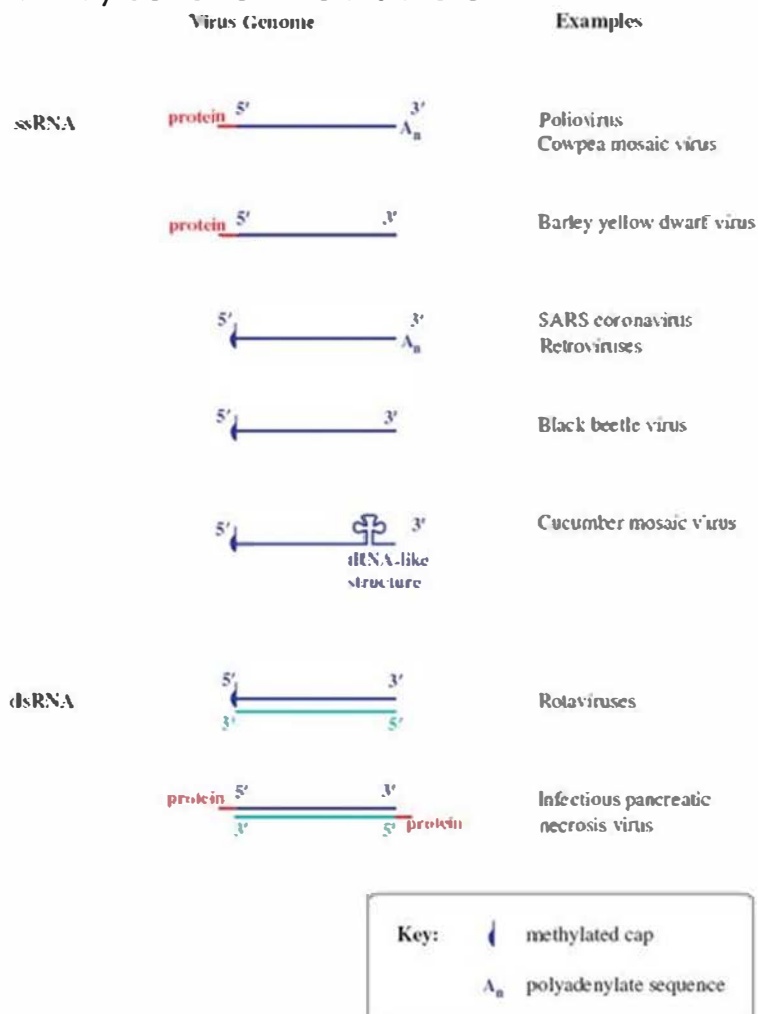


Figure 3.5 RNA virus genomes with one or both ends modified. The 5' end may be linked to a protein or a methylated nucleotide cap. The 3' end may be polyadenylated or it may be folded like a transfer RNA.



Some genome RNAs have one or both of the modifications that occur in eukaryotic messenger RNAs (mRNAs): a methylated nucleotide cap at the 5' end (Section 6.3.4) and a sequence of adenosine residues (a polyadenylate tail; poly(A) tail) at the 3' end (Section 6.3.5).

The genomes of many ssRNA viruses, including most of those in Figure 3.5, function as mRNAs after they have infected host cells. It is worth noting that very few of these RNAs have both a cap and a poly(A) tail, and some have neither.

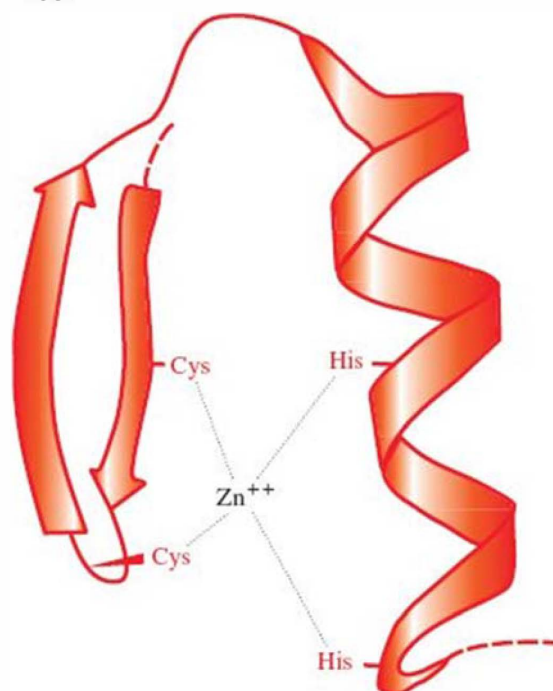
The genomes of some ssRNA plant viruses are base paired and folded near their 3' ends to form structures similar to transfer RNA. These structures contain sequences that promote the initiation of RNA synthesis.

3.2.4 Proteins non-covalently associated with virus genomes

Many nucleic acids packaged in virions have proteins bound to them non-covalently. These proteins have regions that are rich in the basic amino acids lysine, arginine, and histidine, which are positively charged and able to bind strongly to the negatively charged nucleic acids.

Papillomaviruses and polyomaviruses, which are DNA viruses, have cell histones bound to the virus genome. Most proteins associated with virus genomes, however, are virus coded, such as the HIV-1 nucleocapsid protein that coats the virus RNA; 29% of its amino acid residues are basic. As well as their basic nature, nucleic acid-binding proteins may have other characteristics, such as zinc fingers (Figure 3.6); the HIV-1 nucleocapsid protein has two zinc fingers.

Figure 3.6 A zinc finger in a protein molecule. A zinc finger has recurring cysteine and/or histidine residues at regular intervals. In this example there are two cysteines and two histidines.



In some viruses, such as tobacco mosaic virus (Section 3.4.1), the protein coating the genome constitutes the capsid of the virion.

3.2.5 Segmented genomes

Most virus genomes consist of a single molecule of nucleic acid, but the genes of some viruses are encoded in two or more nucleic acid molecules. These segmented genomes are much more common amongst RNA viruses than DNA viruses. Examples of ssRNA viruses with segmented genomes are the influenza viruses (Chapter 16), which package the segments in one virion, and brome mosaic virus, which packages the segments in separate virions. Most dsRNA viruses, such as members of the family Reoviridae (Chapter 13), have segmented genomes.

The possession of a segmented genome provides a virus with the possibility of new gene combinations, and hence a potential for more rapid evolution (Section 20.3.3.c). There are, however, prices to be paid for this advantage. Viruses with the genome segments packaged in one virion need an efficient mechanism to ensure that each virion packages a full set of segments. For viruses with the segments packaged in separate virions at least one of each virion category must enter a cell in order to initiate an infection.

3.2.6 Repeat sequences

The genomes of many viruses contain sequences that are repeated. These sequences include promoters, enhancers, origins of replication, and other elements that are involved in the control of events in virus replication. Many linear virus genomes have repeat sequences at the ends (termini), in which case the sequences are known as terminal repeats (Figure 3.7). If the repeats are in the same orientation they are known as direct terminal repeats (DTRs), whereas if they are in the opposite orientation they are known as inverted terminal repeats (ITRs). Strictly speaking, the sequences referred to as ITRs in single-stranded nucleic acids are not repeats unless the molecule becomes double stranded during replication. These ITRs are, in fact, repeats of the complementary sequences (see ssDNA and ssRNA (bT) in Figure 3.7).

3.3 VIRUS PROTEINS

The virion of tobacco mosaic virus contains only one protein species and the virions of parvoviruses contain two to four protein species. These are viruses with small genomes. As the size of the genome increases, so the number of protein species tends to increase; 50 protein species have been reported in the virion of herpes simplex virus 1, and over 140 in the virion of the algal virus *Paramecium bursaria* Chlorella virus 1.

Proteins that are components of virions are known as structural proteins. They have to carry out a wide range of functions, including:

- protection of the virus genome;

- attachment of the virion to a host cell (for many viruses);
- fusion of the virion envelope to a cell membrane (for enveloped viruses).

Virus proteins may have additional roles, some of which may be carried out by structural proteins, and some by non-structural proteins (proteins which are synthesized by the virus in an infected cell but which are not virion components).

These additional roles include:

- enzymes, e.g. protease, reverse transcriptase;
- transcription factors;
- primers for nucleic acid replication;
- formation of ion channels in viral and cell membranes;
- interference with the immune response of the host.

NOMENCLATURE OF VIRUS PROTEINS

There is no standard system of nomenclature for virus proteins, with different systems having evolved for different groups of viruses. For quite a number of viruses the following system has been adopted, the proteins being numbered in decreasing order of size:

- structural proteins: VP1, VP2, VP3 . . . (VP = virus protein)
- non-structural proteins: NSP1, NSP2, NSP3 . . .

Many virus proteins are known by an abbreviation of one or two letters, which may indicate:

<ul style="list-style-type: none"> • a structural characteristic: 	G (glycoprotein) P (phosphoprotein)
<ul style="list-style-type: none"> • or a function: 	F (fusion) P (polymerase) RT (reverse transcriptase).

In a virion the genome is enclosed in a protein coat, known as a capsid. For some viruses the genome and the capsid constitute the virion, while for other viruses there are additional components. There may be an envelope at the surface of the virion, in which case there may be protein between the envelope and the capsid; alternatively there may be an internal lipid membrane. A few viruses produce protein occlusion bodies in which virions become embedded. We shall consider each of these components in turn.

3.4 CAPSIDS

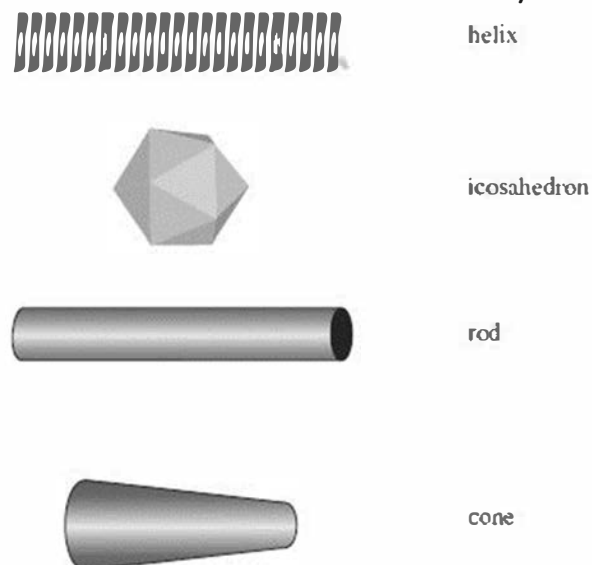
Virus genomes removed from their capsids are more susceptible to inactivation, so a major function of the capsid is undoubtedly the protection of the genome. A second function of many capsids is to recognize and attach to a host cell in which the virus can be replicated. A third function may be to ensure that the virus genome is transported to the location within the cell where the genome can be transcribed and replicated. Although the capsid must be stable enough to survive in

the extracellular environment, it must also be able to alter its conformation so that, at the appropriate location in the host cell, it can release its genome.

For many viruses the capsid and the genome that it encloses constitute the virion. For other viruses a lipid envelope (Section 3.5.1), and sometimes another layer of protein, surrounds this structure, which is referred to as a nucleocapsid.

Capsids are constructed from many molecules of one or a few species of protein. The individual protein molecules are asymmetrical, but they are organized to form symmetrical structures. Some examples of symmetrical structures are shown in Figure 3.8. A symmetrical object, including a capsid, has the same appearance when it is rotated through one or more angles, or when it is seen as a mirror image. For the vast majority of viruses the capsid symmetry is either helical or icosahedral.

Figure 3.8 Symmetrical structures. All these types of symmetry are seen amongst viruses. The most common are helical and icosahedral symmetries.

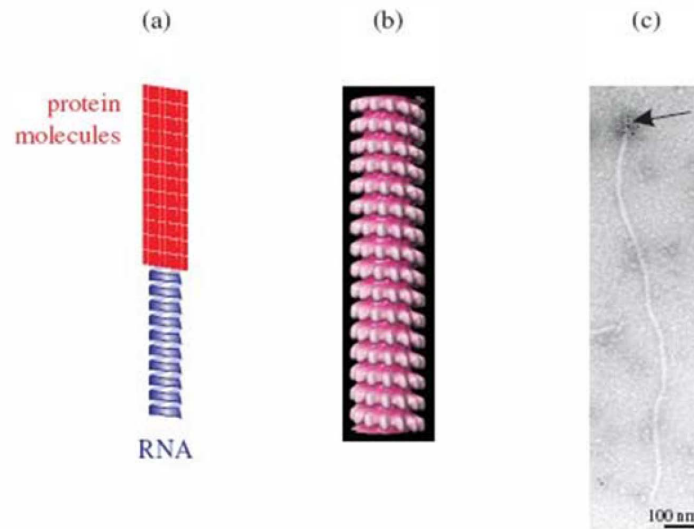


3.4.1 Capsids with helical symmetry

The capsids of many ssRNA viruses have helical symmetry; the RNA is coiled in the form of a helix and many copies of the same protein species are arranged around the coil (Figure 3.9(a), (b)). While the RNA is within the coil for many viruses, there is evidence that the RNA of some viruses is on the surface of the coil. The coiled ribonucleoprotein forms an elongated structure, which may be a rigid rod if strong bonds are present between the protein molecules in successive turns of the helix, or a flexible rod (Figure 3.9(c)) if these bonds are weak. The length of the capsid is determined by the length of the nucleic acid.

Figure 3.9 Helical symmetry. (a) Structure of a capsid with helical symmetry. The ssRNA coil is coated with repeated copies of a protein. (b) Part of measles virus nucleocapsid. The complete nucleocapsid is folded and enclosed within an envelope. (c) Beet yellows virus particle. The virion is a long flexible rod, at one end of which there is a tail (arrow) composed of a minor capsid protein, detected here by specific antibodies labeled with gold.

Sources: (b) Reconstructed image from cryo-electron microscopy, courtesy of Dr David Bhella (MRC Virology Unit, Glasgow). Reinterpretation of data in Bhella et al. (2004) *Journal of Molecular Biology*, 340, 319 (by permission of Elsevier Limited). (c) Courtesy of Professor Valerian Dolja, originally published in Alzhanova et al. (2001) *The EMBO Journal*, 20, 6997. Reproduced by permission of Nature Publishing Group.



For many ssRNA viruses, such as measles and influenza viruses, the helical nucleic acid coated with protein forms a nucleocapsid, which is inside an envelope (see Figure 16.1). The nucleocapsid may be coiled or folded to form a compact structure.

The virions of some plant viruses that have helical symmetry (e.g. tobacco mosaic virus) are hollow tubes; this allows the entry of negative stain, making the center of the virion appear dark in electron micrographs. The rod-shaped tobacco rattle virus has a segmented genome with two RNAs of different sizes packaged in separate virions, resulting in two lengths of virion. Helical symmetry is observed in some bacteriophages; these viruses have filamentous virions with genomes of either dsDNA or ssDNA (Section 20.4.2).

3.4.2 Capsids with icosahedral symmetry




Before proceeding further, a definition of the term icosahedron is required.

An icosahedron is an object with:

- 20 faces, each an equilateral triangle;
- 12 vertices, each formed where the vertices of five triangles meet;
- 30 edges, at each of which the sides of two triangles meet.

An icosahedron has five-, three-, and two-fold axes of rotational symmetry (Figure 3.10).

Figure 3.10 The three axes of symmetry of an icosahedron.

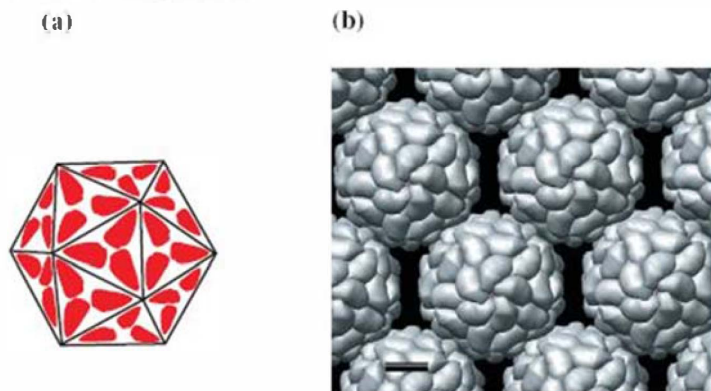
Viewing toward:		Axis of rotational symmetry
vertex		5-fold
triangular face		3-fold
edge		2-fold

Capsids with icosahedral symmetry consist of a shell built from protein molecules that appear to have been arranged on scaffolding in the form of an icosahedron. They have less contact with the virus genome than the capsid proteins of viruses with helical symmetry.

To construct an icosahedron from identical protein molecules the minimum number of molecules required is three per triangular face, giving a total of 60 for the icosahedron (Figure 3.11(a)). The capsid of satellite tobacco mosaic virus is constructed in this way (Figure 3.11(b)).

Figure 3.11 Capsid constructed from sixty protein molecules. (a) Arrangement of protein molecules, with three per triangular face. (b) Virions of satellite tobacco mosaic virus. The bar represents 5 nm. Image created with the molecular graphics program UCSF Chimera from the Resource for Biocomputing, Visualization, and Informatics, at the University of California, San Francisco.

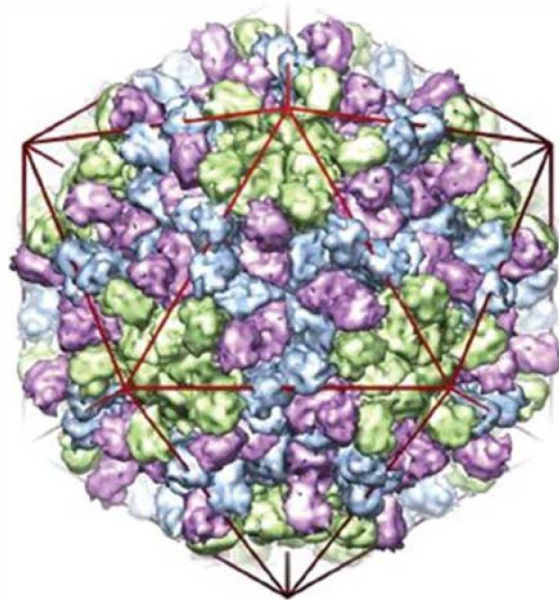
Source: Courtesy of Tom Goddard.



In capsids composed of more than 60 protein molecules it is impossible for all the molecules to be arranged completely symmetrically with equivalent bonds to all their neighbors. In 1962 Donald Caspar and Aaron Klug proposed a theory of quasi-equivalence, where the molecules do not interact equivalently with one another, but nearly equivalently. Hence, the capsid of a virus built from 180 identical protein molecules, such as turnip crinkle virus, contains three types of bonding between the molecules (Figure 3.12).

Figure 3.12 Turnip crinkle virus capsid. The capsid is built from 180 copies of the coat protein in three quasi-equivalent conformations; some protein molecules (green) are around the fivefold symmetry axes, while the remainder (pink and blue) are around the threefold symmetry axes. An icosahedron is superimposed.

Source: Bakker et al. (2012) *Journal of Molecular Biology*, 417, 65–78. Reproduced by permission of the authors and Elsevier Limited.

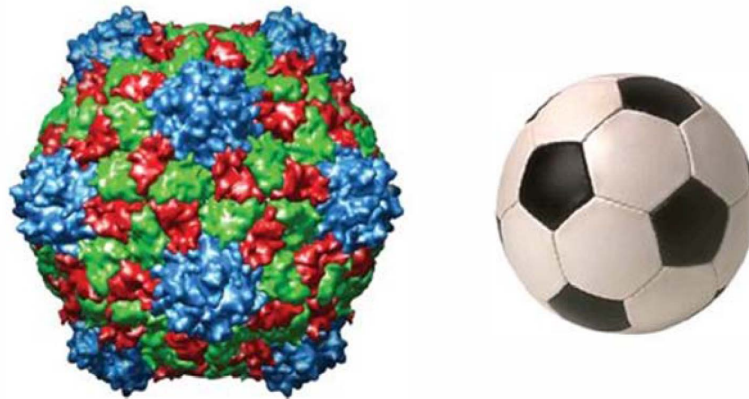


The cowpea mosaic virus capsid is composed of two protein species (Figure 3.13): one is present as pentamers at the vertices of the icosahedron (12 \times 5 = 60 copies) and the other is present as hexamers on the faces. Each hexamer is composed of three copies of a protein with two domains. The arrangement is similar to that of the panels on the surface of the football in Figure 3.13.

Figure 3.13 Capsid constructed from two protein species. The cowpea mosaic virus capsid is constructed from one protein species (blue) that forms 12 pentamers and from a second protein species with two domains (green and red) that forms 20 hexamers. The football is similarly constructed from 12 pentamers and 20 hexamers. The cowpea mosaic virus image is from the VIPER database (Shepherd et al., 2006).

Source: The image was reconstructed using the data of Lin et al. (1999) *Virology*, 265, 20. Reproduced by permission of Elsevier Limited.

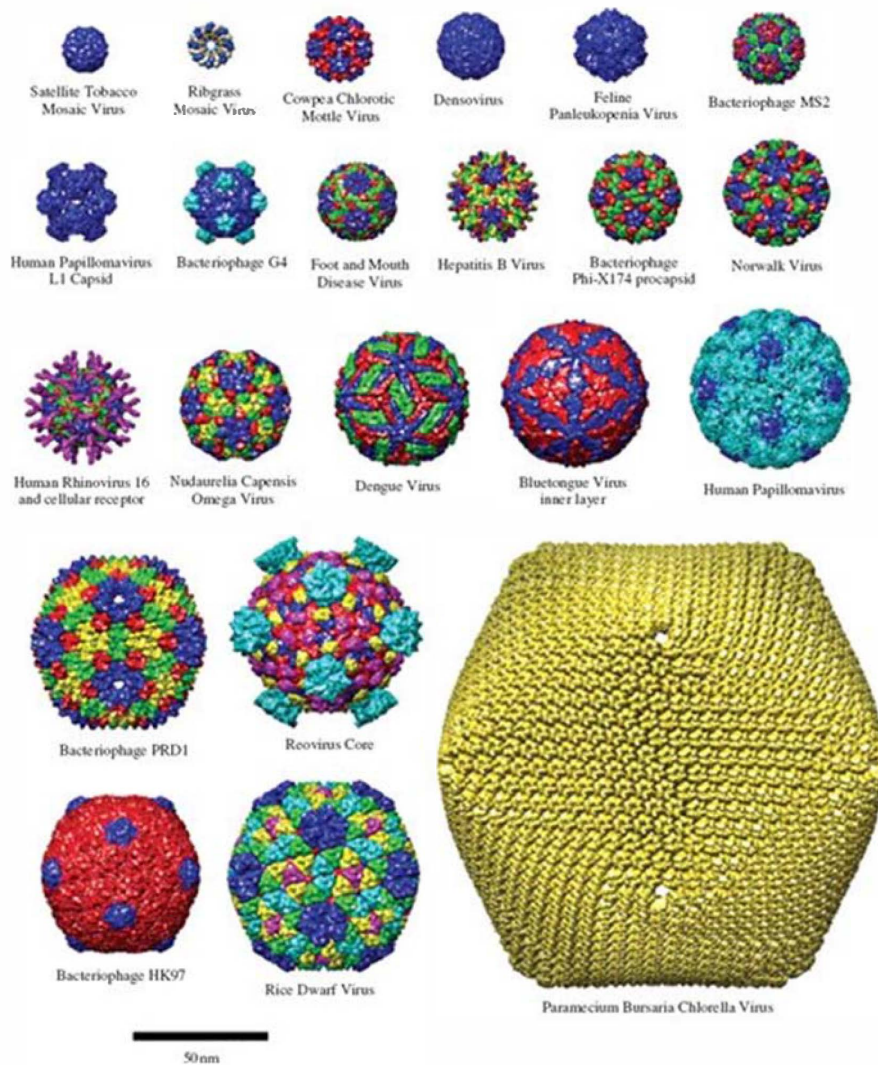
Cowpea mosaic virus capsid



The huge range in the sizes of virus genomes (Section 3.2.1) is mirrored by a huge range in the sizes of icosahedral capsids. The satellite tobacco mosaic virus capsid is about 17 nm in diameter, and the diameter of the *Paramecium bursaria* *Chlorella* virus 1 capsid is about ten times greater than this (Figure 3.14), while the mimivirus capsid is about 500 nm in diameter. Increasing size of the capsid is usually accompanied by an increase in the number of the component protein species.

Figure 3.14 Capsids with icosahedral symmetry. Some of the wide ranges of capsid architectures and sizes are illustrated. The images were created with the molecular graphics program UCSF Chimera using data from cryo-electron microscopy and X-ray diffraction.

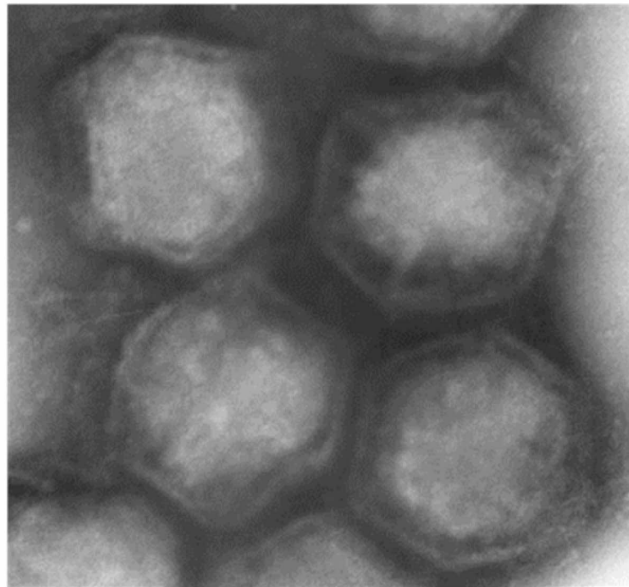
Source: Goddard et al. (2005) *Structure*, 13, 473. Reproduced by permission of Elsevier Limited.



3.4.2.a Capsid shapes

It is clear from the images in Figure 3.14 that capsid surfaces vary in their topography; there may be canyons, hollows, ridges, and/or spikes present. It is also clear that some capsids actually have the shape of an icosahedron, such as that of Paramecium bursaria Chlorella virus 1, which is 165 nm across when measured along the two- and threefold axes and 190 nm across when measured along the fivefold axes. Capsids that have an icosahedral shape have an angular outline in electron micrographs (Figure 3.15).

Figure 3.15 Transmission electron micrograph of negatively stained virions of invertebrate iridescent virus 1.



An icosahedral shape is not an inevitable outcome of icosahedral symmetry; the football in Figure 3.13 is constructed in the form of icosahedral symmetry, but the structure is spherical. Many small viruses that have capsids with icosahedral symmetry appear to be spherical, or almost spherical, and their virions are often described as isometric, such as those of densoviruses and foot and mouth disease virus (Figure 3.14).

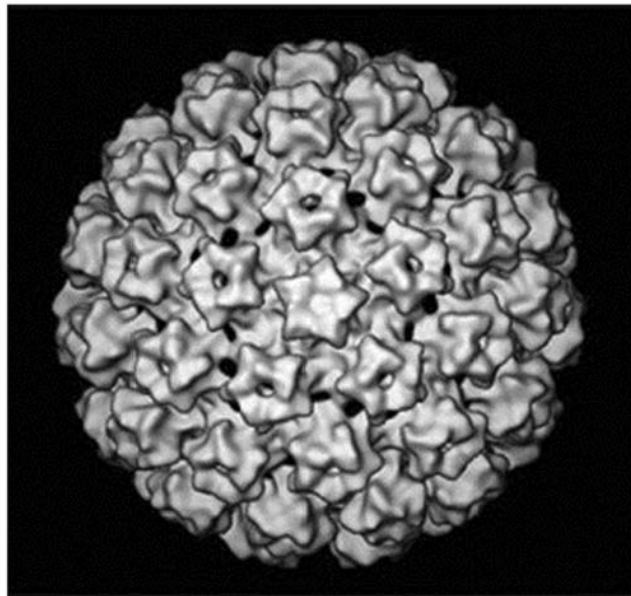
Some capsids with icosahedral symmetry are elongated. The capsids of geminiviruses (plant viruses) are formed from two incomplete icosahedra. Another plant virus, alfalfa mosaic virus, has four sizes of virion; all are 19 nm diameter, but three are elongated as a result of insertions of a protein lattice between a half-icosahedral structure at each end of the capsid.

3.4.2.b Capsomeres

The capsids of some viruses, such as papillomaviruses (Figure 3.16), are clearly constructed from discrete structures. These structures are called capsomeres and each is built from several identical protein molecules.

Figure 3.16 Papillomavirus capsid reconstruction.

Source: Trus et al. (1997) *Nature Structural Biology*, 4, 413, with the permission of the authors and Nature Publishing Group.

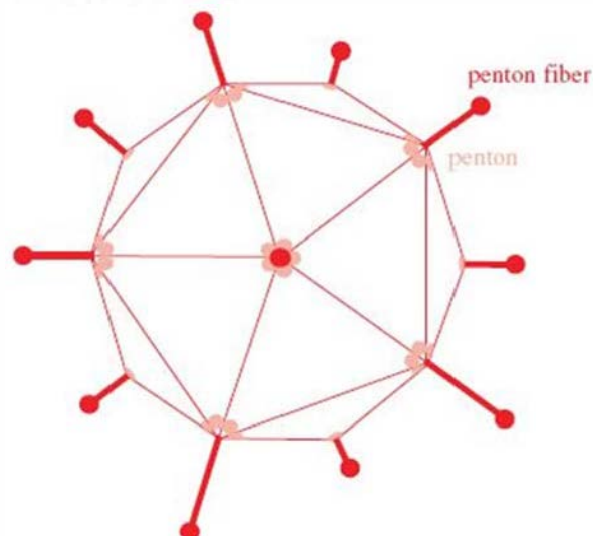


The capsids of papillomaviruses are constructed from 72 capsomeres, which are all identical, but the capsids of some viruses are constructed from two types of capsomere: pentons, found at the vertices of the icosahedron, and hexons, which make up the remainder of the capsid. In these viruses there are always 12 pentons (one at each vertex), but the number of hexons varies; for example, the capsids of herpesviruses and adenoviruses contain 150 and 240 hexons, respectively.

3.4.2.c Structures at capsid vertices

Some icosahedral viruses have a structure such as a knob, projection, or fiber at each of the 12 vertices of the capsid. For example, the virions of some phages (e.g. G4; Figure 3.14) have projections, while the adenovirus virion has a fiber, with a knob attached, at each of the 12 pentons (Figure 3.17). These structures at the capsid vertices are composed of distinct proteins that are involved in attachment of the virion to its host cell and in delivery of the virus genome into the cell.

Figure 3.17 Adenovirus virion. At each of the 12 vertices of the virion there is a penton, and attached to each penton there is a protein fiber with a knob at the end. The rest of the capsid is constructed from hexons. See Figure 12.1 for an electron micrograph of an adenovirus.



3.4.2.d Tailed bacteriophages

The majority of the known phages are constructed in the form of a tail attached to a head, which contains the virus genome. All of these phages have dsDNA genomes. The head has icosahedral symmetry and may be isometric, as in phage lambda (λ), or elongated, as in phage T4. The tail is attached to one of the vertices of the head and may be long, as in phage λ , or short, as in phage T7. Attached to the tail there may be specialized structures such as fibers and/or a baseplate.

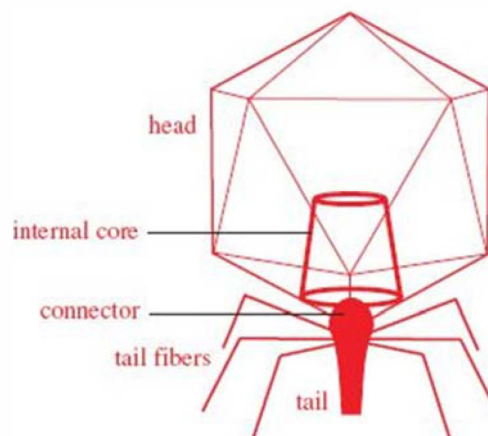
The tail is attached to the head via a structure composed of one or more distinct proteins; this structure is known as a portal or a connector. It contains a powerful motor that moves the virus DNA into the immature head during virus assembly. The connector, head, and tail are each symmetrical, but in different ways:

- the connector has 12-fold rotational symmetry;
- the icosahedron vertex attached to one end of the connector has fivefold rotational symmetry (Section 3.4.2);
- the tail attached to the other end of the connector has sixfold rotational symmetry.

The capsid of herpesviruses also has a portal with roles similar to those of the tailed phage portals/connectors.

Some of the tailed phages have been objects of intensive study and a lot of the detail of their structures has been uncovered. One such phage is T7 (Figure 3.18).

Figure 3.18 Structure of phage T7. Each of the components is composed of one or more distinct proteins.



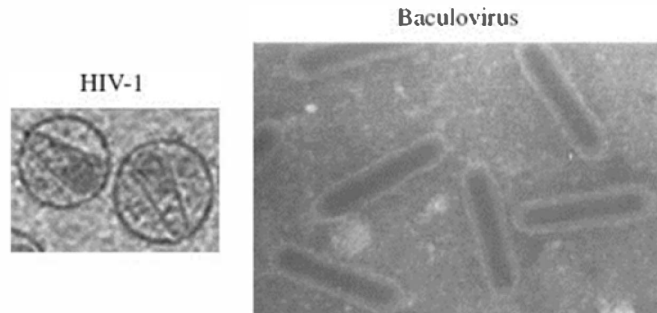
Inside the head of phage T7 is a cylindrical structure (the internal core) around which the DNA is wound. The core is attached to a connector at one of the vertices of the head and the tail is attached to the other end of the connector. The tail is very short and tapers from the connector to the tip; attached to the tail are six tail fibers. Further details about the structure of tailed phages are given in Section 20.5.

3.4.3 Conical and rod-shaped capsids

HIV-1 and baculoviruses have capsids that are conical and rod shaped, respectively (Figure 3.19). Inside each capsid is a copy of the virus genome coated in a highly basic protein. Both of these viruses have enveloped virions (Section 3.5.1).

Figure 3.19 Virions containing conical and rod-shaped capsids.

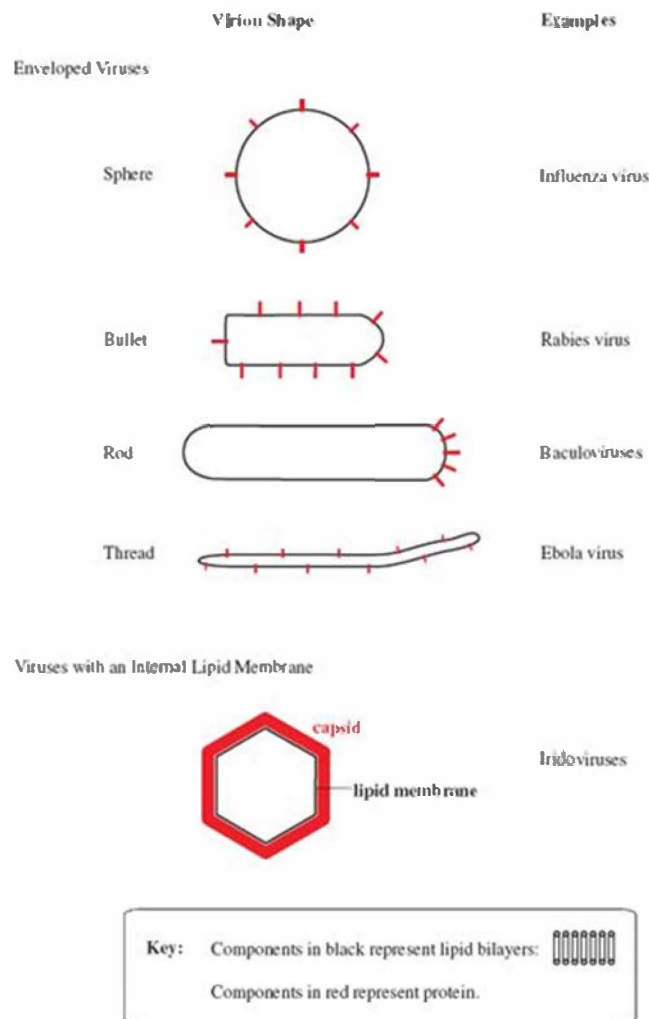
Sources: HIV-1 virions from Wei and Yin (2010) *Journal of Structural Biology*, 172, 211. Reproduced by permission of Elsevier and the authors. Baculovirus virions are those of *Aglaia urticae* nucleopolyhedrovirus. From Harrap (1972) *Virology*, 50, 124. Reproduced by permission of Elsevier Limited.



3.5 VIRION MEMBRANES

Many viruses have a lipid membrane component. In most of these viruses the membrane is at the virion surface and is associated with one or more species of virus protein. This lipid-protein structure is known as an envelope and it encloses the nucleocapsid (nucleic acid plus capsid). The virions of most enveloped viruses, such as herpesviruses, are spherical or roughly spherical, but other shapes exist (Figure 3.20). Some viruses have a membrane located, not at the virion surface, but within the capsid.

Figure 3.20 Lipid-containing viruses.



3.5.1 Enveloped virions

Many animal viruses are enveloped, including all those with helical symmetry, e.g. influenza viruses, and a significant number of those with icosahedral symmetry, e.g. herpesviruses. Enveloped virions are much less common amongst the viruses that infect plants (e.g. potato yellow dwarf virus) and they are extremely rare amongst the viruses that infect prokaryotes (e.g. Pseudomonas phage $\Pi+6$).

Associated with the membranes of an enveloped virus are one or more species of protein. Most of these proteins are integral membrane proteins and most are O- and/or N-glycosylated (Section 6.4.3.a). Many of the glycoproteins in virion envelopes are present as multimers. The envelope of influenza A virus (Chapter 16), for example, contains two glycoprotein species: a hemagglutinin present as trimers and a neuraminidase present as tetramers. There is also a third protein species (M2) that is not glycosylated.

Some envelope proteins span the membrane only once, but some, such as those of hepatitis B virus (Section 19.3.4), span the membrane several times. The polypeptide chain is highly hydrophobic at each membrane anchor.

Some surface glycoproteins of enveloped viruses perform the function of fusing the virion membrane to a cell membrane during the infection process (Section 5.2.4.b).

These fusion proteins have an additional hydrophobic region that plays a major role in the fusion process.

Many enveloped viruses, including influenza viruses and retroviruses, have a layer of protein between the envelope and the nucleocapsid. This protein is often called a matrix protein. In some viruses, however, such as yellow fever virus, there is no such layer and the nucleocapsid interacts directly with the internal tails of the integral membrane protein molecules (Figure 8.3).

3.5.2 Virions with internal membranes

There are several viruses that have a lipid membrane within the virion rather than at the surface (Figure 3.20). These membranes have proteins associated with them and are present in the virions of *Paramecium bursaria* Chlorella virus 1 (Figure 3.14) and the iridoviruses (Figure 3.15).

3.5.3 Membrane lipids

Most virion membranes are derived from host cell membranes that undergo modification before incorporation into virions. For example, the HIV-1 envelope is derived from the plasma membrane of the host cell, but the envelope contains more cholesterol and sphingomyelin, and less phosphatidylcholine and phosphatidylinositol, than the plasma membrane.

Some viruses are able to replicate in more than one kind of host cell; for example, alphaviruses replicate in both mammalian cells and insect cells. When progeny virions are released from cells the lipid composition of the envelope may reflect that of the cell. Virions of the alphavirus Semliki Forest virus produced in hamster kidney cells contain about five times more cholesterol than virions produced in mosquito cells.

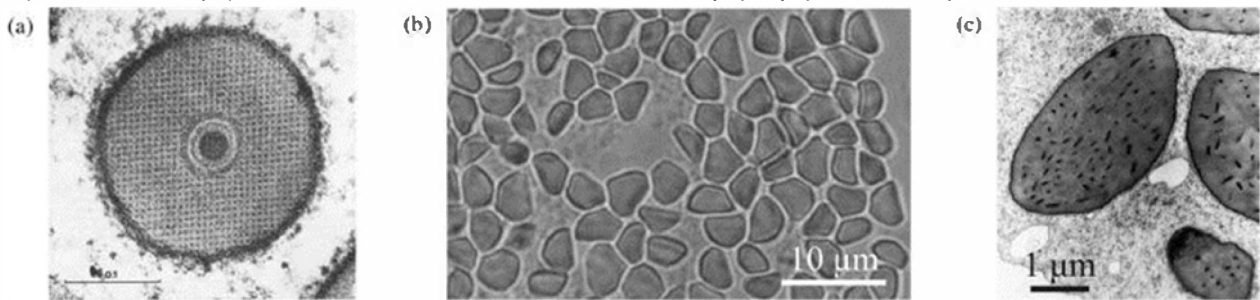
3.6 OCCLUSION BODIES

Some viruses provide added protection to the virions whilst outside their hosts by occluding them in protein crystals. These occlusion bodies, as they are known, are produced by many invertebrate viruses, including most baculoviruses. There are two major types of occlusion body in which baculoviruses embed their rod-shaped virions (Figure 3.19): the granuloviruses form small granular occlusion bodies, generally with a single virion in each (Figure 3.21(a)), while the nucleopolyhedroviruses form large occlusion bodies with many virions in each (Figure 3.21(b), (c)).

Figure 3.21 Baculovirus occlusion bodies. (a) Transverse section through an occlusion body of a granulovirus of the Indian meal moth (*Plodia interpunctella*). The nucleocapsid surrounded by its membrane can be seen at the center, surrounded by the crystalline protein that forms the occlusion body. The bar represents 0.1 μm . (b), (c) Nucleopolyhedrovirus of the leatherjacket (*Tipula paludosa*). (b) Light micrograph of occlusion bodies. (c) Electron micrograph of a

thin section of an occlusion body. The virions are randomly embedded in the crystalline protein.

Sources: (a) Arnott and Smith (1967) *Journal of Ultrastructure Research*, 21, 251. Reproduced by permission of Elsevier Limited. (b), (c) Courtesy of Dr Liz Boslem.



3.7 OTHER VIRION COMPONENTS

3.7.1 Virus RNA in DNA viruses

Virus RNA is present in the virions of a number of DNA viruses. The hepadnaviruses and the caulimoviruses have short RNA sequences covalently attached to their DNAs. These RNAs were primers for DNA synthesis and remain attached to the genome in the mature virion.

There is evidence that virus mRNAs are packaged in the virions of herpesviruses, but the roles of these molecules are not clear.

3.7.2 Cell molecules in virions

The incorporation of cell lipids into virions has already been discussed. Other cell molecules that become incorporated into virions include the following:

- Transfer RNA molecules. These are present in the virions of retroviruses.
- Proteins. When a virion is assembled, some cell protein may be incorporated. There are reports of HIV-1 incorporating several cell proteins, including cyclophilin A in association with the capsid, and human leukocyte antigens in the envelope.
- Polyamines. Spermidine and other polyamines have been reported in a variety of viruses.
- Cations. Na^+ , K^+ , Zn^{2+} , and Mg^{2+} have been reported as components of virions.

One likely role for polyamines and cations is the neutralization of negative charges on the virus genome.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- describe the components of virions;
- illustrate the variety of virus genomes;
- outline the functions of virus structural and non-structural proteins;

- define the terms helical symmetry and icosahedral symmetry;
- describe the virions of a selection of naked and enveloped viruses.

SOURCES OF FURTHER INFORMATION

Books

Fischer, W. B. (ed.) (2005) *Viral Membrane Proteins: Structure, Function and Drug Design*, Kluwer/Plenum

Flint, S. J. et al. (2009) Chapters 3 and 4 in *Principles of Virology*, Volume 1, 3rd edition, ASM Press

Harrison, S. C. (2007) Principles of virus structure. Chapter 3 in *Fields Virology*, 5th edition (Knipe D. M. and Howley, P. M., editors-in-chief), Lippincott, Williams and Wilkins

Historical paper

Caspar, D. L. D. and Klug, A. (1962) Physical principles in the construction of regular viruses. *Cold Spring Harbor Symposia Quantitative Biology*, 27, 1-24

Recent papers

Amos, L. A. and Finch, J. T. (2004) Aaron Klug and the revolution in biomolecular structure determination. *Trends In Cell Biology*, 14, 148-152

Clare, D. K. and Orlova, E. V. (2010) 4.6 Å Cryo-EM reconstruction of tobacco mosaic virus from images recorded at 300 keV on a 4k × 4k CCD camera. *Journal of Structural Biology*, 171, 303-308

Huiskonen, J. H. and Butcher, S. J. (2007) Membrane-containing viruses with icosahedrally symmetric capsids. *Current Opinion in Structural Biology*, 17, 229-236

Rey, F. A. (2010) Virology: one protein, many functions. *Nature*, 468, 773-775

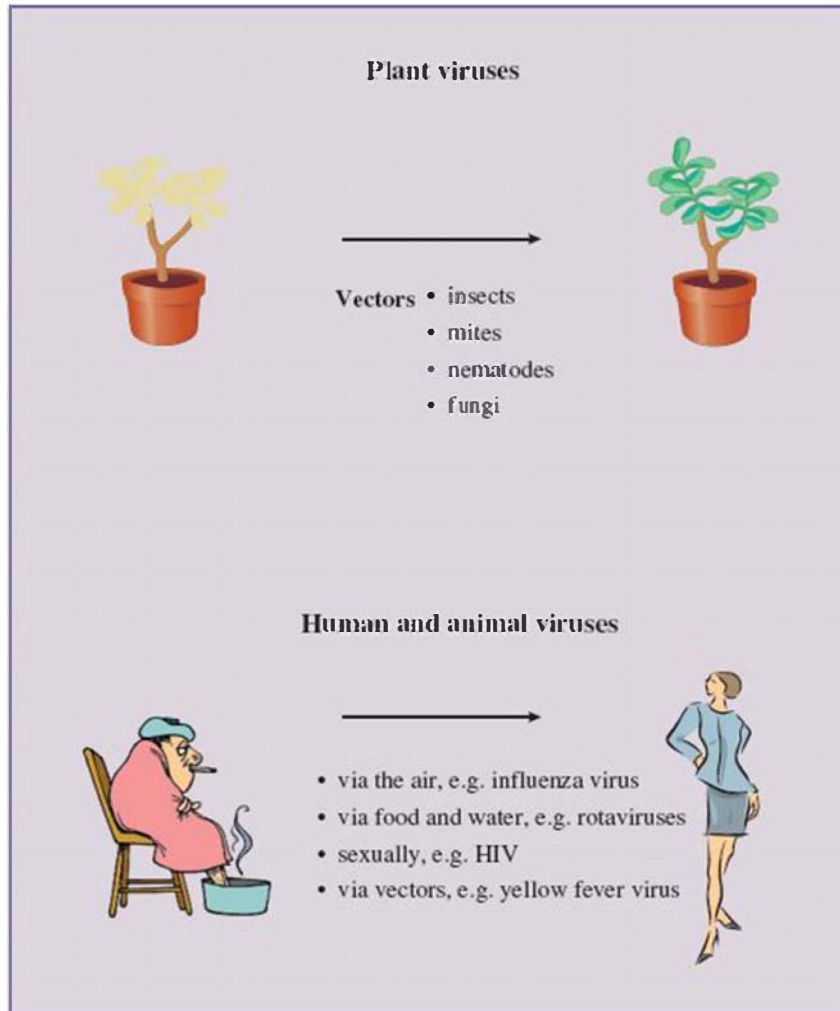
Ruigrok, R. W. H., Czapka, T., and Kolakofsky, D. (2011) Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Current Opinion in Microbiology*, 14, 504-510

Schneemann, A. (2006) The structural and functional role of RNA in icosahedral virus assembly. *Annual Review of Microbiology*, 60, 51-67

CHAPTER 4

Virus Transmission

CHAPTER 4 AT A GLANCE



4.1 INTRODUCTION TO VIRUS TRANSMISSION

A minimum proportion of the virions produced in infected hosts must be transmitted to new hosts in which more virions can be manufactured. If this does not happen the virus will die out. The only other possibility for the survival of virus genes is for them to be maintained in cells as nucleic acids, which are replicated and passed on to daughter cells when the cells divide.

Viruses of bacteria and other microbial hosts are released from infected cells into the environment of the host, where further susceptible cells are likely to be present. These viruses are dependent upon chance encounters with susceptible cells, to

which they may bind if receptors on the surface of those cells come into contact with virus attachment sites.

Viruses of multicellular animals and plants must also find new cells to infect. An infection may spread to adjacent cells, or to cells in a distant part of the host after transport in the blood of an animal or in the phloem of a plant, but ultimately a virus must find new hosts to infect if it is to survive.

Some virus infections modify the behavior of their hosts in order to increase the probability of transmission. Mammals infected with rabies virus often become aggressive; this change in behavior increases the likelihood of the host biting another individual and transmitting virus in the saliva. Some plant-feeding insect larvae infected with baculoviruses become more mobile in the late stages of infection, thus aiding virus dispersal. If the virus-infected insect migrates toward the top of a plant, then lower leaves, where other larvae will feed, will become contaminated when the host dies and putrefies.

On their journeys between hosts, viruses may have to survive adverse conditions in an environment such as air, water, or soil; the survival of virus infectivity outside the host is discussed in Chapter 24. There are some viruses, however, that can be transmitted to new hosts without seeing the light of day, in other words without exposure to the outside environment. These are viruses that can be transmitted directly from host to host, for example during kissing or sexual intercourse, and viruses that are transmitted via vectors. Also included in this category are viruses that can be transmitted directly from a parent to members of the next generation. Transmission in these cases is said to be vertical; otherwise it is described as horizontal.

Viruses may be moved over long distances in a variety of ways. Rivers and winds can move viruses to new areas. The foot and mouth disease outbreak on the Isle of Wight, UK, in 1981 was initiated by virus that had spread via the air from Brittany, more than 250 km away. Viruses of birds, fishes, humans, and other hosts are transported within their hosts to other parts of the planet as a result of migration, travel, and animal export, for example:

- bird migration (avian influenza viruses);
- human travel (SARS coronavirus);
- animal export (monkeypox virus).

Furthermore, viruses can be transported on inanimate materials; for example, foot and mouth disease virus on straw and farm vehicles.

Once virions have entered a multicellular organism they may have further to travel before suitable host cells are encountered. During this final stage of their journey the virions will encounter hazards in the form of host defense mechanisms, which must be survived if the virions are to remain infective when they reach their destination.

In theory, a single virion can initiate an infection, but in practice it is often found that a host must be inoculated with a minimum number of virions in order for that host to become infected. The reasons for this are probably many and varied, and may include some virions being defective and some being inactivated by the host's immune systems. This minimum amount of virus required for infection of a host is known as the minimum infective dose.

In this chapter we shall look at the ways that viruses of plants and animals are transmitted from host to host, but first we must look at some general aspects of virus transmission by vectors.

4.1.1 Transmission via vectors: general principles

Many viruses of plants and animals are transmitted between hosts by organisms that feed on them; these organisms act as vectors. There are several concepts concerning vector transmission that are common to viruses of both plants and animals so, before discussing these host categories separately, it is useful to consider these concepts.

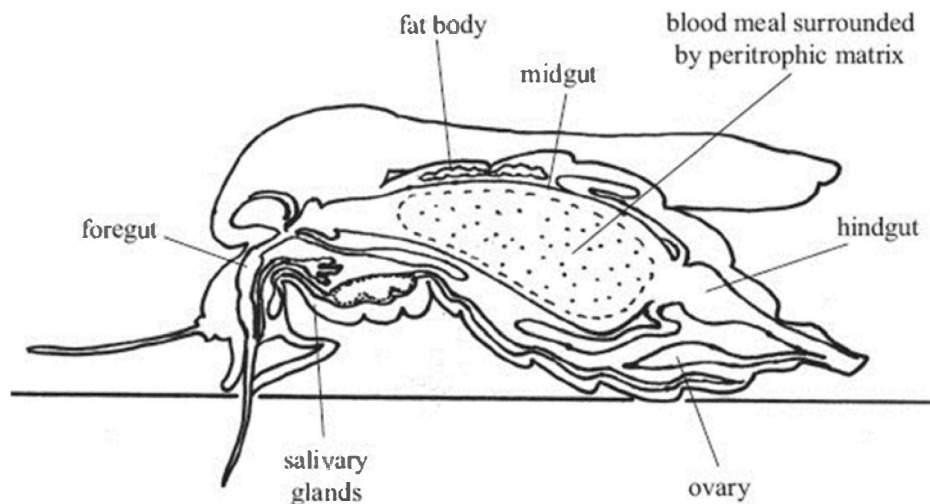
Most vectors of viruses are arthropods (insects, mites, ticks); the arthropod-transmitted viruses of vertebrates are sometimes referred to as arboviruses (arthropod-borne viruses).

The principle of vector transmission is that the vector acquires a virus when it feeds on an infected host and subsequently transmits the virus to one or more new hosts. Some viruses are transmitted after virions have become attached to the mouthparts of their vectors during feeding. Transmission in this way may occur within seconds or minutes of the vector acquiring the virus. Many vector-transmitted viruses, however, cross the gut wall of the vector and enter its circulatory system. The virus ultimately reaches the salivary glands and is secreted into the saliva, which may transport virus into new hosts when the vector feeds. This mode of transmission is said to be circulative, and transmission does not occur until hours or days after the vector has acquired the virus.

Some circulative viruses replicate in one or more tissues and organs of their vectors (Figure 4.1); thus there are many viruses that can replicate in both invertebrates and vertebrates, and a few viruses that can replicate in both invertebrates and plants. It is fascinating to contemplate the existence of viruses with such broad host ranges. Many (but not all) of the invertebrates appear to suffer little or no harm when they are infected, so it has been hypothesized that some modern vertebrate viruses and plant viruses are descended from viruses of invertebrates that later extended their host ranges to vertebrates or plants.

Figure 4.1 Cross-section of a mosquito. Some of the organs and tissues that may become infected by a virus acquired in a blood meal are indicated.

Source: Higgs (2004) Chapter 7, 63rd Symposium of the Society for General Microbiology, Cambridge University Press, by permission of the author and the Society for General Microbiology.



If the reproductive organs of the vector are infected there may be possibilities for vector-to-vector transmission. Some viruses are sexually transmitted (male to female and vice versa) and some are transmitted to the next generation within the egg; the latter is known as transovarial transmission.

There is a high degree of specificity between most vectors and the viruses that they transmit. Generally, specificity is greater between plant viruses and their vectors than between vertebrate viruses and their vectors.

4.2 TRANSMISSION OF PLANT VIRUSES

Plant cells are surrounded by thick cell walls that present significant barriers to virus entry; most plant viruses are carried across these barriers by vectors. A wide variety of organisms use plants as sources of nutrition and some of these organisms, especially invertebrates, act as virus vectors (Figure 4.2). Many of the vectors (e.g. aphids, nematodes) feed by piercing cell walls and ingesting the contents, while beetles feed by biting.

Figure 4.2 Examples of invertebrate vectors of plant viruses, and examples of viruses transmitted.

Sources:

• aphids

• left: *Myzus persicae*, courtesy of Magnus Gammelgaard Nielsen;

• right: *Aphis fabae*, courtesy of Robert Carter;

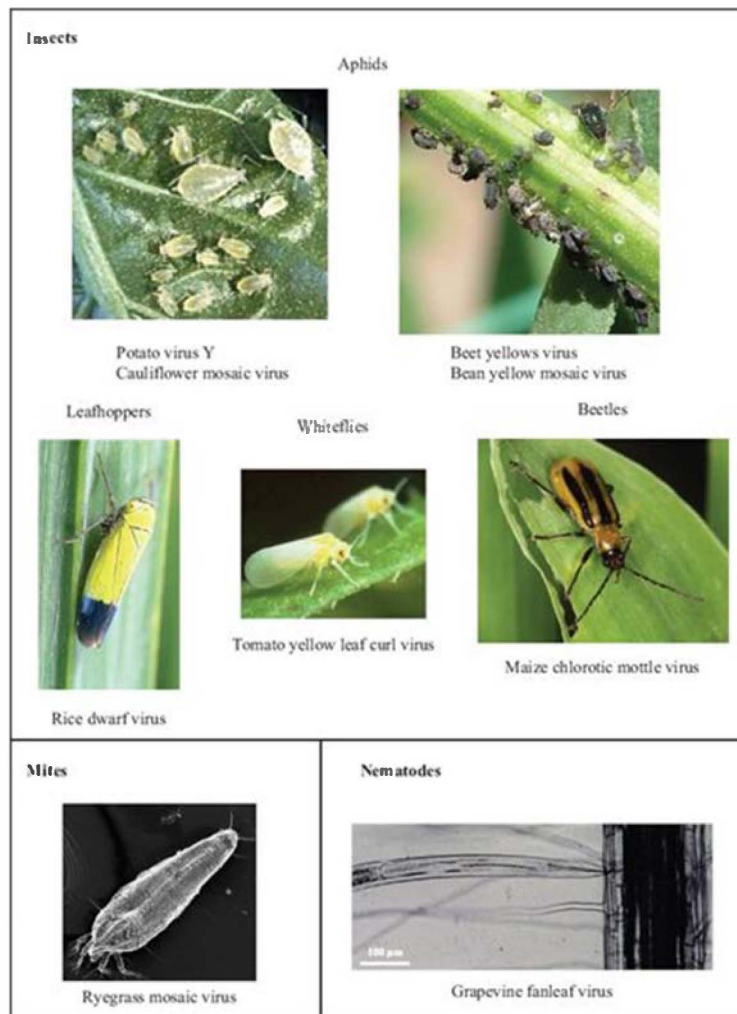
• leafhopper (*Nephotettix cincticeps*), courtesy of Keiji Morishima;

• whitefly (*Bemisia tabaci*), courtesy of Dr J. K. Brown, University of Arizona;

• beetle (*Diabrotica virgifera virgifera*), courtesy of Anthony Zukoff;

• mite (*Abacarus hystrix*), courtesy of Dr Anna Skoracka, A. Mickiewicz University of Poznan;

• nematode (*Xiphinema index* feeding on a root), from MacFarlane and Robinson (2004) Chapter 11, 63rd Symposium of the Society for General Microbiology, Cambridge University Press, reproduced by permission of the Society for General Microbiology.



The most common vectors of plant viruses are aphids, which, as we have already pointed out, feed by ingesting the contents of cells. The reader may question how an invertebrate that feeds by removing the contents of plant cells can introduce virus into cells that subsequently support replication of the virus. The answer appears to be that these vectors probe a number of cells before selecting one on which to feed, so virus may be transmitted into cells that are probed, but not significantly damaged by the vector.

The nematodes that transmit viruses are soil-dwelling animals that pierce root cells and then ingest their contents. Virus transmission ceases after a nematode molts, indicating that virus does not move from its gut into its body. An interesting feature of virus-vector specificity concerns transmission of viruses with different shapes by different types of nematode. Tobraviruses have rod-shaped virions and are transmitted by nematodes in the family Trichodoridae, while nepoviruses have isometric virions and are transmitted by nematodes in the family Longidoridae.

The basis of some cases of plant virus-vector specificity lies in specific amino acid sequences in virus proteins. Some viruses bind specifically to structures in the mouthparts of their vectors via sequences on the surfaces of capsids. For some other viruses, such as potyviruses and caulimoviruses, important roles are played by non-structural proteins (helper components) that are synthesized in infected

cells. Helper components act as bridges between virions and their vector mouthparts.

Some plant-parasitic fungi can also act as virus vectors: for example, *Spongospora subterranea*, which infects potato causing powdery scab disease, is a vector of potato mop-top virus. If a plant is infected with both a fungus and a virus, then virions may be taken into developing fungal spores. The virus may survive in a spore for months or years until the spore germinates on a new host, which then becomes infected not only with the fungus but also with the virus. Some fungus-transmitted plant viruses have been classified in a genus the name of which reflects their mode of transmission and the virion shape: the genus *Furovirus* (fungus-transmitted rod-shaped).

About 20% of plant viruses can be transmitted vertically; in other words, seed can be infected, leading to infection of the next generation. Most seed-transmitted viruses are carried in the embryo, which may have acquired its infection from either an infected ovule or an infected pollen grain. Examples of viruses that can be transmitted via seed include the nepoviruses and the tobnaviruses, two groups of nematode-transmitted viruses mentioned above.

Viruses may enter plants that become damaged; for example, rain may splash virus in soil onto a region of the plant damaged by wind. Many plant viruses can also be transmitted by artificial means; for example, grafting material from a virus-infected plant can introduce virus into a new host.

4.3 TRANSMISSION OF VERTEBRATE VIRUSES

4.3.1 Non-vector transmission of vertebrate viruses

Many viruses of vertebrates (including humans) infect their hosts via the mucous membranes of the upper respiratory tract. An infected individual may shed virus-containing droplets into the air during sneezing, speaking, and coughing, and new hosts may become infected by inhaling this material. Virus present in genital secretions can be transmitted during sexual contact, while viruses infecting the intestinal tract are shed in feces and may enter new hosts ingesting feces-contaminated food or water.

Some viruses may be released from lesions; for example, foot and mouth disease virus from lesions on the feet and in the mouth, papillomaviruses from warts, and herpes simplex virus from lesions on the lips. These viruses could be transmitted directly by contact between an infected host and a non-infected host, or they could be transmitted indirectly through contamination of the environment. Transmission of rabies virus, for example from dog to human, requires a bite from an infected mammal to introduce virus-containing saliva into the body.

In most of the cases discussed so far the host becomes infected when a virus infects epithelial cells and/or lymphoid cells at a body surface such as the respiratory tract, the intestinal tract, the urogenital tract, the skin, or the eye. Some infections (e.g. common cold, rotavirus) remain limited to these tissues, while others (e.g. measles, polio, HIV) commonly cross the epithelial surface and spread to other organs and tissues.

Modes of vertebrate virus transmission not involving vectors are summarized in Table 4.1, along with some examples.

Table 4.1 Non-vector transmission of vertebrate viruses





Transmission route	Examples of viruses transmitted
Horizontal transmission	
B B Respiratory tract	Influenza viruses (mammals)
	Common cold viruses
	Measles virus
B B Intestinal tract	Influenza viruses (birds)
	Rotaviruses
B B Abrasions and wounds	Papillomaviruses
	Rabies virus
B B Genital tract	HIV
	Papillomaviruses
Vertical transmission	
B B Mother to fetus via the placenta	Rubella virus
B B Mother to baby via milk	HIV

4.3.2 Vector transmission of vertebrate viruses

Vectors that transmit vertebrate viruses are mainly blood-feeding arthropods that acquire their viruses when they take blood meals from infected animals (Figure 4.3). Obviously for this to occur virus must be present in the blood of infected animals, a situation known as viremia.

Figure 4.3 Vectors of vertebrate viruses.

Sources: Mosquito (*Aedes aegypti*) courtesy of James Gathany and the US Centers for Disease Control and Prevention; midge (*Culicoides* sp.) from Defra website. Crown copyright material, reproduced with the permission of the Controller of HMSO and Queen's Printer for Scotland; tick (*Ixodes ricinus*) courtesy of Per Arvid F...sen.

Examples of Vertebrate Virus Vectors		Examples of Viruses Transmitted	Virus Hosts
Living Vectors			
Mosquitoes		Yellow fever virus West Nile virus Chikungunya virus	Humans
Midges		Bluetongue virus	Sheep
Ticks		Louping ill virus	Sheep
Inanimate Vectors			
Syringes and Needles		Hepatitis B virus HIV	Humans

Some arthropod parasites that acquire virus infections from the vertebrates on which they feed remain infected for life. This is the case with ticks that become infected with tick-borne encephalitis virus. Transovarial transmission has been demonstrated for a number of viruses, including yellow fever virus in mosquitoes. Viruses in blood or blood products can be transmitted during medical procedures by inanimate vectors such as blood transfusion equipment, syringes, and needles. Viruses can also be transmitted when needles for drug injection are shared, when needlestick injuries occur, and when hemophiliacs are injected with blood factors containing viruses.

4.3.3 Virus survival in a new vertebrate host

A virus has not been successfully transmitted until it has initiated an infection in a new host. In a vertebrate host a virus must evade several defense systems before this is achieved. A virus in the lower respiratory tract must avoid removal by the mucociliary escalator, while a virus in the gastro-intestinal tract must avoid inactivation by extremes of pH. The mucosal surfaces of the respiratory, intestinal, and urogenital tracts present significant barriers to infection. The surfaces are bathed in fluids that contain anti-viral substances, such as complement proteins, and the fluids are viscous, limiting virus access to the cells.

Viruses may be phagocytosed by a range of cell types, including macrophages, which are large cells. Viruses that are introduced into the blood by a vector may be phagocytosed by neutrophils, which constitute the majority of the white blood cells.

These aspects of host immunity are not virus-specific; they are components of the innate immune system. The virus may also have to contend with antibodies and other components of the adaptive immune system if the host has previously been infected with, or vaccinated against, the virus. The roles of the innate and adaptive immune systems in immunity to viruses are covered more thoroughly in Chapter 9.

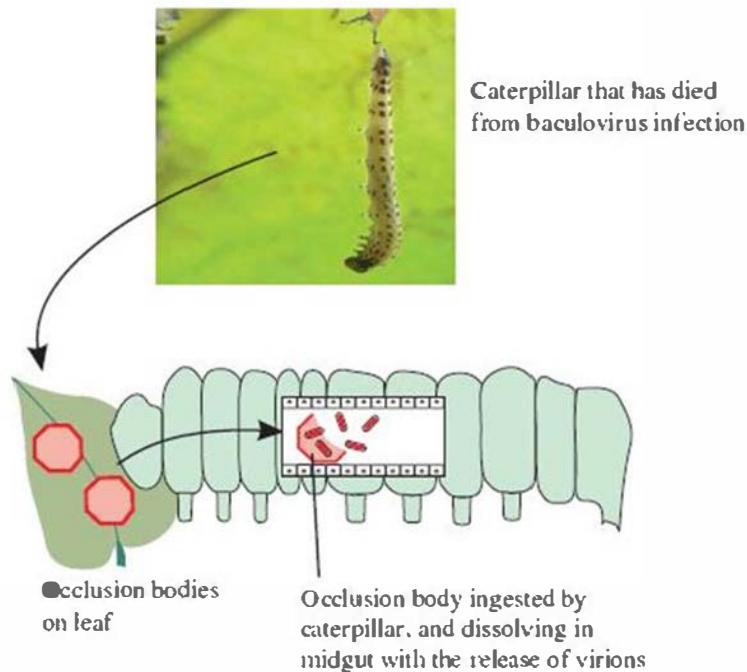
4.4 TRANSMISSION OF INVERTEBRATE VIRUSES

We have discussed the replication of вЂўplantвЂў and вЂўvertebrateвЂў viruses in invertebrate vectors, and it could be argued that these viruses are as much вЂўinvertebrateвЂў viruses as they are viruses of these other host types. In addition to these viruses, invertebrates are hosts to many other viruses that replicate only in invertebrates; most of the known invertebrate viruses have insect hosts.

Many of these viruses have evolved occlusion bodies, which are large protein structures in which virions become embedded in the infected cell (Section 3.6). Viruses that produce occlusion bodies include cypoviruses (Section 13.1) and baculoviruses. Occlusion bodies may be expelled in the feces of a virus-infected insect or they may remain in the host until it is ingested by a predator or it dies as a result of the virus infection. They are robust structures, able to survive putrefaction of the host and providing protection for virions in the outside environment. New insects become infected when occlusion bodies become ingested along with food, which is often plant material. Enzymes and high pH in the gut break down occlusion bodies, releasing the virions. The transmission of a baculovirus via occlusion bodies is depicted in Figure 4.4.

Figure 4.4 Occlusion bodies as vehicles for baculovirus transmission. Putrefaction of insects that have been killed by virus infection results in the deposit of occlusion bodies on leaves. Occlusion bodies ingested by insects are dissolved in the midgut, releasing virions.

Source: Photograph from Castro et al. (2009) *Journal of Invertebrate Pathology*, 102, 149. Reproduced by permission of Elsevier and the authors.



Vertical transmission and vector transmission of a number of invertebrate viruses have been reported. Vertical transmission may occur either within eggs (transovarially) or on the surface of eggs. Vectors involved in virus transmission between insect hosts include parasitic wasps.

4.5 PERMISSIVE CELLS

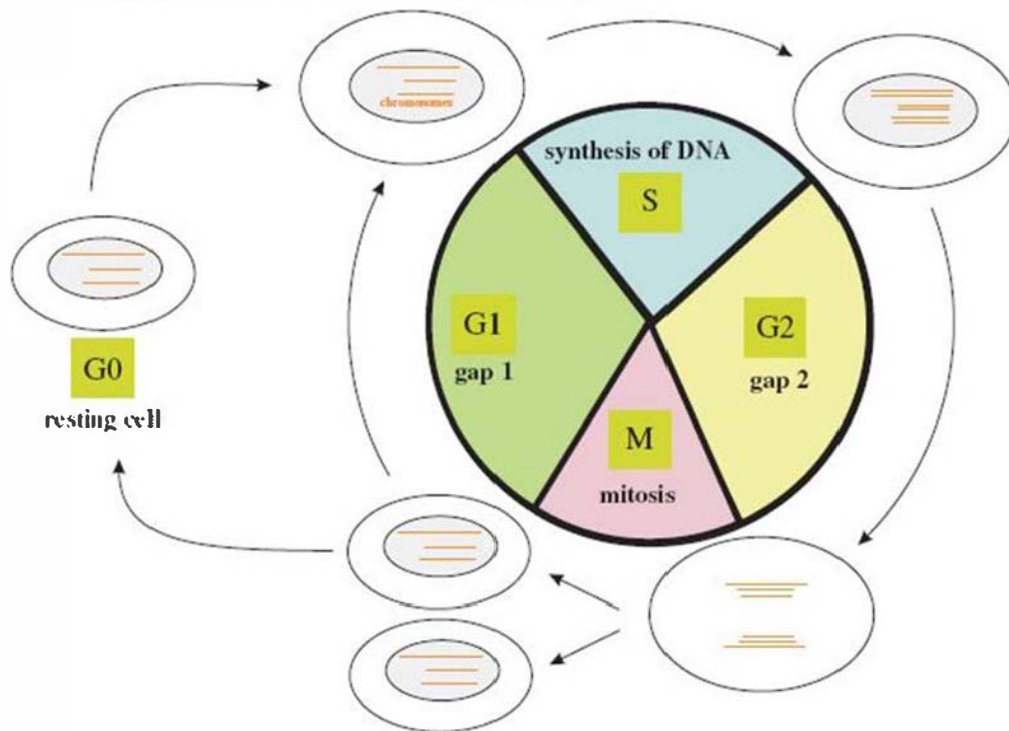
As already stated, virus transmission has not been achieved until a virus has been transferred to a new host and infected a cell that will permit the replication of the virus (a permissive cell). For viruses that bind to the host cell surface as the first step in infection, the cell must have appropriate receptors that the virus can bind to. Furthermore, all the requirements of the virus must either be present in the cell or they must be inducible; these requirements include proteins, such as transcription factors and enzymes.

Some viruses are restricted to a narrow range of permissive cell types; for example, hepatitis B viruses are restricted almost exclusively to hepatocytes. Many other viruses are much less specific and, as we have seen in this chapter, a few can replicate in both animal and plant hosts.

Some viruses of eukaryotes need the host cell to be in a particular phase of the cell cycle (Figure 4.5). For example, retroviruses require access to the cell nucleus and most can enter the nucleus only when the nuclear envelope has broken down in the M (mitosis) phase. Papillomaviruses and parvoviruses are small DNA viruses that are heavily dependent upon their host cells. In order to replicate their DNA, both groups of viruses require the host DNA replicating enzymes present in the S (synthesis of DNA) phase of the cell cycle. Papillomaviruses produce proteins that enable them to manipulate the cell cycle.

Figure 4.5 The eukaryotic cell cycle. The cell synthesizes DNA and histones during the S (synthesis) phase and divides during the M (mitosis) phase. These phases are

separated by two gap phases (G1 and G2). Non-dividing cells suspend the cycle in the G1 phase and are said to be in the G0 state.



Cells have evolved many mechanisms to defend themselves against infection with viruses and cellular microbes. Prokaryotic cells produce restriction endonucleases to degrade foreign DNA, including that of phages. The cells modify their own DNA so that it is unaffected by the restriction enzymes. The cells of vertebrate animals respond to the presence of dsRNA, and other virus-specific molecules, by synthesizing cytokines, including interferons, interleukins, and tumor necrosis factor. These proteins can trigger a range of anti-viral defenses. A cell can be permissive for a virus only if the virus has the ability to overcome the defenses of the cell.

In order to successfully infect a cell, a virus must deliver its genome into the cell, the genome must survive intact, and all attempts by the host to terminate virus replication must be thwarted. In the next four chapters we shall consider the various stages that must be completed before a virus has successfully replicated itself.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- describe the modes of transmission of plant viruses and animal viruses;
- evaluate the roles of vectors in virus transmission;
- outline the immune mechanisms that an animal virus must evade before it can initiate infection.

SOURCES OF FURTHER INFORMATION

Books

- Gillespie, S. H., Smith, G. L., and Osbourn, A. (2004) *Microbe-Vector Interactions in Vector-Borne Diseases*, 63rd Symposium of the Society for General Microbiology, Cambridge University Press
- Taylor, C. E. and Brown, D. J. F. (1997) *Nematode Vectors of Plant Viruses*, CAB Journals
- Bergelson, J. M. (2003) Virus interactions with mucosal surfaces: alternative receptors, alternative pathways. *Current Opinion in Microbiology*, 6, 386-391
- Blanc, S., Uzest, M., and Drucker, M. (2011) New research horizons in vector-transmission of plant viruses. *Current Opinion in Microbiology*, 14, 483-491
- Hogenhout, S. A. et al. (2008) Insect vector interactions with persistently transmitted viruses. *Annual Review of Phytopathology*, 46, 327-359
- Lambrechts, L. and Scott, T. W. (2009) Mode of transmission and the evolution of arbovirus virulence in mosquito vectors. *Proceedings of the Royal Society B*, 276, 1369-1378

Pereira, L. et al. (2005) Insights into viral transmission at the uterine-placental interface. Trends in Microbiology, 13, 164-174

CHAPTER 5

Attachment and Entry of Viruses into Cells

CHAPTER 5 AT A GLANCE

Overview Of Virus Replication

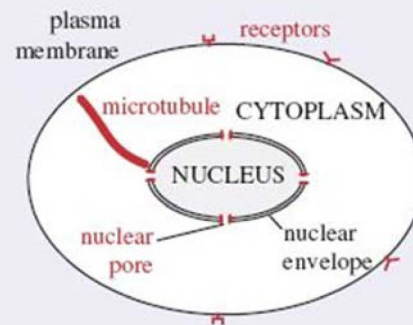
1. Attachment
2. Entry
3. Transcription
4. Translation
5. Genome replication
6. Assembly
7. Exit

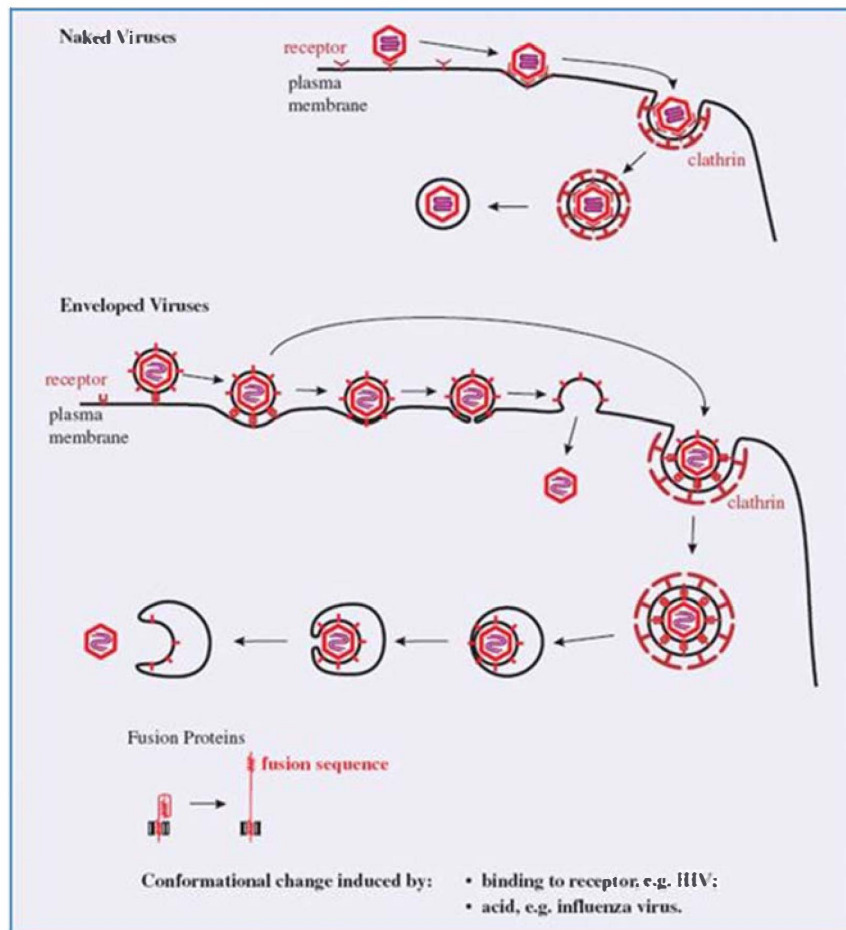
Animal Viruses

All animal viruses must cross the **plasma membrane**.

Some are transported in the cytoplasm via **microtubules**.

Some must cross the **nuclear envelope**, usually via a **nuclear pore**.





5.1 OVERVIEW OF VIRUS REPLICATION

The aim of a virus is to replicate itself, and in order to achieve this aim it needs to enter a host cell, make copies of itself, and get the new copies out of the cell. In general the process of virus replication can be broken down into seven steps:

1. Attachment of a virion to a cell;
2. Entry into the cell;
3. Transcription of virus genes into messenger RNA molecules (mRNAs);
4. Translation of virus mRNAs into virus proteins;
5. Genome replication;
6. Assembly of proteins and genomes into virions;
7. Exit of the virions from the cell.

The first letter from each step gives the abbreviation AETTGAE, which may provide a memory aid.

When trying to understand the modes of replication of different types of virus these steps provide a useful template. It is, however, a general template, because not all of the seven steps are relevant to all viruses, the steps do not always occur in the same order, and some viruses have an additional step! In the later stages of replication several steps occur concurrently. For many viruses, transcription, translation, genome replication, virion assembly, and exit can all be in progress at the same time.

In the dialogue between the cell and the intruder, the cell provides critical cues that allow the virus to undergo molecular transformations that lead to successful internalization, intracellular transport, and uncoating.

Smith and Helenius (2004)

In this chapter we discuss the first two steps of this generalized replication cycle: attachment and entry into the host cell. We concentrate mainly on the mechanisms used by animal viruses to gain entry into host cells, but bacteriophages are also considered. The following three chapters deal with the remaining five steps of the generalized replication cycle.

The first step, attachment of a virion to a cell, applies to viruses that infect animal and bacterial hosts. Before these viruses can cross the outer membrane or wall of the host cell they must first bind to specific molecules on the cell surface. Most plant viruses, on the other hand, are delivered directly into a cell by a vector (Section 4.2).

5.2 ANIMAL VIRUSES

5.2.1 Cell receptors and co-receptors

A virion attaches via one or more of its surface proteins to specific molecules on the surface of a host cell. These cellular molecules are known as receptors and the recognition of a receptor by a virion is highly specific, like a key fitting in its lock. It has been found that some viruses need to bind to a second type of cell surface molecule (a co-receptor) in order to infect a cell. In at least some cases, binding to a receptor causes a conformational change in the virus protein that enables it to bind to the co-receptor.

Receptors and co-receptors are cell surface molecules with a wide range of functions that include:

- acting as receptors for chemokines and growth factors;
- mediating cell-to-cell contact and adhesion.

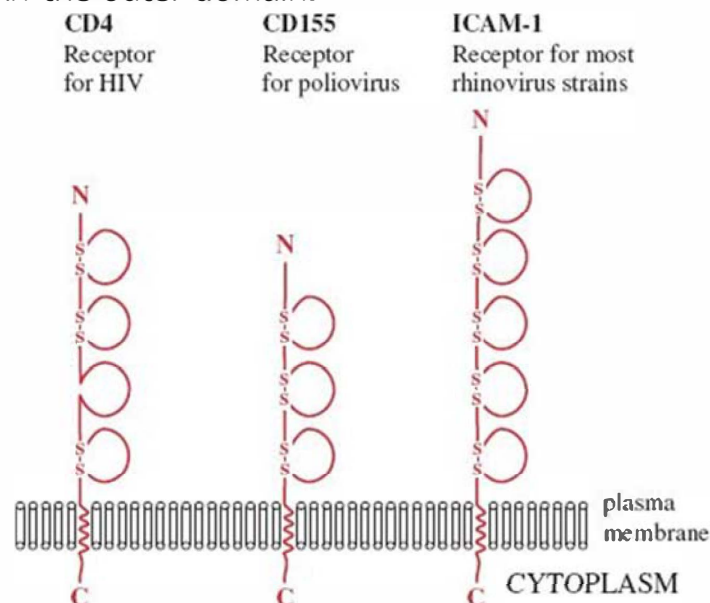
Some cell surface molecules used by viruses as receptors are sugars, but most are glycoproteins; some examples are given in Table 5.1.

Table 5.1 Examples of cell receptors, virus proteins involved in attachment, and (for enveloped viruses) fusion proteins

Virus	Cell receptor	Virus protein(s) Involved In	
		attachment to receptor	fusion
Naked viruses			
Approx. 90% of human rhinoviruses	Intercellular adhesion molecule-1 (ICAM-1)	VP1 + VP3	
Approx. 10% of human rhinoviruses	Low-density lipoprotein receptors	VP1	
Poliovirus	CD155	VP1	
Enveloped viruses			
Murine leukemia viruses	Mouse cationic amino acid transporter	SU (surface glycoprotein)	TM (transmembrane glycoprotein)
HIV-1	CD4	gp120	gp41
Influenza viruses A and B	Sialic-acid-containing glycoproteins	Hemagglutinin	Hemagglutinin
Measles virus	Signaling lymphocyte activation molecule (CD150)	Hemagglutinin	Fusion

A number of receptors are glycoproteins that are folded into domains similar to those found in immunoglobulin molecules. Three examples are shown in Figure 5.1; in each molecule the binding site for the virus is located in the outermost domain.

Figure 5.1 Cell receptors with immunoglobulin-like domains. Each of these molecules is a glycoprotein; the sugar side-chains are not shown. Each loop indicates an immunoglobulin-like domain. Most of the domains are stabilized by one or two disulfide (S-S) bonds. In each receptor the site that the virus binds to is located in the outer domain.



Many of the cell surface molecules used by viruses as receptors are in regions of the plasma membrane that are coated on the inner surface with one of the proteins clathrin or caveolin.

5.2.1.a Evidence that a cell surface molecule is a virus receptor

A search for the receptor for a particular virus might begin with the production of a panel of mono-clonal antibodies against cell surface proteins. If one of the

antibodies blocks virus binding and infectivity then this is strong evidence that the corresponding antigen is the receptor.

Further evidence that a molecule is a virus receptor might come from experiments that show one or more of the following:

- soluble derivatives of the molecule block virus binding/infectivity;
- the normal ligand for the molecule blocks virus binding/infectivity;
- introduction of the gene encoding the molecule into virus-resistant cells, and expression of that gene, makes those cells susceptible to infection.

5.2.2 Virus attachment sites

Each virion has multiple sites that can bind to receptors, and each site is made up of regions of one or more protein molecules. Some examples of these proteins are given in Table 5.1. The attachment sites of naked viruses are on the capsid surface, sometimes within depressions (e.g. poliovirus), and sometimes on ridges (e.g. foot and mouth disease virus). Poliovirus and foot and mouth disease virus are picornaviruses, and it has been demonstrated for picornaviruses in general that binding to receptors induces major structural changes in the virion (Section 14.4.1). The attachment sites of some naked viruses are on specialized structures, such as the fibers and knobs of adenoviruses (Section 3.4.2.c) and the spikes of rotaviruses (Chapter 13), while the attachment sites of enveloped viruses are on glycoproteins in the envelope.

Some virion surface proteins that bear the attachment sites are able to bind strongly to red blood cells of various species and cause them to clump, a phenomenon known as hemagglutination. The proteins responsible for hemagglutination are called hemagglutinins. Examples of viruses that can hemagglutinate are influenza viruses (Figure 16.1) and measles virus.

5.2.3 Attachment of virions to receptors

The forces that bind virions to receptors include hydrogen bonds, ionic attractions, and van der Waals forces. Sugar moieties on the receptor and/or on the virion are commonly involved in these forces. No covalent bonds are formed between virions and receptors.

Initially, a virion is weakly bound to a cell at only one or a few receptors. At this stage the attachment is reversible and the virion may detach, but if it remains attached there are opportunities for more receptors to bind, and if sufficient bind then the attachment to the cell becomes irreversible.

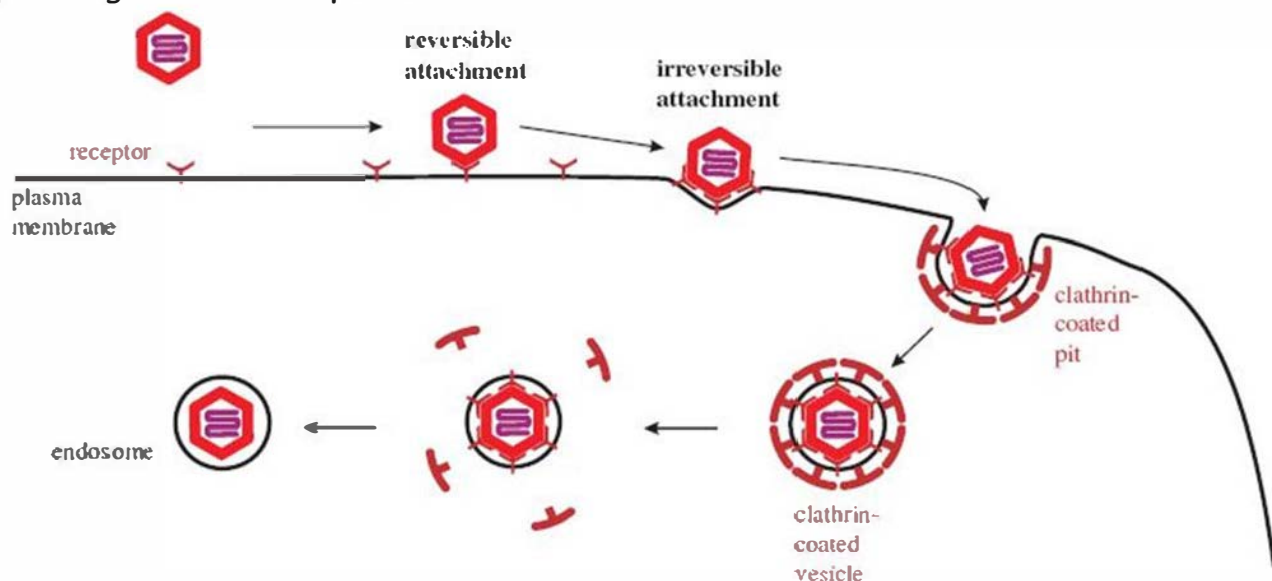
5.2.4 Entry of animal viruses into cells

After binding to receptors animal viruses must cross the plasma membrane to gain entry to the host cell. They may do this either at the cell surface or they may cross the membrane of an endosome, which is a vesicle formed by part of the plasma membrane pinching off into the cytoplasm. This process (endocytosis) is used by

cells for a variety of functions, including nutrient uptake and defense against pathogens. There are a number of endocytic mechanisms, including clathrin-mediated endocytosis and caveolin-mediated endocytosis; clathrin and caveolin are cell proteins. Most animal viruses hijack one or more of these mechanisms in order to gain access to their host cells.

Many viruses, such as adenoviruses and vesicular stomatitis virus, are taken into the host cell by clathrin-mediated endocytosis. Molecules of clathrin accumulate on the inner side of the plasma membrane at the site where the virion has attached. The clathrin forces the membrane to bend around the virion, forming a pit, which is pinched off to form a clathrin-coated vesicle (Figure 5.2). The clathrin is lost from the vesicle leaving the virion in an endosome. Some viruses, such as simian virus 40, are endocytosed at caveolin-coated regions of the plasma membrane and the virions end up in caveolin-coated endosomes. Other viruses are taken up by mechanisms that are independent of clathrin and caveolin.

Figure 5.2 Attachment and entry of a naked virion. As more receptors bind to the virion its attachment to the cell becomes irreversible. The process shown here is clathrin-mediated endocytosis. The membrane of the endosome is formed by pinching off from the plasma membrane.



An endosome may fuse with other vesicles such as lysosomes, which have a pH of 4.8 to 5.0, thus lowering the pH within the vesicle. The pH may be further lowered by a process that pumps hydrogen ions across the membrane; the process involves the hydrolysis of ATP to ADP. This acidification is important for those enveloped viruses that need to carry out acid-triggered fusion of the envelope with the vesicle membrane (Section 5.2.4.b).

5.2.4.a Entry of naked viruses

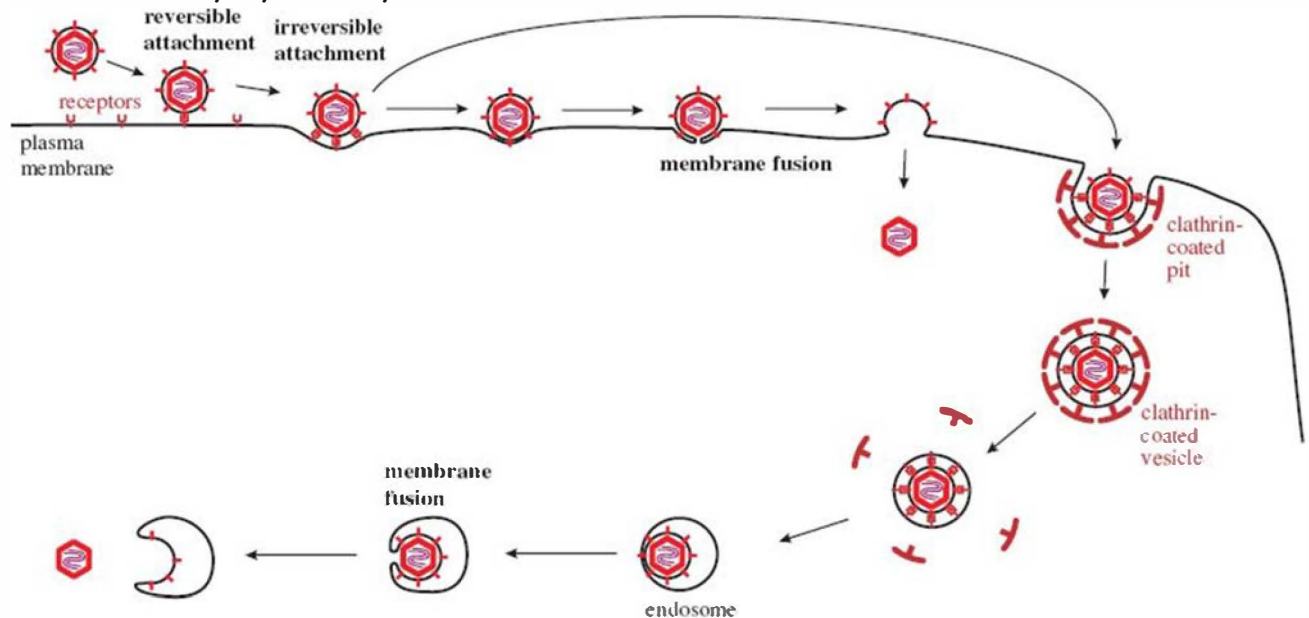
It is possible that some naked viruses deliver their genomes into their host cells through a pore formed in the plasma membrane, but for most naked viruses irreversible attachment of the virion to the cell surface leads to endocytosis. The plasma membrane flows around the virion, more receptors bind, and

eventually the virion is completely enclosed in membrane, which pinches off as an endosome (Figure 5.2). The endosome contents, however, are part of the external environment and the virus is not yet in the cytoplasm. The mechanisms by which virions, or their genomes, are released from endosomes are not fully understood.

5.2.4.b Entry of enveloped viruses

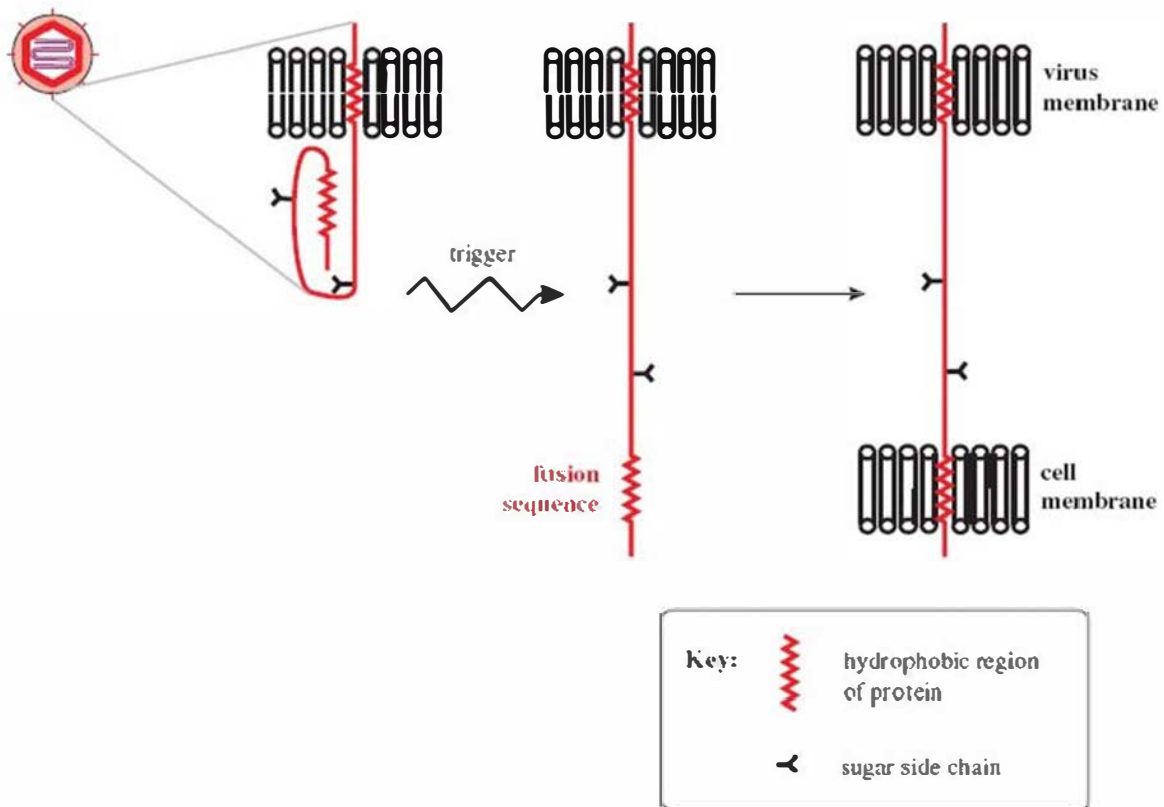
Reversible attachment of an enveloped virion may lead to irreversible attachment, as for naked viruses. There are then two processes whereby infection of the cell may occur: either fusion of the virion envelope with the plasma membrane, or endocytosis followed by fusion of the virion envelope with the endosome membrane (Figure 5.3).

Figure 5.3 Attachment and entry of an enveloped virion. Some enveloped viruses can enter a host cell both by membrane fusion at the cell surface and by endocytosis (clathrin-mediated endocytosis shown here). Other enveloped viruses can enter only by endocytosis.



Both processes involve the fusion of the virion envelope with a cell membrane, either the plasma membrane or a vesicle membrane. Lipid bilayers do not fuse spontaneously and each enveloped virus has a specialized glycoprotein responsible for membrane fusion (Figure 5.4). Some examples of these fusion proteins are given in Table 5.1.

Figure 5.4 Characteristics of fusion proteins of enveloped viruses. Fusion proteins are envelope glycoproteins that have a fusion sequence that is normally hidden. A trigger induces a conformational change in the protein that allows the fusion sequence to insert into a membrane.



Many fusion proteins are synthesized as part of a larger protein which is cleaved after synthesis. The cleavage products remain linked, either through non-covalent bonds (e.g. retroviruses) or through a disulfide bond (e.g. influenza viruses), and are found in the virion envelope as trimers. Each monomer has at least two hydrophobic sequences: a transmembrane sequence and a fusion sequence that initiates membrane fusion by inserting into a cell membrane.

The fusion sequence normally lies hidden and must become exposed in order for fusion to take place. For some fusion proteins the conformational change that exposes the fusion sequence is triggered by the virus binding to a receptor. Other fusion proteins undergo their conformational change once the virus is within an endosome. The low pH of the endosome may directly alter the conformation of the fusion protein, or it may activate a cell protease which cleaves the fusion protein.

If the fusion sequence becomes exposed while the virion is at the cell surface, then infection could occur either by fusion with the plasma membrane or by endocytosis. If, however, the trigger is a low pH, then endocytosis is the only option. There are therefore two categories of membrane fusion:

- pH-independent fusion, e.g. herpesviruses, HIV;
- acid-triggered fusion, e.g. influenza viruses, rhabdoviruses.

Once the fusion sequence is exposed it can insert into the target membrane (plasma membrane or endosome membrane). A further conformational change in the fusion protein pulls the two membranes together. The outer layers of the virus membrane and the cell membrane fuse, a stage known as hemifusion, and then the inner layers fuse, completing the fusion process. The fusion protein is now in a lower energy state.

5.2.5 Intracellular transport

Once in the cell, the virus, or at least its genome, may be delivered to a particular destination, such as the nucleus. The destination may be reached using cell transport systems, including microtubules.

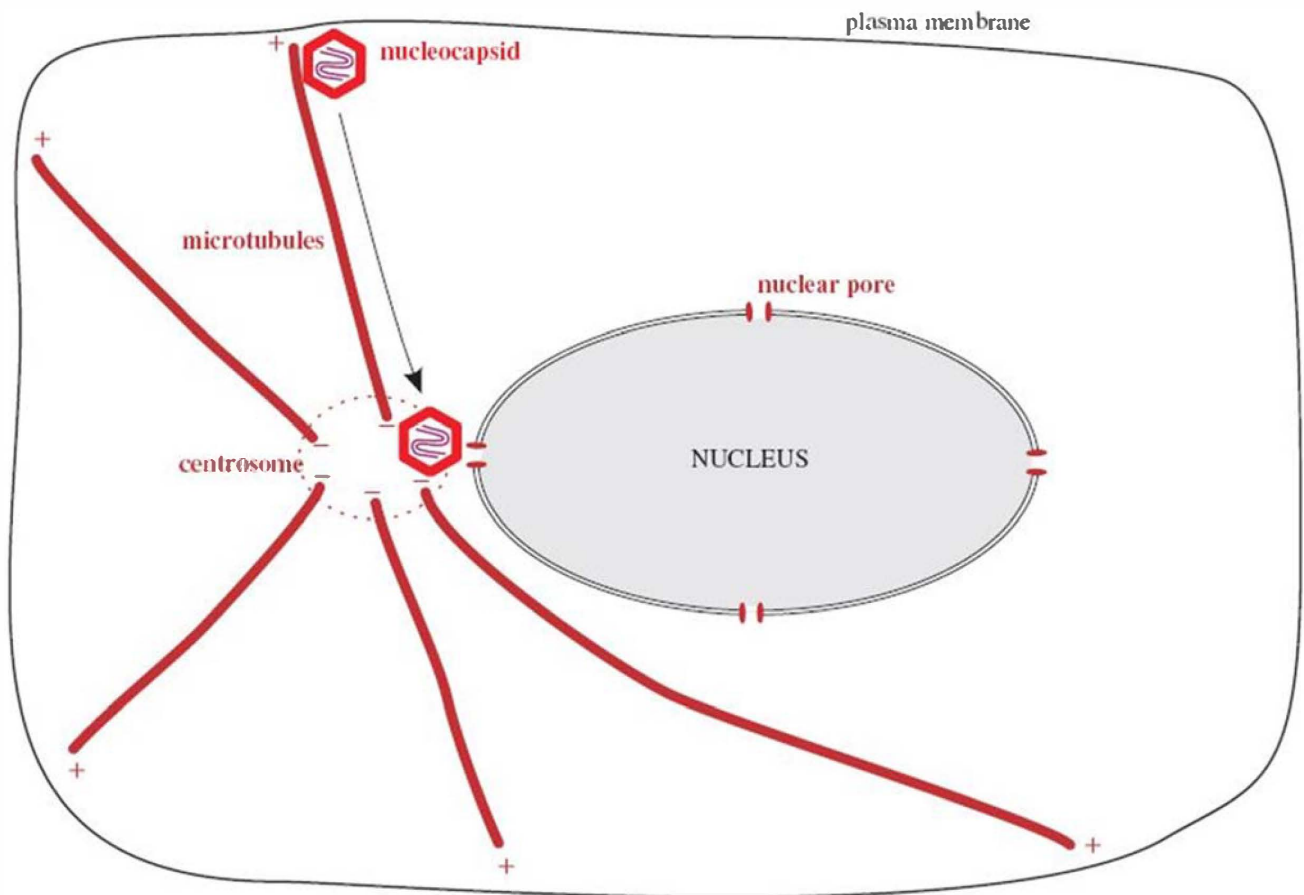
Most RNA viruses of eukaryotes replicate in the cytoplasm and have little or no requirement for the nucleus. The influenza viruses, however, are exceptions as they require the cell splicing machinery, so their genomes must be delivered into the nucleus.

Retroviruses too are RNA viruses that replicate their genomes in the nucleus. They copy their genomes to DNA in the cytoplasm, but for most retroviruses the replication cycle is unable to proceed further until mitosis begins. During mitosis the nuclear envelope is temporarily broken down and the virus DNA (with associated proteins) is able to enter the nuclear compartment. These viruses therefore can replicate only in cells that are dividing. The DNA (with associated proteins) of the lentiviruses (a group of retroviruses that includes HIV), however, can be transported into an intact nucleus. The lentiviruses can therefore replicate in non-dividing cells.

Some DNA viruses, such as iridoviruses and poxviruses, replicate in the cytoplasm of eukaryotic cells, but most DNA viruses replicate in the nucleus. For these viruses (and the influenza viruses and the lentiviruses) the virus genome must be transported to the nuclear envelope and then across it. The proteins of some of these viruses have sequences that allow them to attach to microtubules.

Microtubules are components of the cytoskeleton, providing support and acting as tracks for the transport of materials to particular sites in the cell. Microtubules consist of hollow cylinders, 25 nm in diameter, composed of the proteins α - and β -tubulin. The ends of each microtubule are designated as plus and minus. In most animal cells, the plus ends are located near the plasma membrane, while the minus ends are attached to a structure called the centrosome (also referred to as the microtubule organizing center) near the nucleus (Figure 5.5).

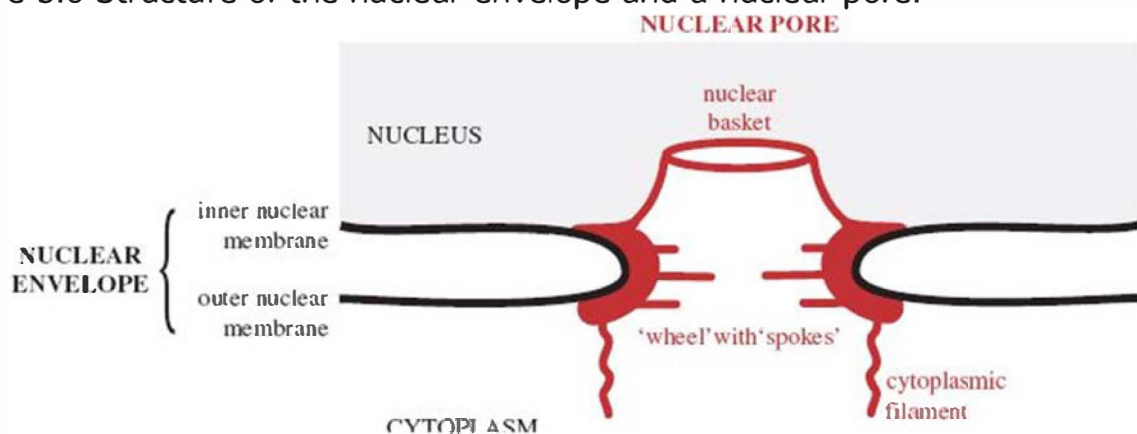
Figure 5.5 Transport of a nucleocapsid along a microtubule. The nucleocapsid is transported from the plus end of a microtubule, near the plasma membrane, to the centrosome, which is located close to the nucleus.



Proteins, known as motor proteins, move themselves and any cargo along the microtubules. A number of viruses (including herpesviruses, adenoviruses, parvoviruses, and retroviruses) exploit this transport system to take their nucleocapsid, or a structure derived from it, from the periphery of the cytoplasm to a location close to the nucleus.

The nuclear envelope is composed of two membranes (Figure 5.6), each of which is a lipid bilayer. Embedded within the nuclear envelope are nuclear pores, which number between 3000 and 5000 in a typical growing cell. The nuclear pores act as gatekeepers, controlling the transport of materials in and out of the nucleus. Traffic into the nucleus includes ribosomal proteins and histones, while traffic out of the nucleus includes ribosomes, mRNAs, and transfer RNAs.

Figure 5.6 Structure of the nuclear envelope and a nuclear pore.



Each nuclear pore is constructed from a complex of about 30 protein species that control the movement of materials in and out of the nucleus. The protein molecules are assembled into structures, including a basket protruding into the nucleus and eight filaments that protrude into the cytoplasm (Figure 5.6). At the center of the pore is a channel through which molecules can diffuse, though most molecules and particles pass through the pore complexed with carrier proteins (importins and exportins). The channel can open to allow the passage of particles up to 25 nm across, and there have been reports of larger particles being moved through nuclear pores. Small nucleocapsids/virions, such as the parvovirus virion, can pass through, but larger viruses must either shed some of their load to form slimmer structures or uncoat at a nuclear pore.

5.2.6 Genome uncoating

Uncoating can be defined as the complete or partial removal of the capsid to release the virus genome. For an animal virus this process takes place in the cytoplasm, at a nuclear pore or in the nucleus.

It should be noted that successful entry of a virion into a cell is not always followed by virus replication. The host's intracellular defenses, such as lysosomal enzymes, may inactivate infectivity before or after uncoating. In some cases the virus genome may initiate a latent infection (Section 9.3.1) rather than a complete replication cycle.

In at least some cases, however, provided that the uncoated virus genome survives intact, and is at the correct location in the cell, then its transcription (or in some cases translation) can begin. These aspects are dealt with in the next chapter.

5.3 BACTERIOPHAGES

Like animal viruses, phages bind specifically to cell surface molecules that function as receptors and co-receptors. For many phages these molecules are on the surface of the host cell wall, but for some phages they are on the surface of other structures (pili, flagella, or capsules). Many phages attach via specialized structures, such as the tail fibers of phage T4. As with animal viruses, the attachment of a phage to its host is initially reversible, and becomes irreversible after attachment to further receptors and/or co-receptors.

Infection of a cell with a phage involves entry into the cell of the virus genome and perhaps a few associated proteins, the capsid and any associated appendages remaining at the cell surface. This is in contrast to most animal and plant viruses, where the entire virion, or at least the nucleocapsid, enters the host cell. Unless the host is a mycoplasma (mycoplasmas do not have cell walls), delivery of a phage genome into a cell requires penetration of a cell wall and perhaps also a slime layer or capsule; the virions of many phages carry enzymes, such as lysozymes, to aid this process.

The attachment and entry of a number of phages are discussed in Chapter 20.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- outline a generalized scheme of virus replication involving seven steps;
- describe how animal viruses attach to and enter their host cells;
- differentiate between the entry mechanisms of naked and enveloped animal viruses;
- describe the roles of cell components in the delivery of some viral genomes to the nucleus;
- outline the infection mechanisms of phages.

SOURCES OF FURTHER INFORMATION

Books

Johnson, J. E. (ed.) (2010) Cell Entry by Non-Enveloped Viruses, Current Topics in Microbiology and Immunology series, Volume 343, Springer

Marsh, M. (ed.) (2005) Membrane Trafficking in Viral Replication, Springer

Journals

Cohen, S., Au, S., and PantHC, N. (2010) How viruses access the nucleus. *Biochimica et Biophysica Acta (BBA) вЂ” Molecular Cell Research*, 1813, 1634вЂ”1645

Haywood, A. M. (2010) Membrane uncoating of intact enveloped viruses. *Journal of Virology*, 84, 10946вЂ”10955

Inoue, T., Moore, P., and Tsai, B. (2011) How viruses and toxins disassemble to enter host cells. *Annual Review of Microbiology*, 65, 287вЂ”305

Lozach, P.-Y., Huotari, J., and Helenius, A. (2011) Late-penetrating viruses. *Current Opinion in Virology*, 1, 35вЂ”43

Mercer, J., Schelhaas, M., and Helenius, A. (2010) Virus entry by endocytosis. *Annual Review of Biochemistry*, 79, 803вЂ”833

Molineux, I. J. (2006) Fifty-three years since Hershey and Chase; much ado about pressure but which pressure is it? *Virology*, 344, 221вЂ”229

Moyer, C. J. and Nemerow, G. R. (2011) Viral weapons of membrane destruction: variable modes of membrane penetration by non-enveloped viruses. *Current Opinion in Virology*, 1, 44вЂ”49

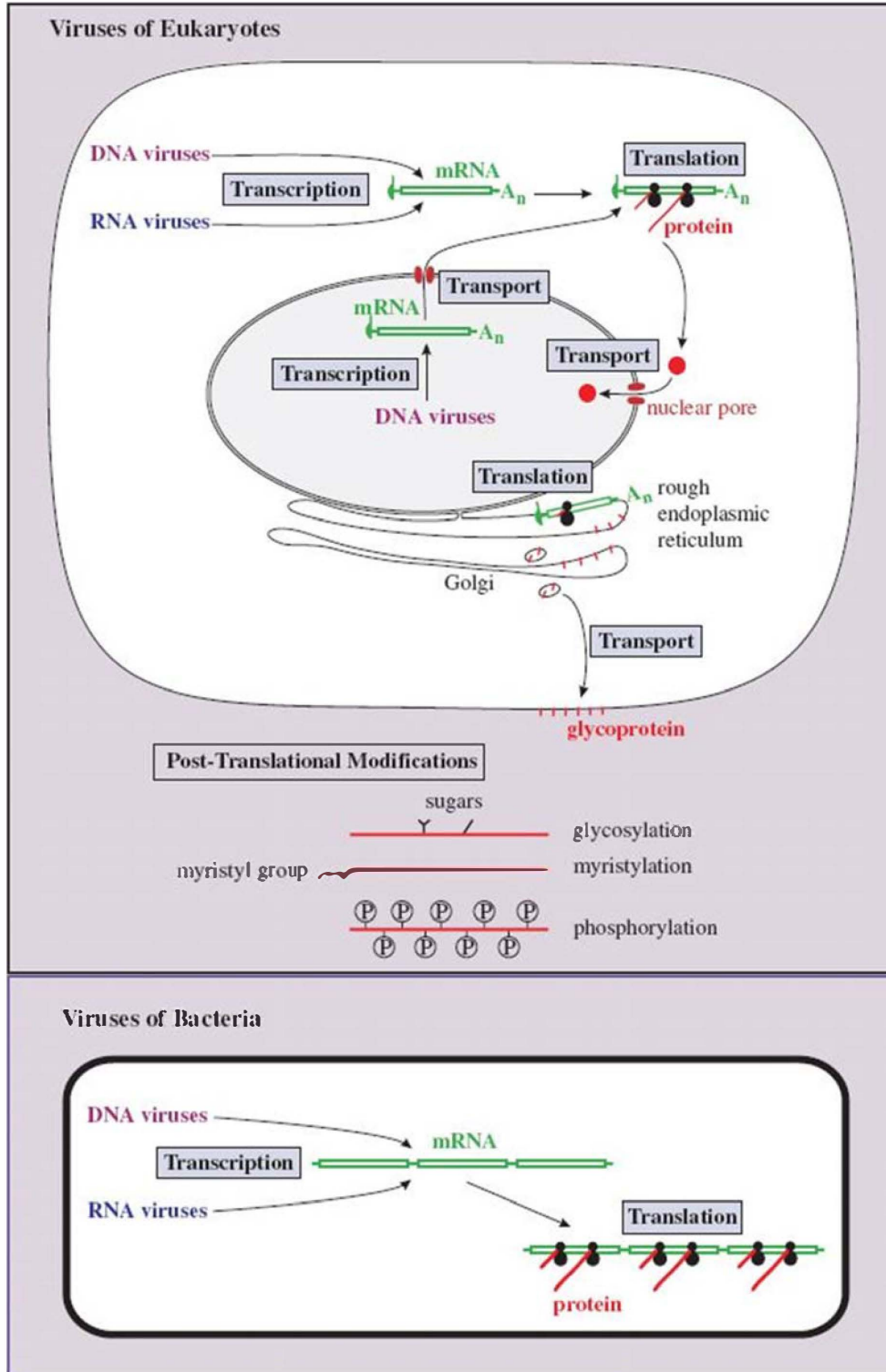
Plempner, R. K. (2011) Cell entry of enveloped viruses. *Current Opinion in Virology*, 1, 92вЂ”100

Smith, A. E. and Helenius, A. (2004) How viruses enter animal cells. Science, 304, 237B-F242

Chapter 6

Transcription, Translation, and Transport

CHAPTER 6 AT A GLANCE



6.1 INTRODUCTION TO TRANSCRIPTION, TRANSLATION, AND TRANSPORT

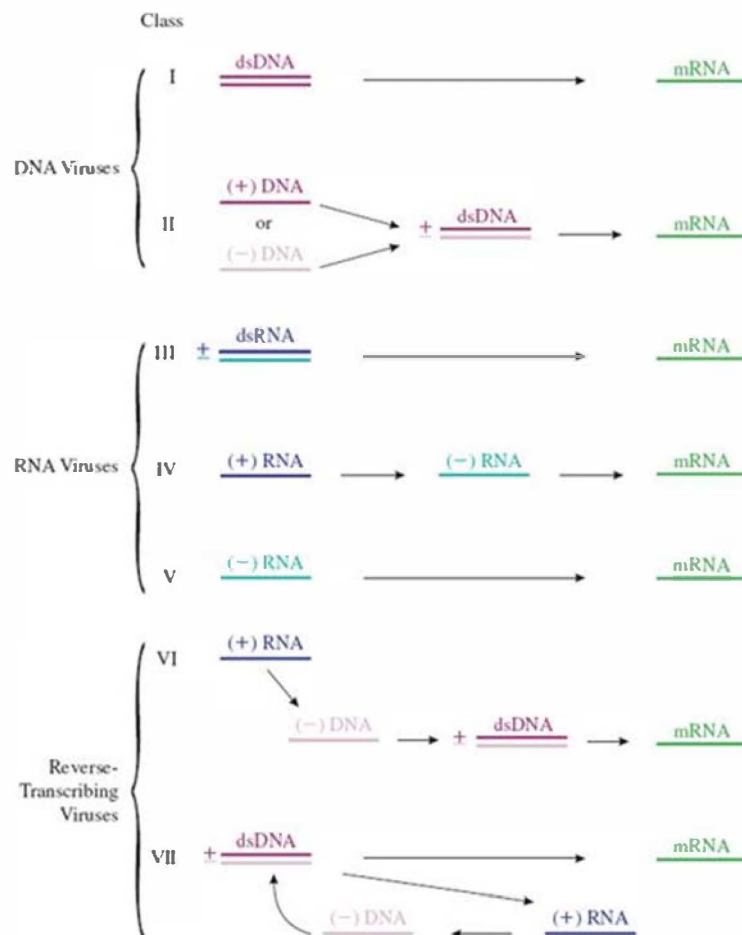
In this chapter we deal with three terms derived from the Latin *trans*, meaning across. For our purposes, transcription refers to the writing across of genetic information from a sequence of bases in a nucleic acid to the complementary sequence in messenger RNA (mRNA), while translation converts the genetic information from the language of bases in nucleic acids to the language of amino acids in proteins. Transcription and translation are steps 3 and 4 of our generalized replication cycle (Section 5.1).

We also discuss in this chapter the transport of virus proteins and RNAs to particular locations in infected cells. We start with an overview of virus transcription, and then we discuss these three “trans” processes in eukaryotic cells. At the end of the chapter we point out some aspects of the processes that are different in bacterial cells.

6.2 TRANSCRIPTION OF VIRUS GENOMES

We have seen how there are four main categories of virus genome: dsDNA, ssDNA, dsRNA and ssRNA (Section 3.2). Because of distinct modes of transcription within the dsDNA and ssRNA categories a total of seven classes of viruses can be recognized (Figure 6.1).

Figure 6.1 Transcription of virus genomes. (+) RNA and (+) DNA have the same sequence as the mRNA, and (–) RNA and (–) DNA have the sequence complementary to the mRNA (except that in DNA thymine replaces uracil). (+) and (–) strands are not indicated for the dsDNA of the Class I viruses as the genomes of most of these viruses have open reading frames (ORFs) in both directions. (+) and (–) strands are indicated for the ssDNA of the Class II viruses. Most of these viruses have either a (+) or a (–) strand genome. A (+) RNA genome (dark blue) has the same sequence as the corresponding mRNA (green); the molecules are shown in different colors to indicate their different functions. In Class VII viruses the (+) RNA shown in blue (pregenome RNA) functions as a template for synthesis of DNA, some of which is used as a template for further transcription. Some ssDNA viruses and some ssRNA viruses have ambisense genomes. This means that that the polarity of the genome is part (+) and part (–).



This division of the viruses into classes based on genome type and mode of transcription was first suggested by David Baltimore and this scheme of virus classification is named after him. He initially proposed six classes.

In the summary of the scheme depicted in Figure 6.1 most of the nucleic acid strands are labeled (+) or (-). This labeling is relative to the virus mRNA, which is always designated (+). A nucleic acid strand that has the same sequence as mRNA is labeled (+) and a nucleic acid strand that has the sequence complementary to the mRNA is labeled (-).

The viruses with (+) RNA genomes (Classes IV and VI) have the same sequence as the virus mRNA. When these viruses infect cells, however, only the Class IV genomes can function as mRNA. These viruses are commonly referred to as plus-strand (or positive-strand) RNA viruses. The Class V viruses are commonly referred to as minus-strand (or negative-strand) RNA viruses.

Class VI viruses must first reverse-transcribe their ssRNA genomes to dsDNA before mRNA can be synthesized. Because they carry out transcription in reverse (RNA to DNA) Class VI viruses are known as retroviruses. The ability of some DNA viruses to carry out reverse transcription was discovered later; these viruses became known as pararetroviruses and Class VII was formed to accommodate them.

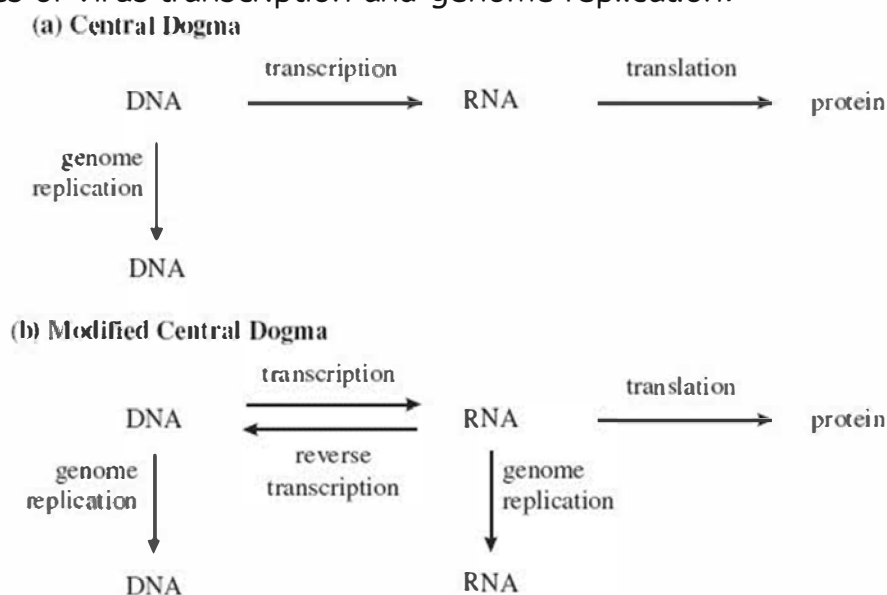
There are a few single-stranded nucleic acids of viruses where there is a mixture of (+) and (-) polarity within the strand. Genomes of this type are known as ambisense, a word derived from the Latin ambi, meaning "on both sides" (as in

ambidextrous). Examples of ambisense genomes include the ssDNA genomes of the geminiviruses, which are plant viruses, and the ssRNA genomes of the arenaviruses, which are animal viruses and include the causative agent of Lassa fever.

6.2.1 Modifications to the central dogma

In 1958 a “central dogma of molecular biology” was proposed by Francis Crick, along with James Watson, one of his collaborators in deducing the structure of DNA. The dogma stated that the flow of genetic information is always from DNA to RNA and then to protein, with genetic information transmitted from one generation to the next through copying from DNA to DNA (Figure 6.2(a)). Increasing understanding of how viruses replicate their genomes necessitated some modifications to this dogma in 1970 (Figure 6.2(b)). Many viruses have RNA genomes that are copied to RNA, some viruses copy from RNA to DNA, and it is now known that cells can copy from RNA to DNA.

Figure 6.2 (a) Francis Crick’s central dogma of molecular biology, which proposed that genetic information is transmitted from DNA to RNA to protein, and from DNA to DNA. (b) Modifications to the central dogma, required mainly because of the various modes of virus transcription and genome replication.

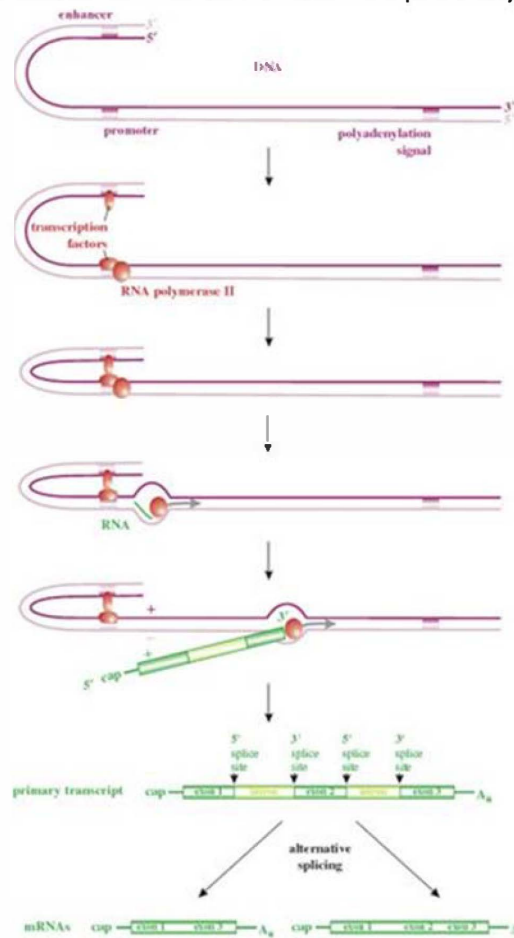


6.3 TRANSCRIPTION IN EUKARYOTES

We start this section with a brief summary of transcription from eukaryotic cell genes, as many viruses transcribe their genes by similar processes, some of them using parts of the cell transcription machinery (Figure 6.3). The expression of a gene is controlled by sequences in the DNA, including:

- a promoter – binds transcription factors and the RNA polymerase;
- enhancers – bind transcription factors that significantly boost the rate of transcription.

Figure 6.3 Transcription from dsDNA in eukaryotes. Transcription is initiated after transcription factors bind to sequences within the promoter and enhancer(s). The primary transcript is capped at the 5' end and polyadenylated at the 3' end. The mRNAs are formed by removal of introns from the primary transcript.



6.3.1 Promoters and enhancers

The promoters of many eukaryotic cell and virus genes contain the consensus sequence:

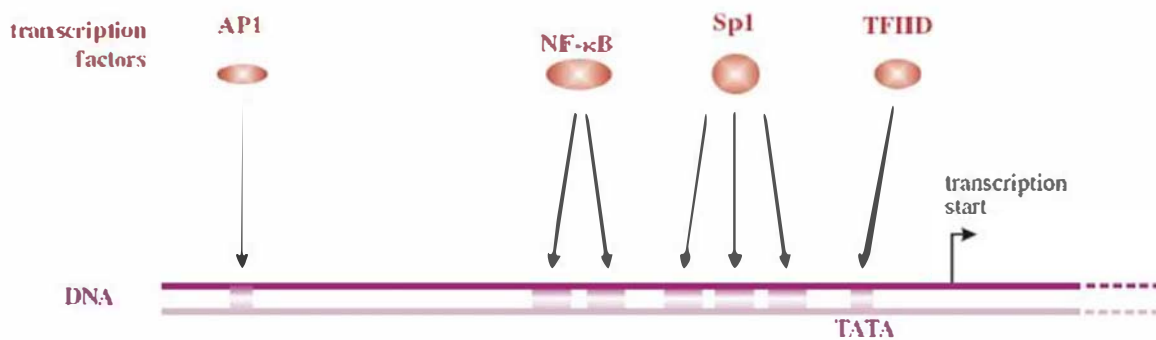
T A T A A/T A A/T A/G

A/T and A/G indicate alternative nucleotides at those sites.

The sequence is known as a TATA box and is usually located 25–30 bp upstream from the transcription start site. A TATA box is present in the single promoter of HIV-1 (Figure 6.4), but in only one of the four promoters of hepatitis B virus (Section 19.8.3).

Figure 6.4 Part of the long terminal repeat in the HIV-1 genome, indicating the locations of some of the sequences that control transcription, and the transcription factors that bind to those sequences. The promoter (TATA box) binds TFIID, and enhancers bind:

- AP1 (activator protein 1)
- NF-κB (nuclear factor κB)
- Sp1 (stimulatory protein 1).



Enhancers contain sequences that bind transcription factors that may then interact with transcription factors bound to the promoter. These interactions may increase the rate of transcription starts by RNA polymerase II. Many enhancers are found only in specific cell types, for example liver cells.

6.3.2 Transcription factors

Transcription factors are proteins that bind to specific promoter and enhancer sequences to control the expression of genes. Many transcription factors activate genes, but some play other roles, including repression of gene expression.

Transcription factor IID (TFIID) is an example of a cell transcription factor. TFIID is a complex of 13 proteins, one of which is the TATA box binding protein. After TFIID has bound to the TATA box other transcription factors (TFIIA, IIB, IIE, IIF, and IIH) and RNA polymerase II bind.

Many viruses use cell transcription factors to activate or repress transcription of their genes. An example is shown in Figure 6.4, which depicts part of the HIV-1 genome in its DNA form (Chapter 18). The promoter includes a TATA box and there is a large number of enhancers.

Tissue-specific transcription factors are required by certain viruses that are specific to a particular tissue, such as liver. Some viruses produce their own transcription factors, such as herpes simplex virus VP16, which is a component of the virion (Section 11.5.2), and human T-lymphotropic virus 1 Tax protein (Section 23.10.5).

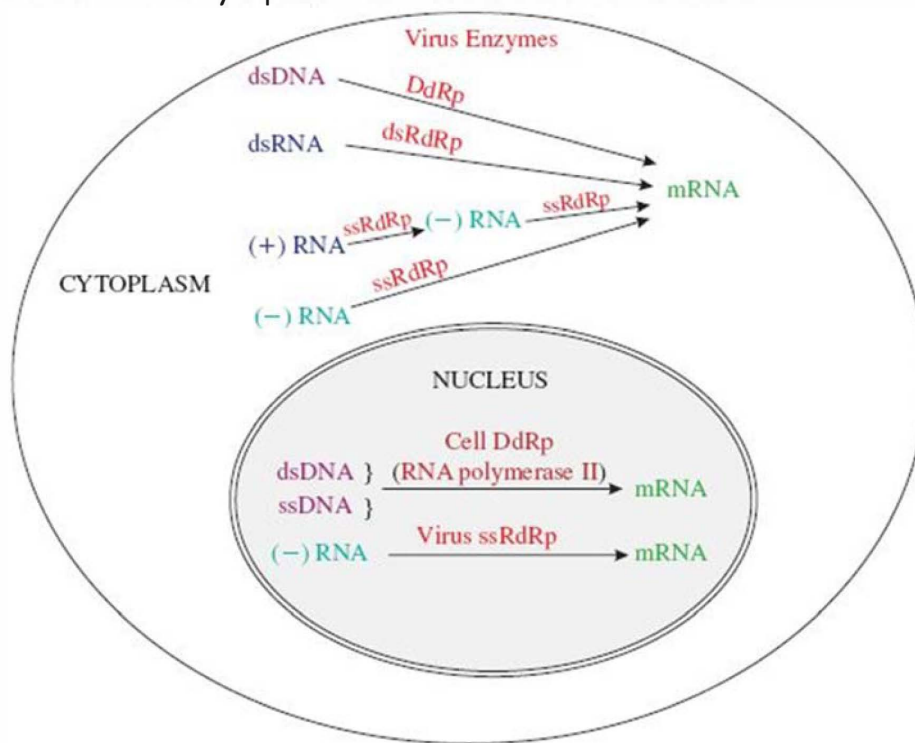
All organisms regulate expression of their genes. A frog has different genes switched on depending on whether it is in the embryo, tadpole or adult stage. Similarly, a virus may have different genes active at different stages of its replication cycle. For some viruses there are two phases of gene expression (early and late), while for others three phases (e.g. herpesviruses; see Section 11.5.2) or four phases (e.g. baculoviruses) can be distinguished.

6.3.3 Transcriptases

Transcriptase is a general term for an enzyme that carries out transcription. Most viruses encode their own transcriptase (Figure 6.5).

Figure 6.5 Enzymes used by viruses of eukaryotes to transcribe their genomes to mRNA. A (+) RNA genome has the same sequence as the virus mRNA. A (-) RNA genome has the sequence complementary to that of the virus mRNA. Transcription from DNA in the nucleus applies not only to those DNA viruses that replicate in the

nucleus, but also to the reverse transcribing viruses. Some minus-strand RNA viruses transcribe in the cytoplasm and some in the nucleus.



DdRp: DNA-dependent RNA polymerase
 dsRdRp: double-stranded RNA-dependent RNA polymerase
 ssRdRp: single-stranded RNA-dependent RNA polymerase

A DNA virus needs a DNA-dependent RNA polymerase to transcribe its genes into mRNA. Viruses that carry out transcription in the nucleus generally use the cell RNA polymerase II; these include the retroviruses and many DNA viruses. DNA viruses that replicate in the cytoplasm use a virus-encoded enzyme because there is no appropriate cell enzyme in the cytoplasm.

An RNA virus (apart from the retroviruses) needs an RNA-dependent RNA polymerase to transcribe its genes into mRNA. Each virus in classes III, IV, and V encodes its own enzyme, in spite of the fact that the cells of many eukaryotes encode ssRNA-dependent RNA polymerases.

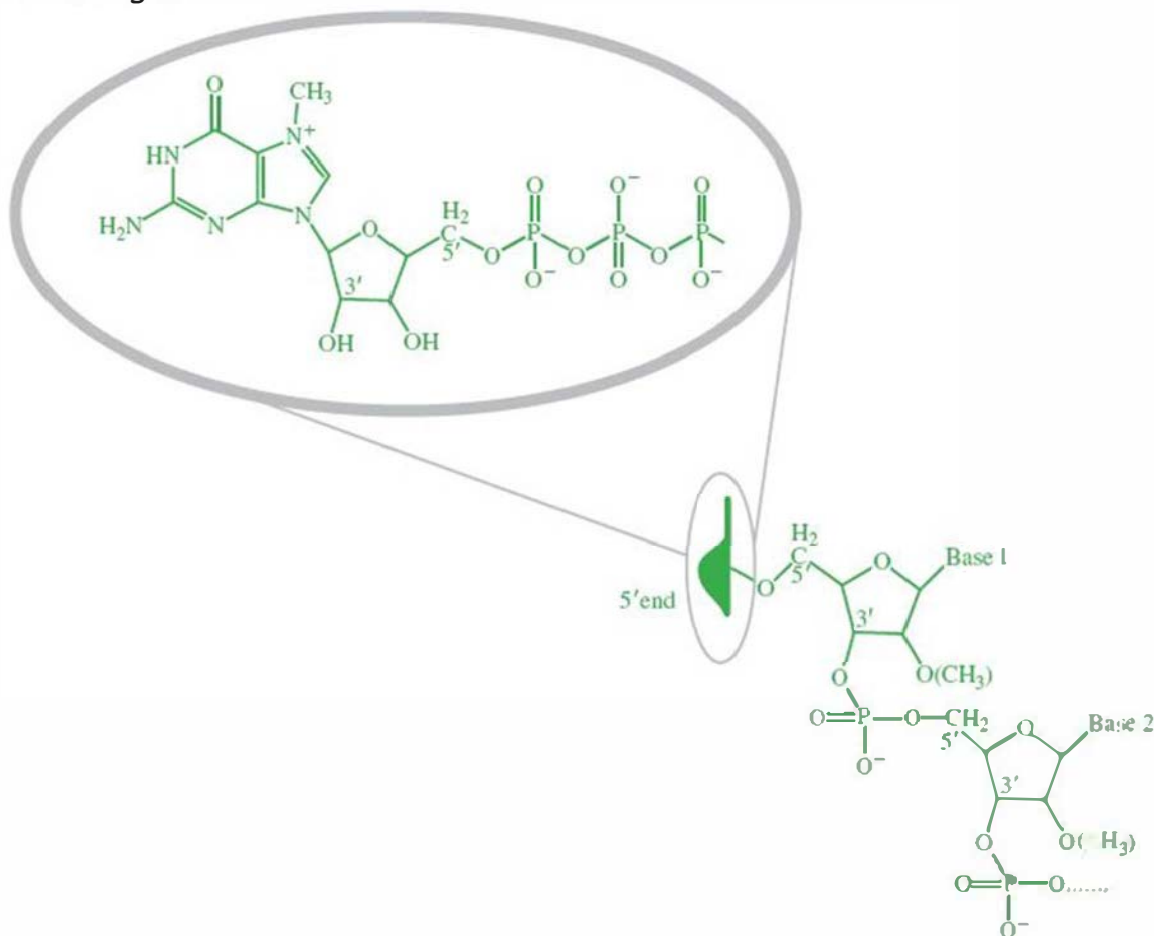
Most viruses that encode their own transcriptase have the enzyme in the virion so that it is available to transcribe the virus genome when a cell is infected. Exceptions are the plus-strand RNA viruses, which must first synthesize the enzyme before transcription can begin.

The retroviruses and the pararetroviruses perform reverse transcription (Section 6.2) using enzymes known as reverse transcriptases. These enzymes are RNA-dependent DNA polymerases, but they also have DNA-dependent DNA polymerase activity, as the process of reverse transcription involves synthesis of DNA using both RNA and DNA as the template.

6.3.4 Capping transcripts

Soon after RNA synthesis has begun, and while transcription is continuing, most transcripts are "capped" at the 5' end (Figure 6.3). The cap is a guanosine triphosphate joined to the end nucleotide by a 5'–5' linkage, rather than the normal 5'–3' linkage. A methyl group is added to the guanosine, and in some cases to one or both of the ribose residues on the first and second nucleotides (Figure 6.6). Throughout the book we shall use a cartoon cap to depict a cap on the 5' end of an RNA molecule.

Figure 6.6 Cap on 5' end of RNA. The inset shows the structure of the cartoon cap, which depicts a methylated guanosine triphosphate linked to the first nucleotide by a 5'–5' linkage.



Most eukaryotic cell and viral mRNAs have a cap at their 5' end. The cap is involved in:

- transport of mRNA from the nucleus to the cytoplasm;
- protection of the mRNA from degradation by exonucleases;
- initiation of translation.

The enzymes that normally cap RNA are listed in Table 6.1. The capping enzymes of the cell are located in the nucleus, and most of the viruses that carry out transcription in the nucleus (Figure 6.5) use these enzymes. Influenza viruses transcribe their genomes in the nucleus, but they do not use the capping enzymes. Instead they "snatch" caps from cell mRNAs; the complex of virus proteins making

up the RNA polymerase binds to cellular mRNA, and then an endonuclease in the complex cleaves the RNA, generally 9–15 nucleotides from the 5' end. The resulting capped oligonucleotide acts as a primer to initiate synthesis of viral mRNA.

Table 6.1 Enzymes involved in capping the 5' ends of RNAs

Enzyme	Activity
RNA triphosphatase	Removes one of the three phosphate groups from the 5' end of the RNA
Guanylyl transferase	Adds guanosine monophosphate to the diphosphate terminus
Methyl transferases	Add methyl group(s)

Of the viruses that replicate in the cytoplasm, many encode their own capping enzymes (e.g. poxviruses, reoviruses, coronaviruses), while some “snatch” caps from cell mRNAs (e.g. bunyaviruses) and others produce transcripts that are not capped (e.g. picornaviruses). Translation of mRNAs without caps is initiated by a mechanism that is not dependent upon a cap (Section 6.4.1).

6.3.5 Polyadenylation of transcripts

A series of adenosine residues (a polyadenylate tail; poly(A) tail) is added to the 3' end of most primary transcripts of eukaryotes and their viruses. Polyadenylation probably increases the stability of mRNAs, and the poly(A) tail plays a role in the initiation of translation (Section 6.4.1). These functions can be provided in other ways, however, as some viruses, such as the reoviruses (Section 13.3.2), do not polyadenylate their mRNAs.

In most cases there is a polyadenylation signal (Figure 6.3) about 10–30 bases upstream of the polyadenylation site. The polyadenylation signal AATAAA was first characterized in simian virus 40 in 1981. It has since been found that this sequence is used by many other animal viruses, such as HIV-1 (Section 18.4.3) and Rous sarcoma virus, as well as by animal cells. Some viruses use other sequences as polyadenylation signals; for example, the mammalian hepadnaviruses (Section 19.8.3) use TATAAA, a sequence that can function as a TATA box in other contexts! In most cases the poly(A) tail is added by the following mechanism. During transcription the RNA polymerase proceeds along the template past the polyadenylation signal and the polyadenylation site. The newly synthesized RNA is then cleaved at the polyadenylation site and the poly(A) tail is added step by step by a complex of proteins, including a poly(A) polymerase. Alternative mechanisms to polyadenylate mRNAs have been evolved by some viruses, including the picornaviruses (Section 14.4.4) and the rhabdoviruses (Section 15.4.2).

6.3.6 Splicing transcripts

Some eukaryotic cell primary transcripts are functional mRNAs, but most contain sequences (introns) that are removed in the nucleus. The remaining sequences (exons) are spliced at specific donor sites and acceptor sites to produce the mRNAs (Figure 6.3). A primary transcript may be cut and spliced in more than one way to produce two or more mRNA species.

Some primary transcripts of viruses that replicate in the nucleus are processed in the same way to produce the virus mRNAs. The first evidence of split genes, as they are known, was reported in 1977 after studies with adenoviruses. Further examples of viruses that have split genes are herpesviruses (Section 11.5.2), parvoviruses (Section 12.4.3) and retroviruses (Section 17.3.4).

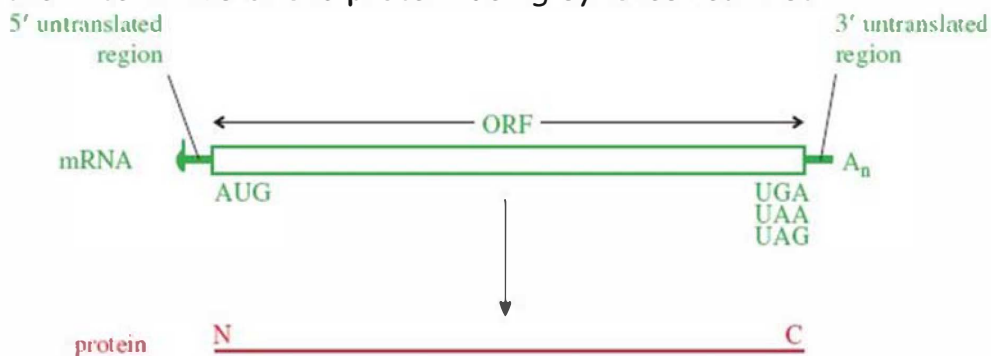
The simplest type of split gene consists of two exons separated by one intron, but some are much more complex; gene K15 of Kaposi's sarcoma-associated herpesvirus has eight exons and seven introns. The HIV-1 genome has a number of splice donor sites and acceptor sites; cutting and splicing the primary transcript results in more than 30 species of mRNA (Section 18.4.3).

6.4 TRANSLATION IN EUKARYOTES

Most viruses synthesize their proteins utilizing the entire translation machinery of the host cell (ribosomes, enzymes, other proteins, and transfer RNAs (tRNAs)). Some large DNA viruses, however, encode their own tRNAs, while the mimivirus (Section 3.2.1) also encodes a number of the enzymes and other protein components of the translation machinery.

A typical eukaryotic mRNA is monocistronic; that is, it has one ORF from which one protein is translated (Figure 6.7). Sequences upstream and downstream of the ORF are not translated. Some virus mRNAs have large ORFs that encode polyproteins, large proteins that are cleaved to form two or more functional proteins.

Figure 6.7 Translation from a monocistronic mRNA. There is one open reading frame (ORF), usually starting at the first AUG codon from the 5' end of the mRNA and ending at a stop codon (UGA, UAA, or UAG). Translation occurs in the 5' to 3' direction, the N terminus of the protein being synthesized first.



6.4.1 Initiation of translation

As we have noted, most eukaryotic cell and virus mRNAs have a methylated cap at the 5' end and a poly(A) sequence at the 3' end. These structures play key roles in the initiation of translation. The cap is especially important; it is the binding site for eukaryotic initiation factors (eIFs), a 40S ribosomal subunit, and a methionine tRNA charged with methionine (Figure 6.8). A poly(A)-binding protein binds to the poly(A) tail and interacts with proteins bound at the 5' end; this circularizes the mRNA and leads to initiation of translation.

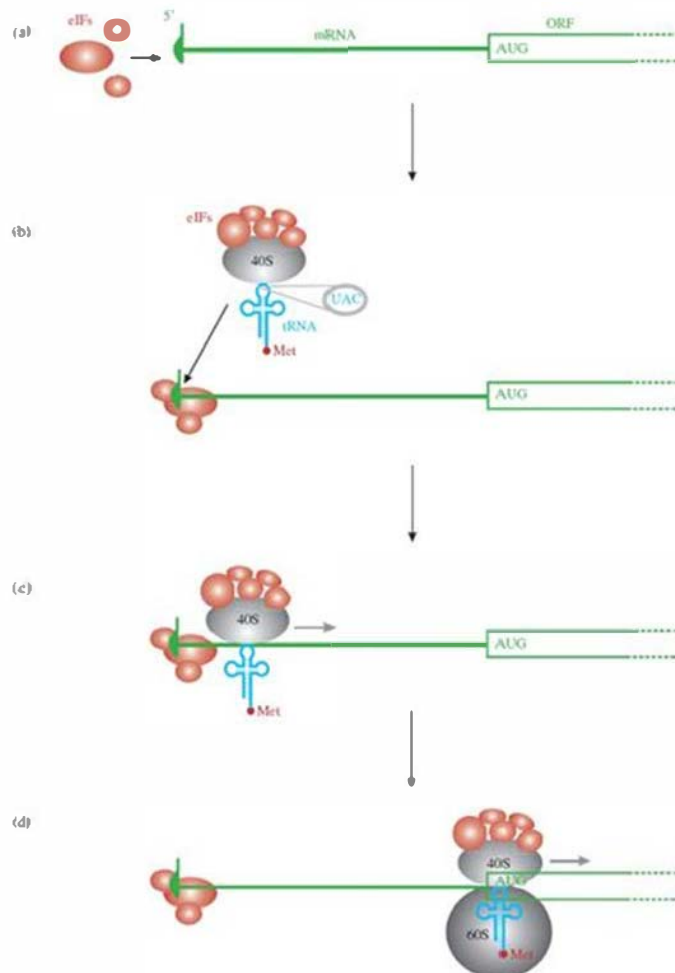
Figure 6.8 Initiation of translation on a capped mRNA.

(a) Eukaryotic initiation factors (eIFs) bind at the 5' end of the mRNA.

(b) A complex of a 40S ribosomal subunit, further eIFs, and a methionine linked to its tRNA bind. The ends of the mRNA become linked (not shown).

(c) The complex scans from the 5' end of the mRNA.

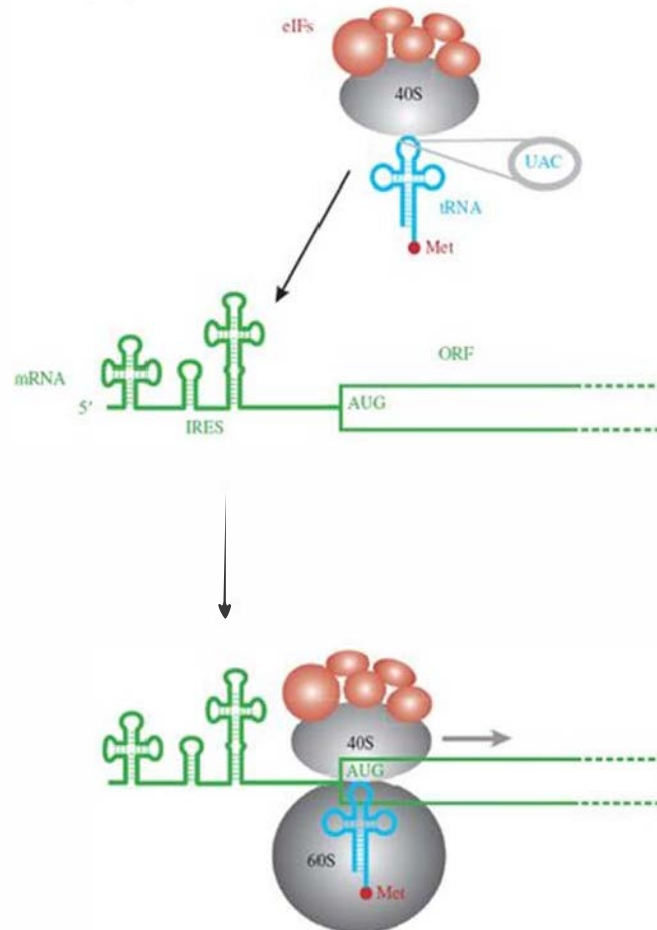
(d) When the first AUG codon is reached it is recognized by the anticodon UAC in the tRNA; a 60S ribosomal subunit is bound.



The 40S ribosomal subunit is moved along the RNA in the 5' to 3' direction, scanning until an initiation codon is encountered in an appropriate sequence context. The initiation codon is normally the first AUG from the 5' end. Some viruses, however, use other initiation codons; Sendai virus uses ACG as the initiation codon for one of its genes.

For mRNAs that are not capped, circularization and initiation of translation occur by different mechanisms. Initiation involves a reduced set of eIFs binding, not at the 5' end, but at an internal ribosome entry site (IRES), which has a high degree of secondary structure (Figure 6.9).

Figure 6.9 Initiation of translation on an mRNA that is not capped. A complex of a 40S ribosomal subunit, eIFs, and a methionine linked to its tRNA bind at an internal ribosome entry site (IRES) upstream of the ORF.



IRESs are present in a number of RNA viruses, including hepatitis C virus and the picornaviruses (Chapter 14). They have also been found in cell mRNAs and in one of the mRNAs of a DNA virus (Kaposi's sarcoma-associated herpesvirus).

6.4.2 Translation from bicistronic mRNAs

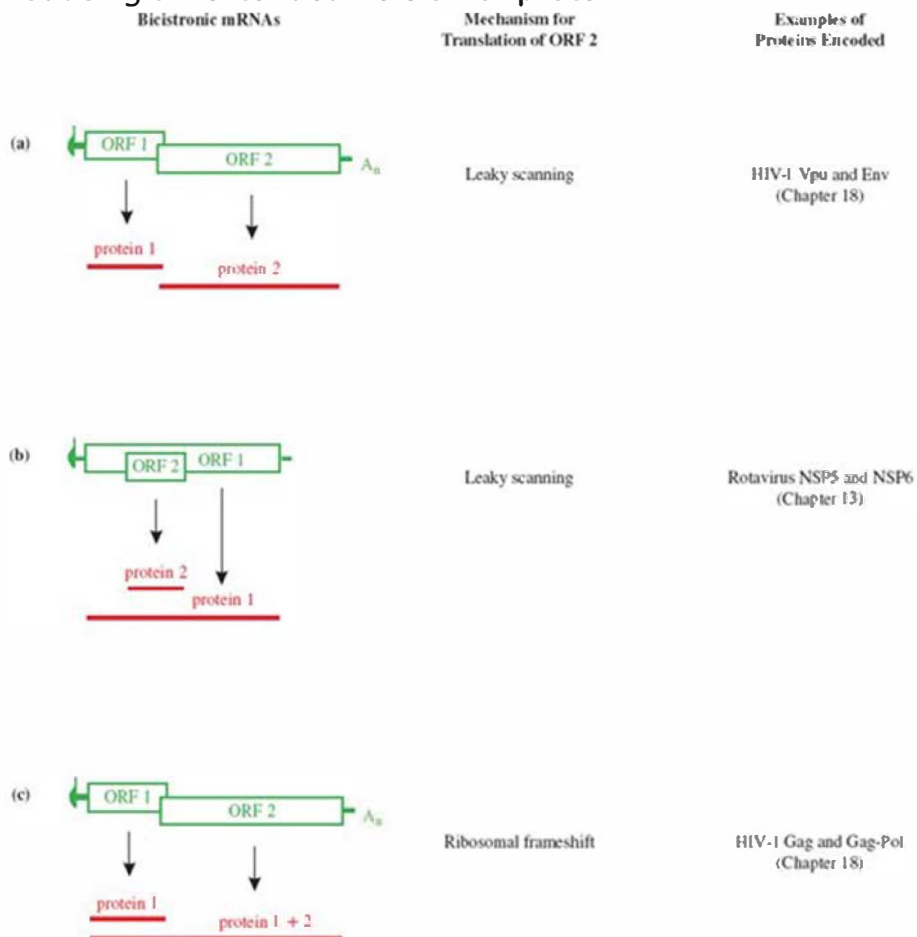
Most eukaryotic cell and virus mRNAs have one ORF, but there are a number of eukaryotic virus mRNAs that have two or more ORFs. Some of these bicistronic and polycistronic mRNAs are functionally monocistronic, i.e. only one of the ORFs, normally the 5' ORF, is translated.

Some structurally bicistronic mRNAs, however, are functionally bicistronic. When both ORFs are translated a difference in their rates of translation provides a mechanism for expressing two genes at different levels.

In many bicistronic mRNAs the ORFs overlap (Figure 6.10(a)); in others there is an ORF within an ORF (Figure 6.10(b)). One mechanism to read the second ORF involves leaky scanning; a 40S ribosomal subunit may overlook the ORF 1 start

codon and initiate translation at the start of ORF 2. The ORFs for the two proteins are in different reading frames, so the proteins that they encode are unrelated. Of course it is essential that the sequence “makes sense” in both reading frames!

Figure 6.10 Translation from bicistronic mRNAs. (a), (b) A ribosome may begin translation at the start of ORF 1, or scanning may be “leaky” and translation may begin at the start of ORF 2. The two start codons are in different reading frames and the two proteins are unrelated. (c) ORF 2 is translated by a ribosomal frameshift, producing an extended version of protein 1.



Another mechanism for reading a second ORF in an mRNA involves ribosomal frameshifting; a ribosome shifts into a different reading frame towards the end of ORF 1. It therefore does not recognize the ORF 1 stop codon, but continues along the mRNA, reading ORF 2 to produce an elongated version of the ORF 1 protein (Figure 6.10(c)). Frameshifting occurs when the ribosome encounters a frameshift signal in the mRNA, followed by a secondary structure, usually a pseudoknot (Section 3.2.2).

6.4.3 Co- and post-translational modification of proteins

During and after their translation, proteins may undergo one or more modifications, including glycosylation, acylation, phosphorylation, and cleavage.

6.4.3.a Glycosylation

Glycosylation involves the addition of oligosaccharide groups to the polypeptide chain. When an oligosaccharide is linked through the —OH group of a serine or threonine residue the process is known as O-glycosylation; when it is linked through the —NH₂ group of an asparagine residue it is known as N-glycosylation.

Proteins destined for glycosylation are synthesized in the rough endoplasmic reticulum, where N-glycosylation commences. They are then transported to the Golgi complex (Section 6.5), where N-glycosylation is completed by enzymes such as α -mannosidases I and II and galactosyl transferase. O-glycosylation takes place in the Golgi complex.

Some glycoproteins have undergone only one type of glycosylation, such as the N-glycosylated gp120 of HIV-1, while many glycoproteins have undergone both O- and N-glycosylation, such as gC and gD of herpes simplex virus.

These glycoproteins of HIV and herpes simplex virus, like most virus glycoproteins, are integral membrane proteins that are components of the virion envelopes. Some glycoproteins, however, are not associated with virion envelopes; rotaviruses have naked virions but their surface protein (VP7) is glycosylated (Chapter 13). A few virus glycoproteins are non-structural proteins, such as the rotavirus protein NSP4.

6.4.3.b Acylation

Acylation is the addition of an acyl group (R—CO—); one that is commonly added to proteins is a myristyl group, where R is CH₃—(CH₂)₁₂. The myristyl group is linked to a glycine residue at the N terminus of the protein. Most viruses that myristylate proteins use the host enzyme N-myristyltransferase.

Many myristylated proteins associate with membranes. This is true for the Gag proteins of most retroviruses (Chapter 17); if these proteins are not myristylated they do not associate with the plasma membrane and virion assembly does not take place. Another example of a myristylated protein is the picornavirus capsid protein VP4 (Chapter 14).

6.4.3.c Phosphorylation

Phosphorylation involves the transfer of a phosphate group from a nucleotide, usually ATP, to the O of an —OH group of a serine, threonine, or tyrosine residue. The transfer is carried out by a protein kinase, which may be of cell or viral origin. The enzyme recognizes short sequences of amino acids that bracket the residue to be phosphorylated.

Phosphorylation can alter the conformation, activity, localization, and/or the stability of a protein, and many cell and viral processes involve protein phosphorylation. Many structural and non-structural virus proteins become phosphorylated; for example, the phosphoproteins of rhabdoviruses (Chapter 15); one-sixth of the amino acid residues in the phosphoprotein of vesicular stomatitis virus are serine and threonine, and many of these are phosphorylated.

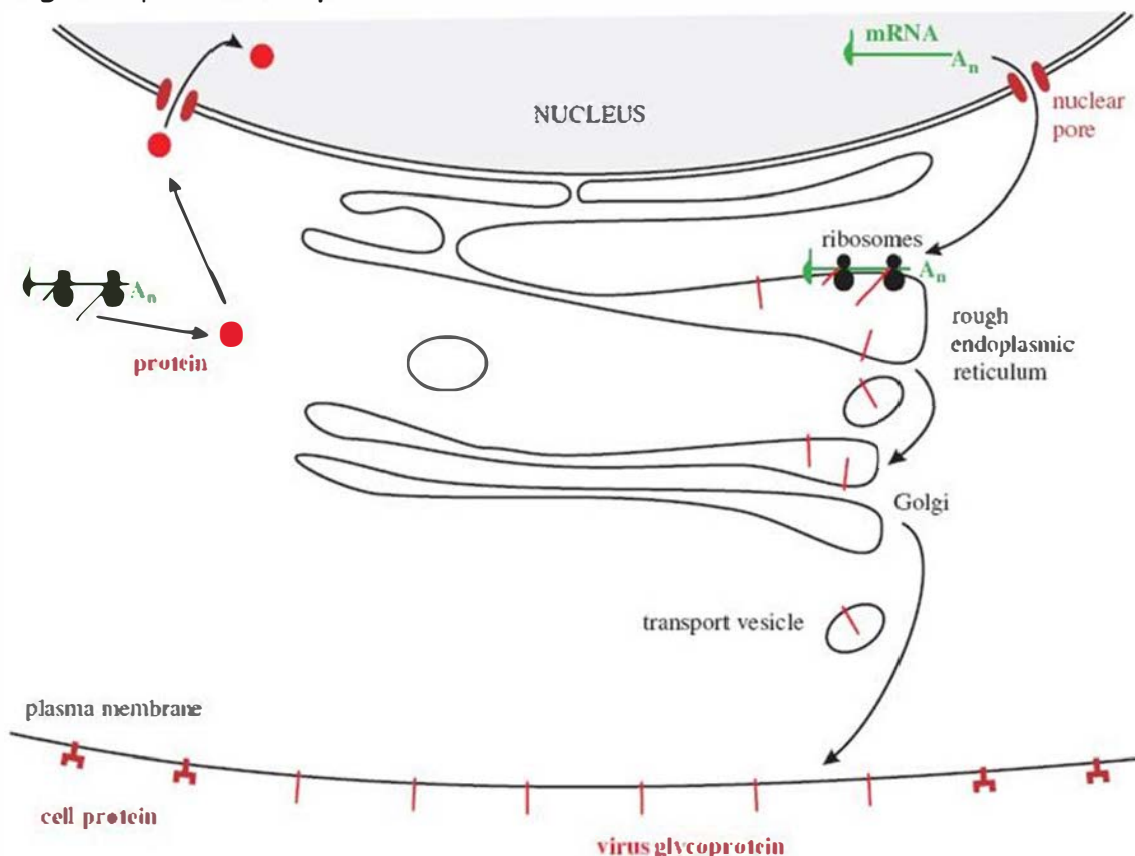
6.4.3.d Cleavage

Many proteins are cleaved by proteases during or after their synthesis. Some viruses encode their own protease; for example, picornaviruses produce proteases that cleave both viral and cell proteins (Chapter 14). Cell proteases are involved in the formation of many viral fusion proteins (Section 5.2.4.b); a precursor protein is cleaved to two subunits, the fusion sequence residing at the N terminus of one of the subunits (e.g. influenza viruses; Chapter 16). Retroviruses use their own protease to cleave their internal structural proteins, while they use a cell protease to cleave their envelope proteins (Chapter 17).

6.5 TRANSPORT OF MOLECULES IN EUKARYOTIC CELLS

We have already discussed how nucleocapsids and other virus structures are transported via microtubules and nuclear pores after entry into a cell (Section 5.2.5). Molecules synthesized by a virus in an infected cell must also be transported to particular sites. Virus mRNAs synthesized in the nucleus are transported to the cytoplasm, and virus proteins are transported to various locations, including the nucleus (Figure 6.11).

Figure 6.11 Synthesis and intracellular transport of molecules. Some proteins translated on free ribosomes are transported via nuclear pores to the nucleus. Some proteins translated in the rough endoplasmic reticulum are transported via the Golgi complex to the plasma membrane.



Many proteins have a sequence of amino acids (a "post code") that specifies their destination. Proteins destined to be incorporated into membranes have a signal sequence, which is a series of hydrophobic amino acid residues, either at the N terminus or internally. Protein synthesis begins on a free ribosome, but when the signal sequence has been synthesized it directs the polypeptide-ribosome complex to the endoplasmic reticulum, where protein synthesis continues. Regions of the endoplasmic reticulum with ribosomes associated are known as a rough endoplasmic reticulum (Figure 6.11).

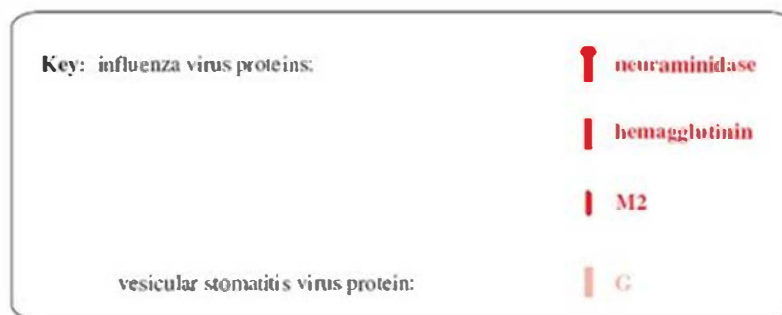
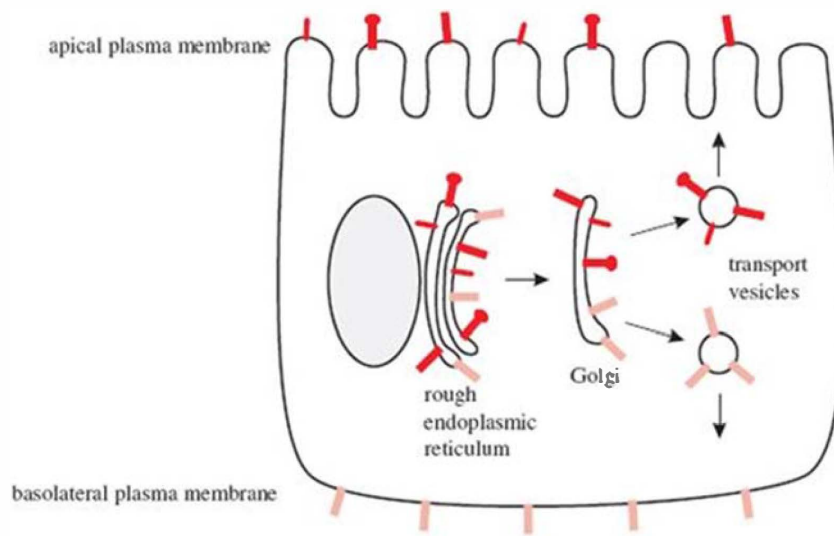
Each integral membrane protein has one or more membrane anchor sequences, which are rich in hydrophobic amino acid residues. For some of these proteins the signal sequence acts as a membrane anchor. Other integral membrane proteins, like the HIV-1 envelope protein, are moved through the membrane until the anchor sequence is reached, and then the signal sequence is removed by a host enzyme.

Many of the proteins synthesized in the rough endoplasmic reticulum are transported via vesicles to the Golgi complex, and most integral membrane proteins become glycosylated in these membrane compartments. From here the glycoproteins may be transported to other membranes, such as the plasma membrane or the nuclear envelope. Progeny virions may bud from these membranes (Section 8.3.1).

Epithelial cells have apical (outer) and basolateral (inner) surfaces, which are composed of different lipids and proteins and are separated by "tight junctions." Budding of enveloped viruses from epithelial cells may be restricted to either the apical surface or the basolateral surface.

Enrique Rodriguez-Boulan and David Sabatini used electron microscopy to examine thin sections of virus-infected epithelial cell cultures that retained their differentiated surfaces. They reported in 1978 that vesicular stomatitis virus buds from the basolateral surface, while influenza virus buds from the apical surface. This probably explains why most influenza virus infections of mammals are localized to the respiratory tract. It can be demonstrated that the glycoproteins of each virus are targeted to the appropriate surface (Figure 6.12). The targeting signal sequences and their locations within the proteins have been determined.

Figure 6.12 Targeting of virus envelope proteins to apical and basolateral surfaces of an epithelial cell. The envelope proteins of influenza A virus and vesicular stomatitis virus are transported to the apical and basolateral surfaces, respectively.



If the virus replicates in the nucleus then most of the virus proteins must be transported into the nucleus. Each of these proteins, like cell proteins that are transported into the nucleus, has a nuclear localization signal, which is rich in one or both of the basic amino acids lysine (K) and arginine (R). A common nuclear localization signal is

PKKKRKV.

This signal was first identified in a simian virus 40 protein known as large T antigen. A nuclear localization signal allows a protein to bind to cell proteins (importins) and subsequently to nuclear pore filaments (Figure 5.6), from where it is transported into the nucleus.

RNAs are also transported within the cell; for example, mRNAs synthesized in the nucleus must be exported through nuclear pores to the cytoplasm. The RNAs are taken to their destinations by proteins. The Rev protein of HIV-1 has both a nuclear localization signal and a nuclear export signal (Figure 6.13). The nuclear localization signal ensures that Rev is transported into the nucleus, where it binds specifically to HIV-1 RNA. The nuclear export signal binds Rev and its RNA cargo to cell proteins (exportins) that transport the complex from the nucleus to the cytoplasm via a nuclear pore.

Figure 6.13 HIV-1 Rev protein. The nuclear localization signal is rich in arginine (R) residues. The nuclear export signal is rich in leucine (L) residues.



6.6 TRANSCRIPTION AND TRANSLATION IN BACTERIA

This section starts with a brief summary of transcription from prokaryotic genes. There is one type of DNA-dependent RNA polymerase for transcription in prokaryotes, and the *E. coli* holoenzyme comprises the catalytic core enzyme containing one β , one β' , and two α subunits, and a regulatory protein, the sigma (σ) factor. The σ factor determines promoter specificity, enabling the polymerase to recognize and bind to specific promoters in the correct orientation to initiate transcription. Bacterial promoters contain characteristic sequences: the -10 region (Pribnow box) and the -35 region, centered 10 and 35 bp upstream of the transcription start site respectively, and recognized by the σ factor. Upstream of the -35 region some promoters additionally contain the UP element, which is recognized by the RNA polymerase α subunit and enhances binding of the enzyme. Once transcription initiation is completed, the σ factor dissociates to be re-used in the process. Different σ factors recognize different promoters. There are two main sigma families: $\sigma 70$ and $\sigma 54$. The $\sigma 70$ family has many different σ factors, including the primary (vegetative) σ factors, directing transcription of genes for bacterial growth and metabolism, and alternative factors, e.g. $\sigma 32$ for transcription of heat shock genes.

Transcription terminates after a terminator is transcribed. Termination may be rho independent (intrinsic) or rho dependent. Rho-independent termination depends on the template DNA sequence, which contains an inverted repeat and a string of adenine residues, located downstream of the rho-independent genes. The inverted repeat forms a stem-loop structure in the RNA transcript and adenine-uracil bonds form in the DNA-RNA hybrid, such that the RNA transcript can dissociate from the template. Rho-dependent termination involves a rho factor, which has helicase and ATPase activity, to unwind DNA-RNA hybrids when the polymerase is stalled at an inverted repeat. This leads to release of the transcript.

Bacterial cells and the viruses that infect them operate transcription and translation mechanisms that differ in a number of respects from those in eukaryotes. Introns are rarely present in the genes of bacteria and their viruses. Some phages use the host DNA-dependent RNA polymerase for their transcription, while others encode their own.

Whereas typical mRNAs in eukaryotes are monocistronic, in bacteria they are polycistronic; that is, each mRNA has several ORFs (Figure 6.14). The mRNAs of

bacteria and their viruses are never capped at the 5' end, and are rarely polyadenylated at the 3' end.

Figure 6.14 Characteristics of bacterial and phage mRNA. There are typically several ORFs, all of which may be translated at the same time. The 5' end is not capped and the 3' end is rarely polyadenylated.



Bacterial translation differs from eukaryotic translation in a number of features:

- Translation may start before transcription is complete. The lack of a nucleus allows transcription and translation to be coupled.
- The ribosomes are smaller (the ribosomal subunits have sedimentation coefficients of 30S and 50S).
- The 30S ribosomal subunit binds directly to a translation initiation region on the mRNA. Initiation generally involves interaction of a mRNA sequence (consensus in *E. coli*: 5'–AGGAGGU–3'), known as the Shine–Dalgarno (SD) sequence, and the complementary sequence (the anti-SD sequence) at the 3' end of 16S ribosomal RNA in the 30S subunit. The SD sequence (ribosome binding site) is located just upstream of the AUG start codon.
- The methionine of the initiator methionyl tRNA is generally formylated.
- A much smaller number of initiation factors is involved.
- All ORFs within an mRNA are translated and several may be translated concurrently.

A few phages have overlapping genes, which may be translated by reading through a stop codon or by ribosomal frameshifting (Section 6.4.2).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- explain how virus genes are transcribed and translated;
- describe the post-translational modifications that some virus proteins undergo;
- highlight differences in transcription and translation between prokaryotic and eukaryotic cells;
- discuss the transport of virus proteins and RNA within cells.

SOURCES OF FURTHER INFORMATION

Books

Brown, T. A. (2006) *Genomes 3*, 3rd edition, Garland Science

Flint, S. J. et al. (2009) Chapters 6–12 in *Principles of Virology, Volume 1*, 3rd edition, ASM Press

Klug, W. S. et al. (2009) *Essentials of Genetics*, Pearson

Sugrue, R. J. (2007) *Viruses and glycosylation*. Chapter 1 in *Glycoviropology Protocols* (Sugrue R. J., editor), Humana Press

Historical papers

Crick, F. (1970) Central dogma of molecular biology. *Nature*, 227, 561–563

Rodriguez-Boulan, E. and Sabatini, D. D. (1978) Asymmetric budding of viruses in epithelial monolayers: a model system for study of epithelial polarity. *Proceedings of the National Academy of Sciences USA*, 75, 5071–5075

Recent papers

Ahlgvist, P. (2006) Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nature Reviews Microbiology*, 4, 371–382

Balvay, L. et al. (2009) Structural and functional diversity of viral IRESes. *Biochimica et Biophysica Acta – Gene Regulatory Mechanisms*, 1789, 542–557

Decroly, E. et al. (2012) Conventional and unconventional mechanisms for capping viral mRNA. *Nature Reviews Microbiology*, 10, 51–65

Hsieh, M. J., White, P. J., and Pouton, C. W. (2010) Interaction of viruses with host cell molecular motors. *Current Opinion in Biotechnology*, 21, 633–639

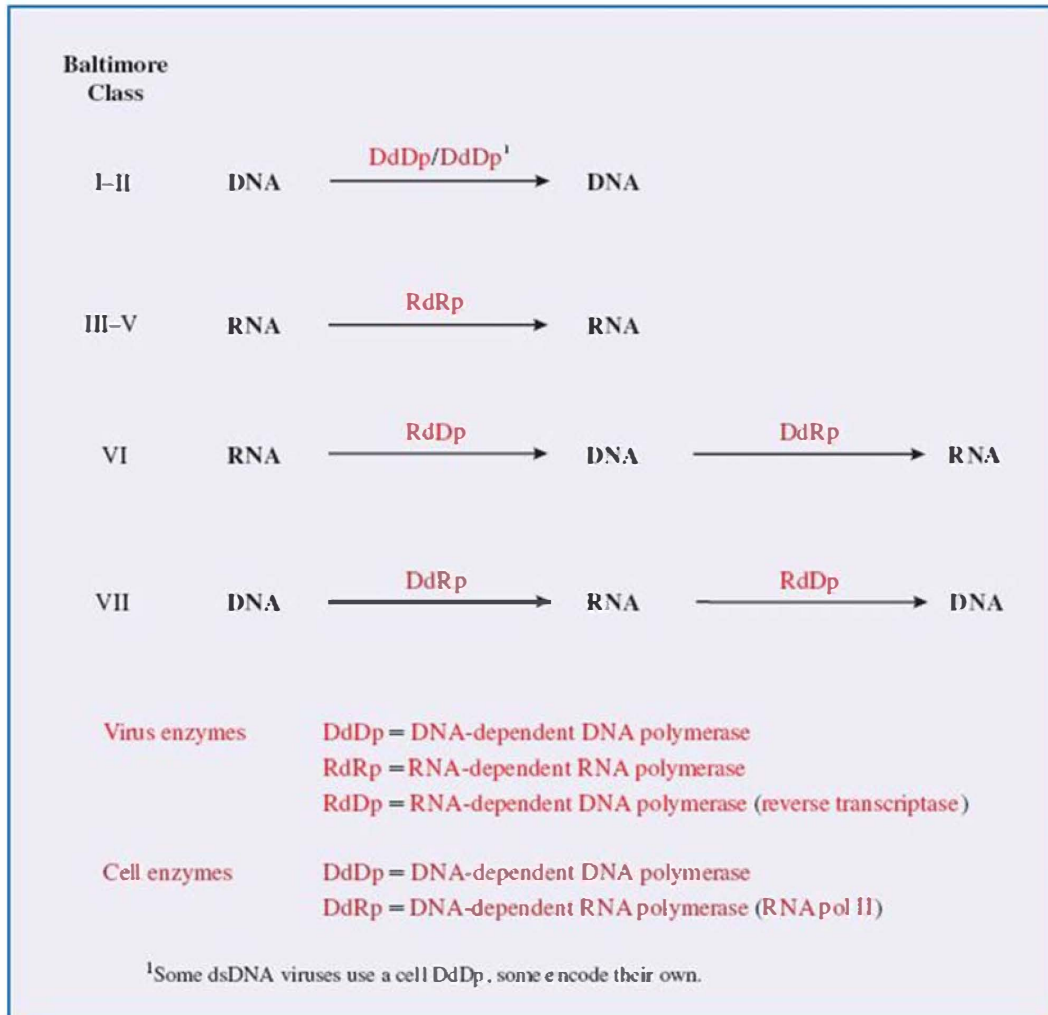
López-Lastra, M. et al. (2010) Translation initiation of viral mRNAs. *Reviews in Medical Virology*, 20, 177–195

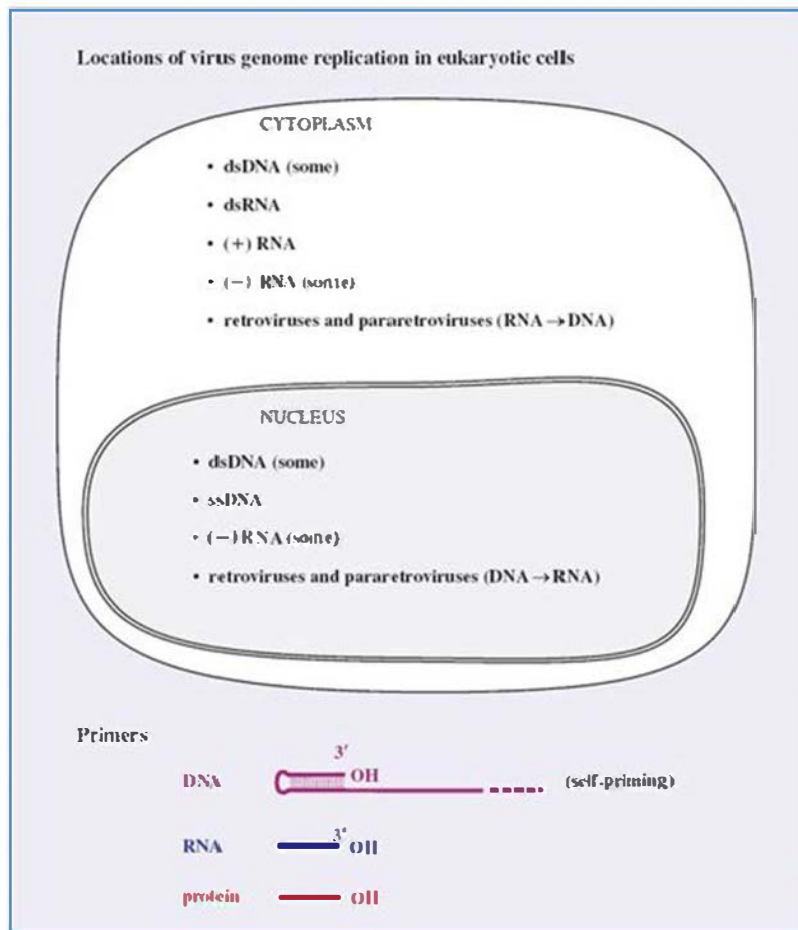
Reineke, L. C. and Lloyd, R. E. (2011) Animal virus schemes for translation dominance. *Current Opinion in Virology*, 1, 363–372

CHAPTER 7

Virus Genome Replication

CHAPTER 7 AT A GLANCE



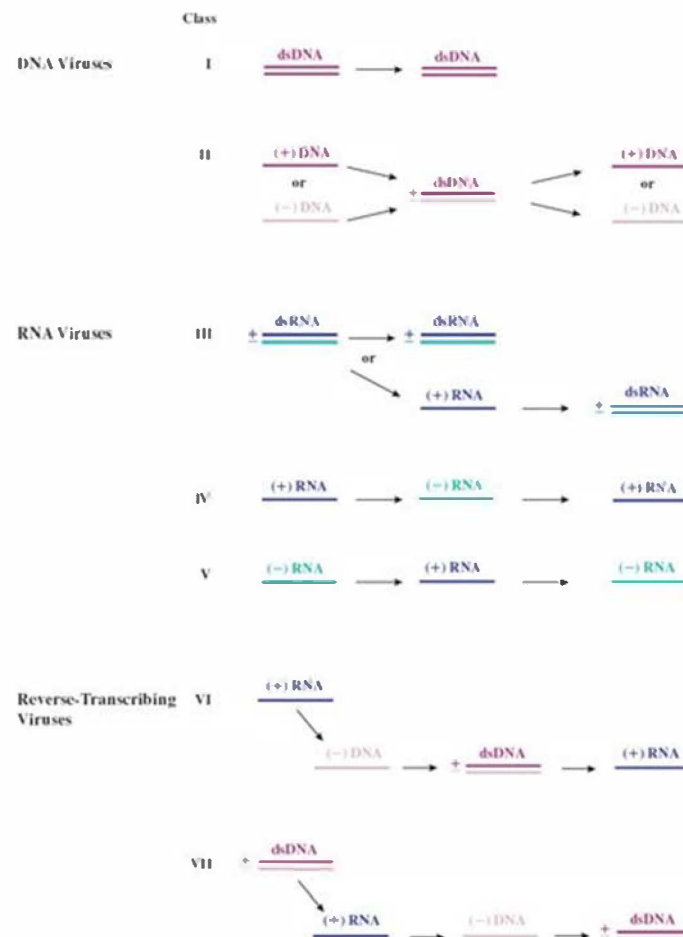


7.1 OVERVIEW OF VIRUS GENOME REPLICATION

In this chapter we consider the fifth step of our generalized replication cycle: genome replication. The genome of the infecting virus is replicated so that viral transcription can be amplified and to provide copies of the genome for progeny virions.

Generally, DNA viruses copy their genomes directly to DNA, and RNA viruses copy their genomes directly to RNA. There are, however, some DNA viruses that replicate their genomes via an RNA intermediate and some RNA viruses that replicate their genomes via a DNA intermediate. The various replication modes of virus genomes are summarized in Figure 7.1.

Figure 7.1 Replication of virus genomes in the seven Baltimore Classes. (+) RNA and (+) DNA have the same sequence as the mRNA (except that in DNA thymine replaces uracil). (-) RNA and (-) DNA have the sequence complementary to the mRNA (except that in DNA thymine replaces uracil). (+) and (-) strands are not indicated for the dsDNA of the class I viruses because the genomes of most of these viruses have ORFs in both directions. (+) and (-) strands are indicated for the ssDNA of the class II viruses. Most of these viruses have either a (+) or a (-) strand genome. Some ssDNA viruses and some ssRNA viruses have ambisense genomes.



Single-stranded genomes are designated as plus or minus depending on their relationship to the virus mRNA. Plus-strand genomes have the same sequence as the mRNA (except that in DNA thymine replaces uracil), while minus-strand genomes have the sequence complementary to the mRNA. Single-stranded DNA is converted to dsDNA prior to copying.

There are two classes of viruses with (+) RNA genomes (Figure 7.1). Class IV viruses copy their (+) RNA genomes via a (-) RNA intermediate, while class VI viruses replicate via a DNA intermediate. The synthesis of DNA from an RNA template (reverse transcription) is also a characteristic of class VII viruses.

In this chapter we shall look at some general aspects of virus genome replication, and then we shall give individual attention to the DNA viruses, the RNA viruses and the reverse-transcribing viruses.

7.2 LOCATIONS OF VIRUS GENOME REPLICATION IN EUKARYOTIC CELLS

As we saw in Chapter 5, when viruses infect eukaryotic cells the genomes of some are delivered to the cytoplasm and some are conveyed to the nucleus. The destination of a virus genome, and hence the location in which it is replicated, varies with the type of genome (Table 7.1).

Table 7.1 Locations of virus genome replication in eukaryotic cells

Virus genome	Cytoplasm	Nucleus
dsDNA (except pararetroviruses)	Some	Some
ssDNA		All
dsRNA	All	
(+) RNA (except retroviruses)	All	
(-) RNA	Some	Some
(+) RNA (retroviruses)}		
dsDNA (pararetroviruses) }	ssRNA dsDNA	dsDNA ssRNA

The genomes of most DNA viruses are replicated in the nucleus, but those of some dsDNA viruses are replicated in the cytoplasm. The genomes of most RNA viruses are replicated in the cytoplasm, but there are exceptions amongst the minus-strand RNA viruses (Table 7.1). The retroviruses and pararetroviruses are special cases: each replicates RNA to DNA in the cytoplasm and DNA to RNA in the nucleus.

For those viral genomes that are replicated in the cytoplasm, nucleic acid synthesis usually takes place inside structures within the cytoplasm. For the dsRNA viruses, the retroviruses and the pararetroviruses these structures are composed of viral proteins; for the remaining plus-strand RNA viruses they are composed of cell-derived membranes. These viral and cell-derived structures are described in later chapters. One probable reason why viruses avoid replicating their nucleic acids free in the cytoplasm is to protect them from defense mechanisms of the host cells. Obviously there must be an adequate supply of nucleotides into these structures in order for nucleic acid synthesis to proceed.

7.3 INITIATION OF GENOME REPLICATION

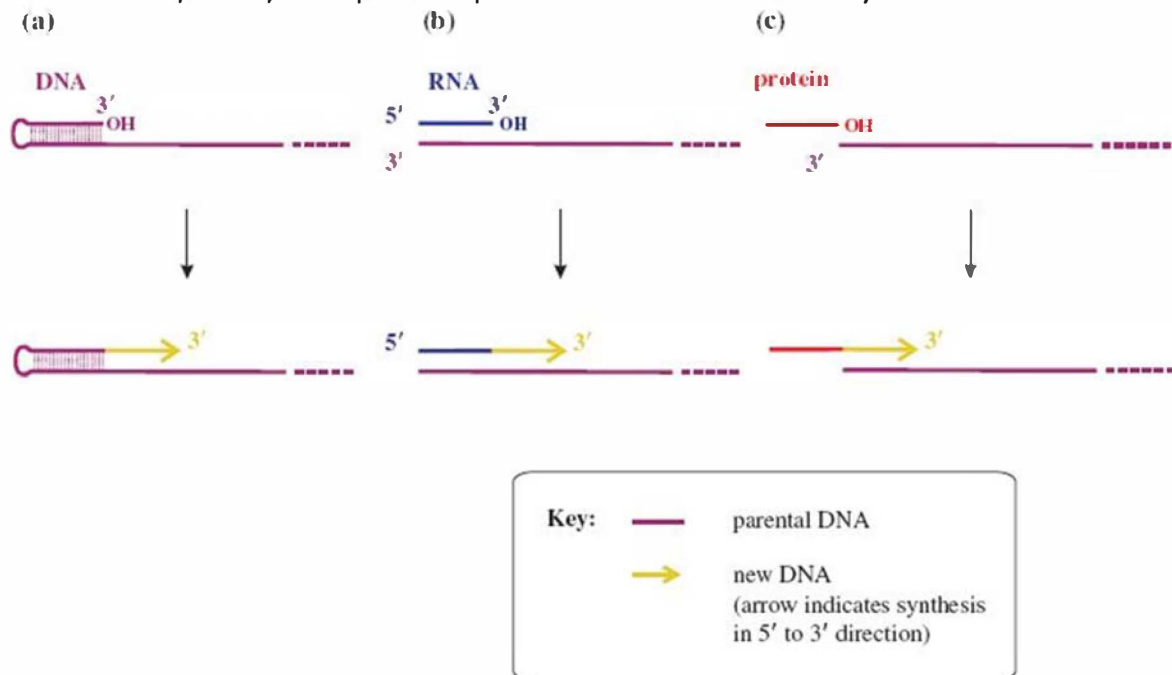
The replication of a virus genome is initiated at a specific nucleotide sequence that is recognized by the proteins that initiate replication. For most linear genomes this sequence is at one of the ends of the nucleic acid.

Replication of DNA, and of some viral RNAs, requires priming, which is the reaction of a nucleotide with an —OH group on a molecule at the initiation site. Initiation of genome replication for many RNA viruses, including rotaviruses and rhabdoviruses, does not require a primer; replication initiates when the first ribonucleotide of the new strand base pairs with a ribonucleotide in the template RNA. This first ribonucleotide effectively acts as a primer when its 3' —OH group becomes linked to the second ribonucleotide.

Some ssDNA viruses, such as parvoviruses, use self-priming. At the 3' end of the DNA there are regions with complementary sequences that can base pair (Figure

7.2(a)). The —OH group of the nucleotide at the 3' end forms a linkage with the first nucleotide, then DNA synthesis proceeds to copy the whole genome.

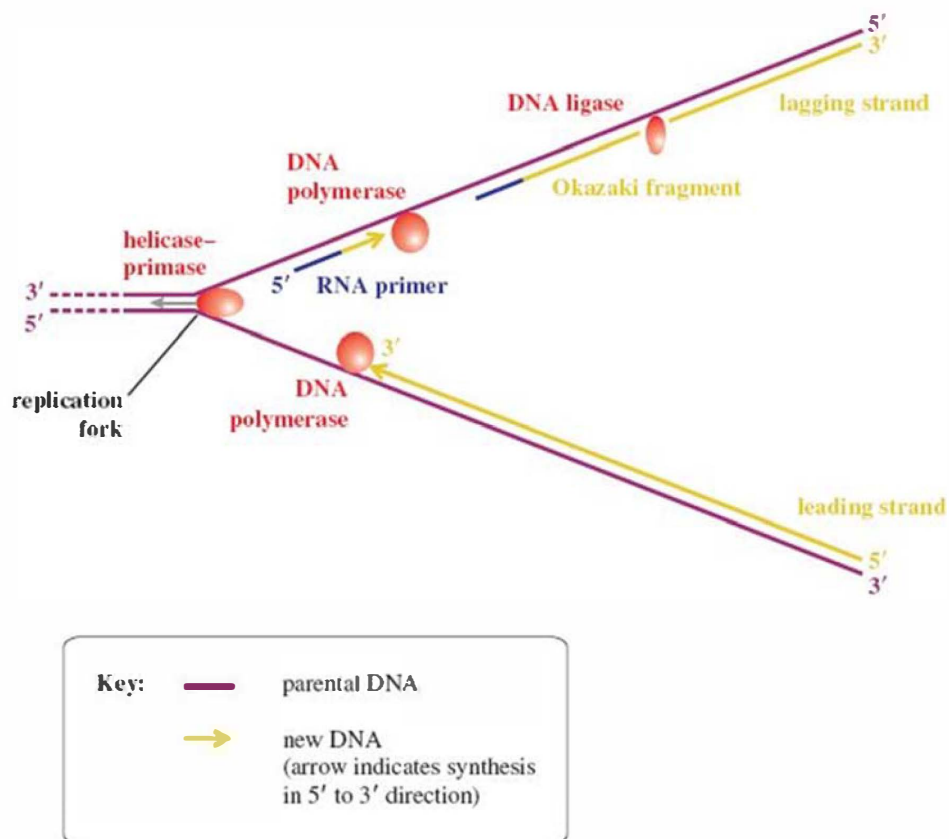
Figure 7.2 DNA, RNA, and protein primers for nucleic acid synthesis.



7.3.1 RNA and protein primers

In order to initiate the replication of many DNA genomes, and some RNA genomes, a molecule of RNA or protein is required to act as a primer (Figure 7.2(b) and (c)). Synthesis of cell DNA commences after a region of the double helix has been unwound by a helicase and after a primase has synthesized short sequences of RNA complementary to regions of the DNA. These RNAs act as primers; one is required for the leading strand, while multiple primers must be synthesized for the Okazaki fragments of the lagging strand (Figure 7.3). The first nucleotide of a new sequence of DNA is linked to the 3' —OH group of the primer RNA.

Figure 7.3 DNA replication. A helicase–primase unwinds the dsDNA and synthesizes RNA primers that are used by the DNA polymerase to initiate DNA synthesis. The leading strand is synthesized continuously, while the lagging strand is synthesized as Okazaki fragments that are joined together by a DNA ligase.



Some DNA viruses also use RNA primers during the replication of their genomes. Some viruses, such as polyomaviruses, use the cell primase to synthesize their RNA primers, while others, such as herpesviruses and phage T7, encode their own primases.

During their replication cycle the retroviruses synthesize DNA from a (+) RNA template (Section 17.3.2). They use a cell transfer RNA to prime (-) DNA synthesis, then they use the 3' —OH group in a polypurine tract of the partly degraded (+) RNA template to prime (+) DNA synthesis. The retrovirus DNA becomes integrated into a cell chromosome. If the infection is latent and the cell subsequently divides (Section 9.3.1), then the virus DNA is copied along with the cell DNA, using RNA primers synthesized by the cell primase.

For some viruses the primer for initiation of nucleic acid replication is the —OH group on a serine or tyrosine residue in a protein (Figure 7.2(c)). DNA viruses that use protein primers include some animal viruses (e.g. adenoviruses) and some phages (e.g. tectiviruses). RNA viruses that use protein primers include some animal viruses (e.g. picornaviruses) and some plant viruses (e.g. luteoviruses).

Hepadnaviruses are DNA viruses that use a protein primer to initiate (-) DNA synthesis and an RNA primer to initiate (+) DNA synthesis (Section 19.8.6).

Protein primers (and the RNA primers of hepadnaviruses) are not removed once their role is performed and they are found linked to the 5' ends of the genomes in virions (Section 3.2.3).

7.4 POLYMERASES

The key enzymes involved in virus genome replication are DNA polymerases and RNA polymerases. Many viruses encode their own polymerase, but some use a host cell enzyme (Figure 7.4).

Figure 7.4 Enzymes used by viruses to replicate their genomes. Many viruses encode a polymerase to replicate their genome, but some use a cell enzyme.

Virus enzymes

DNA $\xrightarrow{\text{DNA-dependent DNA polymerase}}$ DNA

RNA $\xrightarrow{\text{RNA-dependent RNA polymerase}}$ RNA

RNA $\xrightarrow{\text{RNA-dependent DNA polymerase (reverse transcriptase)}}$ DNA

Cell enzymes

DNA $\xrightarrow{\text{DNA-dependent DNA polymerase}}$ DNA

DNA $\xrightarrow{\text{DNA-dependent RNA polymerase (RNA pol II)}}$ RNA

A DNA virus requires a DNA-dependent DNA polymerase. Amongst the DNA viruses that replicate in the nuclei of eukaryotic cells, viruses with small genomes (e.g. papillomaviruses) use the cell enzyme, while viruses with large genomes (e.g. herpesviruses) encode their own enzyme. Those DNA viruses that replicate in the cytoplasm must encode their own enzyme.

The enzyme that replicates the genome of an RNA virus is often referred to as a replicase; for many RNA viruses this is the same enzyme as that used for transcription (Section 6.3.3). The retroviruses and the pararetroviruses encode reverse transcriptases to transcribe from RNA to DNA, and use the host cell RNA polymerase II to transcribe from DNA to RNA.

Many viral polymerases form complexes with other viral and/or cell proteins to produce the active enzyme. Some of these additional proteins are processivity factors; for example, an *Escherichia coli* thioredoxin molecule functions as a processivity factor for the DNA polymerase of phage T7.

7.5 DNA REPLICATION

The viruses of Class I (dsDNA) and Class II (ssDNA) replicate their genomes via dsDNA. The ssDNA viruses first synthesize a complementary strand to convert the genome into dsDNA.

Each viral DNA has at least one specific sequence (ori; replication origin) where replication is initiated. The proteins that initiate DNA replication bind to this site, and amongst these proteins are:

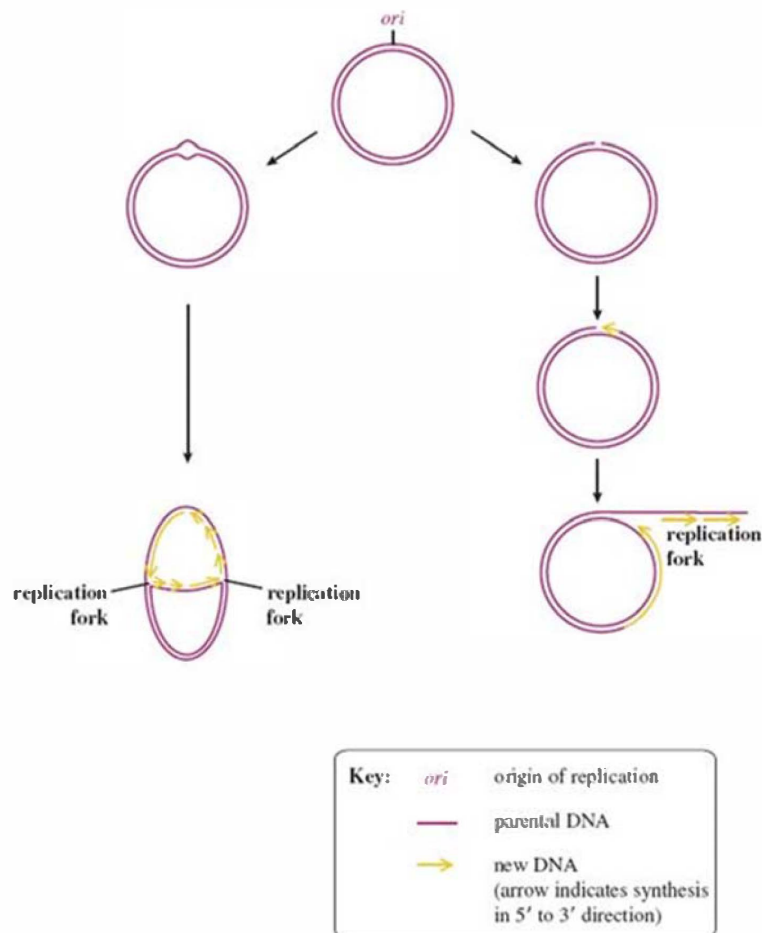
- a helicase (unwinds the double helix at that site);
- a ssDNA binding protein (keeps the two strands apart);
- a DNA polymerase.

Viral dsDNA is generally replicated by a process similar to that used by cells to copy their genomes. The basic process and the enzymes involved are outlined in Figure 7.3. Fewer proteins are involved in bacterial systems than in eukaryotic systems; for example, the helicase–primase of phage T7 is a single protein molecule, while that of herpes simplex virus is a complex of three protein species.

DNA synthesis takes place near a replication fork. One of the daughter strands is the leading strand and the other is the lagging strand, synthesized as Okazaki fragments, which become joined by a DNA ligase. After a dsDNA molecule has been copied each of the daughter molecules contains a strand of the original molecule. This mode of replication is known as semi-conservative, in contrast to the conservative replication of some dsRNA viruses (Section 7.6).

Some DNA genomes are linear molecules, while some are covalently closed circles (Section 3.2). Some of the linear molecules are circularized prior to DNA replication, hence many DNA genomes are replicated as circular molecules, for which there are two modes of replication, known as theta and sigma (Figure 7.5). These terms refer to the shapes depicted in diagrams of the replicating molecules, which resemble the Greek letters θ (theta) and σ (sigma). The sigma mode of replication is also known as rolling circle replication. Some DNA viruses may carry out theta replication early in infection and sigma replication late in infection.

Figure 7.5 Theta and sigma modes of DNA replication. The theta structure is shown with two replication forks as a result of bidirectional replication from ori; unidirectional replication can also give rise to a theta structure.



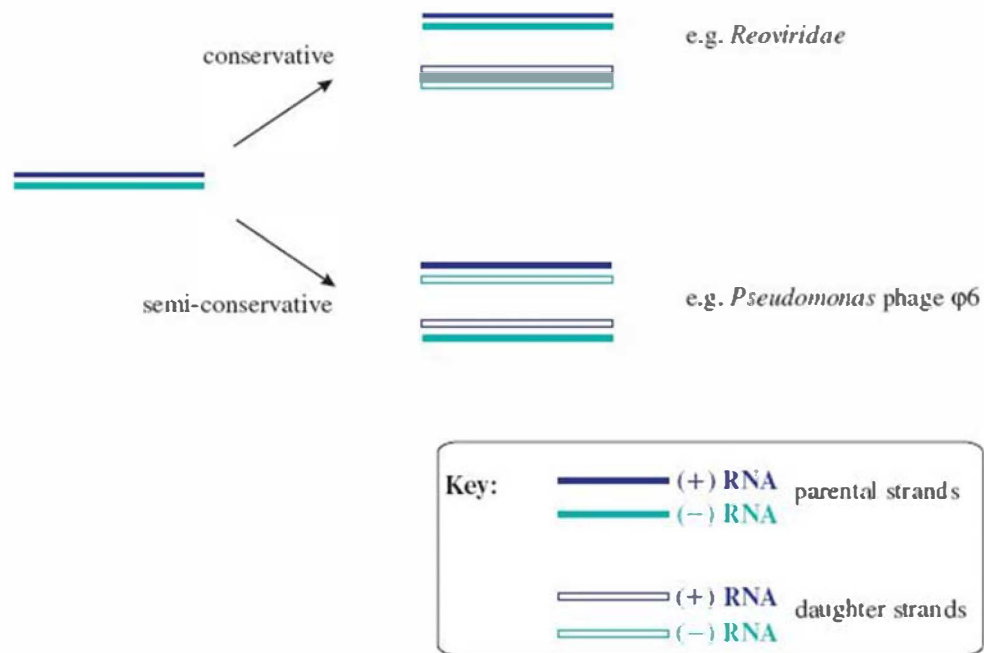
Rolling circle replication of the DNA of some viruses, such as herpesviruses (Section 11.5.3), results in the formation of very large DNA molecules called concatemers. Each concatemer is composed of multiple copies of the virus genome and the concatemers of some viruses are branched. When DNA is packaged during the assembly of a virion an endonuclease cuts a genome length from a concatemer.

7.6 DOUBLE-STRANDED RNA REPLICATION

Double-stranded RNA, like dsDNA, must be unwound with a helicase in order for the molecule to be replicated.

Some dsRNA viruses, e.g. *Pseudomonas* phage $\phi 6$ (ϕ =Greek letter phi), replicate their genomes by a semi-conservative mechanism, similar to dsDNA replication (Section 7.5); each of the double-stranded progeny molecules is made up of a parental strand and a daughter strand. Other dsRNA viruses, including members of the family Reoviridae (Chapter 13), replicate by a mechanism designated as conservative because the double-stranded molecule of the infecting genome is conserved (Figure 7.6).

Figure 7.6 Conservative and semi-conservative replication of dsRNA.



7.7 SINGLE-STRANDED RNA REPLICATION

The ssRNA genomes of viruses in Classes IV and V are replicated by synthesis of complementary strands of RNA that are then used as templates for synthesis of new copies of the genome (Figure 7.1). The synthesis of each RNA molecule requires the recruitment of an RNA-dependent RNA polymerase to the 3' end of the template, therefore both plus- and minus-strand RNA must have a binding site for the enzyme at the 3' end.

An interesting point to note here is that all plus-strand viruses of eukaryotes replicate their RNA in association with membranes derived from cytoplasmic membranous structures. For many groups of viruses, including picornaviruses (Section 14.4.4), these membranes are derived from the Golgi and the endoplasmic reticulum, but other membranous structures are used, including endosomes and lysosomes (by togaviruses) and chloroplasts (by turnip yellow mosaic virus). Viral proteins, including the RNA polymerases, are bound to the membranes. Minus-strand RNA viruses, on other hand, do not replicate their RNA in association with membranes, but the RNA templates are coated with viral protein.

During the replication of ssRNA both (+) and (-) strands of RNA accumulate in the infected cell, but not in equal amounts. Plus-strand RNA viruses accumulate an excess of (+) RNA over (-) RNA, and for minus-strand RNA viruses the reverse is true.

7.8 REVERSE TRANSCRIPTION

Some RNA viruses replicate their genomes via a DNA intermediate, while some DNA viruses replicate their genomes via an RNA intermediate (Figure 7.1). Both of these modes of genome replication involve reverse transcription, which has two major steps: synthesis of (-) DNA from a (+) RNA template followed by synthesis of a

second DNA strand (Figure 7.7). Both steps are catalyzed by a reverse transcriptase that is encoded by the virus.

Figure 7.7 Reverse transcription. ssDNA is synthesized from an RNA template, then the ssDNA is converted to dsDNA. A reverse transcriptase (RT) carries out both steps.



Reverse transcription takes place within a viral structure in the cytoplasm of the infected cell. In later chapters the process is considered in more detail for the retroviruses (Section 17.3.2) and for hepatitis B virus (Section 19.8.6). No viruses of prokaryotes are known to carry out reverse transcription.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- state the locations within eukaryotic cells where different categories of virus genome are replicated;
- explain the role of primers in virus nucleic acid synthesis;
- discuss the roles of virus and host proteins in virus genome replication;
- outline the replication mechanisms of virus DNAs and RNAs;
- explain the term “reverse transcription.”

SOURCES OF FURTHER INFORMATION

Books

Brown, T. A. (2002) Chapter 13 in *Genomes*, 2nd edition, BIOS

Cameron, C. E., Götte, M., and Raney, K. D. (eds) (2009) *Viral Genome Replication*, Springer

Flint, S. J. et al. (2009) Chapters 6–10 in *Principles of Virology*, Volume 1, 3rd edition, ASM Press

Journals

Netherton, C. L. and Wileman, T. (2011) Virus factories, double membrane vesicles and viroplasm generated in animal cells. *Current Opinion in Virology*, 1, 381–387

Ortín, J. and Parra F. (2006) Structure and function of RNA replication. *Annual Review of Microbiology*, 60, 305–326

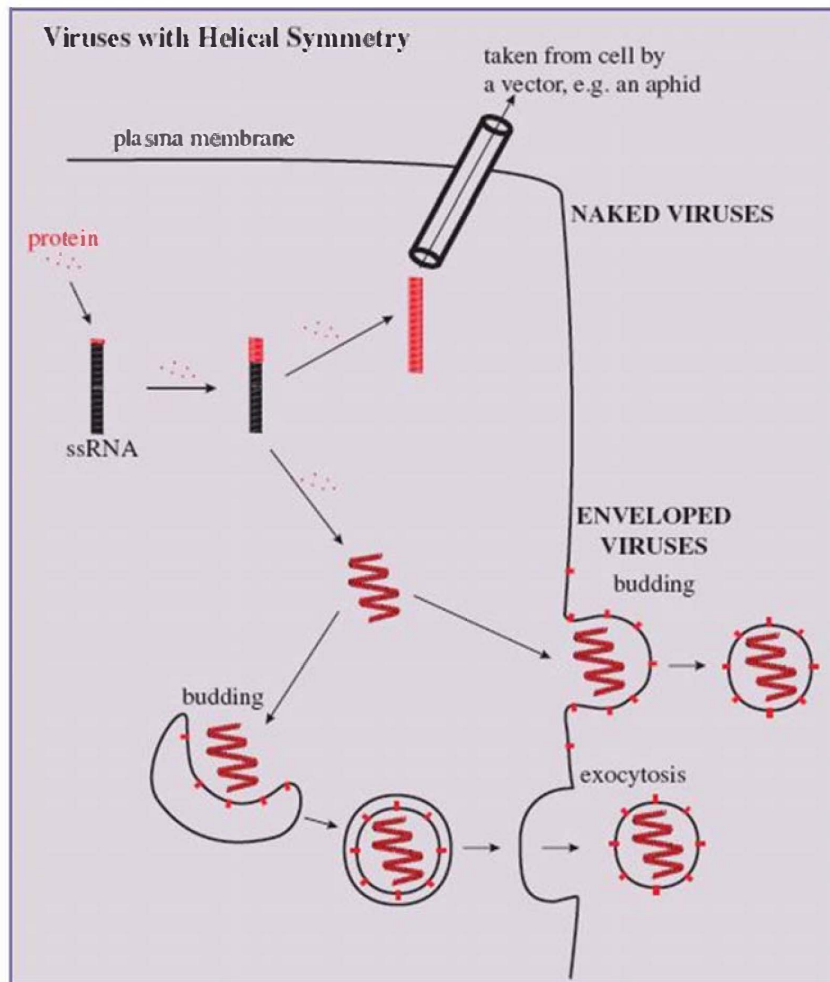
Tao, Y. J. and Ye, Q. (2010) RNA virus replication complexes. *PLoS Pathogens*, 6 (7), e1000943

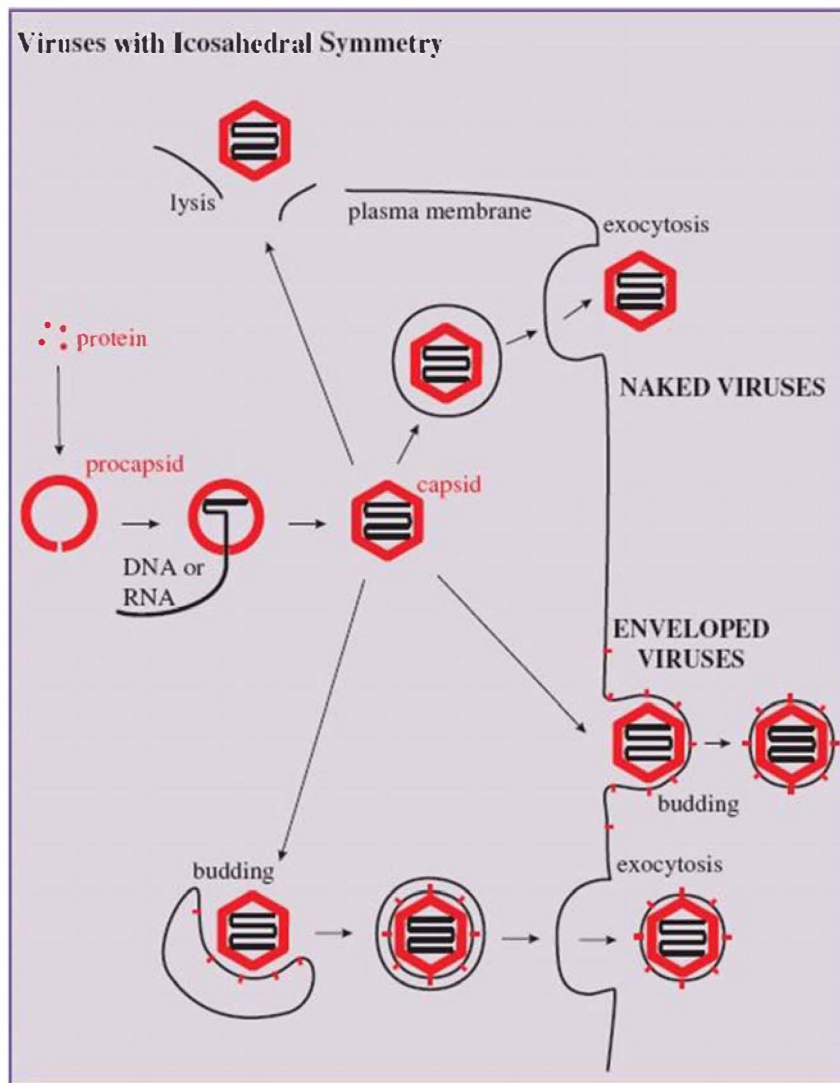
van Dijk, A. A. et al. (2004) Initiation of viral RNA-dependent RNA polymerization. Journal of General Virology, 85, 1077–1093

CHAPTER 8

Assembly and Exit of Virions from Cells

CHAPTER 8 AT A GLANCE





8.1 INTRODUCTION TO ASSEMBLY AND EXIT OF VIRIONS FROM CELLS

In this chapter we deal with the final stages of the replication cycle: assembly of virions and their exit from the cell.

A virion has to be a stable structure as it has to survive in the environment as an infectious entity (though there is huge variation amongst viruses in the degree of stability, as discussed in Chapter 24). There is, however, also a requirement for a virion to become unstable during infection of a host cell so that the viral genome can be released. The virion must therefore have a built-in "switch" that can initiate the change from stability to instability when the appropriate environment is encountered. The "trigger" that activates the "switch" might be binding to a receptor and/or a change in the concentration of H^+ or other ions.

Once threshold quantities of progeny virus genomes and structural proteins have accumulated in the infected cell, assembly of virions can commence. These components are assembled into nucleocapsids. If the virion also contains lipid then

the assembly process also includes the acquisition of this component, either as an internal membrane or as an envelope.

8.2 NUCLEOCAPSID ASSEMBLY

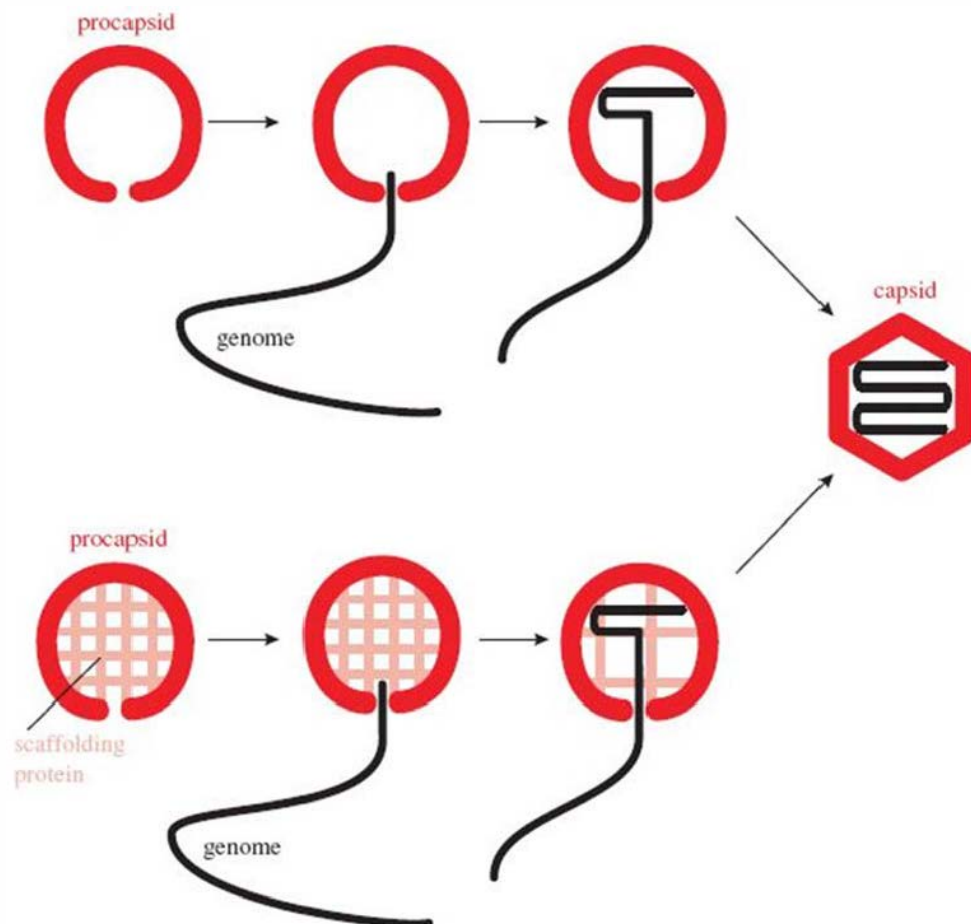
8.2.1 Helical viruses

The assembly of virions and nucleocapsids of ssRNA viruses with helical symmetry (Section 3.4.1) involves coating copies of the genome with multiple copies of a protein, e.g. coat protein for tobacco mosaic virus, nucleoprotein for influenza viruses. Detailed studies of the assembly of tobacco mosaic virus have shown that discs of coat protein are formed first, and then these build up on the RNA to form the rod-shaped virion. The genomes of minus-strand RNA viruses become associated with additional proteins, such as RNA polymerases.

8.2.2 Icosahedral viruses

The assembly of virions and nucleocapsids of many viruses with icosahedral symmetry (Section 3.4.2) involves the construction of an empty protein shell, known as a procapsid, or a prohead in the case of a tailed bacteriophage. The procapsid is filled with a copy of the virus genome (Figure 8.1); during or after this process it may undergo modification to form the mature capsid. Modification of the procapsid may result in a change from a spherical to an icosahedral shape. For some viruses, including adenoviruses and picornaviruses, modification of the procapsid involves cleavage of one or more of the structural proteins.

Figure 8.1 Assembly of an icosahedral nucleocapsid. A protein shell (procapsid) is constructed. A copy of the virus genome enters the procapsid, which undergoes modifications to form the capsid. The procapsid of some viruses is built around scaffolding protein, which is removed as the genome is packaged.



The genome enters the procapsid through a channel located at a site that will become one of the twelve vertices of the icosahedron. Any enzymes involved in packaging the genome are located at this site. In the procapsids of herpesviruses and tailed phages the channel consists of a “portal” composed of one or more protein species (Section 3.4.2.d). The portal contains a powerful motor that moves the virus DNA into the procapsid.

8.2.3 Genome packaging

In the context of packaging virus genomes during virion assembly the question “How are virus genomes selected from all the cell and virus nucleic acids?” is an intriguing one. In a retrovirus-infected cell, for example, less than 1% of the RNA is virus genome. In fact, retroviruses do package a variety of cell RNAs, including tRNA, which plays a key role when the next host cell is infected (Chapter 17). Some other viruses also package cell nucleic acids; for example, tailed phages sometimes package host DNA.

For the most part, however, cell nucleic acids are not packaged. For some viruses it has been shown that this is achieved through the recognition by a virus protein of a specific virus genome sequence, known as a packaging signal. In single-stranded genomes the packaging signal is within a region of secondary structure. Most viruses with single-stranded genomes package either the plus-strand or the minus-strand, so the packaging signal must be present only in the strand to be packaged.

It is important that some copies of the virus genome are not packaged: those that are serving as templates for synthesis of DNA, RNA, or protein. The packaging signal must be masked in these copies of the genome.

Virus genomes are packaged into small volumes. The dsDNA of large icosahedral viruses, such as herpesviruses and the tailed phages, is packed so tightly that the pressure within the capsid is about ten times greater than the pressure inside a champagne bottle.

The tight packaging of viral genomes means that repulsion between the negative charges on their phosphate groups must be overcome. This may be achieved by packaging basic proteins, which are positively charged, along with the genome. Some ssRNA viruses (e.g. rhabdoviruses, influenza viruses, and retroviruses) coat their genomes with basic protein, while some dsDNA viruses (e.g. adenoviruses and baculoviruses) package a basic protein in close association with the genome. All of these basic proteins are virus encoded, but papillomaviruses and polyomaviruses coat their DNA with cell histones, the basic proteins associated with the DNA in eukaryotic cells. Some viruses package other positively charged materials, such as polyamines and cations.

8.2.4 Assembly mechanisms

In the laboratory, virions can be disassembled into their component molecules. For some viruses infectious virions can be reassembled from the purified components (protein and nucleic acid), under appropriate conditions of pH and in the presence of certain ions. Virions that can self-assemble in this way are those with a relatively simple structure, composed of a nucleic acid and one or a small number of protein species. Viruses that can self-assemble in a test tube are assumed to undergo self-assembly in the infected cell. Examples are tobacco mosaic virus (helical symmetry) and the ssRNA phages (icosahedral symmetry). Self-assembly is economical because no additional genetic information is needed for the assembly process.

The virions of more complex viruses (e.g. herpes-viruses and the tailed phages) do not reassemble from their components in a test tube. The environment within the infected cell is required and the virions are constructed by a process of directed assembly.

Directed assembly of icosahedral viruses may involve proteins that are temporarily present while the virion is under construction, but are not present in the mature virion. These proteins are known as scaffolding proteins (Figure 8.1). Once their job is completed they are removed from the procapsid. Some are removed by proteolysis, while others are removed intact and can be recycled. Scaffolding proteins of the tailed phages play roles in determining the size and shape of the phage head.

Some viruses synthesize polyproteins from which the individual virus proteins are cleaved by proteases; many of the cleavages take place during the assembly

process. Picornaviruses and retroviruses are two groups of viruses that utilize this strategy. The final cleavage of a picornavirus polyprotein does not occur until a copy of the genome enters the procapsid (Section 14.4.5), while the internal structures of the retrovirus virion are formed from polyprotein cleavage products late in the assembly process (Chapter 17).

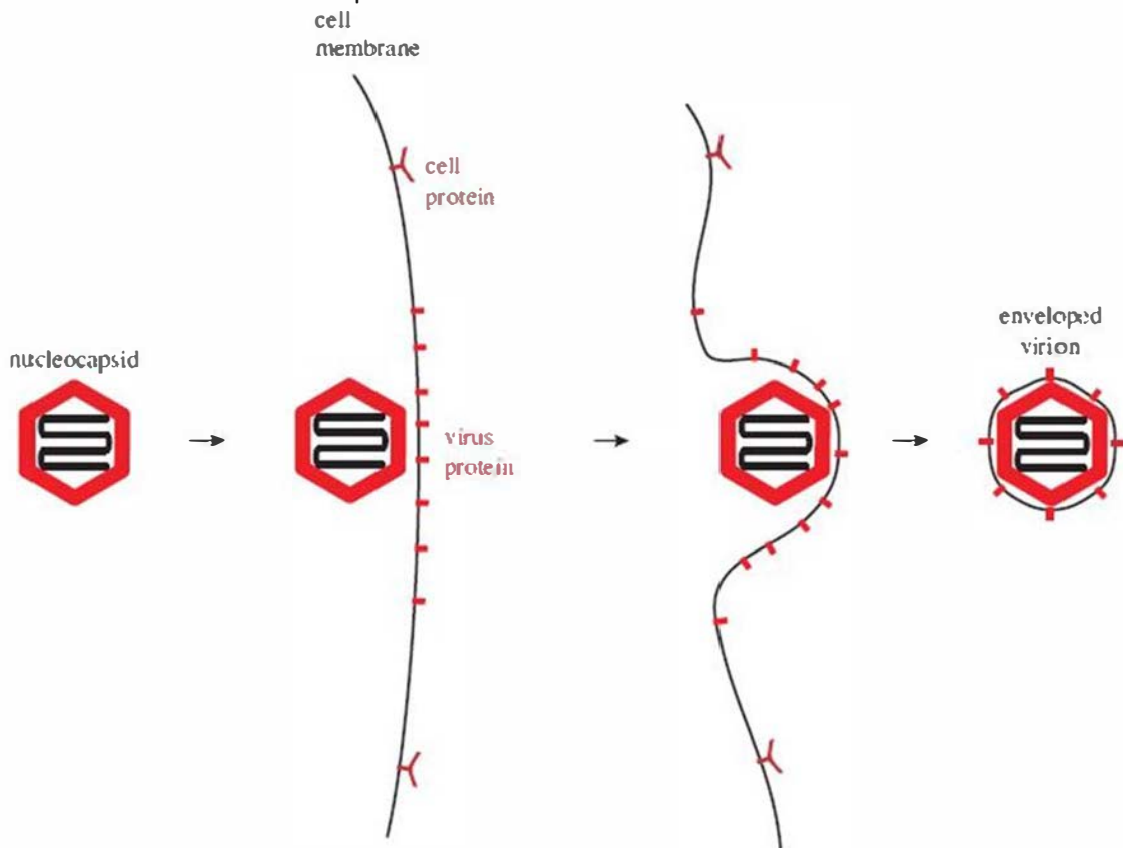
8.3 FORMATION OF VIRION MEMBRANES

Enveloped virions acquire their membrane envelopes by one of two mechanisms; either they modify a host cell membrane and then nucleocapsids bud through it, or the virus directs synthesis of a new membrane, which forms around the nucleocapsids.

8.3.1 Budding through cell membranes

Most enveloped viruses acquire their envelopes by budding through a membrane of the host cell (Figure 8.2). For viruses with eukaryotic hosts this membrane is often the plasma membrane; the virions of most retroviruses and rhabdoviruses acquire their envelopes in this way. Membrane regions where budding will occur become modified by the insertion of one or more species of virus protein, the vast majority of which are glycoproteins (Section 6.4.3.a). Nucleocapsids accumulate adjacent to the regions of membrane containing virus protein.

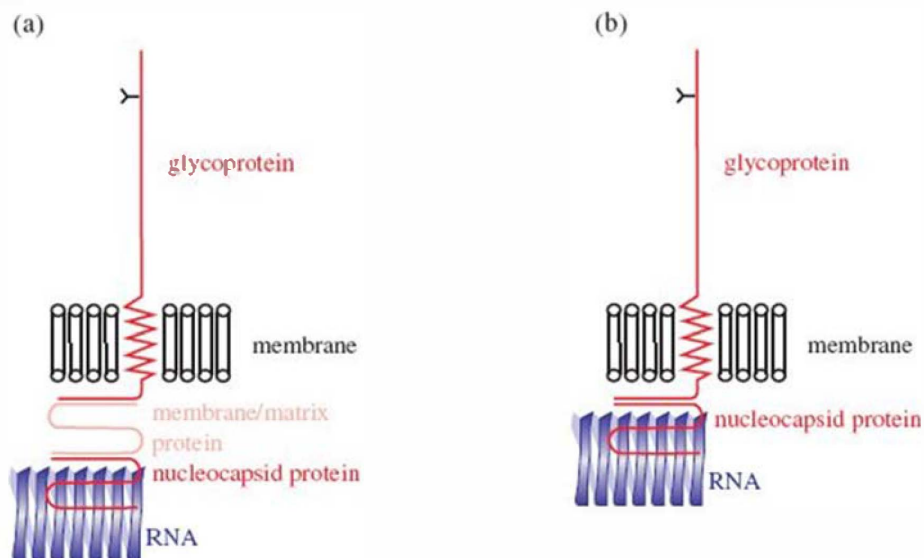
Figure 8.2 Acquisition of a virion envelope by budding. Virus membrane proteins, usually glycoproteins, become incorporated into regions of a cell membrane, often the plasma membrane. A nucleocapsid buds through the membrane, which pinches off to form the virion envelope.



Integral proteins in a lipid bilayer have a degree of mobility as they are “floating in a sea” of lipid. Virus proteins accumulate in regions of membrane from which cell proteins, to a large extent, become excluded; either the virus proteins repel the cell proteins, and/or the virus protein molecules have affinities for each other. Cell proteins may not be totally excluded from these regions and may become incorporated into virus envelopes; for example, HIV envelopes contain major histocompatibility complex class II proteins.

Budding of virions involves interaction between the cytoplasmic tail of a virus glycoprotein in the membrane and another virus protein. In a number of virus groups, including paramyxoviruses and rhabdoviruses, this protein is the M (membrane, matrix) protein (Figure 8.3(a)). M proteins have an affinity for membranes, and bind to nucleocapsids as well as to the virus glycoproteins, “stitching” the two together during budding.

Figure 8.3 Interactions between proteins of enveloped viruses. Viruses (a) with, and (b) without a membrane/matrix protein layer between the envelope and the nucleocapsid.



The vital role played by the M protein in the budding process has been demonstrated using mutants of measles virus (a paramyxovirus) and rabies virus (a rhabdovirus). Mutants defective in the M protein gene either do not bud at all or produce severely reduced amounts of budded virus from infected cells. Roles similar to that of the M protein are played by the M1 protein of influenza A virus (Section 16.3.4) and the MA (matrix) domain of the retrovirus Gag protein (Section 17.3.6).

Not all enveloped viruses have a layer of protein between the envelope and the nucleocapsid (Figure 8.3(b)). In the virions of alphaviruses (e.g. yellow fever virus) the surface of the nucleocapsid interacts directly with the cytoplasmic tails of the glycoproteins in the membrane.

Viruses that bud from the cell do so from particular regions of the plasma membrane, and if the cell is polarized then budding may take place primarily from

one surface. Body surfaces, such as the respiratory tract, are lined with epithelial cells that are polarized. Each cell has an apical (outer) and a basolateral (inner) surface and many viruses bud preferentially from one of these. Influenza A virus buds almost exclusively from the apical surface, while vesicular stomatitis virus buds almost exclusively from the basolateral surface. Experiments in which the envelope protein genes of these viruses were expressed in cells indicated that each protein has a signal that targets it to the surface from which the virus buds (Section 6.5).

The final stage of budding involves the membrane pinching off to release the virion. For many viruses host cell proteins bind to virus proteins and play vital roles in severing the membrane between the cell and the virion. These cell proteins are known as ESCRT (endosomal sorting complex required for transport) proteins. Their roles in the cell are in budding processes that give rise to organelles called multivesicular bodies. Binding sites for ESCRT proteins have been identified in the Gag proteins of most retroviruses and in the matrix proteins of rhabdoviruses, paramyxoviruses, and filoviruses. The binding sites in these virus proteins are known as late (L) domains because they are involved in the late stage of budding. Mutations in these domains can result in failure of virions to be released from the cell surface.

The release of influenza virus particles from cells requires the activity of a virion enzyme, a neuraminidase activity associated with an envelope glycoprotein. Cleavage of neuraminic acid, the substrate for the enzyme, is essential for the release of virions. Some anti-influenza virus drugs have been developed that act by inhibiting the neuraminidase (Section 26.3.6).

Some enveloped viruses bud from membranes other than the plasma membrane. Herpesvirus nucleocapsids are constructed in the nucleus and begin their journey to the cytoplasm by budding through the inner membrane of the nuclear envelope; the virion envelope is acquired in the cytoplasm by budding into vesicles derived from the Golgi complex (Section 11.5.4). Hepadnaviruses bud into a membrane compartment between the endoplasmic reticulum and the Golgi complex (Section 19.8.7). Many viruses that bud through the membranes of organelles are transported in vesicles to the plasma membrane, where they are released from the cell.

One of the rare enveloped phages, *Acholeplasma laidlawii* virus L2, acquires its envelope by budding through the cell membrane of its mycoplasma host.

8.3.2 De novo synthesis of viral membranes

A minority of viruses direct the synthesis of lipid membrane late in the replication cycle. In some cases the membrane forms a virion envelope (e.g. pox-viruses); in other cases the membrane forms a layer below the surface of the capsid (e.g. iridoviruses; Section 3.5.2).

Baculoviruses produce two types of enveloped virion during their replication. One type of virion has the function of spreading the infection to other cells within the host, and this virion acquires its envelope by budding from the plasma membrane. The function of the other type of virion is to infect new host individuals. Its envelope is laid down around nucleocapsids within the nucleus and the resulting virions become incorporated into occlusion bodies (Section 3.6).

Among the viruses of prokaryotes, the enveloped *Pseudomonas* phage $\phi 6$ acquires its envelope in the cytoplasm. The lipids are derived from the host cell membrane and the virion is released by lysis of the cell. The virion of *Alteromonas* phage PM2 is not enveloped, but it has a membrane between outer and inner protein shells. The membrane is synthesized in the infected cell.

The various origins of virion membranes are summarized in Table 8.1.

Table 8.1 Summary of virion membrane acquisition

Site of virion assembly	Origin of virion membrane	Examples	
		Virion envelope	Internal virion membrane
Eukaryotic cell cytoplasm	Plasma membrane	Most retroviruses Most rhabdoviruses	
	Between endoplasmic reticulum and Golgi	Hepadnaviruses	
	<i>De novo</i> synthesis	Poxviruses	Iridoviruses
Eukaryotic cell nucleus	Inner nuclear membrane	Nucleorhabdoviruses	
	<i>De novo</i> synthesis	Baculovirus virions that will be occluded	
Prokaryotic cell	Plasma membrane	<i>Acholeplasma laidlawii</i> virus L2	
	<i>De novo</i> synthesis	<i>Pseudomonas</i> phage $\phi 6$	<i>Alteromonas</i> phage PM2

8.4 VIRION EXIT FROM THE INFECTED CELL

The virions of many viruses are released from the infected cell when it bursts (lyses), a process that may be initiated by the virus. Many phages produce enzymes (lysins, such as lysozymes) that break bonds in the peptidoglycan of the host bacterial cell walls. Other phages synthesize proteins that inhibit host enzymes with roles in cell wall synthesis; this leads to weakening of the cell wall and ultimately to lysis.

Average yields of infectious virions per cell (burst sizes; Section 2.8.3) vary considerably. A typical burst size for phage T4 (a large virus in a small host cell) is 200, while a typical burst size for a picornavirus (a small virus in a large host cell) is 100 000.

Many viruses do not lyse their host cells; instead, progeny virions are released from the cells over a period of time. Obviously the concept of burst size is not relevant for this situation. We have already seen how virions that acquire an envelope from the plasma membrane leave the cell by budding. Virions that acquire envelopes

from internal membranes of the cell, however, are still within the cytoplasm after acquiring their envelopes. These viruses, and some naked viruses, leave the cell by exocytosis and similar mechanisms. To reach the plasma membrane they are transported in vesicles, which fuse with the plasma membrane to release the virions. Other viruses, such as vaccinia virus, attach via motor proteins to microtubules (Section 5.2.5) and use this transport system to reach the cell surface. The filamentous phages are naked viruses with a unique mechanism for exiting from their bacterial hosts: extrusion through the cell surface. The ssDNA genomes of the phages are coated with protein within the cell; as they are extruded from the cell they lose this protein and acquire their capsids (Section 20.4.2.c).

After leaving the host cell most virions remain inert unless they encounter a new host cell. There are some cases, however, where maturation of virions is completed after they have left the cell. In the retrovirus virion, for example, polyprotein cleavages commence during budding and continue after the virion has been released from the cell (Section 17.3.6). In another example an unusual lemon-shaped phage of an archaeon (*Acidianus convivator*) develops two long tails after it has left its host.

Because of their modes of transmission, most plant viruses leave their host cells in ways that differ from those of animal and bacterial viruses. Plant cells are separated from each other by thick cell walls, but in many of the cell walls there are channels, called plasmodesmata, through which the plant transports materials. Viruses are able to spread within the host by passing from cell to cell through plasmodesmata (Section 9.4.1). For spread to new hosts many plant viruses leave the host cell as a component of the meal of a vector (e.g. an aphid or a nematode) that feeds by ingesting the contents of cells (Section 4.2).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- describe the assembly mechanisms for nucleocapsids with (a) helical symmetry and (b) icosahedral symmetry;
- discuss the origins of internal virion membranes and of virion envelopes;
- explain the roles played by membrane/matrix proteins in the budding of some enveloped viruses;
- describe mechanisms used by viruses to exit from cells.

SOURCES OF FURTHER INFORMATION

Book

Hunter, E. (2007) Virus assembly. Chapter 6 in Fields Virology, 5th edition (Knipe, D. M. and Howley, P. M., editors-in-chief), Lippincott, Williams and Wilkins

Journals

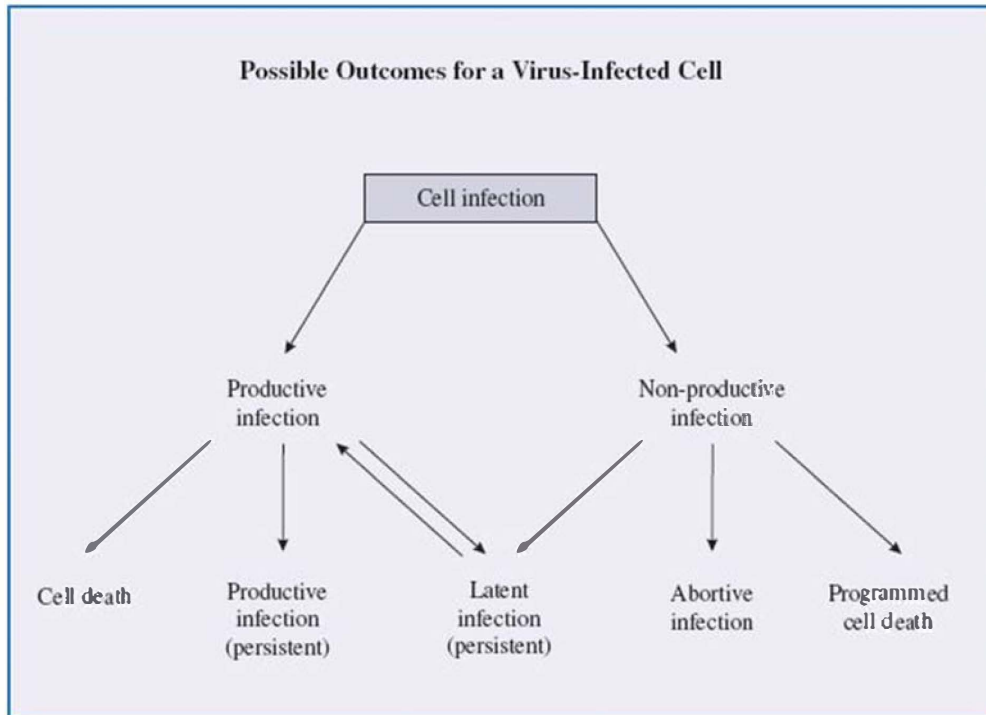
- Aksyuk, A. A. and Rossmann, M. G. (2011) Bacteriophage assembly. *Viruses*, 3, 172–203
- Bieniasz, P. D. (2006) Late budding domains and host proteins in enveloped virus release. *Virology*, 344, 55–63
- Calistri, A. et al. (2009) Role of multivesicular bodies and their components in the egress of enveloped RNA viruses. *Reviews in Medical Virology*, 19, 31–45
- Mothes, W. et al. (2010) Virus cell-to-cell transmission. *Journal of Virology*, 84, 8360–8368
- Ren, X. and Hurley, J. H. (2011) Proline-rich regions and motifs in trafficking: from ESCRT interaction to viral exploitation. *Traffic*, 12, 1282–1290
- Ward, B. M. (2011) The taking of the cytoskeleton one two three: how viruses utilize the cytoskeleton during egress. *Virology*, 411, 244–250
- Welsch, S., Müller B., and Kräusslich, H. G. (2007) More than one door—budding of enveloped viruses through cellular membranes. *FEBS Letters*, 581, 2089–2097

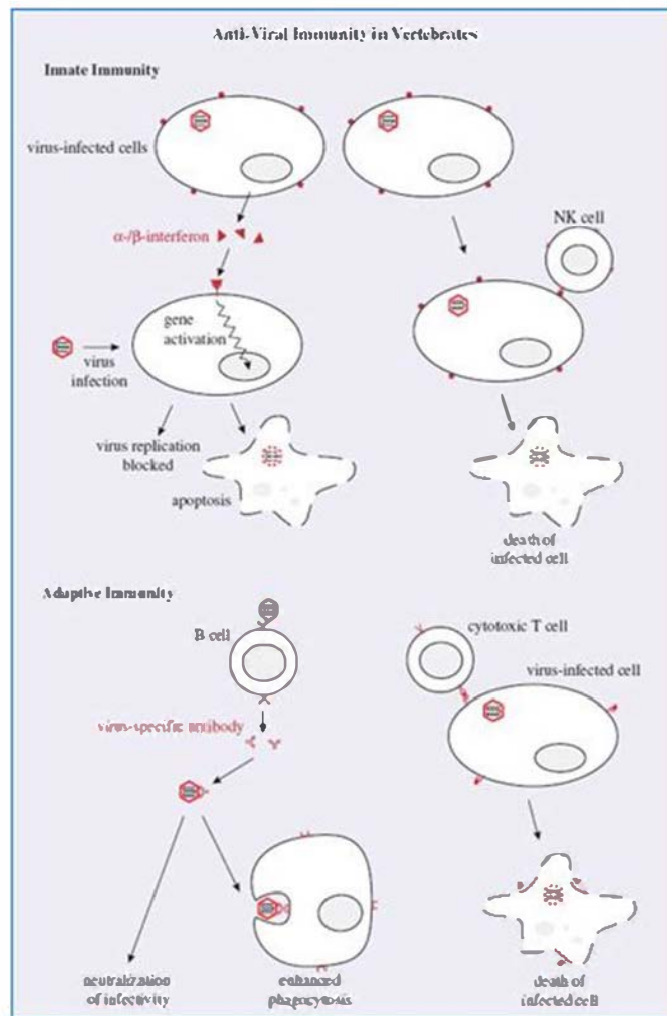
Zlotnick, A. and Mukhopadhyay, S. (2011) Virus assembly, allostery and antivirals. Trends in Microbiology, 19, 14–23

CHAPTER 9

Outcomes of Infection for the Host

CHAPTER 9 AT A GLANCE





9.1 INTRODUCTION TO OUTCOMES OF INFECTION FOR THE HOST

In the previous few chapters we have looked at aspects of the virus replication cycle that culminate in the exit of infective progeny virions from an infected cell. When this is the outcome the infection is said to be productive. Virions may be released when the host cell lyses. Alternatively, the cell may survive releasing virions for a period, which may be short, as in the case of HIV infection, or long, as in the case of hepatitis B virus infection.

Some virus infections, however, are not productive, for a variety of reasons.

- An infection may become latent with the virus genome persisting, perhaps for the lifetime of the cell, and perhaps in the daughter cells if the cell divides.
- An infection may be abortive, in which case neither a latent infection nor a productive infection is established. One cause of abortive infection is a virus with a mutated genome; the virus is said to be defective. It is unable to undergo a complete replication cycle, unless the cell is also infected with a virus that can provide the missing function(s). A virus that is able to provide functions for a defective virus is known as a helper virus.

Some virus infections persist in their hosts for long periods, sometimes for life. In some cases persistent infections are productive, e.g. HIV; in other cases persistence may take the form of periods of latency alternating with periods of productive infection, e.g. many herpesviruses. Long-term infection with some viruses may lead to the development of cancer (Chapter 23).

Some phages, including the filamentous phages (Section 20.4.2), initiate persistent productive infections. Under appropriate conditions the host cells can survive for long periods releasing progeny virions.

The outcome of a virus infection from the point of view of the host may vary from harmless at one end of the scale, through deleterious effects of varying severity, to death at the other end of the scale. The outcome may hinge on complex interplay between a variety of host, viral, and environmental factors.

Over time the organisms that are hosts to viruses have evolved anti-viral defenses, but the viruses have not sat idly by. The relationship between a virus and its host usually involves an arms race and viruses have developed countermeasures against host defenses. Some of these countermeasures are indicated in boxes throughout this chapter.

In this chapter we concentrate mainly on virus infections in vertebrate hosts, and we start by considering some aspects of host immunity.

9.2 FACTORS AFFECTING OUTCOMES OF INFECTION

A major factor affecting the outcome of a virus infection is the efficiency of the host's immune systems. In animals there are both innate and adaptive immune systems, each made up of a large number of molecules and cell types. There are links between the innate and adaptive immune systems; for example, aspects of the innate system stimulate the development of an adaptive immune response.

9.2.1 Innate immunity in vertebrates

Some aspects of innate immunity that viruses might encounter prior to infecting cells were introduced in Section 4.3.3. Further components of innate immunity that might be encountered after cells have become infected include complement, interferons, natural killer (NK) cells, APOBEC3 proteins, and tetherin.

9.2.1.a Complement

The complement system consists of cell surface receptors and soluble factors that can detect and respond to foreign invaders, including viruses. Complement can be activated by three pathways that result in modifications to complement proteins, with a number of possible outcomes for a virus. Complexes of modified complement proteins can insert themselves into membranes containing virus proteins, damaging enveloped virions and lysing infected cells. Some modified complement proteins can coat virions, which then attach to phagocytes (neutrophils and macrophages) via

complement receptors on the phagocyte surface. Complement activation can also enhance adaptive immune responses.

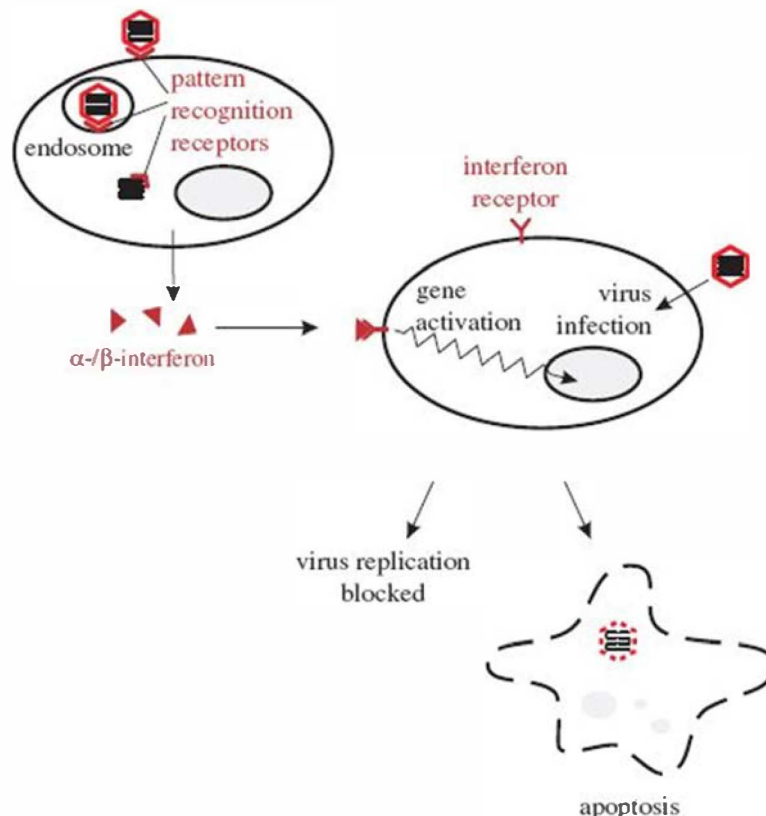
VIRAL COUNTERMEASURES AGAINST COMPLEMENT

Many viruses have evolved defenses against the activities of complement; for example, herpes-viruses and poxviruses produce proteins that reproduce the activities of host proteins that regulate complement activity.

9.2.1.b Interferons

Interferons are proteins synthesized and secreted by cells in response to virus infection. The roles of interferons are to protect adjacent cells from infection and to activate T cell-mediated immunity (Section 9.2.2.b). Interferon synthesis is triggered by virus proteins and nucleic acids binding to pattern recognition receptors, of which there are various types located at the cell surface, within endosomes, and in the cytoplasm (Figure 9.1). One group of these receptors is known as Toll-like receptors; some of them are able to recognize virus nucleic acids as foreign.

Figure 9.1 Interferons. Virus components at the cell surface, within endosomes, and in the cytoplasm can bind to pattern recognition receptors, triggering the cell to synthesize interferons. Secreted interferons bind to receptors on other cells, activating a wide spectrum of genes. If these cells become infected the gene products may either block virus replication or cause the infected cells to die (apoptosis).



There are a number of types of interferon. Alpha- and beta- (O_{\pm} - and O_{I} -) interferons are produced by most cell types when they become infected with

viruses. After secretion the interferons may bind to receptors on nearby cells (Figure 9.1), activating a large number of genes and triggering various anti-viral activities, including the following:

- activation of genes that encode antiviral proteins. One example of an interferon-induced antiviral protein is dsRNA-dependent protein kinase R. This enzyme is synthesized in an inactive form, but when it binds to dsRNA it can phosphorylate several substrates, resulting in either a block to virus replication or the death of the infected cell. Tetherin (Section 9.2.1.e) is another example of an interferon-induced antiviral protein;
- production of major histocompatibility (MHC) class I molecules and proteasome proteins; these molecules enhance the presentation on the infected cell surface of viral peptides to T cells;
- activation of NK cells (Section 9.2.1.c);
- induction of apoptosis (Section 9.2.4).

Another type of interferon, gamma- (O_i-) interferon, is produced mainly by T cells and NK cells when triggered by certain molecules (e.g. interleukin-2) released during immune responses. O_i-interferon has a number of effects, including stimulation of antigen presentation and activation of phagocytes and NK cells.

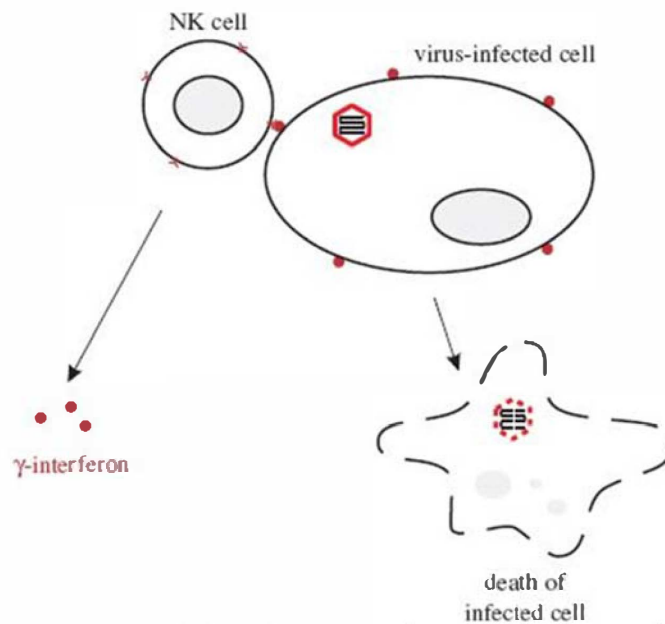
VIRAL COUNTERMEASURES AGAINST INTERFERONS

Most vertebrate viruses produce proteins that inhibit either the production of interferons or their activities. Many RNA viruses, such as poliovirus prevent the synthesis of interferons as a result of a general inhibition of cell gene expression. Some virus proteins, e.g. the NS1 protein of influenza A virus and the NS3-4A protein of hepatitis C virus, block pathways involved in interferon production. Other viral countermeasures include inducing the breakdown of interferon system components.

9.2.1.c Natural killer (NK) cells

NK cells are present throughout the body, but mainly in the blood. They are lymphocyte-like cells that can recognize changes in the surface molecules of virus-infected cells as a result of infection, though they do not recognize specific antigens, unlike B cells and T cells (Section 9.2.2). After recognizing virus-infected cells as target cells, NK cells are able to bind to them and then, as their name implies, kill them (Figure 9.2).

Figure 9.2 Activities of natural killer (NK) cells. After binding to a virus-infected cell, an NK cell releases O_i-interferon and kills the infected cell.



NK cells kill their target cells either by releasing perforins, which are proteins that are inserted into the plasma membrane of the virus-infected cell, or by inducing apoptosis (Section 9.2.4). Also, on binding to infected cells, NK cells release γ -interferon.

VIRAL COUNTERMEASURES AGAINST NK CELLS

Many viruses produce proteins and/or small RNA molecules (known as microRNAs) that reduce the NK cell response. The presence of HIV particles in the blood alters the expression of a number of molecules on the surface of NK cells, reducing the ability of the cells to kill virus-infected cells and to secrete γ -interferon. Human cytomegalovirus produces a microRNA that prevents NK cells from recognizing virus-infected cells.

9.2.1.d APOBEC3 proteins

There are enzymes in the cells of humans and animals that can interfere with the replication of reverse transcribing viruses (retroviruses and hepadna-viruses). These enzymes can induce lethal mutations by deaminating deoxycytidine to deoxyuridine during reverse transcription. The enzymes delight in the name apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3 proteins; the abbreviation APOBEC3 proteins is easier to cope with!

Several of these proteins in human cells can interfere with the replication of HIV. Two of them (APOBEC3F and APOBEC3G) can be incorporated into HIV virions and taken into the next cell, where they can wreak havoc during reverse transcription.

VIRAL COUNTERMEASURES AGAINST APOBEC3 PROTEINS

In the cytoplasm of an HIV-infected cell the virus protein Vif can bind APOBEC3G, triggering its degradation and hence preventing its incorporation into virions.

9.2.1.e Tetherin

Tetherin is an interferon-induced protein that is expressed at the cell surface as dimers anchored in the plasma membrane. Tetherin can prevent the budding of

enveloped viruses from the cell surface by forming a tether between the lipid membranes of the cell and the virion.

VIRAL COUNTERMEASURES AGAINST TETHERIN

Several enveloped viruses produce proteins that counteract the effect of tetherin by downregulating its expression and/or by inducing its degradation. An example of an anti-tetherin protein is the Vpu protein of HIV-1.

9.2.2 Adaptive immunity in vertebrates

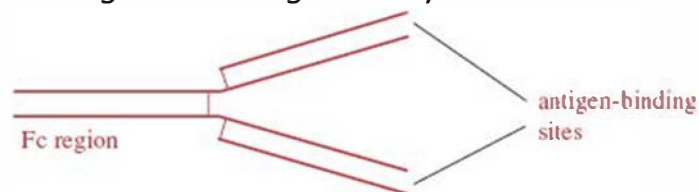
An important outcome of virus infection in a vertebrate host is the development of a virus-specific immune response triggered by the virus antigens. Regions of antigens known as epitopes bind to specific receptors on lymphocytes, activating cascades of events that result in the adaptive immune response.

The key cells in specific immune responses are lymphocytes, of which there are two classes: B lymphocytes (B cells), which develop in the Bursa of Fabricius in birds and in the bone marrow in mammals, and T lymphocytes (T cells), which develop in the thymus. Each lymphocyte has epitope-specific receptors on its cell surface, making it specific for a particular epitope. Naive lymphocytes are those that have not encountered their specific epitopes; these cells have surface molecules and circulation patterns in the body distinct from lymphocytes that have previously encountered their epitopes.

9.2.2.a Antibodies

Antibodies are glycoproteins of a type known as immunoglobulins. The basic structure of an antibody molecule is shown in Figure 9.3. The molecule is constructed from two ν heavy ν and two ν light ν polypeptide chains and contains two antigen-binding sites and a region known as the Fc (Fragment crystallizable) region.

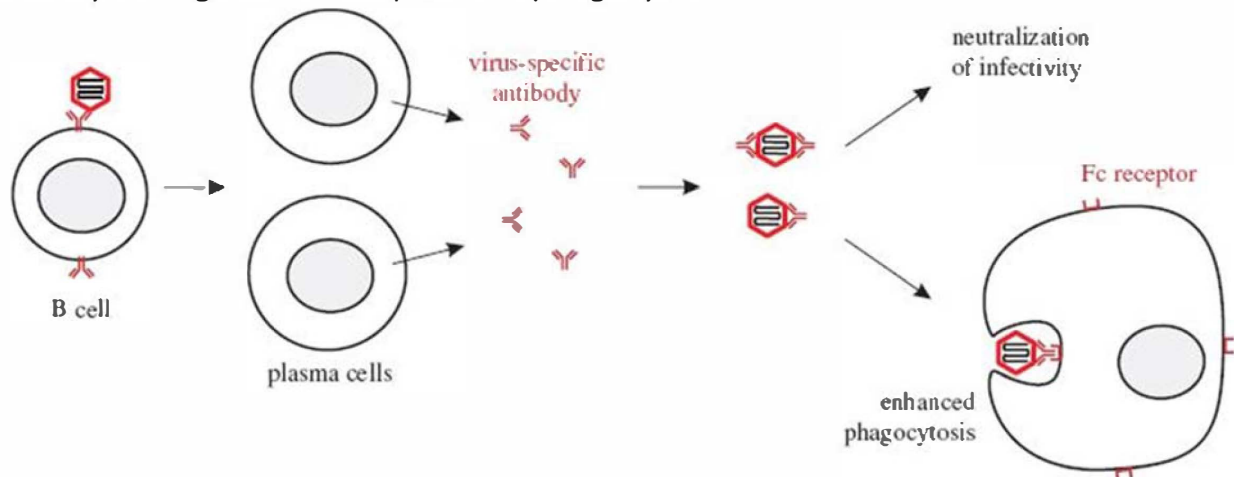
Figure 9.3 An antibody molecule. The molecule consists of four polypeptide chains held together by disulfide bonds. The chains are arranged to form two sites that can bind to a specific antigen. Fc: Fragment crystallizable.



There are several classes of immunoglobulin (Ig), the most significant from the point of view of anti-viral immunity being IgG and IgM in the blood, and IgA at mucosal surfaces. IgG molecules are monomers with the structure depicted in Figure 9.3, while IgA and IgM molecules are normally dimers and pentamers, respectively.

Antigen-specific antibodies are synthesized by plasma cells, which develop from a B cell after it has been stimulated by interaction between the antigen and a specific

immunoglobulin receptor at the cell surface. Antibodies play important roles in several aspects of anti-viral immunity, some of which are summarized in Figure 9.4. Figure 9.4 Production of virus-specific antibody and some of its roles in immunity. Virus-specific antibody is secreted by plasma cells derived from B cells that have recognized virus antigens. The antibody may bind to virions and neutralize their infectivity. Antibody-coated virions may be phagocytosed after attachment of antibody Fc regions to receptors on phagocytes.



Virus-specific antibody can coat both virions and virus-infected cells, and this may lead to their destruction by a variety of mechanisms. A number of cell types in the immune system have receptors for the Fc region of IgG (Figure 9.3), allowing these cells to attach to antibody-coated virions and cells. Cells that have IgG Fc receptors include:

- neutrophils and macrophages (these are phagocytes and may phagocytose the antibody-coated materials; they may also kill cells without phagocytosing them);
- NK cells (may kill virus-infected cells by insertion of perforins into their membranes).

Antibody binding to virions can result in neutralization of infectivity, which may occur by a variety of mechanisms.

- Release of nucleic acid from virions. In studies with several viruses, including poliovirus, it was found that antibodies can attach to virions, and then detach leaving empty capsids devoid of their genomes.
- Prevention of virion attachment to cell receptors. Antibody bound to a virion may mask virus attachment sites. Not all attachment sites, however, are accessible to antibodies; those of most picornaviruses are in deep canyons (Section 14.3.1).
- Release of virions that have attached to cell receptors.
- Inhibition of entry into the cell. Antibody coating fusion proteins on an enveloped virion may inhibit fusion of the envelope with a cell membrane.
- Inhibition of genome uncoating.

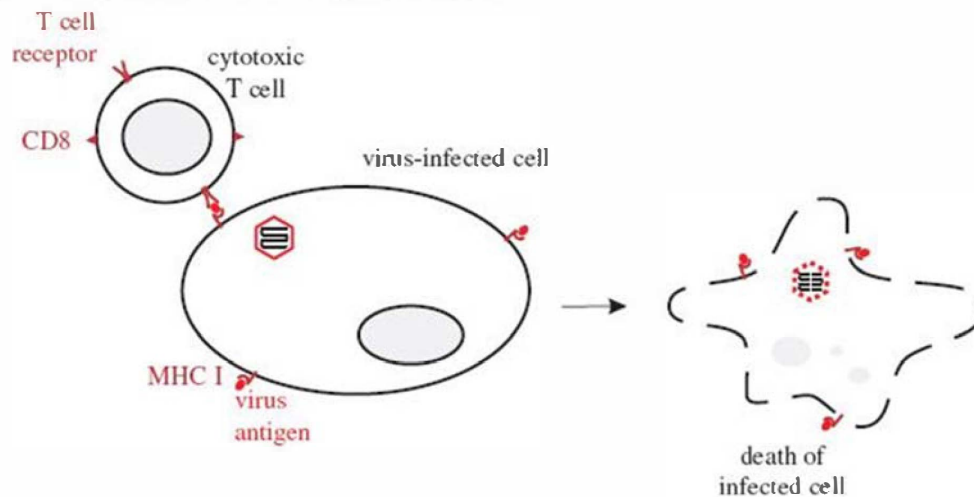
Another outcome of antibody binding to virus antigens is the activation of complement, which can have a number of anti-viral effects (Section 9.2.1.a).

9.2.2.b T cells

Several days after antigenic stimulation, naive or memory T cells develop into effector T cells, of which there are two classes.

- Helper T cells are characterized by the presence of CD4 molecules at the cell surface. These cells secrete cytokines, which play essential roles in the initiation of immune responses, for example in the maturation of cytotoxic T cells and in triggering B cells to develop into antibody-secreting cells and memory B cells.
- Cytotoxic T cells kill virus-infected cells and are characterized by the presence of CD8 molecules at the cell surface (Figure 9.5). There is a requirement for viral antigens to be expressed at the surface of target cells. The antigens may be virion surface proteins (e.g. envelope glycoproteins) though often the target antigens are internal virion components, or even non-structural proteins. Cytotoxic T cells specific for early virus proteins may destroy virus-infected cells before any infectious virus is produced.

Figure 9.5 Recognition and killing of a virus-infected cell by a cytotoxic T cell. A cytotoxic T cell kills a virus-infected cell after the T cell receptors have recognized fragments of virus antigens displayed in association with MHC class I molecules at the cell surface.



The viral antigens are displayed on the surface of infected cells in association with MHC class I molecules. This flags infected cells for destruction by cytotoxic T cells, which can kill target cells by insertion of proteins (perforins) into their membranes or by inducing apoptosis.

VIRAL COUNTERMEASURES AGAINST CYTOTOXIC T CELLS

Some viruses, such as herpesviruses, reduce the level of expression of MHC molecules at the surface of the cells they infect, thereby making it more difficult for cytotoxic T cells to recognize infected cells.

9.2.2.c Immunological memory

The quantity and quality of the adaptive immune response depends on whether or not the host is encountering the virus for the first time. Some B cells and T cells can survive as memory cells long after the first or subsequent encounters with the virus antigens. Memory cells have returned to a resting state, from which they can be reactivated if they encounter the same antigens again. These cells are the basis of immunological memory, which can be formed as a result of a natural infection, but also as a result of encountering antigens in vaccines.

The outcome of infection of a vertebrate animal with a virus may depend on whether or not the host has immunological memory of the virus antigens. If immunological memory is present then signs and symptoms of disease are likely to be less severe, or totally absent.

9.2.3 RNA silencing

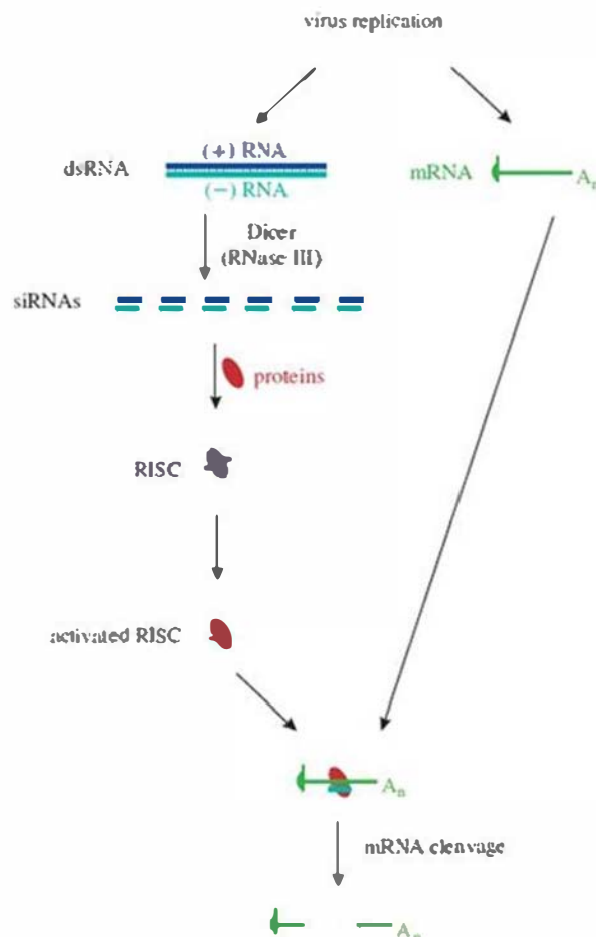
RNA silencing is also known as post-transcriptional gene silencing or RNA interference (RNAi). It is an intracellular process that is induced by dsRNA, which is present in cells infected with dsRNA viruses, and may be present in cells infected with ssRNA viruses as double-stranded regions of ssRNA, and as intermediates during RNA replication. RNA silencing results in the destruction of mRNAs that have the same sequence as the inducing dsRNA; both cellular and viral mRNAs can be destroyed.

The process, which is outlined in Figure 9.6, involves cleavage of the dsRNA into small interfering dsRNAs (siRNAs), 21–25 bp long. The cleavages are performed by a protein complex containing an enzyme called Dicer, which belongs to the RNase III family. These enzymes are unusual in that they show specificity for dsRNAs and cleave them to fragments with 3'–5' overhangs of two or three nucleotides. Each of the siRNAs joins a complex of proteins to form a RISC (RNA-induced silencing complex). The double-stranded siRNA is unwound and the minus-strand RNA remains associated with the complex, which is now an activated RISC. The minus-strand RNA selects the target mRNA by complementary base-pairing, then the mRNA is degraded in that region.

Figure 9.6 RNA silencing. dsRNA is processed by Dicer to siRNAs, which associate with RISC proteins. An activated RISC contains minus-strand RNA, which targets the complex to the specific mRNA. The mRNA is cleaved by a RISC protein.

RISC: RNA-induced silencing complex.

siRNA: small interfering RNA.



RNA silencing has been found in plants and fungi, and in invertebrate and vertebrate animals. In plants and insects it is an important anti-viral defense mechanism, and this may also be true for other organisms. RNA silencing can also be initiated by the introduction of synthetic dsRNA (Section 2.9.3).

VIRAL COUNTERMEASURES AGAINST RNA SILENCING

Most plant viruses encode proteins that can suppress RNA silencing; for example, the ω helper-component proteinase of potyviruses and the P19 protein of tombusviruses are strong suppressors of RNA silencing.

9.2.4 Programmed cell death

Virus infection of a cell may initiate a process that causes the death of the cell before progeny virus has been produced, hence preventing the spread of infection to other cells. In animals this suicide mechanism is known as apoptosis. It can be triggered by virions or virus components binding to cell surface molecules or to molecules within cells; it can also be triggered when the life-span of cells, such as epithelial cells, is complete. Cell proteins that play key roles in apoptosis include Bcl-2 (B-cell lymphoma protein-2) and caspases (proteases that degrade essential cell proteins, such as those that repair DNA and maintain cell structure).

VIRAL COUNTERMEASURES AGAINST APOPTOSIS

Animal viruses have evolved mechanisms that can suppress apoptosis at a variety of points in the process. Several DNA viruses encode proteins related to the Bcl-2 proteins, while many viruses produce inhibitors of caspases. These viral proteins

block apoptosis, resulting in the survival of host cells and the completion of virus replication cycles.

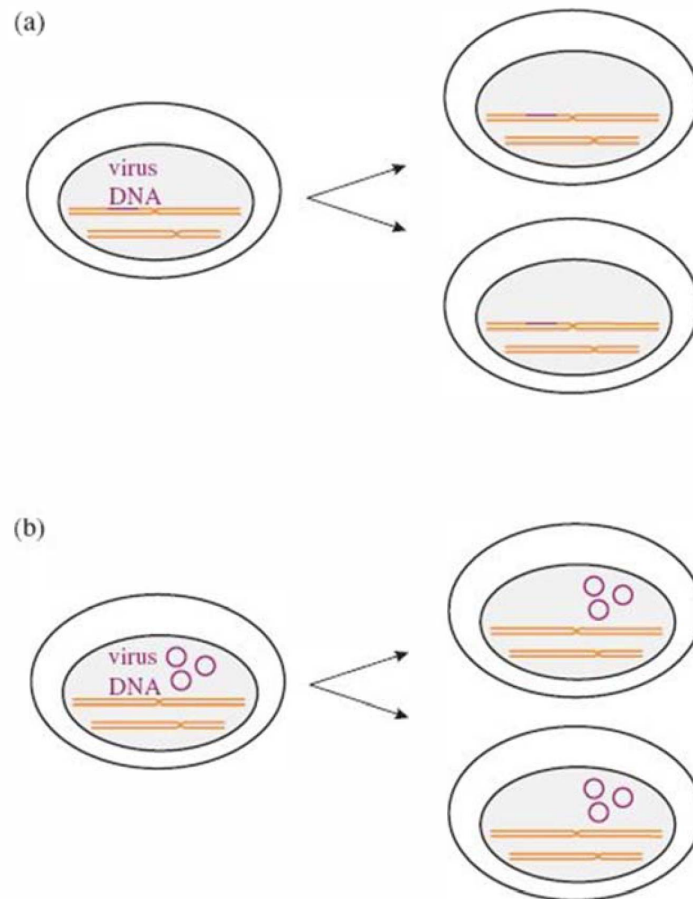
Bacteria have developed similar mechanisms to protect the species from phage infection. The death of a host bacterium before any progeny phage has been produced protects other susceptible cells from infection. These mechanisms have been found in many species, including *Escherichia coli*.

9.3 NON-PRODUCTIVE INFECTIONS

Under some circumstances a virus infects a cell, but the replication cycle is not completed. If the virus genome persists in the cell the infection is said to be latent; otherwise, it is an abortive infection. These situations are discussed in the next two sections.

9.3.1 Latent infections

When a latent infection is initiated the virus genome is maintained in the infected cell either as a sequence of DNA integrated into the cell genome, or as multiple copies of covalently closed circular DNA (Figure 9.7). In eukaryotic cells the virus DNA is associated with cell histones, which play roles in the maintenance of latency. Figure 9.7 Maintenance of virus genomes in cells with latent infections. (a) Virus DNA integrated into a cell chromosome, e.g. retroviruses. (b) Virus DNA present as multiple copies of circular molecules. After infection of the cell the virus genome is circularized and replicated, e.g. most herpesviruses. If a latently infected cell divides, the virus genome is normally replicated along with the cell genome so that each daughter cell contains a copy/copies of the virus genome.



In the case of retroviruses, the integration of a copy of the virus genome into the cell genome occurs at an early stage in the virus replication cycle (Section 17.3.3). When the infection is latent events do not progress beyond this stage, but if the intracellular environment subsequently changes to become favorable for replication of the virus then the latent infection may progress to a productive infection.

The phenomenon of latent phage infection of a bacterium is known as lysogeny and the phage is said to be temperate (Section 20.1). Lysogeny has been observed for many dsDNA phages, and for some ssDNA phages. The phage genome (the prophage) can persist in the cell in ways similar to the persistence of latent animal virus infections (Figure 9.7). In many cases the prophage is integrated into the bacterial genome, but some prophages persist as non-integrated circular DNA. A temperate phage genome may include a gene that can confer a selective advantage on the bacterial host. Some virulence factors of parasitic bacteria are encoded by temperate phages, an example being the Shiga toxin produced by some strains of *E. coli*.

During a latent infection the virus genome may be totally shut down, or a few virus genes may be expressed as proteins and/or non-coding RNAs. If the latently infected cell divides then the virus DNA is replicated and each daughter cell has one or more copies of the virus genome (Figure 9.7). When this happens the virus has achieved its objective of replicating its genes, and it has done so without destroying the host cell!

Many latent infections may be activated to become productive infections, a process known as induction. This may occur if any of the following situations arises.

- A eukaryotic host cell moves into another phase of the cell cycle (Figure 4.5).
- The host cell is irradiated with ultraviolet light. This might trigger, for example, a latent phage infection to replicate and lyse the host cell, or a latent herpes simplex virus infection to replicate and cause a cold sore.
- A host organism becomes immunocompromised. This is another trigger that can reactivate a latent herpes simplex virus infection.
- The host cell becomes infected with a second virus that provides a function that the first virus lacks. In this relationship the latent virus is a satellite virus and the second virus is a helper virus (Section 1.3.4). The adeno-associated viruses are examples of satellite viruses that can initiate latent infections. They require functions of an adenovirus or a herpesvirus for their replication, and are discussed in Chapter 12.

9.3.2 Abortive infections

An abortive infection is one where the virus replication cycle is not completed; the infection is non-productive and the virus genome does not persist in the cell as a latent infection. Some abortive infections can kill the host cell. An infection might be abortive for reasons concerning the cell, the environmental conditions, and/or the virus.

As far as the cell is concerned, virus infection may trigger immune responses, such as interferon production, that abort the infection. Some cell types may not provide all the requirements of the virus (non-permissive cells), so an infection may be aborted, in contrast to permissive cell types that support productive infections.

The virus may contain a mutated genome that may be able to initiate the replication cycle, but may not have the full set of functional genes necessary to complete it. The virus is said to be defective, and if it infects a cell without the means of complementing the mutated or missing genes the result is an abortive infection.

There are a number of types of defective virus. One type is known as a defective interfering particle (DIP). DIPs often arise in the laboratory after animal viruses have been passaged several times in cell culture at high multiplicities of infection. DIPs have also been reported after passage in chick embryos and in mice, and DIPs of plant viruses have been observed.

Most DIPs contain less nucleic acid than the standard virus from which they were derived. By themselves DIPs are either non-infectious or they can initiate only abortive infections. The standard virus can act as a helper virus, so in cells that are co-infected with a DIP and the standard virus the DIP is able to replicate its genome and produce progeny DIPs. When this happens the replication of the DIP

interferes with the replication of the standard virus and reduces its yield, hence an interfering particle.

9.4 PRODUCTIVE INFECTIONS

9.4.1 Spread of infections within multicellular hosts

Progeny virions from the first-infected cell of a multicellular host may infect cells nearby; for example, common cold viruses might infect other epithelial cells lining the respiratory tract and rotaviruses might infect other epithelial cells lining the intestinal tract. Most adjacent animal cells are separated from each other only by their plasma membranes, providing opportunities for direct cell-to-cell spread for some viruses.

Plant viruses are able to spread from cell to cell by passing through plasmodesmata. This is achieved with the aid of specialized proteins called movement proteins (MPs), which act in a variety of ways. Some MPs form complexes either with virus RNA or with virus RNA plus coat protein molecules; other MPs form tubular structures through which fully encapsidated virus RNA is transported (Table 9.1). Each plant virus encodes between one and four MPs, which are multifunctional proteins with other roles in the virus replication cycles.

Table 9.1 Involvement of plant virus movement proteins (MPs) in the spread of infection via plasmodesmata

Mode of transport through plasmodesmata	Virus examples
Virus RNA + MP complexes transported	Tobacco mosaic virus
	Cowpea chlorotic mottle virus
Virus RNA + coat protein + MP complexes transported	Cucumber mosaic virus
Virions transported through tubules composed of MP	Cowpea mosaic virus

In some circumstances progeny virions from infected cells may be transported to distant sites in the host, where susceptible cells may become infected. In the animal body the blood and the nerves may act as transport vehicles, while in plants transport may occur via the phloem.

9.4.2 Disease

Many virus infections result in no disease in the host, while at the other end of the scale a virus infection may result in fatal disease, such as rabies or AIDS. Between these extremes is a huge range of severity in diseases caused by viruses.

Disease may be manifest as symptoms and/or signs. In medical terminology symptoms are subjective features, such as abdominal pain and fatigue, whereas signs are objective features, such as blood in the feces and skin rashes. Symptoms

can be recognized only by the infected individual, whereas signs can be recognized by others. Plant virologists and insect virologists sometimes use the word symptom in connection with objective features, such as leaf lesions in a plant and iridescence in an insect (Chapter 1, At a glance).

The outcome of a complex interplay of factors determines whether or not a virus infection results in disease and, if so, the severity of the disease. These factors may include virus factors, host factors, and human intervention.

Not all viruses are pathogens (disease-causing agents) and some viruses are pathogenic only under certain circumstances. Infection with some viruses (e.g. dependoviruses, some herpesviruses, and some reoviruses) apparently never causes disease, while infection with other viruses (e.g. human parvovirus B19, poliovirus, and hepatitis B virus) may or may not result in disease. Infections that do not result in disease are said to be subclinical, inapparent, or asymptomatic. For pathogenic viruses factors that can affect the outcome of infection include the following.

- The virulence of the virus strain. The virulence of a virus (or any micro-organism) is a measure of the severity of disease it is capable of causing. Influenza A type H5N1 can be said to be more virulent in humans than types H1N1 and H3N2 (Section 16.4), as type H5N1 causes more severe disease than the other two types.
- The dose of virus. A larger dose of virus may result in a shorter incubation period (the time between infection and the first appearance of signs and/or symptoms).

Host factors that can affect the outcome of infection include the effectiveness of immune systems (discussed earlier in this chapter), and these in turn vary with age and nutritional status of the host. It should be noted that a strong immune response by the host does not guarantee the elimination of the virus. HIV can replicate in the presence of high levels of HIV-specific antibody and T cells.

In some cases signs and/or symptoms may result from the host's immune response against the virus. A cytokine storm is responsible for some aspects of AIDS, and CD8 T cells cause some of the liver damage seen as a result of hepatitis B virus infection. Herpes simplex lesions and measles rashes are clinical manifestations of attempts by the human body to destroy virus-infected cells.

Human interventions that could influence the outcome of a virus infection in a human or an animal include the administration of anti-viral antibodies or anti-viral drugs (Chapter 26).

Host recovery from a virus disease may be accompanied by elimination of the virus, as normally occurs after infection with common cold viruses in the respiratory tract and rotavirus infections in the intestine. Alternatively, recovery may be accompanied by establishment of a long-term infection, perhaps for life. Such persistent infections, whether they are productive or latent, may have no further

consequences for the host, or they may cause disease when the host is older. Childhood infection with varicella-zoster virus may result in a lifetime latent infection, which may be reactivated, causing shingles (Section 11.2.2). Long-term persistent infection of humans and animals with certain viruses may lead to the development of cancer (Chapter 23).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- describe the major components of innate and adaptive immunity in vertebrates;
- outline the process of RNA silencing;
- explain programmed cell death;
- describe how viruses can circumvent host immune responses;
- explain the terms:
 - productive infection;
 - non-productive infection;
 - latent infection;
 - abortive infection;
 - defective virus;
- discuss the spread of virus infections within animal bodies and within plants;
- discuss the factors that determine whether virus infection results in disease.

SOURCES OF FURTHER INFORMATION

Books (general immunology)

Delves, P. J. et al. (2011) Roitt's™s Essential Immunology, 12th edition, Wiley-Blackwell

Hannigan, B. M., Moore, C. B. T., and Quinn, D. G. (2009) Immunology, 2nd edition, Scion

Peakman, M. and Vergani, D. (2009) Basic and Clinical Immunology, 2nd edition, Churchill Livingstone

Books (immune responses to virus infections)

Digard, P., Nash, A. A., and Randall, R. E. (eds) (2005) Molecular Pathogenesis of Virus Infections, 64th Symposium of the Society for General Microbiology, Cambridge University Press

Kaufmann, S. H. E., Rouse, B. T., and Sacks, D. L. (2011) The Immune Response to Infection, American Society for Microbiology

Martinez, M. A., (ed.) (2010) RNA Interference and Viruses: Current Innovations and Future Trends, Caister Academic Press

Mossman, K. (ed.) (2011) *Viruses and Interferon: Current Research*, Caister Academic Press

Historical paper

Isaacs, A. and Lindenmann, J. (1957) Virus interference. 1. The interferon. *Proceedings of the Royal Society of London B*, 147, 258-267

Recent papers

Aoshi, T. et al. (2011) Innate and adaptive immune responses to viral infection and vaccination. *Current Opinion in Virology*, 1, 196-203

Arpaia, N. and Barton, G. M. (2011) Toll-like receptors: key players in antiviral immunity. *Current Opinion in Virology*, 1, 447-454

Douglas, J. L. et al. (2010) The great escape: viral strategies to counter BST-2/tetherin. *PLoS Pathogens*, 6 (5), e1000913

Grundhoff, A. and Sullivan, C. S. (2011) Virus-encoded microRNAs. *Virology*, 411, 325-343

Kane, M. and Golovkina, T. (2010) Common threads in persistent viral infections. *Journal of Virology*, 84, 4116-4123

Marriott, A. C. and Dimmock, N. J. (2010) Defective interfering viruses and their potential as antiviral agents. *Reviews in Medical Virology*, 20, 51-62

Schoggins, J. W. et al. (2011) A diverse range of gene products are effectors of the type I interferon antiviral response. *Nature*, 472, 481-485

Stoermer, K. A. and Morrison, T. E. (2011) Complement and viral pathogenesis. *Virology*, 411, 362-373

Vivier, E. et al. (2011) Innate or adaptive immunity? The example of natural killer cells. *Science*, 331, 44-49

Whitmire, J. K. (2011) Induction and function of virus-specific CD4+ T cell responses. *Virology*, 411, 216B–228

CHAPTER 10

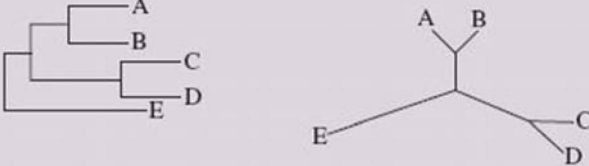
Classification and Nomenclature of Viruses

CHAPTER 10 AT A GLANCE

Authority: International Committee on Taxonomy of Viruses
<http://www.ictvdb.org/index.htm>

Taxonomic Groups: Order
Family
Subfamily
Genus
Species

Phylogenetic trees: – based on genome sequences
– indicate relationships between viruses:



Baltimore Classification: – seven classes of viruses
– based on genome type and transcription.

10.1 HISTORY OF VIRUS CLASSIFICATION AND NOMENCLATURE

Virologists are no different to other scientists in that they find it useful to classify the objects of their study into groups and subgroups. In the early days, when little was known about viruses, they were loosely grouped on the basis of criteria such as the type of host, the type of disease caused by infection, and whether the virus is transmitted by an arthropod vector.

As more was learned about the characteristics of virus particles some of these began to be used for the purposes of classification, for example:

- whether the nucleic acid is DNA or RNA;
- whether the nucleic acid is single-stranded or double-stranded;

- whether or not the genome is segmented;
- the size of the virion;
- whether the capsid has helical symmetry or icosahedral symmetry;
- whether the virion is naked or enveloped.

Various combinations of these criteria produced some useful virus groups, but there was no single approach to the naming of the groups, and names were derived in a variety of ways:

- small, icosahedral, single-stranded DNA viruses of animals were called parvoviruses (Latin parvus = small);
- nematode-transmitted polyhedral (icosahedral) viruses of plants were called nepoviruses;
- phages T2, T4, and T6 were called T even phages.

Serological relationships between viruses were investigated, and distinct strains (serotypes) could be distinguished in tests using antisera against purified virions. Serotypes reflect differences in virus proteins and have been found for many types of virus, including rotaviruses and foot and mouth disease virus.

10.1.1 International Committee on Taxonomy of Viruses

By 1966 it was decided that some order had to be brought to the business of naming viruses and classifying them into groups, and the International Committee on Taxonomy of Viruses (ICTV) was formed. The committee now has many working groups and is advised by virologists around the world.

The ICTV lays down the rules for the nomenclature and classification of viruses, and it considers proposals for new taxonomic groups and virus names. Those that are approved are published in book form (please see Sources of further information at the end of this chapter) and on the web; these sources should be consulted for definitive information. The web site for this book (www.wiley.com/go/carter) has links to relevant web sites.

10.2 MODERN VIRUS CLASSIFICATION AND NOMENCLATURE

For a long time virologists were reluctant to use the taxonomic groups such as family, subfamily, genus, and species that have long been used to classify living organisms, but taxonomic groups of viruses have gradually been accepted and are now established (Table 10.1). Some virus families have been grouped into orders, but higher taxonomic groupings, such as class and phylum, are not used. Only some virus families are divided into subfamilies.

Table 10.1 Taxonomic groups of viruses

Taxonomic group	Suffix	Example 1	Example 2	Example 3
Order	<i>-virales</i>	<i>Caudovirales</i>	<i>Mononegavirales</i>	<i>Nidovirales</i>
Family	<i>-viridae</i>	<i>Myoviridae</i>	<i>Paramyxoviridae</i>	<i>Coronaviridae</i>
Subfamily	<i>-virinae</i>	–	<i>Paramyxovirinae</i>	<i>Coronavirinae</i>
Genus	<i>-virus</i>	<i>T4-like viruses</i>	<i>Morbillivirus</i>	<i>Betacoronavirus</i>
Species	–	Enterobacteria phage T4	Measles virus	Severe acute respiratory syndrome-related virus

Each order, family, subfamily, and genus is defined by viral characteristics that are necessary for membership of that group, whereas members of a species have characteristics in common but no one characteristic is essential for membership of the species. Many species contain variants known as virus strains, serotypes (differences are detected by differences in antigens), or genotypes (differences are detected by differences in genome sequence).

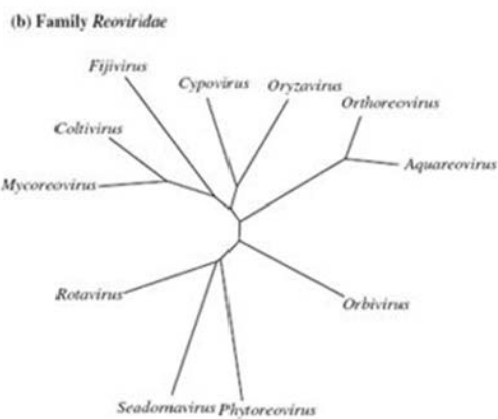
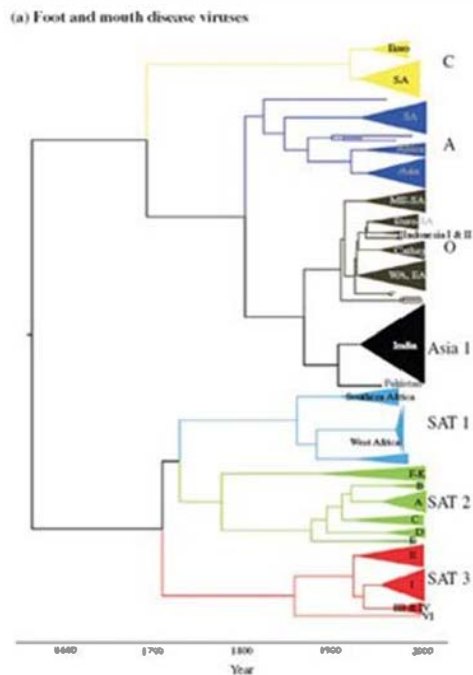
Many of the early names of virus groups were used to form the names of families and genera, e.g. the picornaviruses became the family Picornaviridae. Each taxonomic group has its own suffix and the formal names are printed in italic with the first letter in upper case (Table 10.1), e.g. the genus Morbillivirus. When common names are used, however, they are not in italic and the first letter is in lower case (unless it is the first word of a sentence), e.g. the morbilliviruses.

10.2.1 Classification based on genome sequences

Now that technologies for sequencing virus genomes and for determining genome organization are readily available, the modern approach to virus classification is based on comparisons of genome sequences and organizations. The degree of similarity between virus genomes can be assessed using computer programs, and can be represented in diagrams known as phylogenetic trees because they show the likely phylogeny (evolutionary development) of the viruses. Phylogenetic trees may be of various types (Figure 10.1):

Figure 10.1 Phylogenetic trees. (a) Rooted tree showing relationships between foot and mouth disease virus serotypes based on VP1 sequences. The serotypes evolved in different regions of the world. (b) Unrooted tree showing relationships between genera in the family Reoviridae based on VP1 sequences.

Source: (a) Tully and Fares (2008) *Virology*, 382, 250. Reproduced by permission of Elsevier.



- rooted phylogenetic tree begins at a root which is assumed to be the ancestor of the viruses in the tree:
- unrooted phylogenetic tree no assumption is made about the ancestor of the viruses in the tree.

The branches of a phylogenetic tree indicate how sequences are related. The branches may be scaled or unscaled; if they are scaled, their lengths represent genetic distances between sequences.

For many groups of viruses, analysis of the sequence and organization of their genomes has supported earlier classifications, e.g. the genera of the family Reoviridae (Figure 10.1(b)). Another example is the rhabdoviruses (Chapter 15), which were originally grouped together because of their bullet-shaped morphology, but which have turned out also to be related genetically.

10.2.2 Nomenclature of viruses and taxonomic groups

The naming of individual viruses has been a rather haphazard business, with somewhat different approaches taken for viruses of different host types. Bacterial viruses were simply allotted codes, such as T1, T2, and Φ X174. Viruses of humans and other vertebrates were commonly named after the diseases that they cause,

e.g. measles virus, smallpox virus, foot and mouth disease virus, though some were named after the city, town or river where the disease was first reported, e.g. Newcastle disease virus, Norwalk virus, Ebola virus. Some of these original names have been adopted as the formal names of the viruses.

Some of the place names where viruses were first found have become incorporated into the names of virus families and genera (Table 10.2).

Table 10.2 Names of virus families and genera derived from place names

Place name	Family/genus name
Bunyamwera (Uganda)	Family Bunyaviridae
Ebola (river in Zaire)	Genus Ebolavirus
Hantaan (river in South Korea)	Genus Hantavirus
Hendra (Australia) and Nipah (Malaysia)	Genus Henipavirus
Norwalk (United States)	Genus Norovirus

Many insect viruses were named after the insect, with an indication of the effect of infection on the host. A virus was isolated from *Tipula paludosa* larvae that were iridescent as a result of the large quantities of virions in their tissues (see photograph in Chapter 1, At a glance). The virus was named *Tipula iridescent virus*, and later renamed as invertebrate iridescent virus 1. Most plant viruses were given names with two components: the host and signs of disease, e.g. potato yellow dwarf virus, tobacco rattle virus. Some of these names have been used as the bases for family and genus names (Table 10.3).

Table 10.3 Names of families, subfamilies, and genera of plant viruses based on the host and signs of disease

Host and disease signs	Family/genus name
Brome mosaic	Family Bromoviridae
Cauliflower mosaic	Family Caulimoviridae
Cowpea mosaic	Subfamily Comovirinae
Tobacco mosaic	Genus Tobamovirus
Tobacco rattle	Genus Tobravirus
Tomato bushy stunt	Family Tombusviridae

As in other areas of biology, many names of virus taxonomic groups are based on Latin words, while some have Greek origins; a sample is given in Table 10.4.

Table 10.4 Names of virus families and genera based on Latin and Greek words. Note that there are two Latin words meaning yellow. One was used to

name the flaviviruses (animal viruses) and the other was used to name the luteoviruses (plant viruses).

		Translation	Reason for name	Family/genus name
Latin	<i>Arena</i>	Sand	Ribosomes in virions resemble sand grains in thin section	Family <i>Arenaviridae</i>
	<i>Baculum</i>	Stick	Capsid shape	Family <i>Baculoviridae</i>
	<i>Filum</i>	Thread	Virion shape	Family <i>Filoviridae</i>
	<i>Flavus</i>	Yellow	Yellow fever virus	Family <i>Flaviviridae</i>
	<i>Luteus</i>	Yellow	Barley yellow dwarf virus	Family <i>Luteoviridae</i>
	<i>Parvus</i>	Small	Virion size	Family <i>Parvoviridae</i>
	<i>Tenuis</i>	Thin, fine	Virion shape	Genus <i>Tenuivirus</i>
	<i>Toga</i>	Cloak	Virion is enveloped	Family <i>Togaviridae</i>
Greek	<i>Kloster</i>	Thread	Virion shape	Family <i>Closteroviridae</i>
	<i>Kystis</i>	Bladder, sack	Virion is enveloped	Family <i>Cystoviridae</i>
	<i>Mikros</i>	Small	Virion size	Family <i>Microviridae</i>
	<i>Podos</i>	Foot	Phages with short tails	Family <i>Podoviridae</i>

The student of virology thus gains some grounding in the classical languages! We can note that both Latin and Greek translations of *νῆμα* thread *νῆμα* have been used to name the filoviruses and the closteroviruses, respectively. Both of these families have thread-shaped virions. Similarly, Latin and Greek translations of *μικρὸν* small *μικρὸν* have been used to name the parvoviruses (animal viruses) and the microviruses (phages). The word for *μικρὸν* small *μικρὸν* from a third language was used when devising a name for small RNA viruses; the Spanish *pequeño* was linked to *ARN* RNA to form *pequeño* RNA viruses.

10.3 BALTIMORE CLASSIFICATION OF VIRUSES

We have seen how viruses can be grouped into seven classes on the basis of the type of genome and the way in which the genome is transcribed and replicated (Sections 6.2 and 7.1). This approach to virus classification was first suggested by David Baltimore, after whom the scheme is named. An advantage of the Baltimore classification is its differentiation between plus-strand RNA viruses that do (class VI) and do not (class IV) carry out reverse transcription, and between dsDNA viruses that do (class VII) and do not (class I) carry out reverse transcription.

In the chapters that follow we shall examine in depth a representative family of viruses from each of the seven Baltimore classes.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- evaluate the traditional criteria used to classify viruses into families and genera;

- write family and genus names in the correct format;
- explain how genome sequence data are used to classify viruses;
- evaluate phylogenetic trees;
- explain the basis of the Baltimore classification of viruses.

SOURCES OF FURTHER INFORMATION

Books

King, A. M. Q. et al. (eds) (2012) *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*, Academic Press

Pringle, C. R. (2005) The classification of vertebrate viruses. Chapter 4 in Topley & Wilson's *Microbiology & Microbial Infections, Virology*, 10th edition, Volume 1 (Mahy, B. W. J. and ter Meulen, V., editors), Hodder Arnold

van Regenmortel, M. H. V. (2005) The nature and classification of viruses. Chapter 3 in Topley & Wilson's *Microbiology & Microbial Infections, Virology*, 10th edition, Volume 1 (Mahy, B. W. J. and ter Meulen, V., editors), Hodder Arnold

Journal

van Regenmortel, M. H. V. and Mahy, B. W. J. (2004) Emerging issues in virus taxonomy. *Emerging Infectious Diseases*, 10 (1): <http://www.cdc.gov/ncidod/eid/vol10no1/03-0279.htm>

Web sites

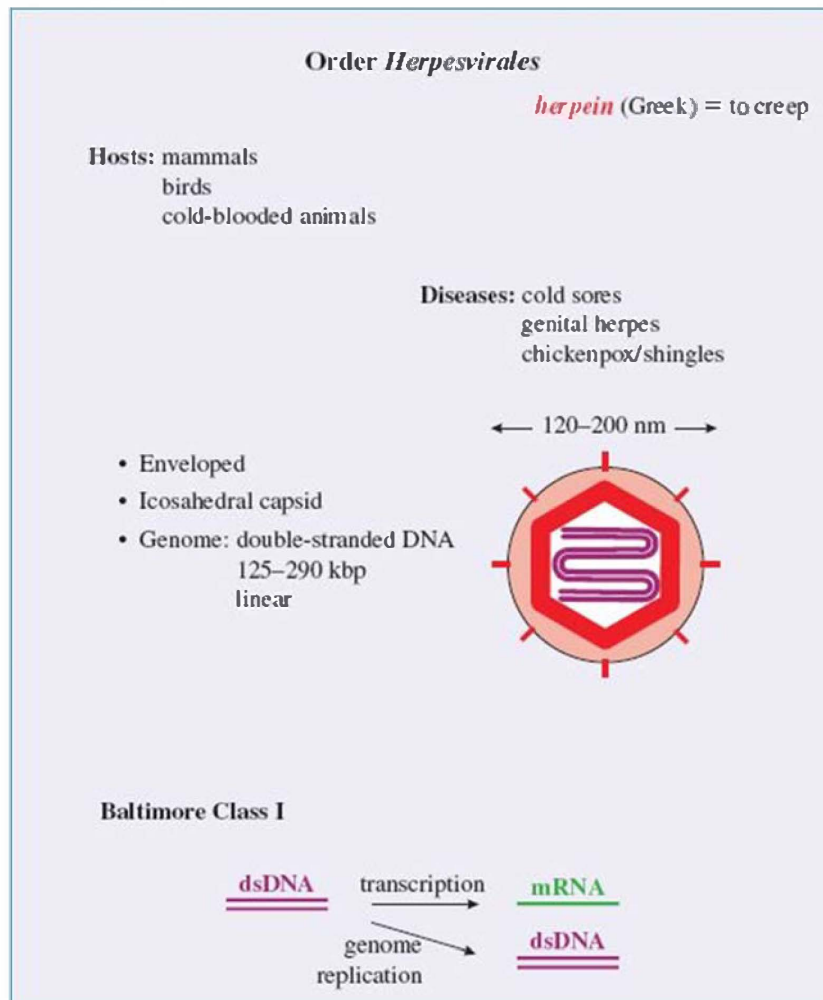
International Committee on Taxonomy of Viruses, <http://www.ncbi.nlm.nih.gov/ICTVdb/index.htm>

US National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/genomes/VIRUSES/viruses.html>

CHAPTER 11

Herpesviruses (and Other dsDNA Viruses)

CHAPTER 11 AT A GLANCE



11.1 INTRODUCTION TO HERPESVIRUSES

The word herpesvirus is derived from the Greek herpein, meaning to creep. More than 100 herpesviruses have been isolated and classified in the order Herpesvirales. Those found in mammals, birds, and reptiles constitute the family Herpesviridae, while other herpesviruses have been found in fish, amphibians, and a mollusk.

Herpesviruses are class I viruses, with genomes of dsDNA. A notable characteristic of herpesviruses is that, once they have infected a host, they commonly remain as persistent infections for the lifetime of the host. These infections are often latent infections, which can be reactivated from time to time, especially if the host becomes immunocompromised. Both primary and reactivated herpesvirus infections can either be asymptomatic or can result in disease of varying severity. The

outcome depends on the interplay between the particular virus and its host, and especially on the immune status of the host.

11.2 THE HUMAN HERPESVIRUSES

There are eight herpesviruses known in man (Table 11.1), and most adults are persistently infected with most of them.

Table 11.1 Human herpesviruses

Herpesvirus	Disease example(s)
Herpes simplex virus 1	Cold sores
Herpes simplex virus 2	Genital herpes
Varicella-zoster virus	Chickenpox, shingles
Epstein-Barr virus	Glandular fever, tumors
Human cytomegalovirus	Congenital defects
Human herpesvirus 6	Exanthem subitum
Human herpesvirus 7	Exanthem subitum
Kaposi's sarcoma-associated herpesvirus	Kaposi's sarcoma

11.2.1 Herpes simplex viruses 1 and 2

Herpes simplex viruses 1 and 2 (HSV-1 and HSV-2) initially infect epithelial cells of the oral or genital mucosa, the skin or the cornea. The virus may spread to neurons in which it may be transported to their nuclei, where it may establish latent infections.

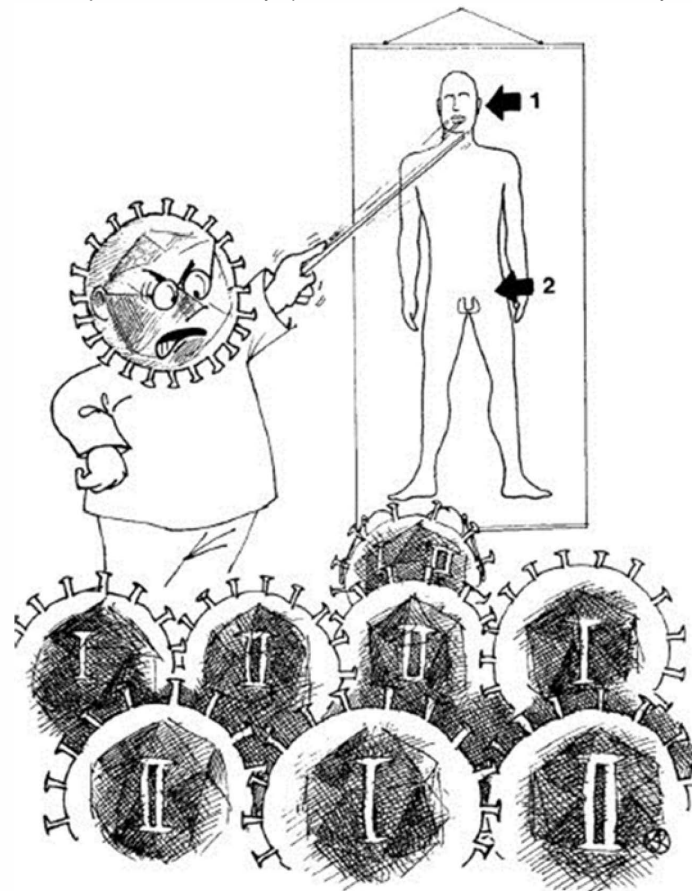
HSV-1 commonly infects via the lips or the nose between the ages of 6 and 18 months. A latent infection may be reactivated if, for example, the host becomes stressed or immunosuppressed. Reactivation results in the production of virions, which in about 20-40% of cases are transported within the neuron to the initial site of infection, where they cause productive infection in epithelial cells, resulting in a cold sore. Occasionally there may be serious complications such as encephalitis, especially in immunocompromised hosts.

HSV-2 is the usual causative agent of genital herpes, which is a sexually transmitted disease. In newborn babies infection can result in serious disease, with a mortality rate of about 54%.

Although the face and the genitals are the normal sites of infection for HSV-1 and HSV-2, respectively, there are increasing numbers of cases where HSV-1 infects the genitals and HSV-2 infects the face (Figure 11.1).

Figure 11.1 Get your priorities right.

Source: Haaheim, Pattison, and Whiteley (2002) A Practical Guide to Clinical Virology, 2nd edition. Reproduced by permission of John Wiley & Sons.



11.2.2 Varicella-zoster virus

Infection with varicella-zoster virus usually occurs in childhood and causes varicella (chickenpox), when the virus spreads through the blood to the skin, causing a rash. It may also spread to nerve cells, where it may establish a latent infection. The nerves most often affected are those in the face or the trunk, and these are the areas most commonly affected in zoster (shingles) when a latent infection is reactivated.

11.2.3 Epstein-Barr virus

Epstein-Barr virus (EBV) is transmitted in saliva. Epithelial cells are infected first then the infection spreads to B cells, which are the main host cell type for this virus. More than 90% of people become infected with EBV, usually during the first years of life, when infection results in few or no symptoms. In developed countries some individuals do not become infected until adolescence or adulthood. A proportion of these individuals develop infectious mononucleosis (glandular fever), commonly called "the kissing disease" by doctors. EBV is associated with a number of tumors in humans (Chapter 23).

11.2.4 Human cytomegalovirus

In the vast majority of infections with human cytomegalovirus, symptoms are either absent or they are mild. In a pregnant woman, however, the virus can infect the placenta and then the fetus, for whom the consequences may be serious. In one study it was estimated that 0.7% of children are born infected with the virus. In some of these there is evidence of virus-induced damage, including small brain size and enlargement of the liver and spleen; 0.5% of these congenitally infected babies die. Up to 20% of the survivors develop problems, include hearing loss and mental retardation.

Human cytomegalovirus can also cause severe disease (e.g. pneumonitis, hepatitis) in immunocompromised patients such as those with AIDS, those who have received treatment for cancer, and those who are immunosuppressed because they have received an organ transplant.

11.2.5 Human herpesvirus 6

There are two types of human herpesvirus 6, known as HHV-6A and HHV-6B. Infection of a child with the latter can cause a fever and the sudden appearance of a rash known as exanthem subitum.

11.2.6 Human herpesvirus 7

Human herpesvirus 7 was first isolated from a culture of CD4 T cells that developed a cytopathic effect; the cells were from a healthy person. The virus has been associated with some cases of exanthem subitum.

11.2.7 Kaposi's sarcoma-associated herpesvirus

Kaposi's sarcoma-associated herpesvirus was discovered in 1994 and is named after the tumor with which the virus is associated (Chapter 23).

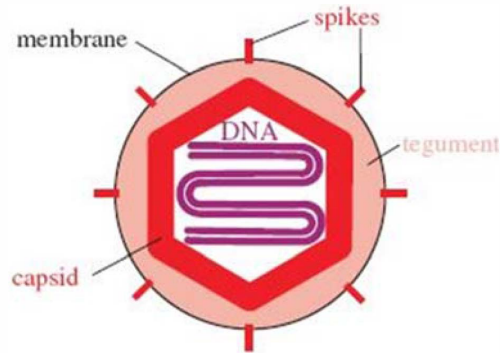
11.3 THE HERPESVIRUS VIRION

Herpesviruses have relatively complex virions composed of a large number of protein species organized into three distinct structures: capsid, tegument, and envelope (Figure 11.2). The virus genome is a linear dsDNA molecule, which is about 125 kbp in the smallest herpesviruses and about 290 kbp in the largest herpesviruses. The DNA is housed in the capsid, which is icosahedral, and the capsid is surrounded by the tegument. Many of the envelope glycoprotein molecules are organized into spikes of various dimensions.

Figure 11.2 The herpesvirus virion.

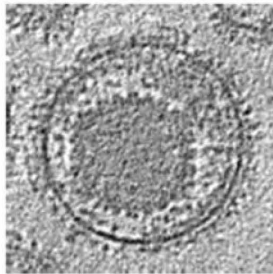
Source: (b) and (c) are images of HSV-1, from Grifnewald and Cyrklaff (2006) *Current Opinion in Microbiology*, 9, 437. Reproduced by permission of the authors and Elsevier Limited.

(a) Virion components

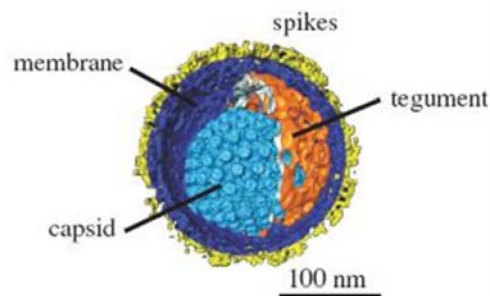


Electron cryo-tomographic visualizations

(b) Central slice through reconstructed volume



(c) Cut-away view

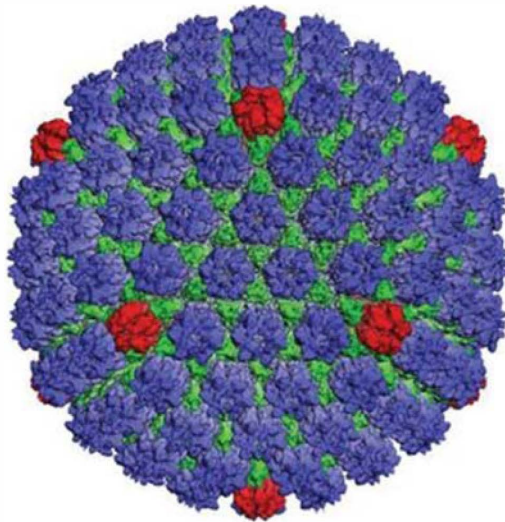


In the HSV-1 virion the capsid sits asymmetrically within the tegument (Figure 11.2(b) and (c)). The HSV-1 tegument contains at least 26 species of virus protein, some cell proteins, and some virus mRNA molecules. The envelope contains at least 16 protein species, most of which are glycoproteins forming 600B~750 spikes.

The capsid is constructed from 162 capsomeres, 150 of which are hexons (Figure 11.3). At one of the vertices of the icosahedron there is a structure called the portal. There is evidence from electron microscopy of HSV-1 capsids treated with antibody specific for pUL6 that this protein forms the portal. The virus DNA passes through the portal when entering the procapsid during packaging, and it is likely that it also passes through the portal when leaving the capsid during the infection process. There is a penton at each of the other eleven vertices of the icosahedron.

Figure 11.3 The HSV-1 capsid. Reconstructed image from cryo-electron microscopy.

Source: Courtesy of Professor Wah Chiu, Baylor College of Medicine, Houston, TX. Reinterpretation of data from Zhou et al. (2000) *Science*, 288, 877, with permission of the American Association for the Advancement of Science.



A number of schemes have evolved for the nomenclature of herpesvirus proteins, with the result that an individual protein may have two or more different names in the literature. Most of the structural proteins are commonly named VP (virus protein). In HSV-1 the most abundant protein in the tegument is VP16, while the pentons and hexons are composed of VP5; a penton is made from five molecules and a hexon from six molecules of VP5. Other proteins make up structures called triplexes, which connect the capsomeres. The envelope glycoproteins are each prefixed with g, for example, gB, gC, and gD.

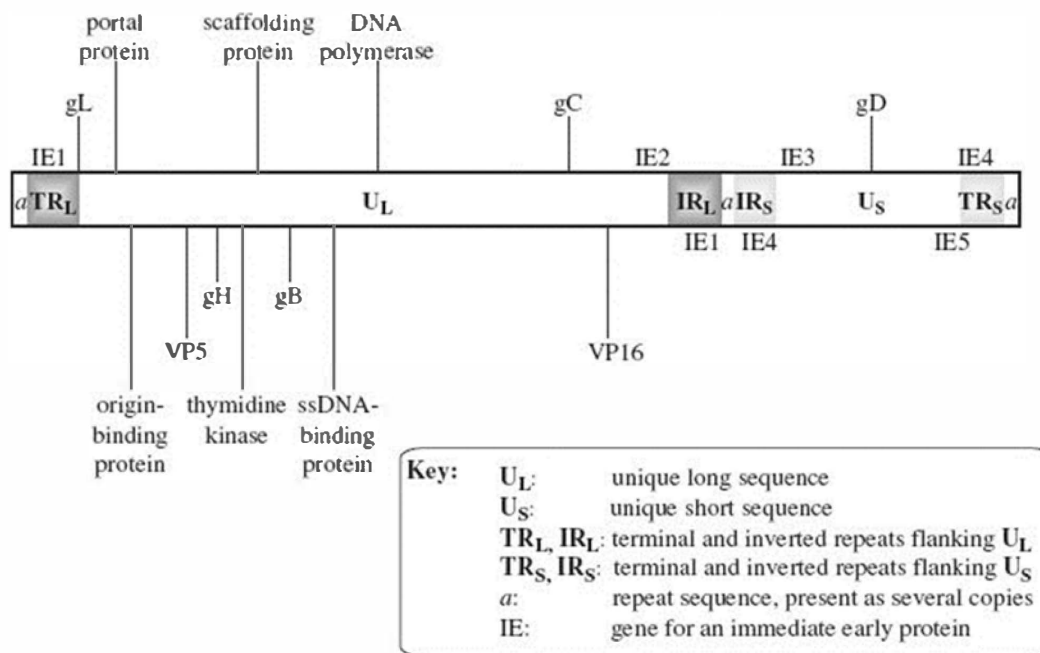
11.4 HSV-1 GENOME ORGANIZATION

HSV-1 is one of the most studied herpesviruses; we shall look at its genome organization and then at its replication. The genome consists of two unique sequences each flanked by repeat sequences (Figure 11.4). The unique sequences are not of equal length: the longer is designated UL and the shorter is designated US. A sequence known as the a sequence is present at each end of the genome, and at the junction of the inverted repeat (IR) sequences. The a sequence plays an important role during replication when genomes are packaged into procapsids.

Figure 11.4 The HSV-1 genome.

The locations of some genes are indicated. Those shown above the genome are in ORFs read left to right; those shown below the genome are in ORFs read right to left.

The locations of the genes for the five immediate early (IE) proteins are indicated. Two of these genes are in repeat sequences, so there are two copies of these genes.



The HSV-1 genome encodes at least 74 proteins and some RNAs that are not translated. The inverted repeats contain some genes, so the genome contains two copies of these genes. Both strands of the DNA are used for coding.

11.5 HSV-1 REPLICATION

Although HSV-1 infects only humans in nature, a variety of animal species and cell cultures can be infected in the laboratory. The replication cycle of the virus has been studied in cell cultures from a number of species, including humans, monkeys, mice, and dogs.

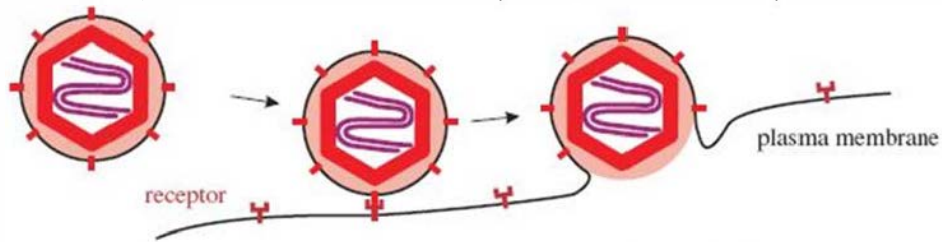
11.5.1 Attachment and entry

The sequence of events at the cell surface usually involves the HSV-1 virion binding initially to heparan sulfate, and then to the main receptor. The latter can be one of several types of molecule including some nectins, which are cell adhesion molecules. Depending on the cell type, either the virion envelope fuses with the plasma membrane (Figure 11.5) or infection occurs by endocytosis, followed by fusion between the envelope and the endosome membrane. At least five of the envelope glycoprotein species are involved in these processes (Table 11.2).

Table 11.2 HSV-1 glycoproteins involved in attachment to, and entry into, the host cell

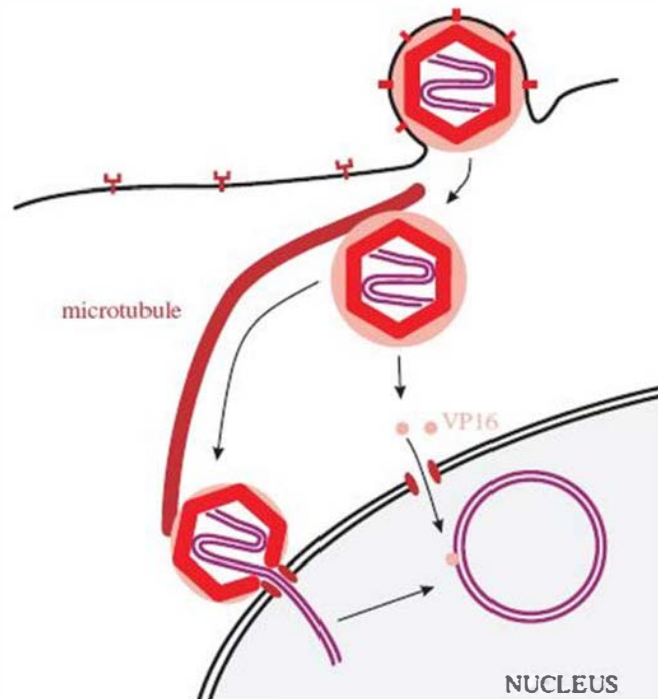
Process	HSV-1 glycoprotein(s) involved
Attachment to heparan sulfate	gB and/or gC
Attachment to main cell receptor	gD
Fusion of virion envelope and plasma membrane or endosome membrane	gB + gH/gL complex

Figure 11.5 HSV-1 attachment and entry into the cell. Virus glycoproteins bind to receptors on the cell, then the virion envelope fuses with the plasma membrane.



After delivery into the cytoplasm some tegument proteins are released. Other tegument proteins remain associated with the nucleocapsid, which must be transported to the nucleus where virus replication takes place. When the host cell is a neuron this journey is a long one; it has been estimated that it would take at least 200 years for the partially tegumented nucleocapsid to reach the nucleus by passive diffusion. In fact it is rapidly transported along microtubules to a nuclear pore (Figure 11.6). After docking at a nuclear pore, the capsid undergoes a conformational change that releases the DNA into the nucleus. The ends of the linear DNA are covalently joined, bringing the two end copies of the genome together and forming a circular molecule. The DNA becomes associated with cell histones.

Figure 11.6 Transport of HSV-1 DNA and VP16 into the nucleus. The nucleocapsid is transported along a microtubule to a site close to the nucleus. Docking of the nucleocapsid at a nuclear pore is followed by release of the genome into the nucleus. Molecules of VP16 released from the tegument are also transported into the nucleus.



Studies with antibodies specific for tegument proteins have provided evidence that these proteins are transported to several sites in the cell. At these sites the proteins play a variety of roles, including the activation of virus genes and the

down-regulation of host DNA, RNA, and protein synthesis. One tegument protein, virion host shutoff (vhs) protein, specifically degrades certain cell mRNAs, while another, VP16, is transported to the nucleus, where it becomes associated with the virus DNA (Figure 11.6) and is involved in the activation of virus genes.

11.5.2 Transcription and translation

Herpesvirus genes are expressed in three phases: immediate early (IE), early (E), and late (L) (Figure 11.7). Some authors refer to these phases by the Greek letters O_{\pm} , O_I , and O_L . There are introns in a few of the HSV-1 genes, mainly in the IE genes.

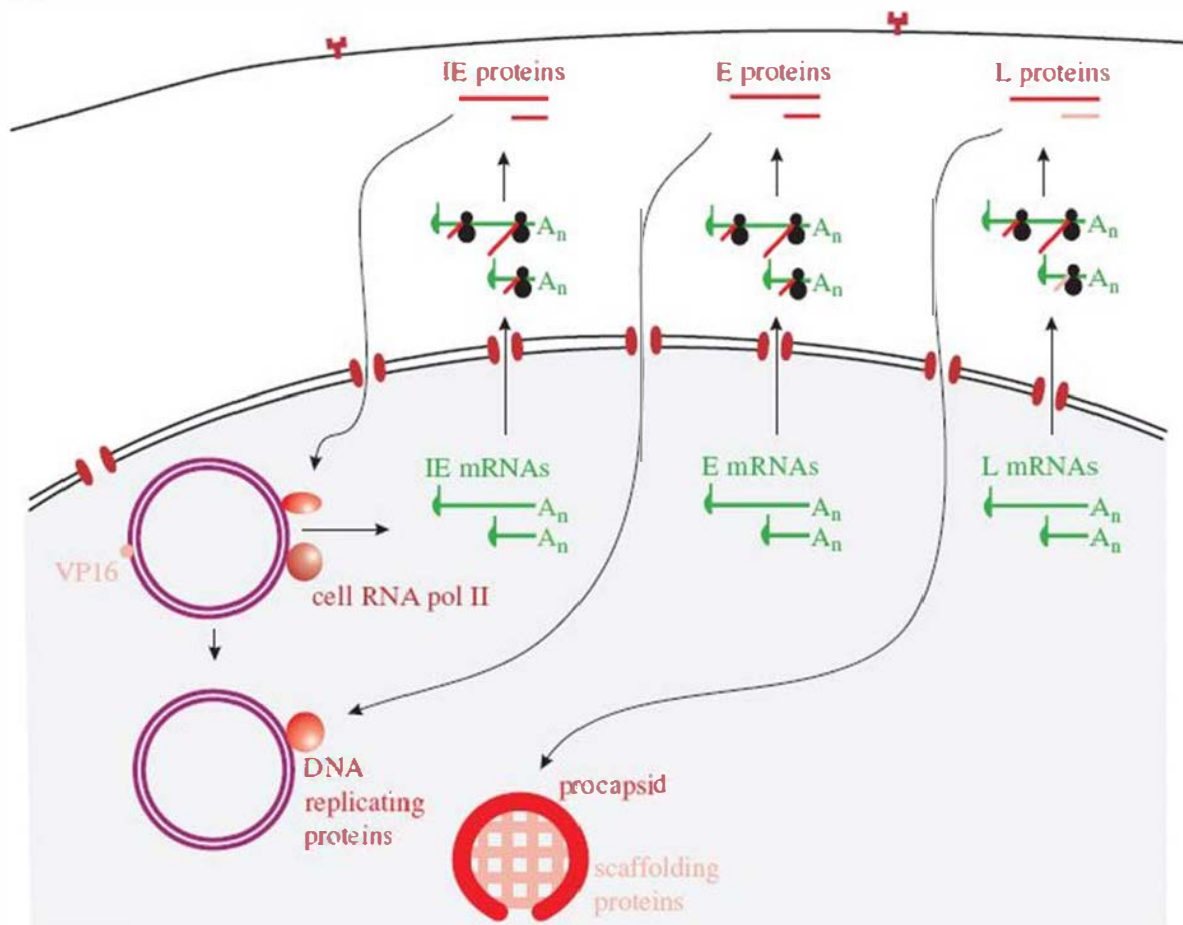
Figure 11.7 HSV-1 transcription and translation.

There are three phases of transcription and translation:

IE: immediate early

E: early

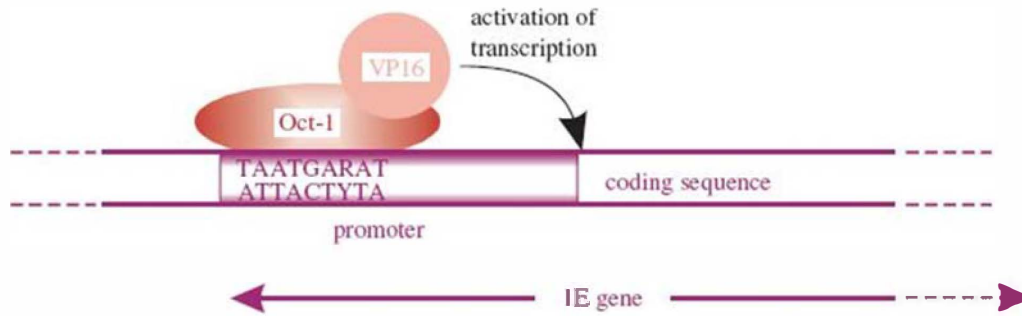
L: late.



The IE genes are activated by VP16. It was noted above that VP16 from infecting virions associates with the virus DNA. It does this by binding to a complex of cell proteins including Oct-1, which binds to the sequence TAATGARAT present in the promoter of each IE gene (Figure 11.8). VP16 then acts as a transcription factor by recruiting the host RNA polymerase II and associated initiation components to the gene.

Figure 11.8 Activation of transcription of HSV-1 IE genes by VP16.

R: purine
Y: pyrimidine



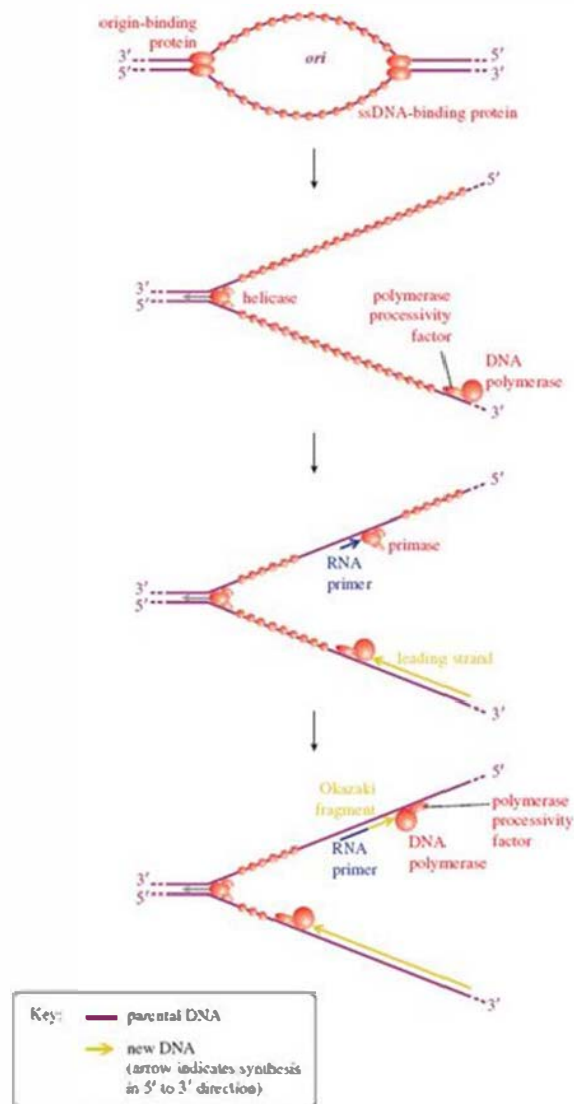
There are five IE proteins; all are transcription factors with roles in switching on E and L genes and in down-regulating the expression of some of these genes. At least some of the IE proteins have more than one role.

Some of the E proteins have roles in virus DNA replication (next section), which takes place in discrete regions of the nucleus known as replication compartments. Virus DNA and proteins accumulate in these compartments along with cell RNA polymerase II. Most of the L proteins are the virus structural proteins.

11.5.3 Genome replication

Replication of the virus DNA requires E proteins, seven of which are essential (Figure 11.9). Copies of an origin-binding protein bind at one of three ori sites in the virus DNA; the protein has helicase activity, causing the DNA to unwind at that site. A ssDNA-binding protein binds to the single strands, preventing the double helix from re-forming.

Figure 11.9 Roles of HSV-1 proteins in DNA replication. The roles of the seven virus proteins essential for DNA replication are shown. The helicase and the primase are complexes of the same three proteins. Please see text for details.

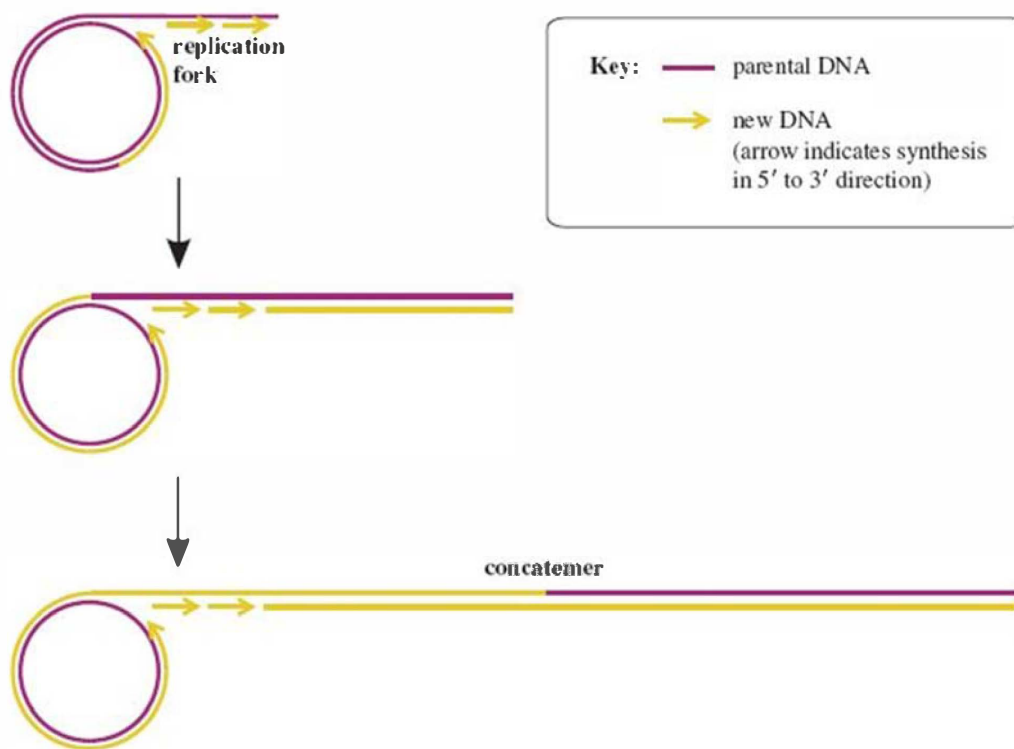


The ori site is then bound by a complex of three proteins which act as a helicase, further unwinding the double helix to form a replication fork. On one of the strands a complex of the same three proteins (now acting as a primase) synthesizes a short sequence of RNA complementary to the DNA. This RNA acts as a primer, and the DNA polymerase, complexed with a polymerase processivity factor, starts to synthesize the leading strand of DNA. On the other genome strand primases synthesize short RNAs, which are primers for the synthesis of the Okazaki fragments of the lagging strand of DNA.

In addition to the activities of the proteins just described, other virus proteins, such as a thymidine kinase, may also be involved in DNA replication.

It is thought that the circular DNA is first amplified by O_{\pm} replication, and that later the replication mode switches to O_{\pm} replication, also known as rolling circle replication (Section 7.5). The latter is the predominant mode of herpesvirus DNA replication, and the products are long DNA molecules known as concatemers (Figure 11.10), each of which consists of multiple copies of the virus genome.

Figure 11.10 Formation of a concatemer. See Figure 7.5 for earlier stages of rolling circle replication of DNA. Color coding indicates the fates of the two parental strands of DNA.

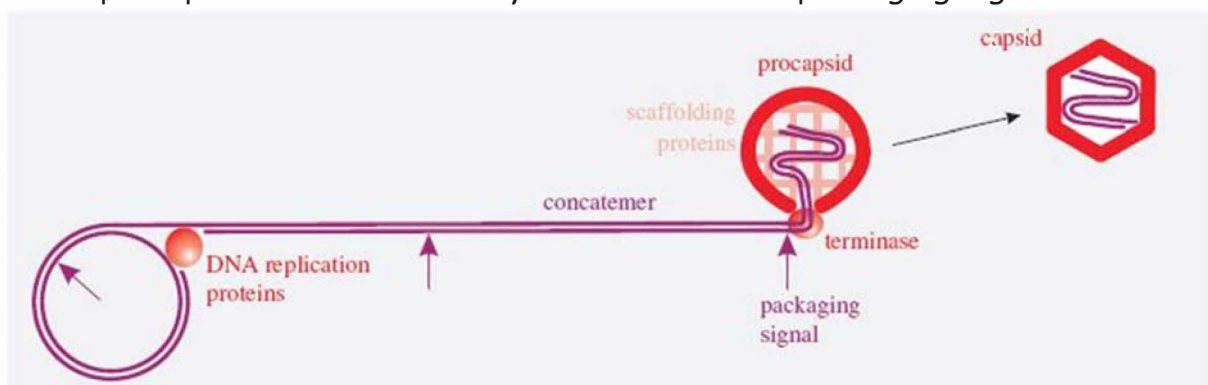


11.5.4 Assembly and exit of virions from the cell

The virus envelope glycoproteins are synthesized in the rough endoplasmic reticulum and are transported to the Golgi complex. Capsid proteins (such as VP5), scaffolding proteins, and some tegument proteins accumulate in the nuclear replication compartments, where procapsids are constructed. A procapsid is more rounded than a mature capsid and its structural integrity is maintained by the incorporation of scaffolding proteins. Before or during DNA packaging the scaffolding proteins are removed by a virus-encoded protease.

The DNA enters the procapsid via the portal. Each procapsid acquires a genome-length of DNA, which is cut from a concatemer (Figure 11.11). Cleavage of the concatemer occurs at a packaging signal which is within a sequence (Figure 11.4) and so is present at the junction of two copies of the genome. Pumping the DNA into a procapsid, recognition of the packaging signal, and cleavage of the DNA are carried out by a complex of virus proteins known as a terminase.

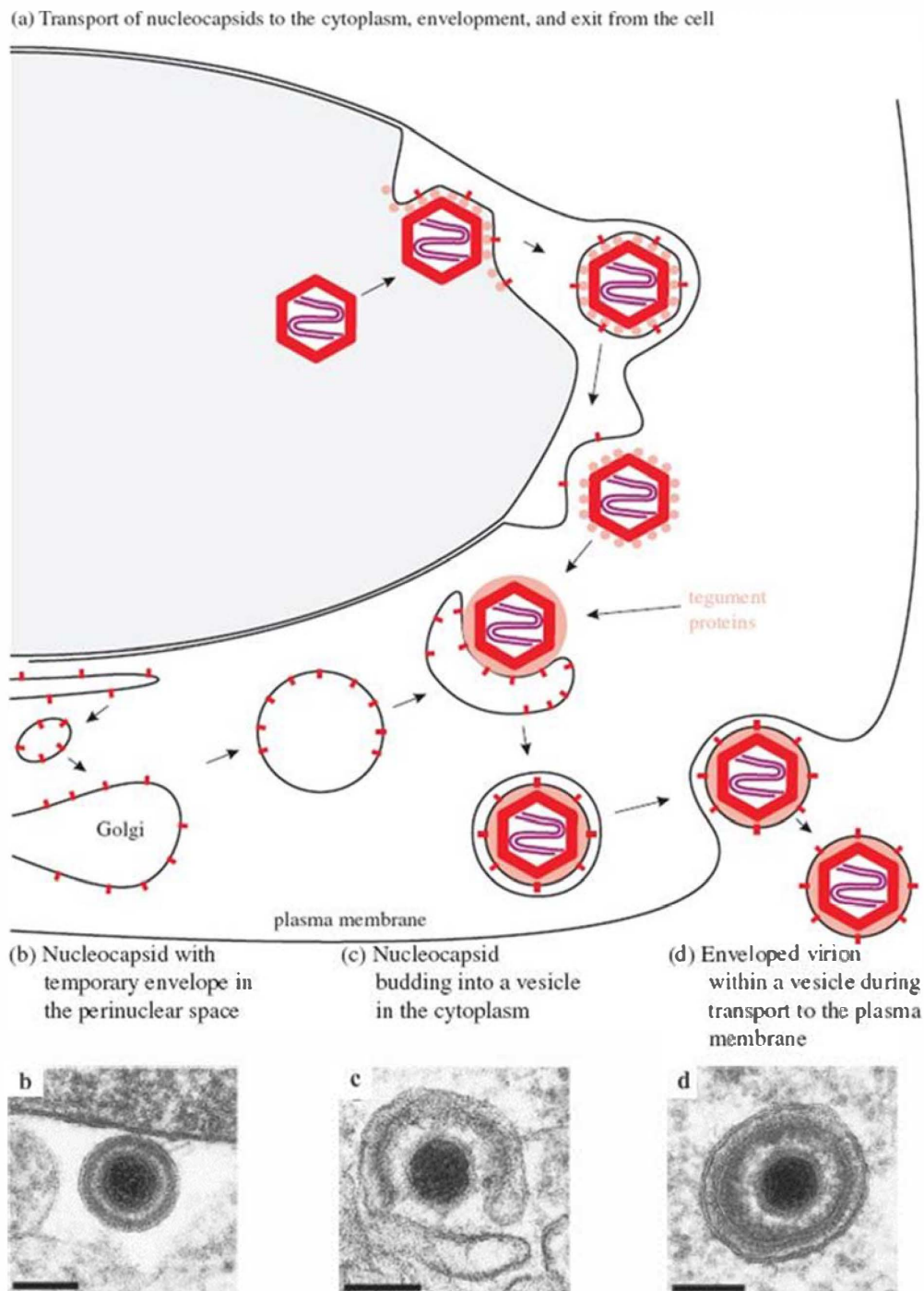
Figure 11.11 Packaging HSV-1 DNA into a procapsid. A genome-length of DNA enters a procapsid and is cleaved by a terminase at a packaging signal.



Once the nucleocapsid has been constructed it must acquire its tegument and envelope, and then the resulting virion must be released from the cell. These are complex processes and many of the details are still unclear, but there is evidence that the sequence of events is as follows (Figure 11.12).

Figure 11.12 Final stages of HSV-1 assembly.

Source: (b)–(d) from Mettenleiter, Klupp, and Granzow (2009) *Virus Research*, 143, 222. Scale bar: 100 nm. Reproduced by permission of Elsevier and the authors.



- The nucleocapsid acquires some tegument proteins in the nucleus.

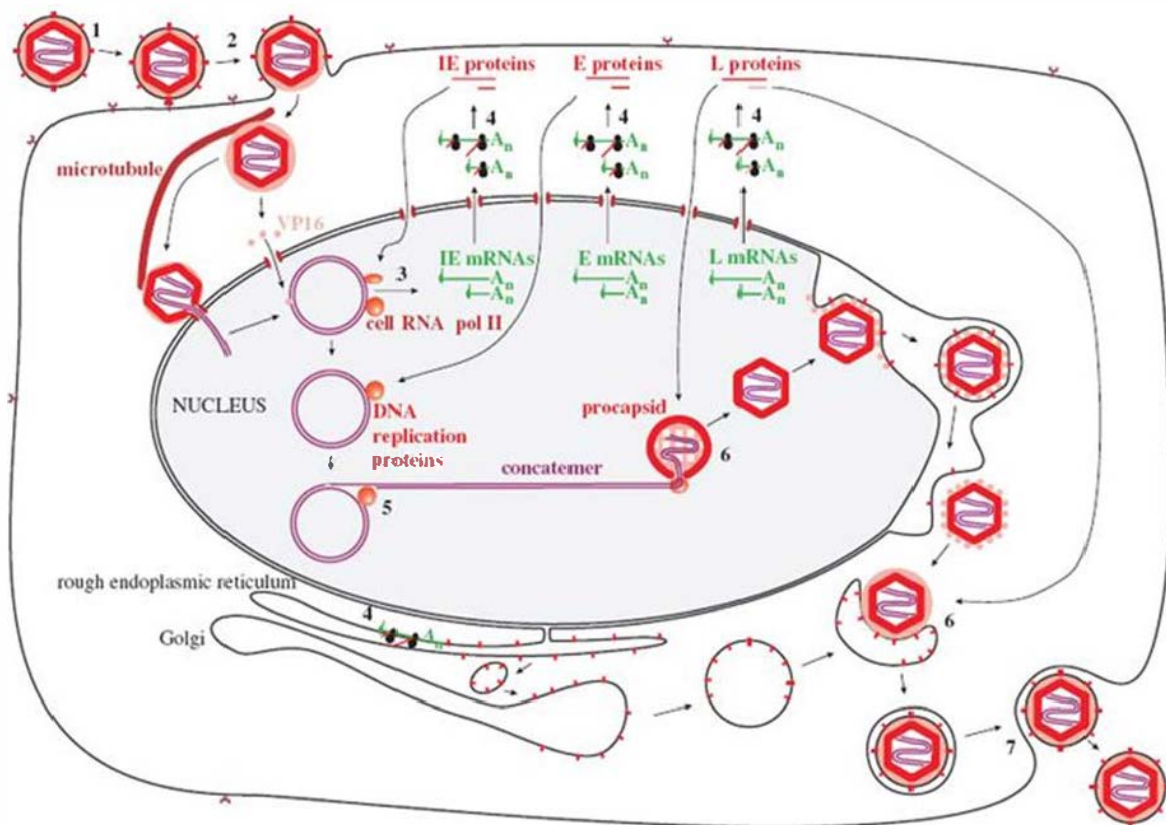
- The nucleocapsid buds through the inner membrane of the nuclear envelope into the perinuclear space, giving the nucleocapsid a temporary envelope (Figure 11.12(b)).
- The temporary envelope fuses with the outer membrane of the nuclear envelope, releasing the nucleocapsid into the cytoplasm.
- The remainder of the tegument proteins are added.
- The virion envelope is acquired by budding into a vesicle derived from the Golgi complex (Figure 11.12(c)).
- The enveloped virion is transported in the vesicle to the plasma membrane (Figure 11.12(d)).
- The vesicle membrane fuses with the plasma membrane, releasing the virion from the cell.

Although the virion envelopes are derived from membranes within the cell, virus glycoproteins are also expressed at the cell surface. These glycoproteins can fuse infected cells with non-infected cells, resulting in the formation of giant cells known as syncytia, which can be observed both in infected cell cultures and in HSV lesions. Infected cell cultures are normally destroyed 18–24 hours after inoculation.

11.5.5 Overview of HSV-1 replication

The HSV-1 replication cycle is summarized in Figure 11.13.

Figure 11.13 The HSV-1 replication cycle. Stages in virion assembly include procapsid construction, packaging a copy of the genome, and acquisition of the envelope by budding into a vesicle within the cytoplasm.



1. Attachment
2. Entry
3. Transcription
4. Translation
5. Genome replication
6. Assembly
7. Exit

11.6 LATENT HERPESVIRUS INFECTION

When infection of a cell with a herpesvirus results in latency, rather than a productive infection, the virus genome is maintained in the nucleus. For most herpesviruses the latently-infected nucleus harbors multiple copies of the genome as circular DNA molecules, but some, such as human herpesvirus 6, integrate the virus genome into a cell chromosome. Most of the virus genome is switched off during latency, but a few regions are transcribed into RNA, and some viruses also synthesize a few proteins.

No virus proteins are required to maintain latency in cells that do not divide, so none are produced in neurons latently infected with HSV-1, but virus RNAs are synthesized. Primary transcripts, known as latency-associated transcripts (LATs), are synthesized from the LAT gene, which is located in the terminal repeats of the genome (Figure 11.4). Processing of the primary transcripts gives rise to two shorter LATs and a series of microRNAs. These RNAs have a variety of roles that include inhibition of apoptosis, which ensures the survival of the neuron with its latent HSV-1 infection.

EBV, in contrast, becomes latent in memory B cells, which divide from time to time; the virus therefore synthesizes proteins needed to maintain the copy number of its genome when the host cell divides.

The likelihood of a latent herpesvirus infection becoming reactivated is increased if the host becomes immunocompromised; the greater the degree to which the host is immunocompromised, the greater the likelihood of reactivation.

11.7 OTHER dsDNA VIRUSES

Some further examples of viruses with dsDNA genomes in Baltimore class I are given in Table 11.3. Baculoviruses infect insects and other invertebrates. Some of them are used as insecticides and some are used as gene vectors in protein expression systems. Papillomaviruses are the causative agents of warts (papillomas) and some of them cause human cancers. The dsDNA phages are considered further in Section 20.5.

Table 11.3 Examples of dsDNA viruses

Animal viruses	Baculoviridae	Autographa californica nucleopolyhedrovirus
	Papillomaviridae	Human papillomaviruses (Section 23.2)
	Polyomaviridae	Simian virus 40 (Section 23.3)
	Poxviridae	Smallpox virus
Bacterial viruses	Myoviridae	Phage T4 (Section 20.5.1)

Most dsDNA viruses, like the herpesviruses, encode genes in both strands of the DNA, so the strands cannot be designated as plus and minus. There are some dsDNA viruses, however, such as the papillomaviruses and phage T7, that encode all of their genes in one strand of the DNA.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- list the human herpesviruses and explain their importance;
- describe the phenomenon of herpesvirus persistence and explain its importance;
- describe the HSV-1 virion;
- outline the main features of the HSV-1 genome;
- describe the replication cycle of HSV-1.

SOURCES OF FURTHER INFORMATION

Books

Krummenacher, C. et al. (2007) Entry of herpesviruses into cells: the enigma variations. In *Viral Entry into Host Cells* (Pohlmann, S. and Simmons, G., editors), Landes Bioscience

Pertel, P. E. and Spear, P. G. (2008) Biology of herpesviruses. Chapter 23 in *Sexually Transmitted Diseases*, 4th edition (Holmes, K. K. et al., editors), McGraw-Hill

Weller, S. K., (ed.) (2011) *Alphaherpesviruses: Molecular Virology*, Caister Academic Press

Whitley, R. J. and Roizman, B. (2009) Herpes simplex virus. Chapter 19 in *Clinical Virology*, 3rd edition (Richman, D. D., Whitley, R. J., and Hayden, F. G., editors), ASM Press

Journals

Brown, J. C. and Newcomb, W. W. (2011) Herpesvirus capsid assembly: insights from structural analysis. *Current Opinion in Virology*, 1, 142-149

Cullen, B. R. (2011) Herpesvirus microRNAs: phenotypes and functions. *Current Opinion in Virology*, 1, 211-215

Davison, A. J. et al. (2009) The order. Herpesvirales. *Archives of Virology*, 154, 171-177

Kelly, B. J. (2009) Functional roles of the tegument proteins of herpes simplex virus type 1. *Virus Research*, 145, 173-186

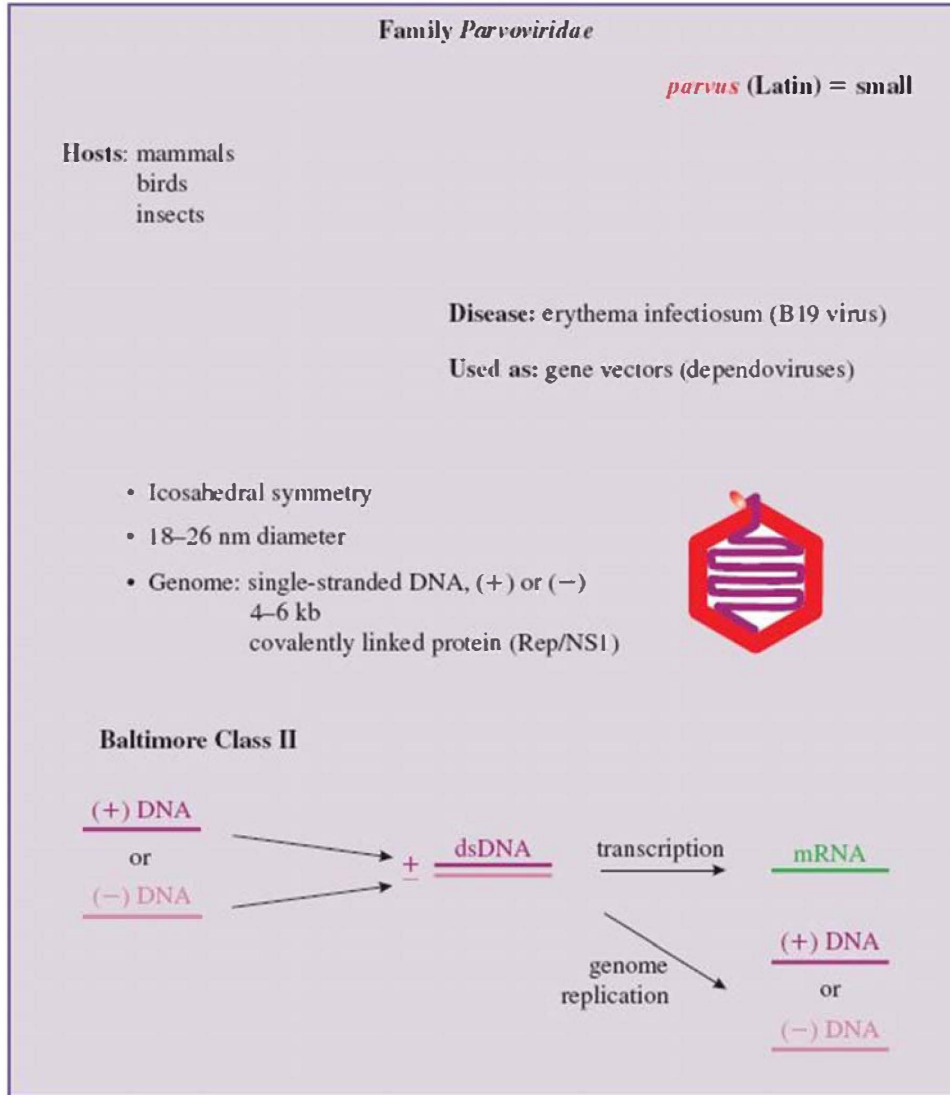
Lee, C.-P. and Chen, M.-R. (2010) Escape of herpesviruses from the nucleus. *Reviews in Medical Virology*, 20, 214-230

Lee, J. H. et al. (2008) Identification of structural protein-protein interactions of herpes simplex virus type 1. *Virology*, 378, 347-354

CHAPTER 12

Parvoviruses (and Other ssDNA Viruses)

CHAPTER 12 AT A GLANCE



12.1 INTRODUCTION TO PARVOVIRUSES

Parvoviruses are among the smallest known viruses, with virions in the range 18–26 nm in diameter. They derive their name from the Latin *parvus* (small). The family *Parvoviridae* has been divided into two subfamilies: the *Parvovirinae* (vertebrate viruses) and the *Densovirinae* (invertebrate viruses). Some of the genera and species of the subfamilies are shown in Table 12.1.

Table 12.1 Some of the viruses in the family *Parvoviridae*

Parvovirinae	Dependovirus	Adeno-associated virus 2
	Parvovirus	Minute virus of mice

		Feline panleukopenia virus
	Erythrovirus	B19 virus
	Bocavirus	Human bocavirus
Densovirinae	Iteravirus	Bombyx mori densovirus

The subfamily Parvovirinae includes the genus Dependovirus, the members of which are defective, normally replicating only when the host cell is co-infected with a helper virus. Other parvoviruses that do not require helper viruses are known as autonomous parvoviruses.

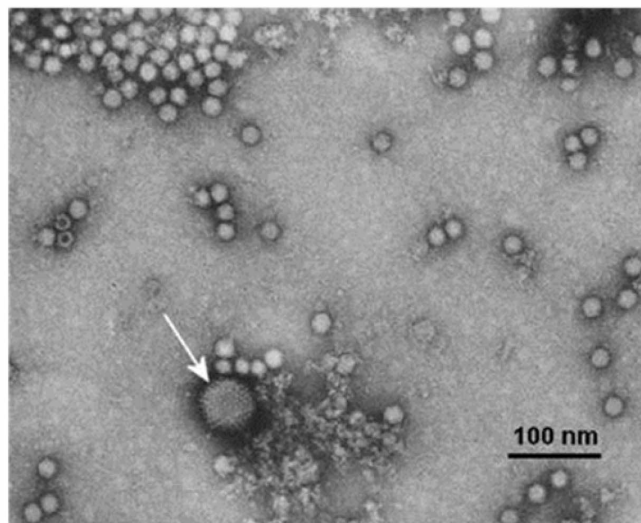
12.2 EXAMPLES OF PARVOVIRUSES

12.2.1 Dependoviruses

The first dependovirus to be discovered was observed in the electron microscope as a contaminant of an adenovirus preparation (Figure 12.1). The contaminant turned out to be a satellite virus (Section 1.3.4): a defective virus dependent on the help of the adenovirus for replication. The satellite virus was therefore called an adeno-associated virus. Other dependoviruses (distinct serotypes) have since been found in adenovirus preparations and in infected humans and other species. Results of surveys using serological methods and PCR detection of virus DNA indicate that dependovirus infections are widespread.

Figure 12.1 Virions of adenovirus (arrowed) and dependovirus.

Source: Reproduced with permission of Professor M. Stewart McNulty and The Agri-Food and Biosciences Institute.



Not all dependoviruses have an absolute requirement for the help of an adenovirus. Other DNA viruses (e.g. herpesviruses) may sometimes act as helpers, and some dependoviruses may replicate in the absence of a helper virus under certain circumstances.

Dependoviruses are valuable gene vectors. They are used to introduce genes into cell cultures for mass-production of the proteins encoded by those genes, and they are being investigated as possible vectors to introduce genes into the cells of patients for the treatment of various genetic diseases and cancers. One of the advantages of dependoviruses for such applications is the fact that they are not known to cause any disease, in contrast to other virus vectors, such as retroviruses (Section 17.5).

12.2.2 Autonomous parvoviruses

A parvovirus that does not require a helper virus was discovered in serum from a healthy blood donor. The virus, named after a batch of blood labeled B19, infects red blood cell precursors. Many infections with B19 are without signs or symptoms, but some result in disease, such as fifth disease (erythema infectiosum), in which affected children develop a “slapped-cheek” appearance (Figure 12.2).

Figure 12.2 Child with fifth disease.

Source: Reproduced by permission of the New Zealand Dermatological Society.



Other diseases caused by B19 virus include:

- acute arthritis;
- aplastic anemia in persons with chronic hemolytic anemia;
- hydrops fetalis (infection may be transmitted from a pregnant woman to the fetus and may kill the fetus).

In 2005 a new human parvovirus was discovered using a technique for molecular screening of nasopharyngeal aspirates from children with lower respiratory tract disease. The virus is related to known parvoviruses in the genus Bocavirus.

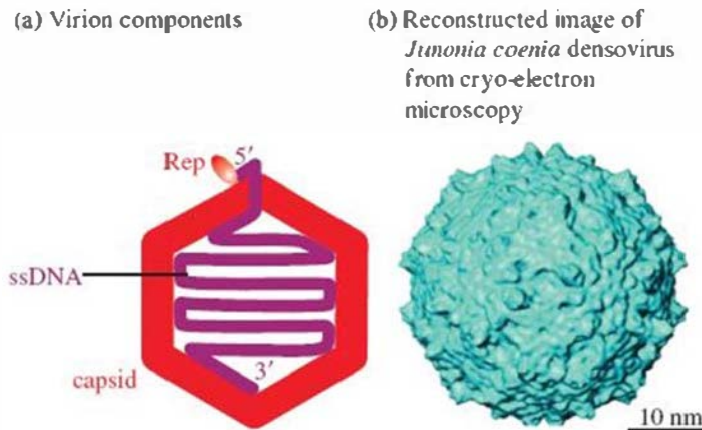
Viruses in the subfamily Densovirinae cause the formation of dense inclusions in the nucleus of the infected cell. Some of these viruses are pathogens of the silkworm (*Bombyx mori*), and can cause economic damage to the silk industry.

12.3 PARVOVIRUS VIRION

Parvoviruses are small viruses of simple structure with the ssDNA genome enclosed within a capsid that has icosahedral symmetry (Figure 12.3). Negatively stained virions of a parvovirus can be seen in Figure 12.1

Figure 12.3 The parvovirus virion.

Source: (b) Bruemmer et al. (2005) *Journal of Molecular Biology*, 347, 791. Reproduced by permission of Elsevier.



12.3.1 Capsid

The parvovirus capsid is built from 60 protein molecules. One protein species forms the majority of the capsid structure and there are small amounts of between one and three other protein species, depending on the virus. The proteins are numbered in order of size, with VP1 the largest; the smaller proteins are shorter versions of VP1. Each protein species contains an eight-stranded β -barrel structure that is common to many viral capsid proteins, including those of the picornaviruses (Chapter 14). Within VP1 there is a phospholipase domain that plays a role in penetrating the endosome membrane during entry into a cell.

The virion is roughly spherical, with surface protrusions and canyons (Figure 12.3(b)). At each of the vertices of the icosahedron there is a protrusion with a pore at the center.

12.3.2 Genome

Parvoviruses have genomes composed of linear ssDNA in the size range 4–6 kb. At the 5' end there is a covalently linked protein molecule; depending on the virus, this protein is known either as Rep (because it is involved in replication) or NS1 (previously defined as non-structural). This protein, together with a short sequence of the DNA to which it is linked, is present on the surface of the virion (Figure 12.3(a)).

At each end of the DNA there are a number of short complementary sequences that can base pair to form a secondary structure (Figure 12.4). Some parvovirus genomes have sequences at their ends known as inverted terminal repeats (ITRs), where the sequence at one end is complementary to, and in the opposite orientation to, the sequence at the other end (Section 3.2.6). As the sequences are complementary, the ends have identical secondary structures (Figure 12.4(a)). Other parvoviruses have a unique sequence, and therefore a unique secondary structure, at each end of the DNA (Figure 12.4(b)).

Figure 12.4 Base pairing at the termini of parvovirus DNA.

(a) "Inverted repeats" at the termini



(b) Unique sequences at the termini



During replication, parvoviruses with ITRs generate and package equal numbers of (+) and (-) strands of DNA, while most viruses with unique sequences at the termini do not. The percentages of virions containing (+) DNA and (-) DNA therefore vary with different viruses (Table 12.2).

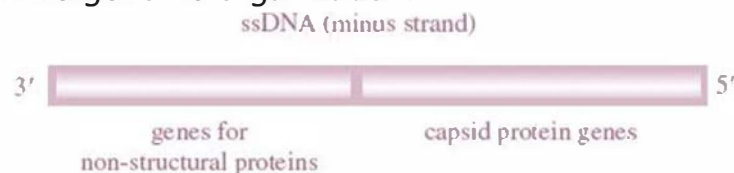
Table 12.2 Percentages of parvovirus virions containing (+) and (-) strand DNA

Virus	Sequences at the DNA termini ¹	% virions containing (+) DNA	% virions containing (-) DNA
B19 virus	ITRs	50	50
Dependoviruses	ITRs	50	50
Minute virus of mice	Unique	1	99

¹The termini of a parvovirus DNA are either "inverted repeats" (ITRs; Figure 12.4(a)) or unique (Figure 12.4(b)).

In a (-) DNA the genes for non-structural proteins are towards the 3' end and the structural protein genes are towards the 5' end (Figure 12.5).

Figure 12.5 Parvovirus genome organization.



12.4 PARVOVIRUS REPLICATION

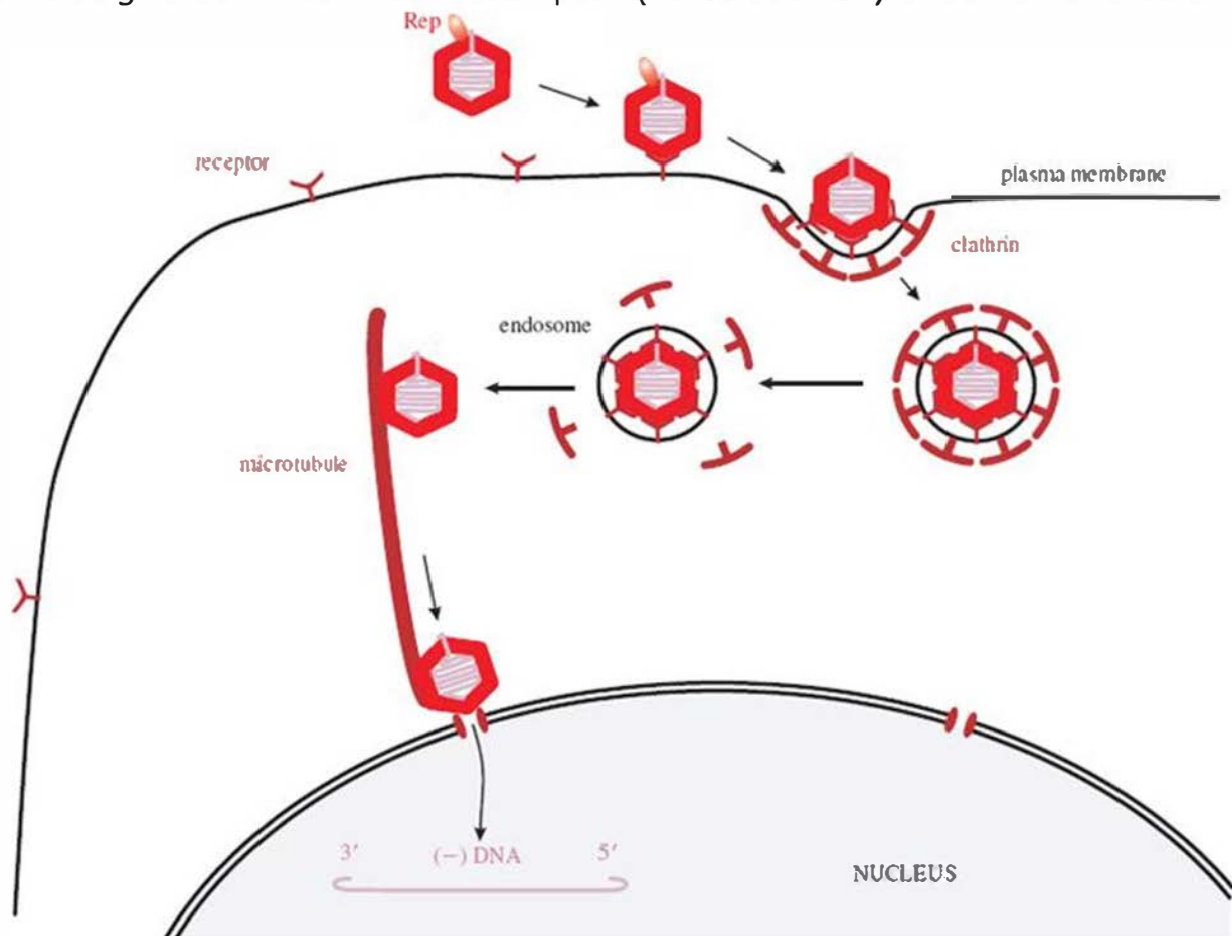
The small genome of a parvovirus can encode only a few proteins, so the virus depends on its host cell (or another virus) to provide important proteins. Some of these cell proteins (a DNA polymerase and other proteins involved in DNA replication) are available only during the S phase of the cell cycle (Figure 4.5), when DNA synthesis takes place. This restricts the opportunity for parvovirus replication to the S phase. Contrast this situation with that of the large DNA viruses, such as the herpesviruses (Chapter 11), which encode their own DNA-replicating enzymes, allowing them to replicate in any phase of the cell cycle.

This account of parvovirus replication is based on studies with several parvoviruses. Some aspects specific to the dependoviruses are covered in Section 12.4.6.

12.4.1 Attachment and entry

A virion attaches to receptors on the surface of a potential host cell (Figure 12.6); it must also bind to one or more co-receptor molecules before the entry process can get under way. In the case of B19 virus the host cell is a red blood cell precursor, the receptor is the blood group P antigen, and the co-receptor is $\alpha 5\beta 1$ integrin. The Rep protein is removed from the 5' end of the viral genome while the virion is still on the surface of the cell.

Figure 12.6 Parvovirus attachment and entry. A virion is taken into the cell by clathrin-mediated endocytosis. After release from the endosome it is transported on a microtubule to a site close to the nucleus. It is uncertain whether uncoating of the virus genome occurs at a nuclear pore (as shown here) or within the nucleus.

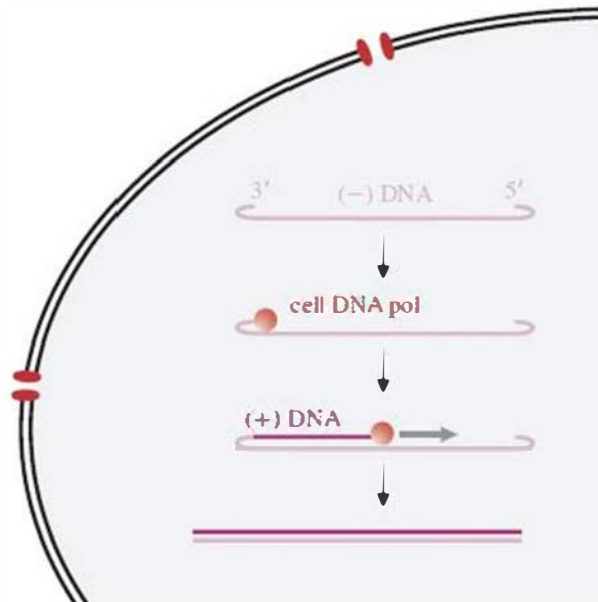


Entry into the cell is by clathrin-mediated endocytosis. The endosome membrane is probably breached by the phospholipase domain of VP1, allowing release of the virion into the cytoplasm. The virion then associates with microtubules and is transported to a nuclear pore. Nuclear localization signals have been found in the VP1 proteins of some parvoviruses and it is likely that the capsid docks with a nuclear pore. It is uncertain whether the virus genome is released at a nuclear pore, or whether it is released in the nucleus after the virion has been transported through a pore. With a diameter of 18–26 nm, the parvovirus virion is small enough to pass through a nuclear pore, unlike the herpesvirus nucleocapsid (Section 11.5.1).

12.4.2 Single-stranded DNA to double-stranded DNA

In the nucleus the single-stranded virus genome is converted to dsDNA by a cell DNA polymerase (Figure 12.7). The ends of the genome are double stranded as a result of base pairing (Figure 12.4), and at the 3' end the -OH group acts as a primer to which the enzyme binds.

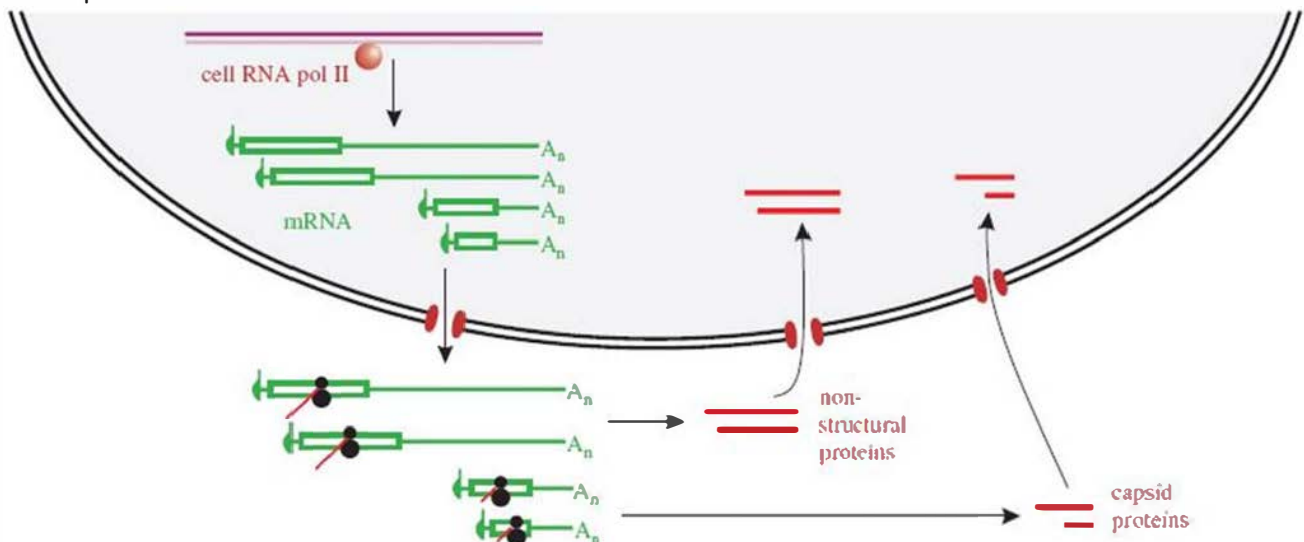
Figure 12.7 Conversion of ssDNA to dsDNA by a cell DNA polymerase. Not all steps are shown.



12.4.3 Transcription and translation

The cell RNA polymerase II transcribes the virus genes and cell transcription factors play key roles. The primary transcript(s) undergo splicing events to produce two size classes of mRNA (Figure 12.8). The larger mRNAs encode the non-structural proteins and the smaller mRNAs encode the structural proteins (see genome organization, Figure 12.5).

Figure 12.8 Parvovirus transcription and translation. The cell RNA polymerase II transcribes the virus genome. After translation much of the virus protein is transported to the nucleus.



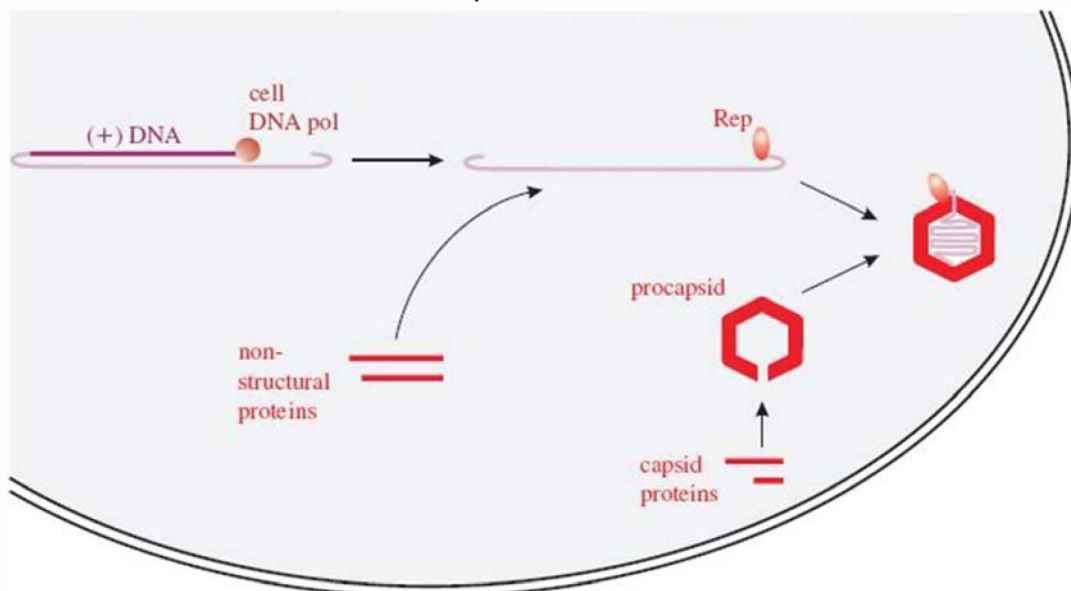
The non-structural proteins are phosphorylated and play a wide variety of roles in the control of gene expression and in DNA replication and packaging.

12.4.4 DNA replication and virion assembly

The viral DNA is replicated by a mechanism called rolling-hairpin replication. This is a leading strand mechanism and sets parvoviruses apart from other DNA viruses, which replicate their genomes through leading and lagging strand synthesis (Section 7.5). Parvoviruses are heavily dependent on the DNA replication machinery of their host cells, though some activities are carried out by the virus Rep/non-structural proteins.

Procapsids are constructed from the structural proteins; each procapsid is filled by a copy of the virus genome, either a (+) DNA or a (-) DNA as appropriate (Figure 12.9). The genome probably enters via one of the twelve pores at the vertices of the procapsid.

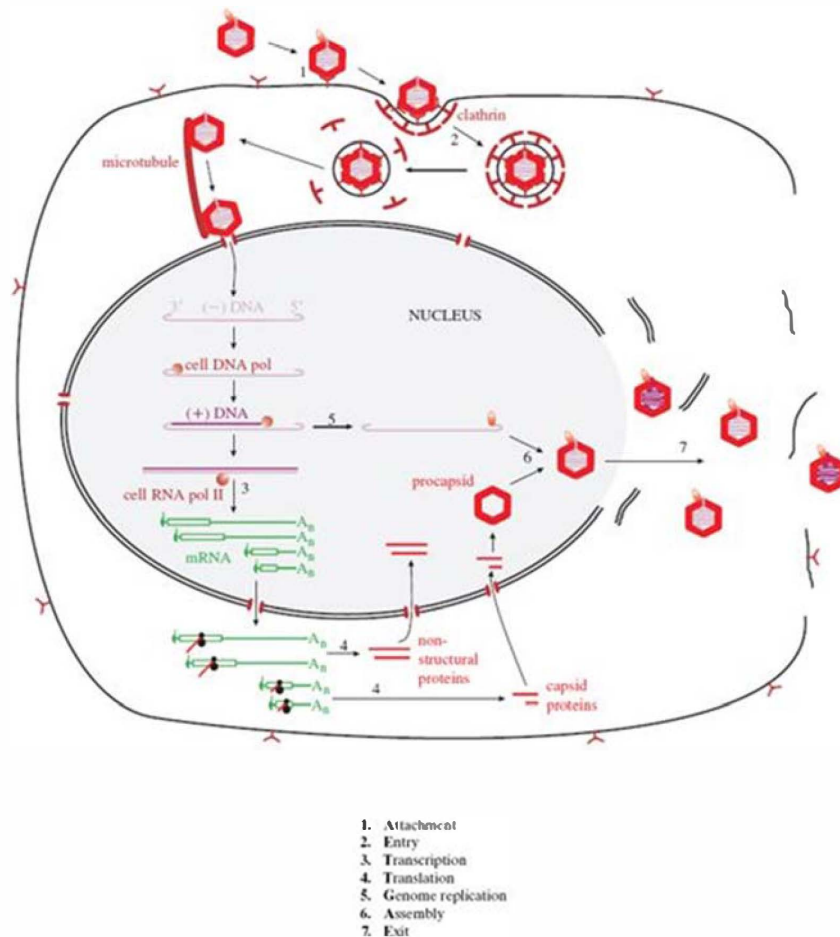
Figure 12.9 Parvovirus virion assembly.



12.4.5 Overview of parvovirus replication

The parvovirus replication cycle is summarized in Figure 12.10.

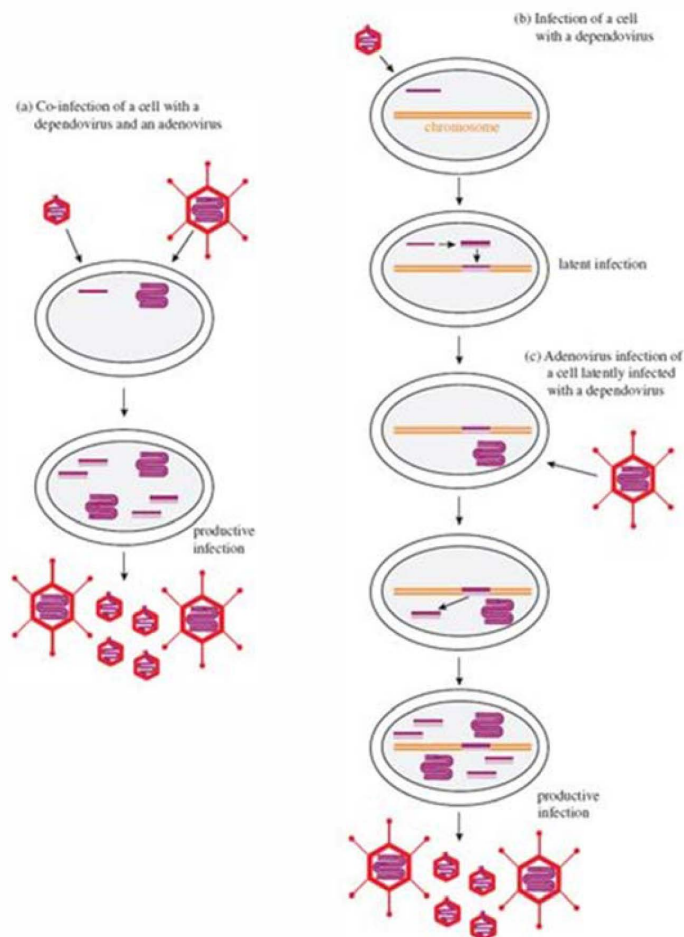
Figure 12.10 The parvovirus replication cycle.



12.4.6 Dependovirus replication

When a cell is co-infected with a dependovirus and an appropriate helper virus, there is a productive infection with both viruses (Figure 12.11(a)).

Figure 12.11 Dependovirus replication. The latent dependovirus infection is shown with the viral DNA integrated into a cell chromosome; in many cases the viral DNA is not integrated. Co-infection with a helper virus provides conditions for a productive dependovirus infection.



In nature it is perhaps more common for a dependovirus to infect a cell in the absence of a helper virus, in which case the virus genome may persist in the cell as a latent infection. It was thought that dependovirus latent infections always require integration of the viral genome (as dsDNA) into a cell chromosome (Figure 12.11(b)), but there is now evidence that in many cases the viral genome is not integrated. If integration occurs in human cells, the virus DNA is inserted at a specific site in chromosome 19. Latent dependovirus infections have been found in a number of cell lines of human and monkey origin. If a cell with a latent dependoviral genome becomes infected with an appropriate helper virus then a productive infection with both viruses can ensue (Figure 12.11(c)).

12.5 OTHER ssDNA VIRUSES

Some further examples of viruses with ssDNA genomes are given in Table 12.3. These viruses, and indeed the majority of the known ssDNA viruses, have circular genomes. The only viruses known with ssDNA linear genomes are the parvoviruses, the main subjects of this chapter. The ssDNA phages are considered further in Section 20.4.

Table 12.3 Examples of ssDNA viruses

	Family	Example
Animal viruses	Circoviridae	Porcine circovirus

Plant viruses	Geminiviridae	Maize streak virus
Bacterial viruses	Microviridae	Phage ϕ X174 (Section 20.4.1)

LEARNING OUTCOMES

After review of this chapter you should be able to:

- give examples of parvoviruses and explain their importance;
- describe the parvovirus virion;
- outline the main features of the parvovirus genome;
- describe the replication cycle of parvoviruses;
- explain the difference between autonomous and defective parvoviruses.

SOURCES OF FURTHER INFORMATION

Books

Kerr, J. R. et al. (eds) (2006) *Parvoviruses*, Hodder Arnold

Parrish, C. R. (2010) Structures and functions of parvovirus capsids and the process of cell infection. In *Current Topics in Microbiology and Immunology*, Volume 343, 149–176 (Johnson, J. E., editor), Springer

Journals

Brown, K. E. (2010) The expanding range of parvoviruses which infect humans. *Reviews in Medical Virology*, 20, 231–244

Cohen, S., et al. (2011) Nuclear envelope disruption involving host caspases plays a role in the parvovirus replication cycle. *Journal of Virology*, 85, 4863–4874

Cotmore, S. F. and Tattersall, P. (2007) Parvoviral host range and cell entry mechanisms. *Advances in Virus Research*, 70, 183–232

dos Santos Coura, R. and Nardi, N. B. (2007) The state of the art of adeno-associated virus-based vectors in gene therapy. *Virology Journal*, 4, 99

Harbison, C. E., Chiorini, J. A., and Parrish, C. R. (2008) The parvovirus capsid odyssey: from the cell surface to the nucleus. *Trends in Microbiology*, 16, 208–214

Lüsebrink, J. et al. (2009) Human bocavirus—insights into a newly identified respiratory virus. *Viruses*, 1, 3–12

Plevka, P. et al. (2011) Structure of a packaging-defective mutant of minute virus of mice indicates that the genome is packaged via a pore at a 5-fold axis. *Journal of Virology*, 85, 4822–4827

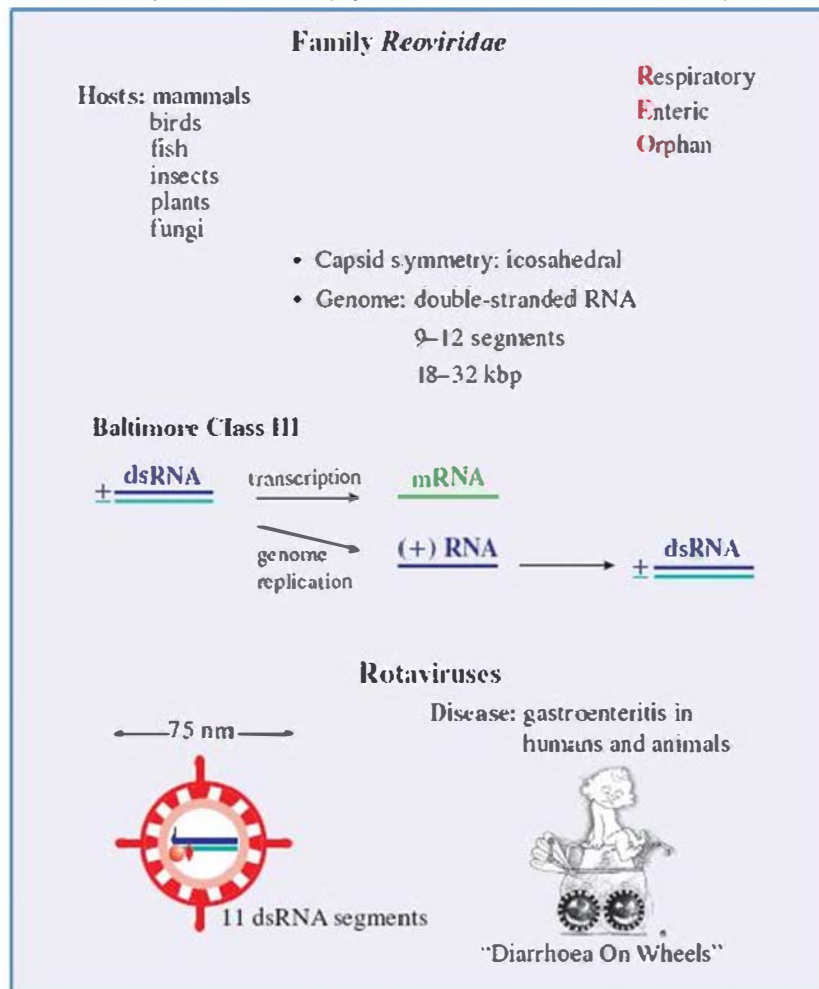
Servant-Delmas, A. et al. (2010) Advances in human B19 erythrovirus biology. Journal of Virology, 84, 9658–9665

CHAPTER 13

Reoviruses (and Other dsRNA Viruses)

Chapter 13 AT A GLANCE

“Diarrhoea On Wheels” from Haaheim et al. (2002) A Practical Guide to Clinical Virology, 2nd edition. Reproduced by permission of John Wiley & Sons



13.1 INTRODUCTION TO REOVIRUSES

Icosahedral viruses with segmented genomes of dsRNA, isolated from the respiratory tracts and enteric tracts of humans and animals, and with which no disease could be associated (orphan), became known as reoviruses. Similar viruses have been found in mammals, birds, fish, invertebrates (including insects), plants, and fungi. Many of these viruses are causative agents of disease, but the original name has been preserved in the family name Reoviridae, and is found in the names of some genera (Table 13.1).

Table 13.1 Some of the genera in the family Reoviridae

Genus	Hosts	Number of dsRNA segments
-------	-------	--------------------------

Cypovirus	Insects	10
Dinovernavirus	Insects	9
Orbivirus	Mammals, birds, invertebrates	10
Orthoreovirus	Mammals, birds	10
Phytoreovirus	Plants, insects	12
Rotavirus	Mammals, birds	11

The original reoviruses are incorporated into the genus Orthoreovirus. The avian viruses are important disease agents, but most orthoreovirus infections of mammals are asymptomatic. The majority of humans become infected with orthoreoviruses early in life and have specific serum antibodies by early adulthood. It is interesting to note that most of the plant-infecting reoviruses are transmitted between plants by insect vectors (Chapter 4). The viruses replicate in both the plant and the insect, generally causing disease in the plant, but little or no harm to the infected insect.

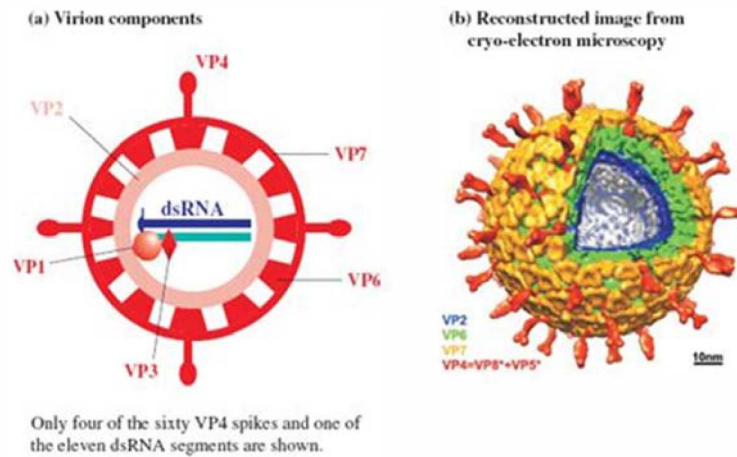
The main focus of this chapter is on the rotaviruses, which have been the subjects of intensive study because they are amongst the most important agents of gastroenteritis in humans and animals. Each year rotaviruses are responsible for the deaths of over half a million children under 5 years old.

13.2 ROTAVIRUS VIRION

Rotaviruses were first described in 1963, when they were observed during electron microscopy of fecal samples from monkeys and mice. Spherical virions about 75 nm in diameter, with structures resembling the spokes of a wheel, were described (Figure 13.1), so the viruses were named after the Latin word *rota* (wheel). Similar viruses were observed 10 years later during electron microscopy of fecal samples from children with diarrhea.

Figure 13.1 The rotavirus virion.

Source: (b) Chen et al. (2009) Proceedings of the National Academy of the USA, 106, 10644. Reproduced by permission of the National Academy of the USA and the authors. (c) by permission of Professor M. Stewart McNulty and The Agri-Food and Biosciences Institute.



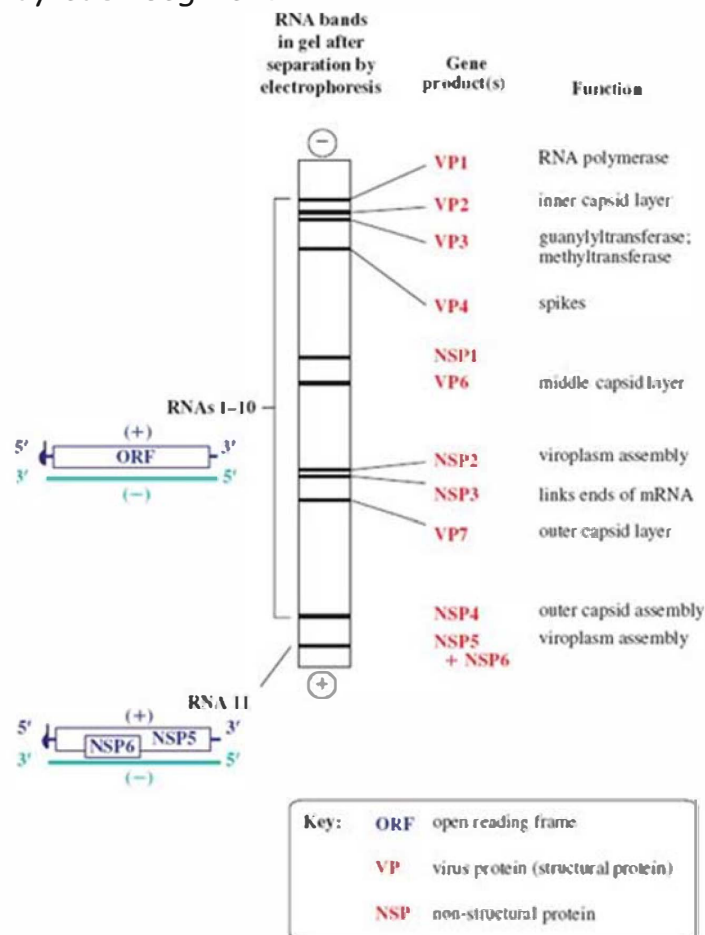
The virion, which has icosahedral symmetry, is also known as a triple-layered particle as the capsid has three layers, each constructed from a distinct virus protein (VP). The inner and middle layers, constructed from VP2 and VP6 respectively, are perforated by channels. The middle layer contains the “spokes” of the “wheel” and is the major component of the virion. The outer layer is constructed from VP7, which is glycosylated. It is unusual to find a glycoprotein in a naked virion, but VP7 is associated with a membrane within the cell before it is incorporated into the virion (Section 13.3.3). Three other protein species are found in the virion: VP1 and VP3 in the core, and VP4, which forms 60 spikes at the surface.

The proteins are numbered in order of their sizes. The three largest proteins are found towards the center of the virion; within the inner capsid layer (VP2), associated with the genome, are 12 copies each of VP1 and VP3. These two proteins are enzymes; VP1 is the RNA-dependent RNA polymerase and VP3 has guanylyl transferase and methyl transferase activities.

The 11 dsRNA segments that make up the rotavirus genome can be separated according to size by electrophoresis in a sodium dodecyl sulfate–polyacrylamide gel (Figure 13.2). Each RNA segment encodes one protein, with the exception of one segment, which encodes two proteins. Hence 12 proteins are encoded: six structural proteins (VP) and six non-structural proteins (NSP). In the diagrams the

RNA strands are color-coded to distinguish the plus-strand (the coding strand) from the minus-strand.

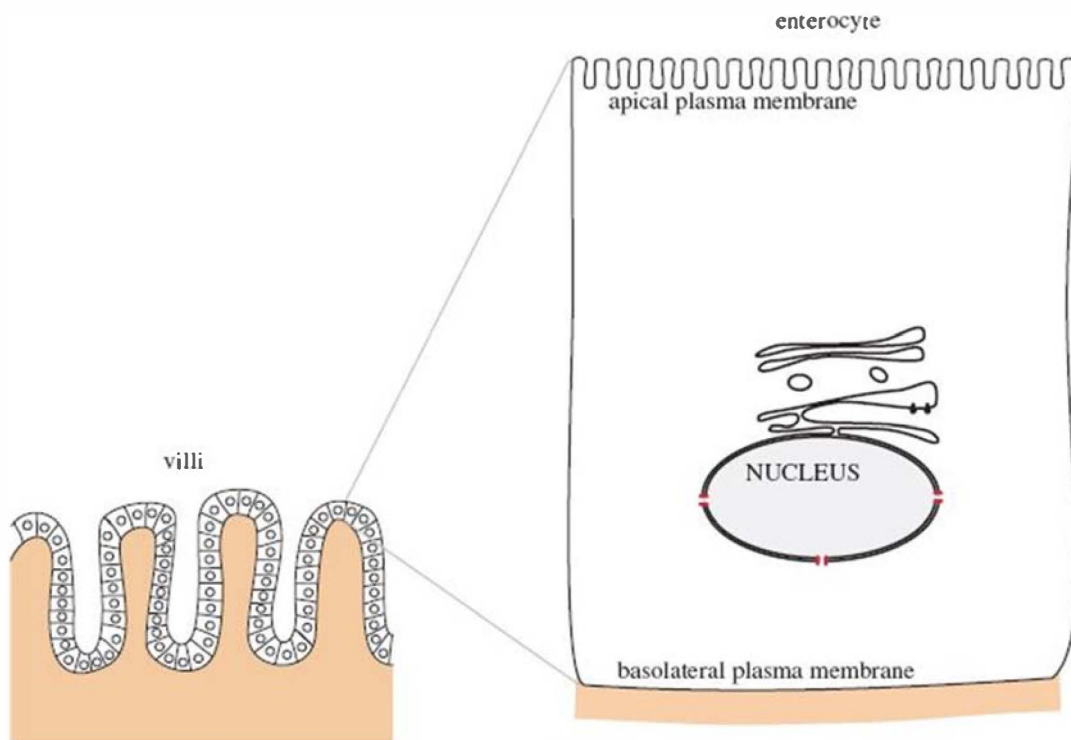
Figure 13.2 The rotavirus genome and gene products. The plus-strand of each of the 11 RNA segments is capped. There is one ORF in each segment, except for the smallest (RNA 11), in which there are two ORFs in different reading frames. NSP5 and NSP6 therefore do not share a common sequence. The diagram shows a typical pattern of rotavirus RNA segments in a gel after electrophoretic separation, and the protein(s) encoded by each segment.



13.3 ROTAVIRUS REPLICATION

Rotaviruses infect cells called enterocytes at the ends of the villi (finger-like extensions) in the small intestine (Figure 13.3).

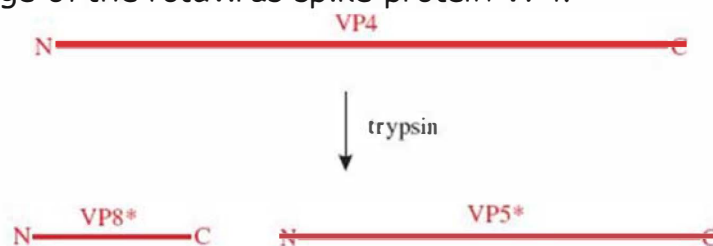
Figure 13.3 Section of the small intestine and detail of an enterocyte.



13.3.1 Attachment and entry

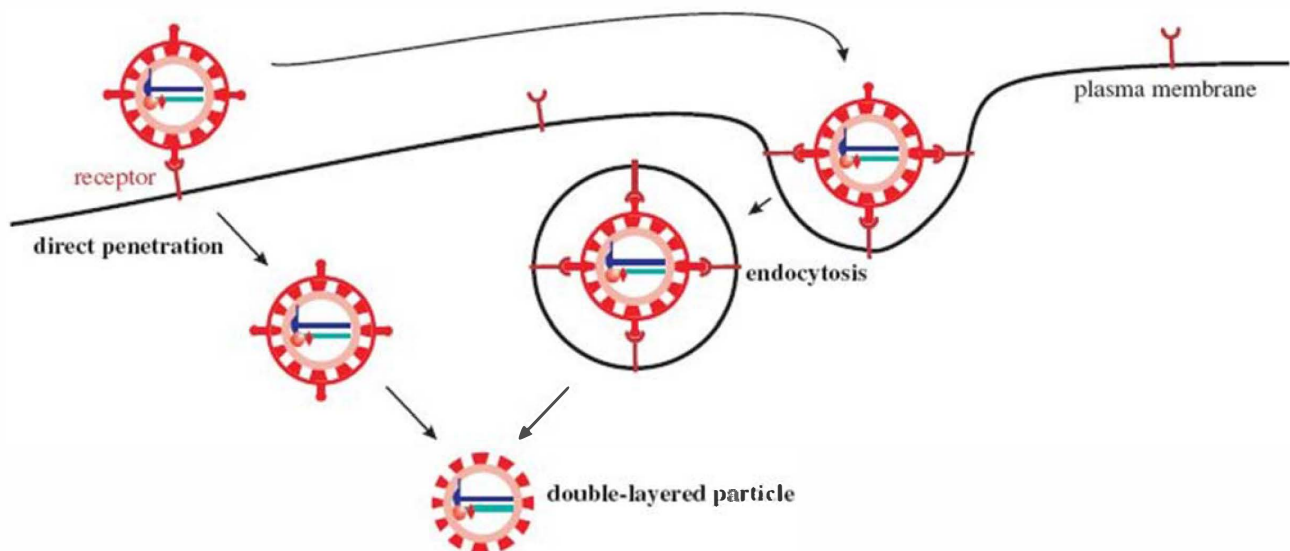
The mechanisms by which rotaviruses attach to and enter their host cells are complex and many details remain uncertain. Cleavage of the spike protein VP4 to the smaller proteins VP8 and VP5 (Figure 13.4) results in much more rapid entry into the cell. This cleavage can be brought about by trypsin-like enzymes in the gut. VP8 and VP5 can bind to several cell surface molecules: VP8 to sialic acid residues on lipids and proteins, and VP5 to integrins. The capsid glycoprotein VP7 can also bind to integrins.

Figure 13.4 Cleavage of the rotavirus spike protein VP4.



There are two possible ways in which a virion can enter the cell: direct penetration across the plasma membrane and endocytosis (Figure 13.5). It is thought that membrane penetration is mediated by a hydrophobic region of VP5; this region is hidden in the uncleaved VP4.

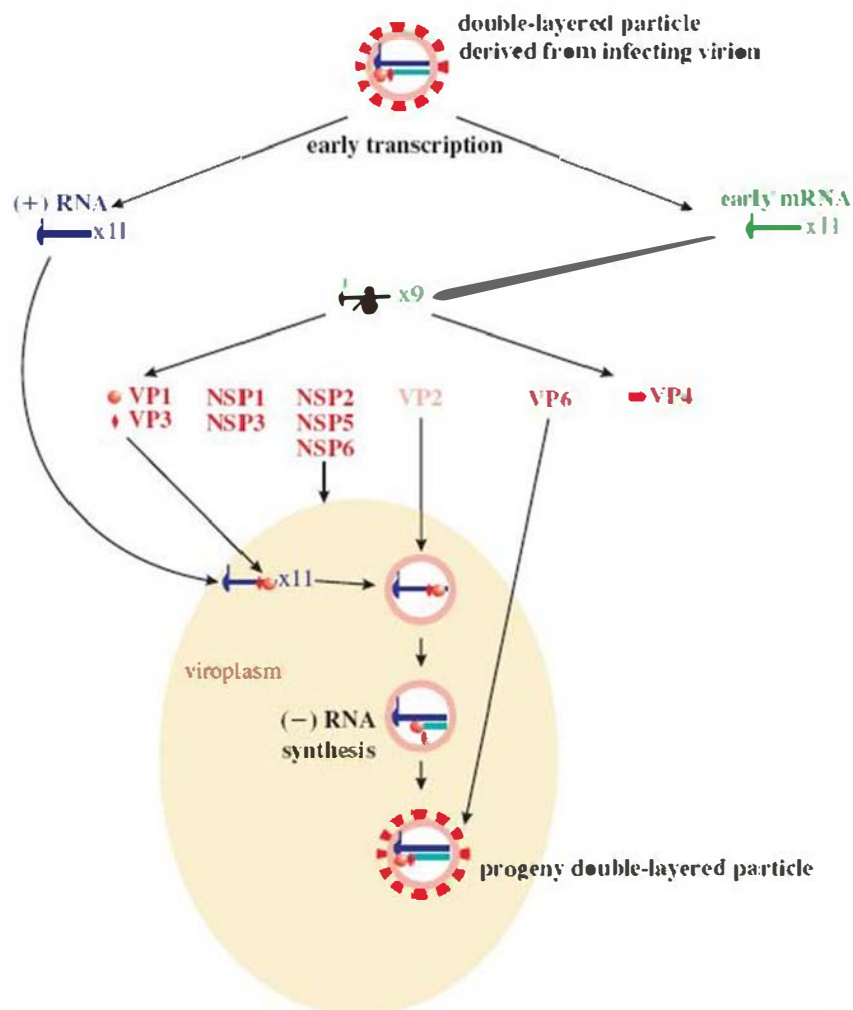
Figure 13.5 Modes of rotavirus entry into the host cell. A rotavirus virion may either penetrate the plasma membrane or it may be endocytosed. Only one of the 11 dsRNA molecules in the virion is shown.



13.3.2 Early events

The outer layer of the virion is removed, leaving a double-layered particle in which transcription is activated (Figure 13.6). It is likely that each of the 11 genome segments is associated with a molecule of VP1 (synthesizes new copies of the (+) RNA) and a molecule of VP3 (caps the 5' ends of the new RNAs). The nucleotides for RNA synthesis enter the particles through the channels in the protein layers and the transcripts are extruded from the particles through the same channels. The transcripts are not polyadenylated.

Figure 13.6 Early events in rotavirus replication. (+) RNA is transcribed from each of the genome segments in the double-layered particle derived from the infecting virion. Initially this RNA functions as mRNA (green) and the 12 virus proteins are translated (translation of two proteins in the rough endoplasmic reticulum is not shown). NSP2 and NSP5 play major roles in the assembly of viroplasm, and other virus proteins accumulate in these structures. VP1 and VP3 bind to (+) RNA molecules (dark blue) and then to VP2, which forms roughly spherical structures. (-) RNA (light blue) is synthesized in these structures and VP6 is added to form progeny double-layered particles.

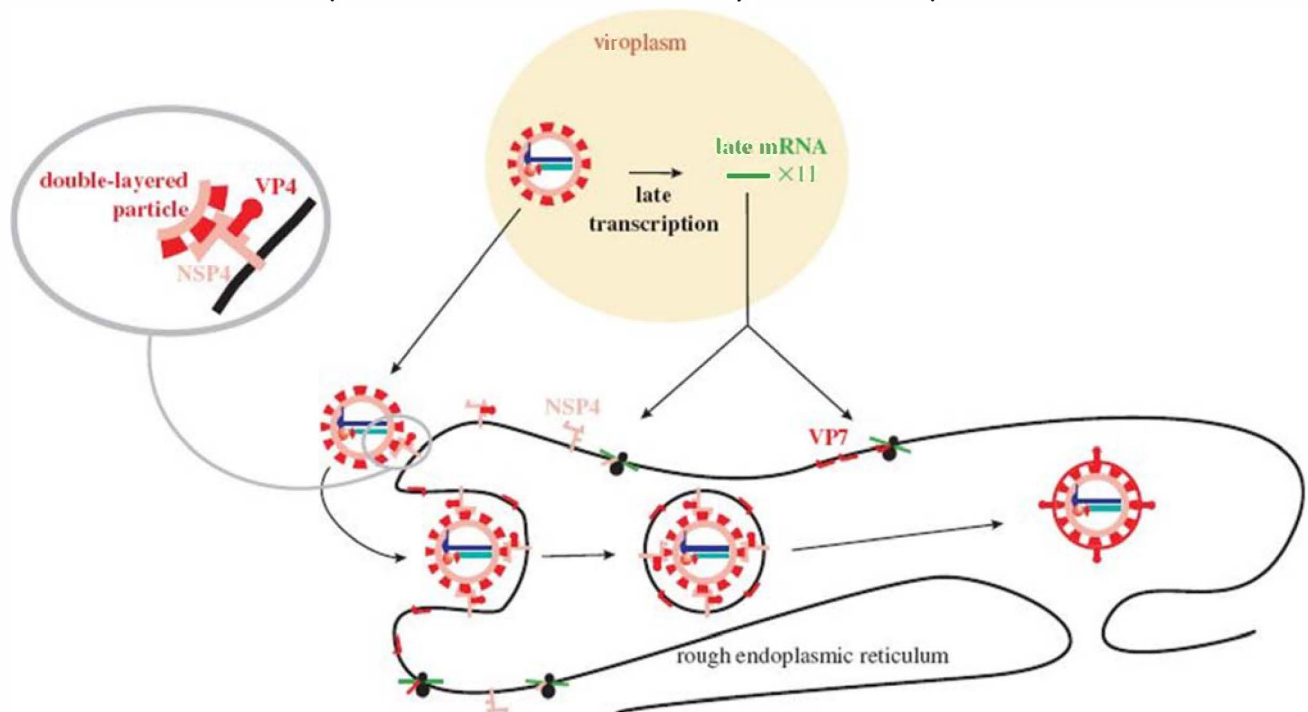


Some of the virus proteins undergo co- and post-translational modifications: VP2 and VP3 are myristylated while NSP5 is phosphorylated. The virus proteins accumulate within the cytoplasm in structures known as viroplasm; there is evidence that NSP2 and NSP5 play major roles in their formation. Transient expression of the genes for these proteins resulted in the formation of viroplasm-like structures, while reduced expression of NSP5 by RNA interference (Section 2.9.3) resulted in fewer and smaller viroplasms. In the viroplasms the newly synthesized (+) RNA molecules bind VP1 and VP3, and then VP2, which forms structures that are roughly spherical. A rigorous selection procedure ensures that each structure receives one each of the 11 RNA species, i.e. a full genome complement. This procedure, which operates for all viruses with multipartite genomes, involves the recognition of a unique sequence in each genome segment. The RNA polymerase activity of VP1 is activated by the interaction with VP2, and the (+) RNA strands act as templates for the synthesis of (-) RNA, producing the 11 dsRNA molecules of the genome in each structure (Figure 13.6). The dsRNA of the infecting virion therefore remains intact and the mode of replication is conservative, in contrast to the replication of the dsRNAs of some other viruses (Section 7.6). VP6 is added to the core, forming the second layer of the capsid. The resulting structure is a double-layered particle similar to that derived from the infecting virion.

13.3.3 Late events

A further round of transcription (secondary transcription) takes place within the progeny double-layered particles. In contrast to the early transcripts, the late transcripts are not capped (Figure 13.7). The translation machinery undergoes changes and now selects uncapped transcripts in preference to capped transcripts; NSP3 plays a role in linking the 5' end of an mRNA to its 3' end. Hence, translation of cell proteins is shut down while translation of virus proteins continues.

Figure 13.7 Late events in rotavirus replication. Secondary transcription takes place in progeny double-layered particles and results in the synthesis of uncapped mRNA. The 12 virus proteins are translated; only NSP4 and VP7 translation are shown here (both are synthesized in the endoplasmic reticulum). NSP4 binds VP4 and a double-layered particle (inset), and this complex buds into the endoplasmic reticulum. The VP4 spikes and the VP7 outer layer of the capsid are added.



The virus requires different quantities of each of its 12 proteins; for example, a large amount of the main capsid protein (VP6) is required, but a relatively small amount of VP1. Although there are equimolar amounts of each of the 11 dsRNAs, the mRNA species are not made in equimolar amounts, and there are also controls at the level of translation. Hence, there are several mechanisms used by the virus to control the quantity of each protein produced.

The final stages of virion assembly involve the addition of the outer layer of the capsid and the spikes. There is some uncertainty concerning the details, but there is evidence suggesting the following series of events.

VP7 and NSP4 are synthesized and N-glycosylated in the rough endoplasmic reticulum, where they accumulate in the membrane. NSP4 has binding sites for both VP4 and double-layered particles (Figure 13.7, inset). After these components

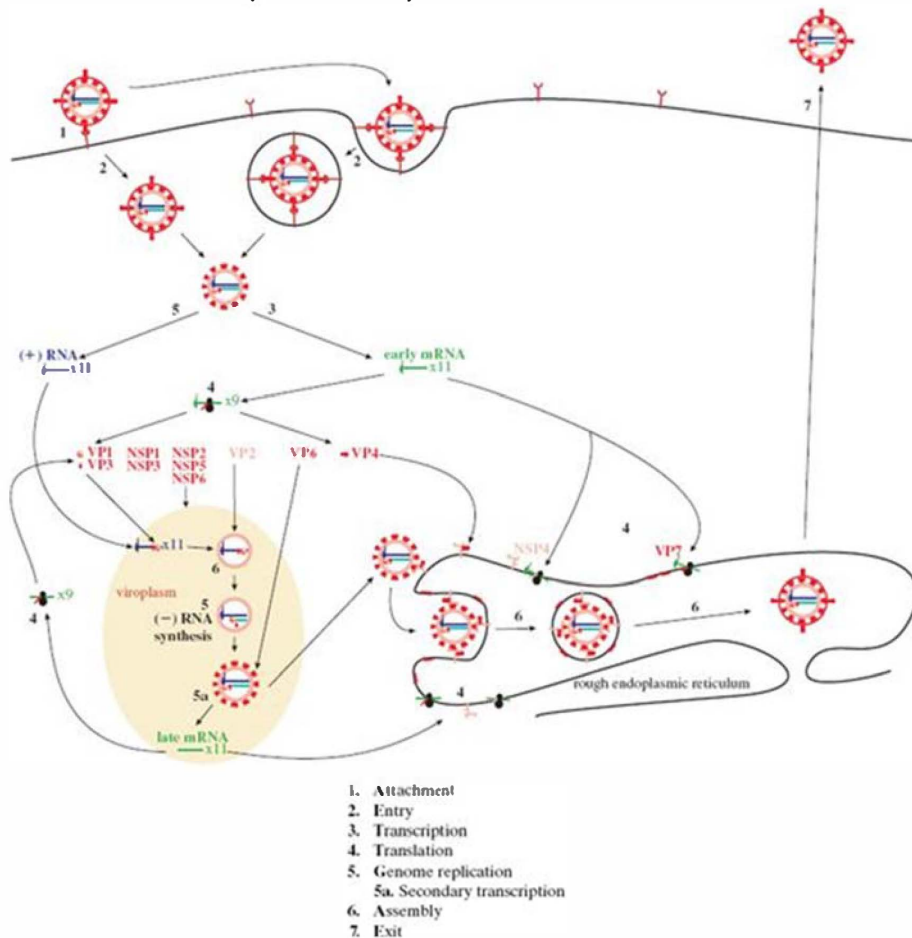
have bound, the particle buds through the membrane into a vesicle within the endoplasmic reticulum. The vesicle membrane forms a temporary “envelope” containing VP7.

VP4 is added to the double-layered particle to form the virion spikes. Cleavage of VP7 molecules releases them from the membrane to build up the outer capsid and lock the spikes in place. Virions are released from the cell either by lysis or by exocytosis.

13.3.4 Overview of rotavirus replication

The rotavirus replication cycle is summarized in Figure 13.8.

Figure 13.8 The rotavirus replication cycle.



13.3.5 Rotavirus disease

Enterocytes on the villi are destroyed as a result of rotavirus infection and this leads to reduced absorption of water, salts, and sugars from the gut. The virus protein NSP4 is responsible for leakage of chloride from cells into the gut lumen. These effects of virus infection, together with the secretion of water and solutes by secretory cells, result in diarrhea and this can lead to dehydration. Treatment is a relatively simple matter of rehydrating the patient with a solution of salts and sugar, but the means for this are not available in some parts of the world; for many, especially young children, the outcome is death.

Although most rotavirus infections are confined to the gut, there is evidence that on occasion the virus may cross the gut and infect other tissues.

13.4 OTHER dsRNA VIRUSES

The virions of most dsRNA viruses have icosahedral symmetry and most of them are naked; an exception is the family Cystoviridae, the members of which are enveloped bacteriophages (Section 20.3). Some further examples of viruses with dsRNA genomes are given in Table 13.2.

Table 13.2 Examples of dsRNA viruses

	Family	Example
Animal viruses	Birnaviridae	Infectious bursal disease virus
Plant viruses	Partitiviridae	White clover cryptic virus 1
Fungal viruses	Chrysoviridae	Penicillium chrysogenum virus
Bacterial viruses	Cystoviridae	Pseudomonas phage $\phi 6$

A problem that dsRNA viruses must overcome is the fact that dsRNA is a potent inducer of a number of cell defense mechanisms, including apoptosis, interferon production, and RNA silencing (Chapter 9). Most of these viruses, including the rotaviruses, have solved this problem by ensuring that their dsRNA is always enclosed within virus protein structures, and is therefore never free in the cytoplasm to trigger these defenses.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- explain how viruses in the family Reoviridae are classified into genera;
- describe the rotavirus virion;
- discuss the main events of the rotavirus replication cycle;
- explain how rotaviruses cause disease.

SOURCES OF FURTHER INFORMATION

Books

Baker, M. and Prasad, B. V. V. (2010) Rotavirus Cell Entry. In Cell Entry by Non-Enveloped Viruses, Current Topics in Microbiology and Immunology, Volume 343, 121–148 (Johnson J. E., editor), Springer

Franco, M. A. and Greenberg, H. B. (2009) Rotaviruses. Chapter 35 in Clinical Virology, 3rd edition (Richman, D. D., Whitley, R. J., and Hayden, F. G., editors), ASM Press

Patton, J. T. (ed.) (2008) Segmented Double-Stranded RNA Viruses: Structure and Molecular Biology, Caister Academic Press

Yang, D. (ed.) (2009) Rotaviruses. Chapter 27 in RNA Viruses: Host Gene Responses to Infections, World Scientific

Journals

Ayala-Breton, C. et al. (2009) Analysis of the kinetics of transcription and replication of the rotavirus genome by RNA interference. *Journal of Virology*, 83, 8819–8831

Chen, J. Z. et al. (2009) Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM. *Proceedings of the National Academy of Sciences of the USA*, 106, 10644–10648

Cheung, W. et al. (2010) Rotaviruses associate with cellular lipid droplet components to replicate in viroplasms, and compounds disrupting or blocking lipid droplets inhibit viroplasm formation and viral replication. *Journal of Virology*, 84, 6782–6798

McClain, B. et al. (2010) X-ray crystal structure of the rotavirus inner capsid particle at 3.8 Å resolution. *Journal of Molecular Biology*, 397, 587–599

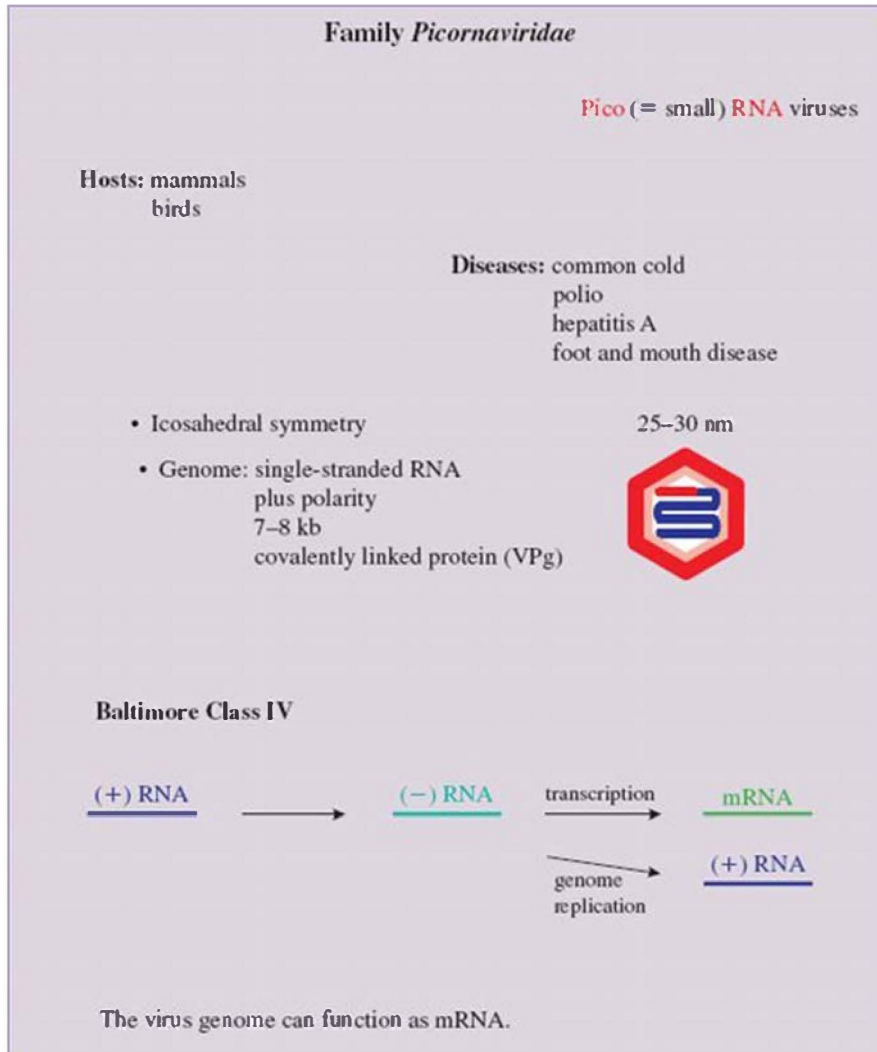
Parashar, U. D. et al. (2006) Rotavirus and severe childhood diarrhea. *Emerging Infectious Diseases*, 12 (2): <http://dx.doi.org/10.3201/eid1202.050006>

Trask, S. D., McDonald S. M., and Patton, J. T. (2012) Structural insights into the coupling of virion assembly and rotavirus replication. *Nature Reviews Microbiology*, 10, 165–177

CHAPTER 14

Picornaviruses (and Other Plus-Strand RNA Viruses)

CHAPTER 14 AT A GLANCE



14.1 INTRODUCTION TO PICORNAVIRUSES

Members of the family Picornaviridae are found in mammals and birds. These viruses have features in common with many plant viruses, and are classified with them in the order Picornavirales. Some of the genera in the family Picornaviridae and some of the important viruses are listed in Table 14.1.

Table 14.1 Examples of picornaviruses

Genus	Name derivation from Greek word	Example(s)
-------	---------------------------------	------------

Hepatovirus	Hepatos = liver	Hepatitis A virus
Enterovirus	Enteron = intestine	Poliovirus
		Coxsackieviruses
		Rhinoviruses
Aphthovirus	Aphtha = vesicles in the mouth	Foot and mouth disease virus

Poliovirus was one of the first viruses to be propagated in cell culture (Enders, Weller, and Robbins, 1949) and was also one of the first to be plaque purified (Dulbecco and Vogt, 1954). Most picornaviruses grow readily in cell culture, are easy to purify, and are stable, making them popular viruses for laboratory studies.

The picornaviruses are class IV viruses; their genome is a plus-strand RNA that functions as mRNA once it is released into a host cell. The first virus molecules to be synthesized in an infected cell are therefore proteins, including those that will replicate the virus RNA. When sufficient quantities of these proteins have accumulated, the role of the RNA switches from mRNA to template for RNA replication. This is in contrast to all other viruses, where transcription of virus genes must occur before virus protein synthesis can start.

14.2 SOME IMPORTANT PICORNAVIRUSES

14.2.1 Hepatitis A virus

Hepatitis A is especially prevalent in developing countries with poor sanitation. In most infants and young children infection is asymptomatic or mild, and leads to lifelong immunity. When adults become infected about 75% develop jaundice; severe hepatitis is a rare complication, which can be fatal.

14.2.2 Poliovirus

Poliovirus has been the subject of much research effort because of the devastating paralysis that it can cause. In fact, the majority of poliovirus infections are relatively harmless infections of the oropharynx and the gut. Serious disease occurs only after other tissues become infected, resulting in viremia (virus in the blood) and spread of infection to the central nervous system. This is a very rare event in babies, who still have anti-poliovirus antibodies acquired from their mothers. If, however, there is less poliovirus in the human environment, so that most infections occur after these antibodies have disappeared, then infection of the central nervous system is more likely. Ironically, polio is a disease primarily associated with improving standards of hygiene and sanitation!

Poliovirus infection of the central nervous system may result in meningitis (from which most patients recover completely), encephalitis, and/or paralytic poliomyelitis. The latter is due to virus replication in motor neurons of the spinal

cord or the brain stem, resulting in paralysis of limbs and/or breathing muscles (Figure 14.1).

Figure 14.1 (a) Children with paralyzed limbs as a result of polio. (b) Respirators (вЂњiron lungsвЂќ) used in the mid-twentieth century to aid the breathing of polio victims.

Source: (a) Reproduced by permission of the World Health Organization. (b) Reproduced by permission of the US Centers for Disease Control and Prevention.

(a)



(b)



The fear of paralysis was a major motivation for research programs to develop polio vaccines in the mid-twentieth century, and these efforts led to the development of inactivated vaccines and live attenuated vaccines (Chapter 25). The use of both types of vaccine has proved very effective in preventing polio, so that by the start of the twenty-first century the annual number of world cases had fallen to 3500, which was 1% of the number of cases in 1988. Polio has been eradicated from many parts of the world and there is hope that the disease will soon be completely eradicated.

14.2.3 Coxsackieviruses

In 1948, investigators in the US town of Coxsackie injected fecal specimens from two suspected polio cases into suckling mice; lesions developed in the skeletal muscles of the mice and the first coxsackieviruses had been isolated. Coxsackieviruses cause a range of medical conditions, including myocarditis (heart disease), meningitis, and rashes.

14.2.4 Rhinoviruses

Rhinoviruses are the most common agents causing upper respiratory tract infections in humans. Most children have had at least one rhinovirus infection by the age of 2 years, and in adults rhinoviruses account for about 50% of common colds. Human rhinoviruses replicate in the epithelium of the upper respiratory tract, where the temperature is around 33вЂњ35 B°C; for at least some of these viruses the optimum temperature for replication is in this range. Many human rhinoviruses are able to replicate at 37 B°C and some of these are probably responsible for disease in the lower respiratory tract.

14.2.5 Foot and mouth disease virus

Foot and mouth disease virus has a much wider host range than other picornaviruses; it infects mammals such as cattle, sheep, goats, pigs, and deer, causing lesions on the feet and in the mouth.

Outbreaks of foot and mouth disease can have serious economic consequences, as milk yields drop, young infected animals may die, and in developing countries animals become unfit for their crucial roles in ploughing and transport. Restrictions on trade in susceptible animal species cause further economic damage. A massive outbreak in the UK in 2001 was brought under control through a series of measures that included the slaughter of more than 6 million animals (Figure 14.2). The estimated cost of the outbreak was over 8 billion pounds. There were smaller outbreaks of foot and mouth disease in the UK in 2007.

Figure 14.2 Foot and mouth disease pyre at a UK farm.

Source: Courtesy of Simon Ledingham.



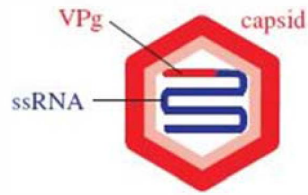
14.3 PICORNAVIRUS VIRION

Picornaviruses are small RNA viruses of relatively simple structure (Figure 14.3). The RNA is enclosed by a capsid, which is roughly spherical and has a diameter of about 25–30 nm.

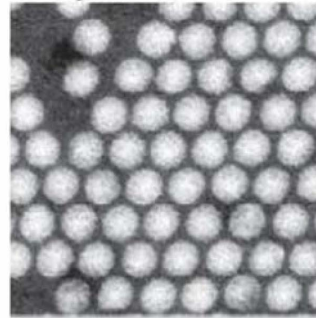
Figure 14.3 The picornavirus virion. VPg: virus protein, genome linked.

Source: Electron micrograph courtesy of J. Esposito (US Centers for Disease Control and Prevention) and Professor Frederick A. Murphy (University of Texas Medical Branch).

(a) Virion components



(b) Electron micrograph of negatively stained virions of poliovirus

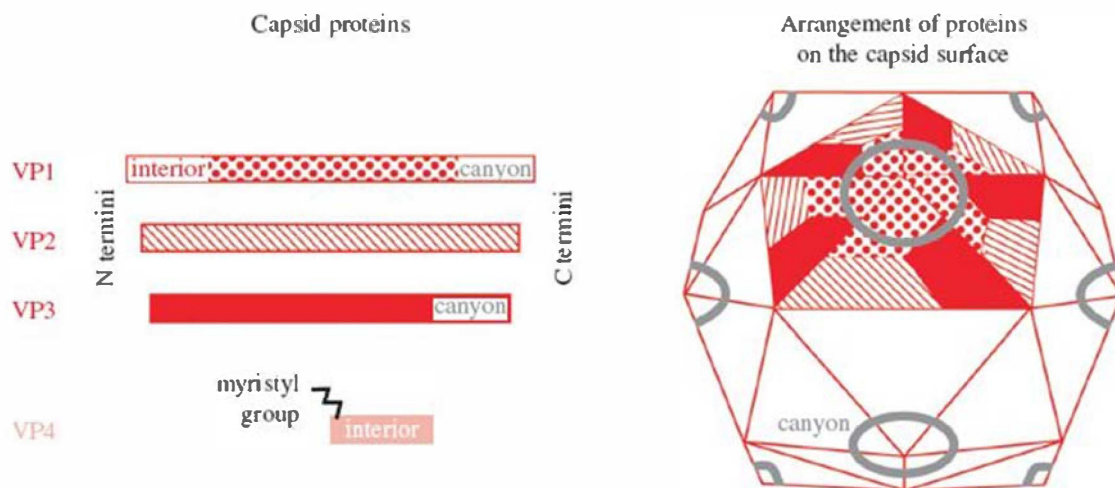


14.3.1 Capsid

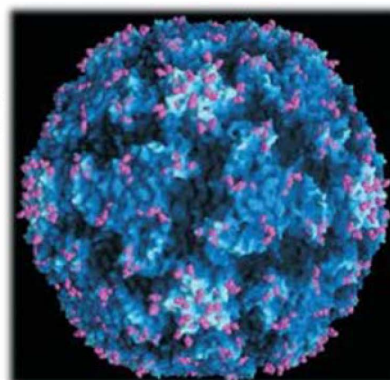
The picornavirus capsid has icosahedral symmetry and is made from 60 identical subunits, each consisting of four virus proteins (VP1-4, numbered from the largest to the smallest). Each of the proteins VP1-3 contains an eight-stranded α -barrel, like many virus capsid proteins, including those of the parvoviruses (Chapter 12). The capsid of human rhinovirus 14 has been studied extensively and some of the results are summarized in Figure 14.4. Parts of VP1-3 are at the surface of the virion, while the N termini of VP1-3 and all 60 VP4 molecules are completely internal.

Figure 14.4 The capsid of human rhinovirus 14. The outer surface of the capsid is composed of regions of VP1, VP2, and VP3. Around each of the vertices is a canyon lined with the C termini of VP1 and VP3. The interior surface of the capsid is composed of VP4 and the N termini of VP1. The canyons are visible in the reconstructed image. The pink regions are neutralizing immunogenic sites (Sherry et al. (1986) *Journal of Virology*, 57, 246).

Source: Reconstructed image courtesy of J. Y. Sgro, University of Wisconsin-Madison. The data for the image are from Arnold and Rossmann (1988) *Acta Crystallographica, Section A*, 44, 270.



Reconstructed image from cryo-electron microscopy



In many picornaviruses, including polioviruses and rhinoviruses, there is a deep cleft around each of the 12 vertices of the icosahedron (Figure 14.4). These clefts, often referred to as canyons, were discovered using X-ray crystallography and cryo-electron microscopy. They are approximately 2 nm deep, which, for a virion of this size, is a significant depth, being approximately 7% of the virion diameter. This is relatively much greater than the depth of the Grand Canyon on the surface of the Earth! The canyons are lined by the C termini of VP1 and VP3 molecules and contain the sites that bind to cell receptors.

Evolution of picornaviruses has generated a lot of variability in the capsid proteins, some of which is reflected in the existence of distinct serotypes (Table 14.2; Figure 10.1(a)). This antigenic variability creates problems for the development of vaccines against these viruses, problems that have been overcome for poliovirus and foot and mouth disease virus, but not for the rhinoviruses.

Table 14.2 Serotypes of picornaviruses

Virus	Number of serotypes
Poliovirus	3
Foot and mouth disease virus	7
Human rhinoviruses	>100

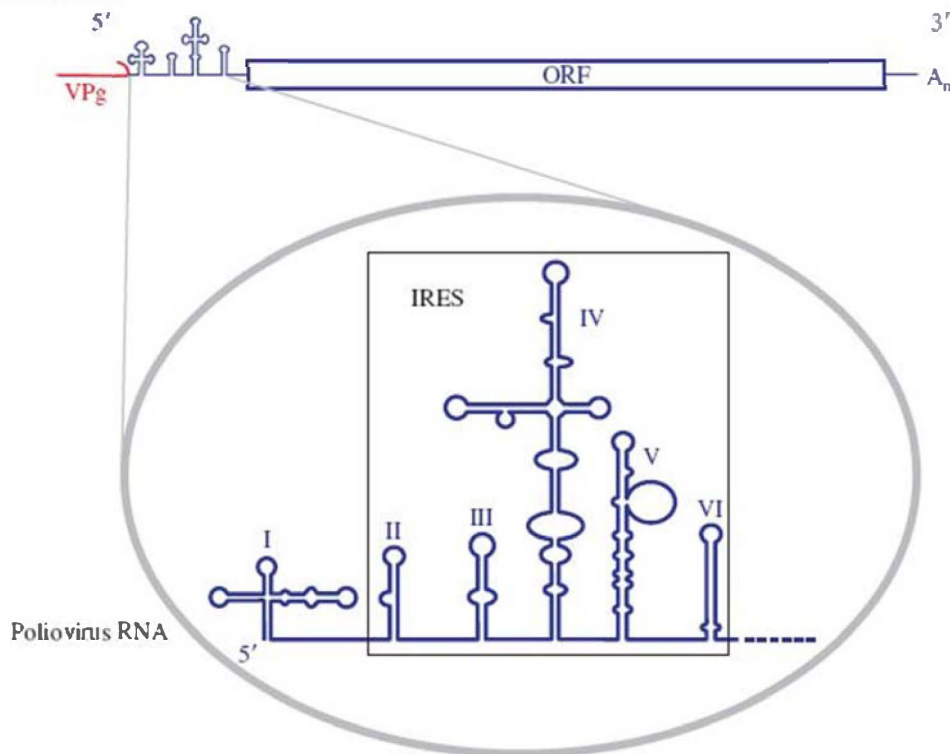
Only . . . by discovering that there were three immunologically distinct forms of polio virus could we move beyond the inadequate iron lungs and truly conquer polio.

Watson J.D. (2000) *A Passion for DNA: Genes, Genomes and Society*, Oxford University Press

14.3.2 Genome

The picornavirus genome is composed of a 7–8 kb ssRNA. Covalently linked to the 5' end of the RNA is a small (2–3 kD) protein known as VPg (virus protein, genome linked) (Figure 14.5). The covalent link is via the OH group of a tyrosine residue at position 3 of VPg. The 3' end of the RNA is polyadenylated.

Figure 14.5 The picornavirus genome. VPg: virus protein, genome linked. Inset: 5' end of poliovirus RNA. IRES: internal ribosome entry site. I–VI: secondary structure domains.



The genome consists of one very large ORF flanked by untranslated regions. Within the untranslated region at the 5' end there is much secondary structure. For example, in poliovirus RNA there are six domains (Figure 14.5), five of which form an internal ribosome entry site (IRES; Section 6.4.1).

14.4 PICORNAVIRUS REPLICATION

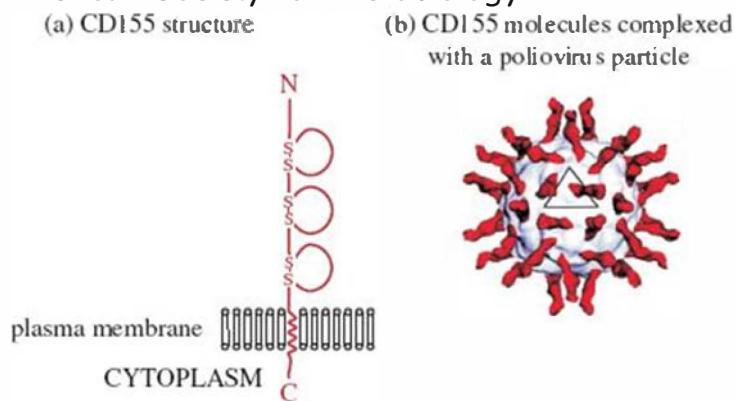
A model of picornavirus replication has been developed, though uncertainty still surrounds many of the details. Many components of the model stem from studies with poliovirus, which replicates in a much wider range of cell types in culture than it does in the body. Information derived from studies with other picornaviruses will also be mentioned.

14.4.1 Attachment

The cell receptors for a number of picornaviruses have been characterized, including those for the rhinoviruses and poliovirus (Section 5.2.1). The poliovirus receptor is CD155, a glycoprotein expressed on many cell types. It is a member of the immunoglobulin superfamily of molecules, with three immunoglobulin-like domains (Figure 14.6(a)); the virus binds to a site in the outer domain.

Figure 14.6 The poliovirus receptor. (a) CD155 is a glycoprotein; the sugar side-chains are not shown. Each loop indicates an immunoglobulin-like domain stabilized by disulfide bonds. (b) Reconstructed image from cryo-electron microscopy.

Source: (b) from He et al. (2003) *Journal of Virology*, 77, 4827. Reproduced by permission of the American Society for Microbiology.

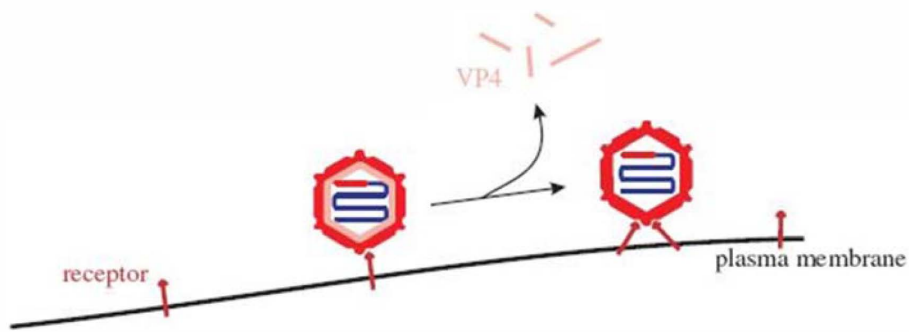


CD155 is found only in humans and some other primate species. In the early days of poliovirus research none of the common laboratory animals could be infected with the virus, but in the early 1990s transgenic mice expressing CD155 were developed. These animals were found to be susceptible to infection with all three serotypes of poliovirus and they have been used in studies of replication and pathogenesis.

For those picornaviruses like poliovirus, with canyons on the virion surface, the attachment sites for cell receptors are located in pockets at the canyon bases. The canyons are too narrow for access by antibodies so the attachment sites are protected from the host's immune surveillance, while the remainder of the virion surface can mutate to avoid the host's immune response. The attachment sites of foot and mouth disease virus, however, are not in canyons but are on surface protrusions.

Binding of a virion to receptors results in major changes to the capsid structure: the N termini of VP1 move from the interior to the exterior surface of the capsid and VP4 is lost from the virion (Figure 14.7). Binding of free CD155 to poliovirus particles (Figure 14.6(b)) also causes loss of VP4.

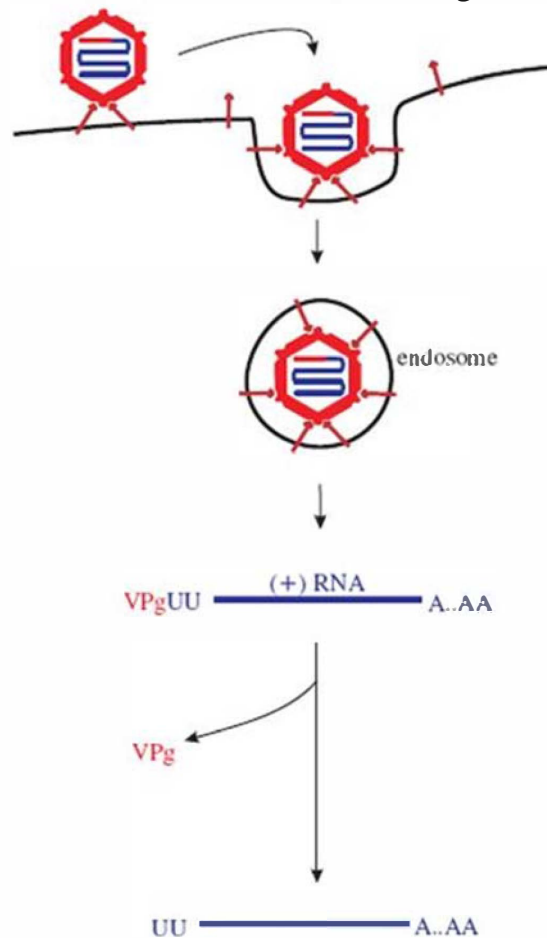
Figure 14.7 Attachment of picornavirus virion to receptors on host cell. The virion undergoes major conformational changes and VP4 is lost.



14.4.2 Entry

Several modes of genome entry into the host cell have been proposed for picornaviruses. One involves transfer of the RNA from the virion into the cytoplasm at the plasma membrane, leaving the capsid at the cell surface. Other modes involve endocytosis (Figure 14.8), followed by either release from the capsid after disruption of the endosome membrane, or release of RNA from the capsid then transport across the endosome membrane. Once the virus genome is free in the cytoplasm the VPg is removed from the 5'UTR end by a cell enzyme.

Figure 14.8 Entry of picornavirus virion by endocytosis. After release from the endosome VPg is removed from the 5'UTR end of the genome.



14.4.3 Translation and post-translational modifications

The virus genome functions as mRNA, but as the RNA is not capped the preliminary stages of translation are different to those for capped mRNAs. The 40S ribosomal

subunit binds not at the 5'UTR end of the RNA, but internally at the IRES (Figure 14.5). Most of the eukaryotic initiation factors (eIFs, normally involved in cap-dependent translation) are involved, but a notable exception is the cap-binding protein eIF4E.

Early investigators found that the number of virus proteins synthesized in the infected cell seemed to be greater than could be encoded by the small genome of a picornavirus. The puzzle was solved when it was found that the genome encodes a single polyprotein, which undergoes a series of cleavages to give rise to all the structural and non-structural proteins. An overview of the process is indicated in Figure 14.9, while some of the detail specific to poliovirus is shown in Figure 14.10.

Figure 14.9 Picornavirus translation and post-translational modifications. A polyprotein is synthesized and rapidly cleaved into P1, P2, and P3. The N terminus of P1 is myristylated. Further proteolytic cleavages take place. The RNA polymerase and VPg are among the products of P3 proteolysis.

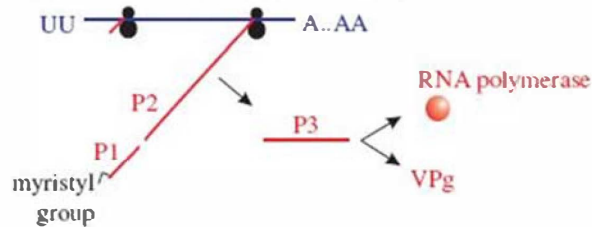
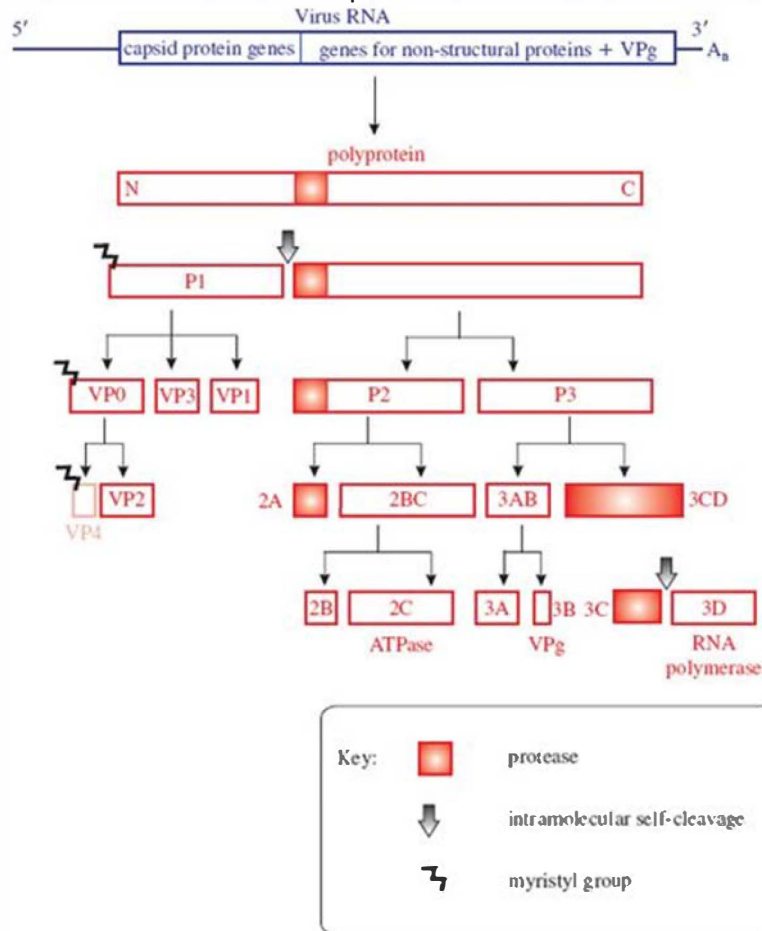


Figure 14.10 Poliovirus translation and post-translational modifications.



Virus-coded proteases (indicated in Figure 14.10) perform the cleavages, some of which are intramolecular self-cleavages. The polyprotein is first cut to yield P1, P2, and P3. P1 becomes myristylated at the N terminus before being cleaved to VP0, VP3, and VP1, the proteins that will form procapsids; VP0 will later be cleaved to produce VP2 and VP4. Other cleavage products include 3B (VPg), 2C (an ATPase), and 3D (the RNA polymerase). The production of the proteins via a polyprotein precursor might suggest that equimolar amounts result, but not all cleavages take place with equal efficiency. The amounts of the proteins therefore vary; for example, greater amounts of the capsid proteins are produced than of 3D (the RNA polymerase).

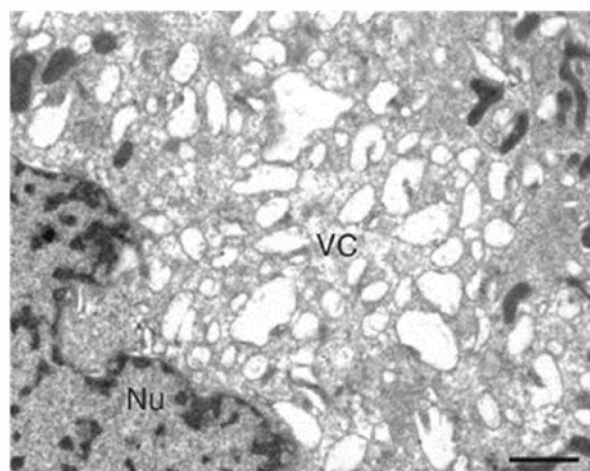
Once RNA replication is under way (Section 14.4.4), significant amounts of (+) RNA are produced; some of this RNA functions as mRNA and large quantities of virus proteins are then synthesized. Because the RNA is large (7.8 kb) there may be up to 60 ribosomes translating from each mRNA.

14.4.4 Transcription/genome replication

RNA synthesis occurs in replication complexes that contain cell proteins, as well as virus proteins and RNA. Antibodies have been used in electron microscopy techniques to determine the cellular locations of these complexes, and it has been shown that they are associated with the membranes of vesicles in the cytoplasm. Large numbers of vesicles develop in infected cells, and it has been found that the lipid composition of the vesicle membranes differs from that of the other cell membranes. All class IV viruses of eukaryotes replicate their RNA in complexes associated with membranes (Figure 14.11).

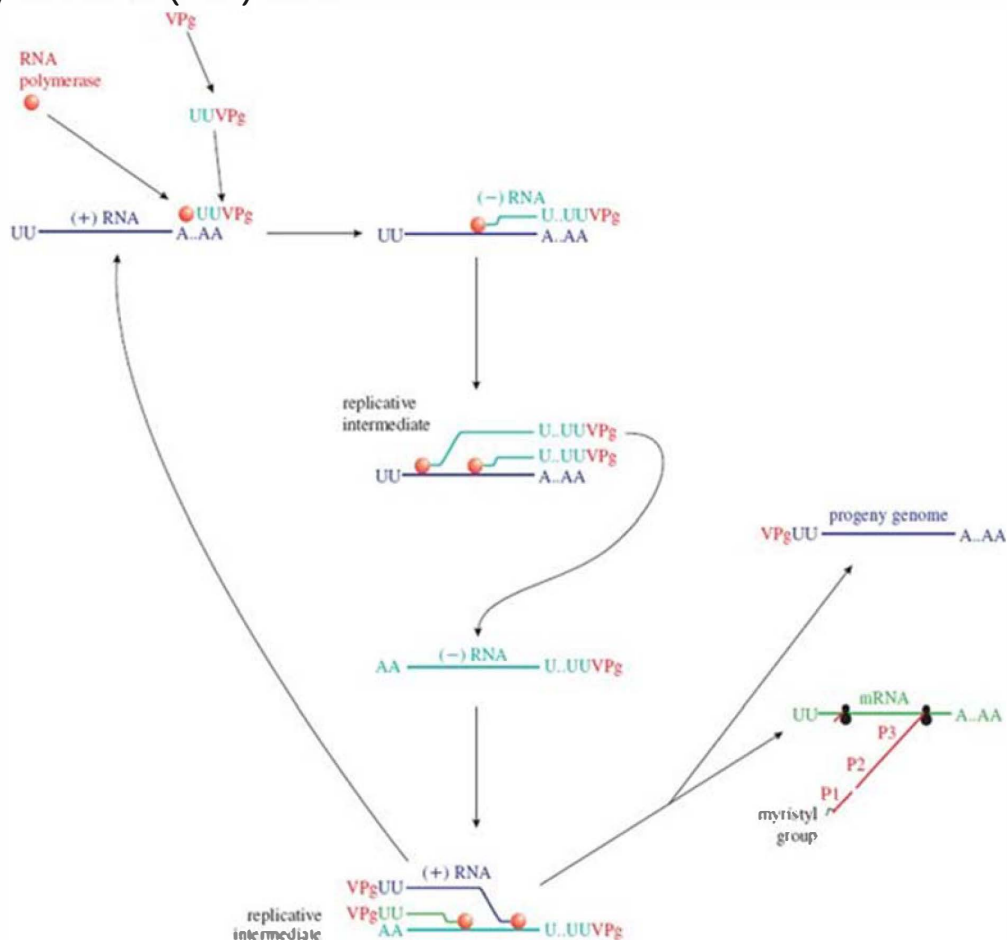
Figure 14.11 Parechovirus-induced vesicles in the cytoplasm. The electron micrograph shows a negatively stained section of a cell 6 hours after infection. VC: vesicle cluster, Nu: nucleus. Scale bar: 1 μ m.

Source: Mackenzie (2005) *Traffic*, 6, 967. Reproduced by permission of John Wiley & Sons.



In order for the infecting (+) RNA to be replicated, multiple copies of (–) RNA must be transcribed and then used as templates for (+) RNA synthesis (Figure 14.12).

Figure 14.12 Picornavirus transcription and genome replication. The genome acts as a template for synthesis of (–) RNA and VPg-UU functions as a primer to initiate synthesis. (–) RNA acts as the template for (+) RNA synthesis, some of which functions as mRNA, some as progeny genomes, and some as templates for further synthesis of (–) RNA.



There is evidence that a uridylylated derivative of the small protein VPg is involved in priming the synthesis of both (–) and (+) RNA. VPg-UU is formed by uridylylation at the hydroxyl group of a tyrosine residue in VPg. This takes place at a stem-loop in the virus genome and is catalyzed by the poliovirus RNA polymerase. Much of the RNA in the replication complexes is present as replicative intermediates (RIs). An RI is an association of RNA molecules consisting of a template RNA and several growing RNAs of varying length. Thus some RIs consist of a (+) RNA associated with growing (–) RNAs, while other RIs consist of a (–) RNA associated with growing (+) RNAs (Figure 14.12). Also associated with RIs are copies of the RNA polymerase.

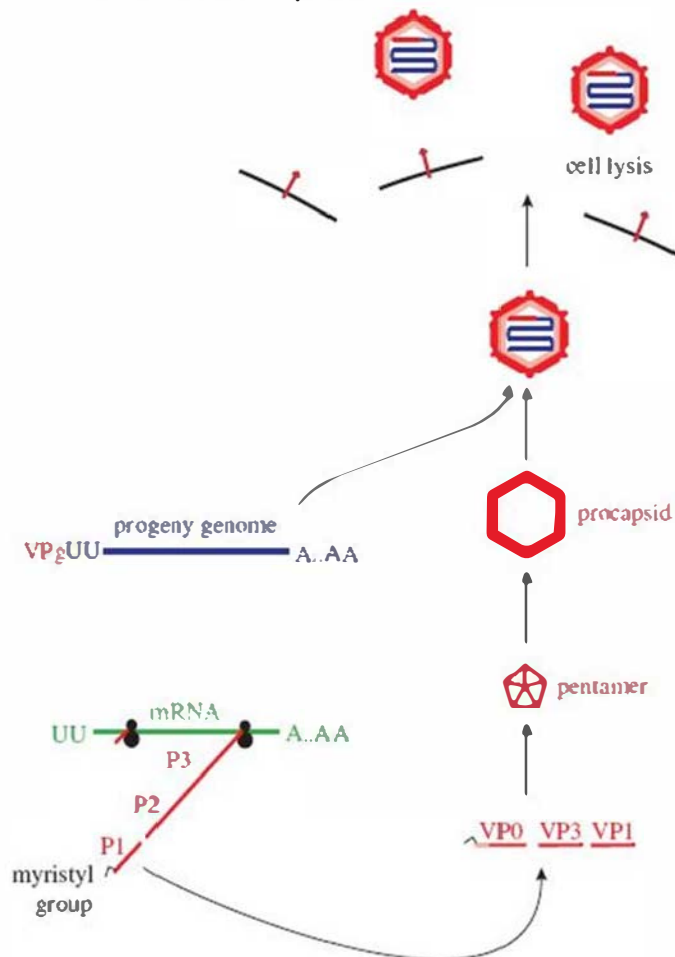
The poly(A) sequence at the 3' end of the plus strand is not synthesized in the usual way. It is transcribed from a poly(U) sequence at the 5' end of the minus strand.

It has been estimated that in a poliovirus-infected cell up to 2500 RNA molecules are synthesized per minute. The virus controls RNA synthesis in such a way that about 50 times more (+) RNA than (−) RNA is made. Some of the (+) RNA molecules are used as templates for further (−) RNA synthesis, some function as mRNA and some are destined to be the genomes of progeny virions (Figure 14.12).

14.4.5 Assembly and exit

Five copies each of VP0, VP3, and VP1 assemble into a pentamer and 12 pentamers form a procapsid (Figure 14.13). Each procapsid acquires a copy of the virus genome, with VPg still attached at the 5' end. At around this time the 60 copies of VP0 are cleaved into VP4 and VP2.

Figure 14.13 Picornavirus assembly and exit. The cleavage products of protein P1 assemble in groups of five to form pentamers. A procapsid is constructed from 12 pentamers, then a copy of the genome enters the procapsid to form a virion. Progeny virions leave the cell when it lyses.



Infected cell cultures develop cytopathic effects; for example, Vero cells (a monkey cell line) infected with poliovirus become rounded. Lysis of the cell releases the virions; approximately 105 virions are produced in each poliovirus-infected cell.

14.4.6 Inhibition of host gene expression

Soon after infection, the expression of host cell genes is inhibited. In poliovirus-infected cells it has been demonstrated that all three host RNA

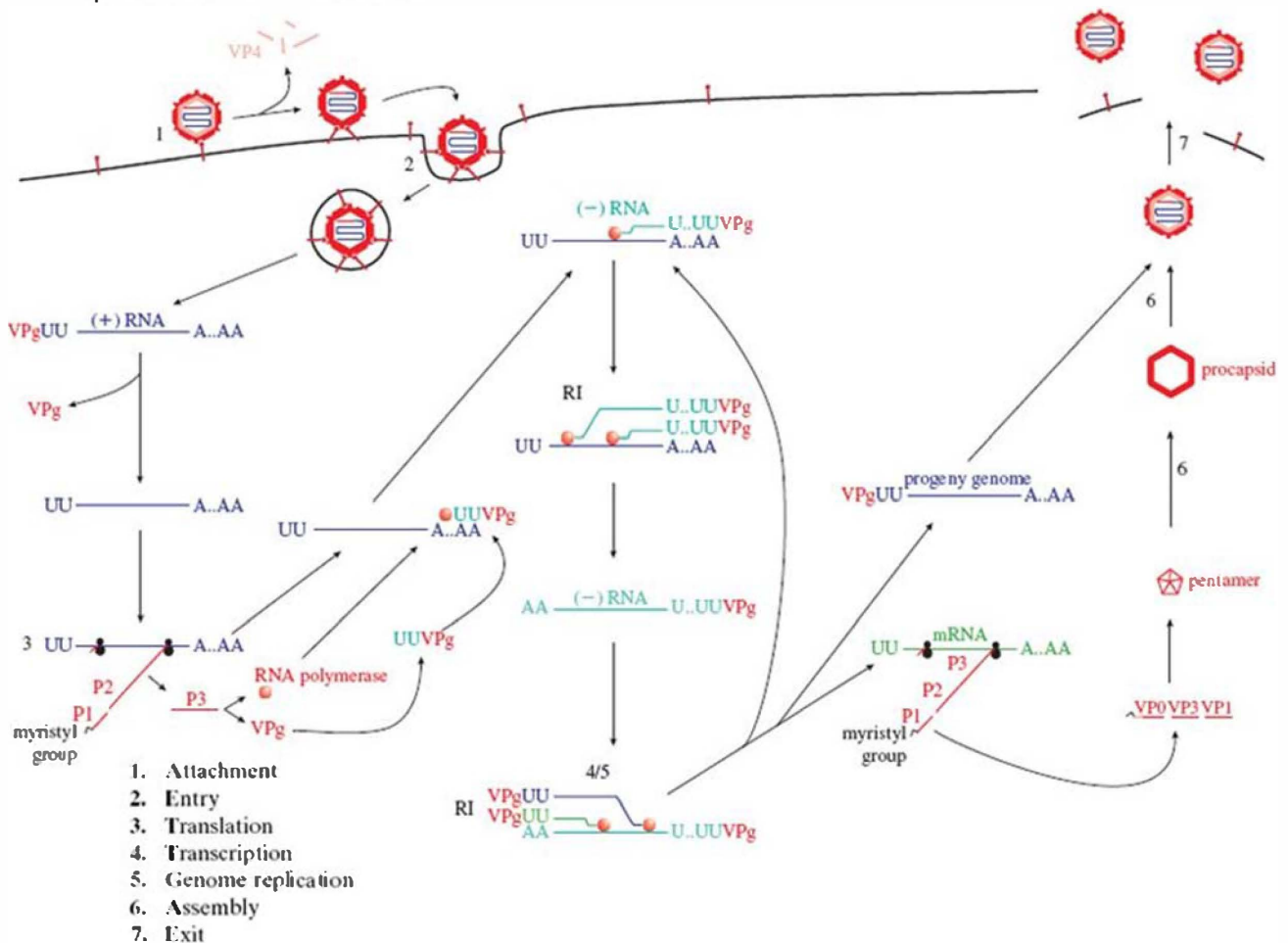
polymerases are rapidly inhibited, thus terminating transcription of cell genes. Translation from pre-existing cell mRNAs is inhibited as a result of modifications to translation initiation factors. These changes do not affect virus protein synthesis, as it is initiated by a cap-independent mechanism. Some of the shut down of host transcription and translation results from cleavage of cell proteins by the viral proteases (Figure 14.10).

14.4.7 Overview of picornavirus replication

A model for picornavirus replication is outlined in Figure 14.14. The cell nucleus is not shown, as all stages of virus replication take place in the cytoplasm. It has been demonstrated that poliovirus can replicate in enucleated cells, indicating that the virus has no requirement for the nucleus. It is almost certain, however, that some virus proteins enter the nucleus and are involved in the inhibition of cell transcription (Section 14.4.6).

Figure 14.14 The picornavirus replication cycle. Note that translation gets under way before transcription, and that transcription and genome replication involve a single process (synthesis of (+) RNA).

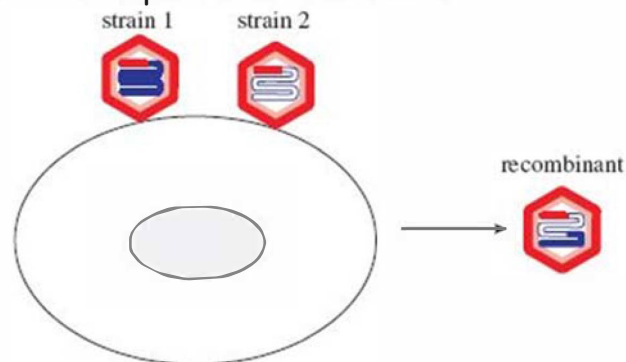
RI: replicative intermediate



14.5 PICORNAVIRUS RECOMBINATION

If a cell becomes infected with two strains of a picornavirus, then recombinant viruses may be present amongst the progeny from that cell (Figure 14.15). A recombinant virus is one that has part of its genome derived from one virus and part derived from another virus. A possible mechanism whereby recombination takes place for a ssRNA virus is described in Section 21.3.3.b.

Figure 14.15 Formation of a picornavirus recombinant. A cell is infected with two strains of a virus. Amongst the progeny virions there may be some that have part of the genome of strain 1 and part that of strain 2.



The following points should be noted for class IV viruses in general.

- Translation occurs before transcription.
- In eukaryotic cells RNA replication takes place on membranes.
- The processes of transcription and genome replication are one and the same, resulting in the production of (+) RNA.
- (+) RNA has three functions:
 - templates for (−) RNA synthesis;
 - mRNA;
 - genomes of progeny virions.

There must be mechanisms to ensure that a (+) RNA molecule engaged in one of these activities is precluded from commencing either of the other two.

The oral polio vaccine contains attenuated strains of all three poliovirus serotypes. Administration of this vaccine can lead to co-infections of gut cells with two serotypes, and recombinants between serotypes may be formed. If a wild type poliovirus is also present in the gut then recombination may occur between the wild type and one of the vaccine strains. Furthermore, new viruses have resulted from recombination between polio vaccine strains and picornaviruses other than poliovirus. There are documented cases where all these types of recombinant have caused polio.

14.6 PICORNAVIRUS EXPERIMENTAL SYSTEMS

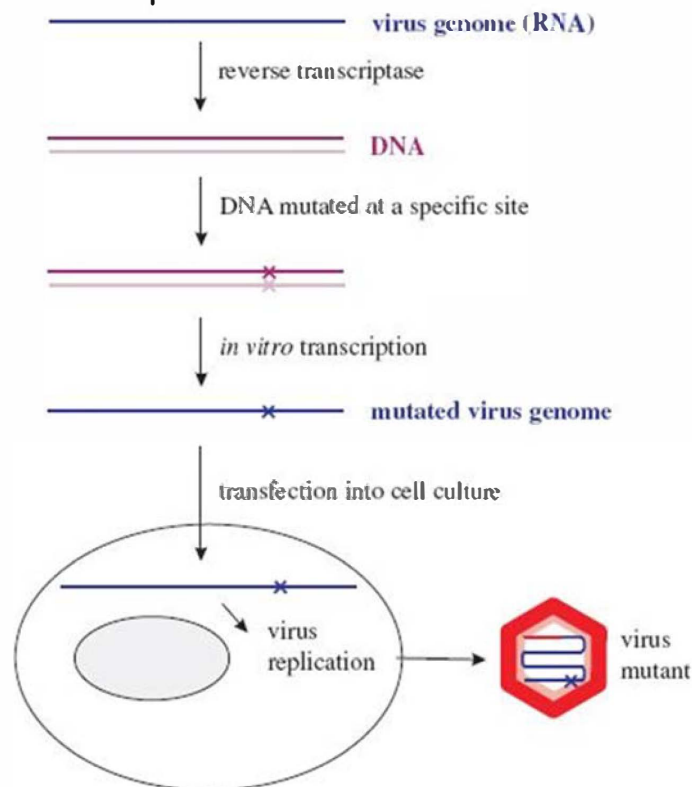
14.6.1 Reverse genetics

This section introduces some techniques that have been developed during work with picornaviruses. Reverse genetics techniques that permit manipulation of RNA genomes are discussed first, followed by techniques that enable the synthesis of infectious virus particles in a cell-free system.

14.6.1 Reverse genetics

Techniques that have been developed for manipulating DNA genomes can be applied to RNA genomes if the RNA is reverse transcribed to DNA. A defined mutation can be introduced into a gene in the DNA to investigate the function of the gene. After mutation the DNA is transcribed back to RNA, which is then transfected into cells in culture. The mutation may block replication of the virus at a specific step, providing evidence that the gene that has been mutated plays a role in that step. If replication proceeds to its conclusion, then the virus produced is a mutant (Figure 14.16).

Figure 14.16 Production of a picornavirus mutant by reverse genetics. Virus RNA is reverse transcribed to DNA, which is genetically manipulated and cloned in bacterial cells. The DNA is transcribed to RNA, which is transfected into cells. A full cycle of virus replication leads to the production of the virus mutant.



In one application of this technique with poliovirus, it was demonstrated that myristylation of VP0 is essential for virion assembly to proceed. A genome was produced in which the glycine residue at the N terminus of VP0 was mutated to

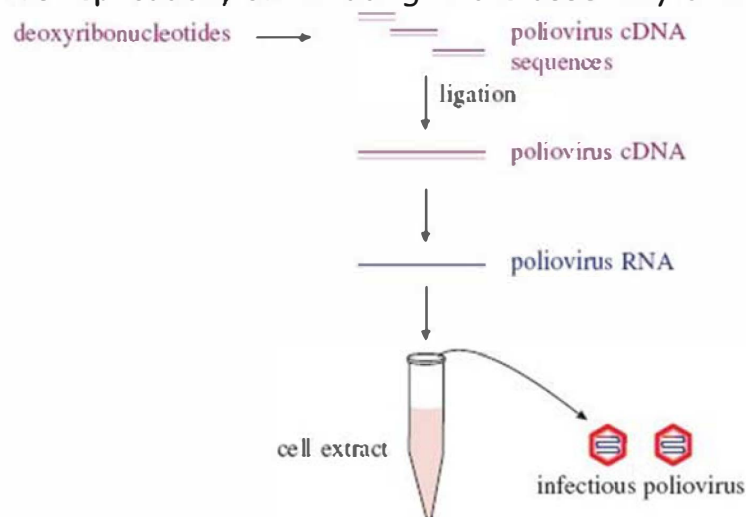
alanine; VP0 of the mutant could not be myristylated and virion assembly was inhibited.

14.6.2 Cell-free synthesis of infectious virus

In 1991 a group of workers recovered infectious poliovirus from a HeLa cell extract to which they had added purified poliovirus RNA. This was the first report of the production of infectious virus outside cells. Cell-free systems based on cytoplasmic extracts have subsequently been used to investigate features of the poliovirus replication cycle that are impossible to dissect using cell-based systems.

In 2002 it was reported that infectious poliovirus had been produced in a process that started with the synthesis of a copy DNA (cDNA) of the virus genome (Figure 14.17).

Figure 14.17 From nucleotides to infectious poliovirus. Short poliovirus cDNA sequences are synthesized and ligated to form a DNA copy of the genome from which poliovirus RNA is transcribed. The RNA is introduced into a cell extract, where it directs virus replication, culminating in the assembly of infectious virus.



The process started with the synthesis of discrete sequences of DNA, which together spanned the sequence of a poliovirus genome, then the sequences were ligated to form a cDNA of the virus genome. RNAs (virus genomes) were transcribed from the cDNA and inoculated into a cell extract; infectious poliovirus was subsequently found in the cell extract. This was the first time that infectious virus had been produced from molecular building blocks in the absence of cells.

14.7 OTHER PLUS-STRAND RNA VIRUSES

The genomes of plus-strand RNA viruses are translated in the infected cell prior to their transcription. Because the ribosomes in eukaryotic cells usually terminate after translation of the first ORF, the viruses need strategies to ensure that all the proteins encoded in the infecting genome are translated. One strategy is that described above for the picornaviruses.

- Polyprotein. All the genetic information is encoded in one ORF; this is translated to produce a polyprotein which is cleaved to produce the individual virus proteins.

Two other strategies are used by plus-strand RNA viruses.

- Subgenomic mRNAs. The genome has two or more ORFs. In many cases the RNA-dependent RNA polymerase is encoded by the ORF at the 5' end of the genome so that it can be translated from the infecting genome; the other ORF(s) are transcribed into subgenomic mRNA(s) that have the same 3' end as the genome.
- Segmented genome. There is one ORF in each RNA segment.

Some viruses use a combination of two of these strategies. Some examples of class IV viruses and their coding strategies are given in Table 14.3.

Table 14.3 Examples of plus-strand viruses and their coding strategies

	Family	Virus example(s)	Coding strategy
Animal viruses	<i>Flaviviridae</i>	West Nile virus (Section 22.3.1)	Polyprotein
		Hepatitis C virus (Section 23.7)	
	<i>Coronaviridae</i>	Severe acute respiratory syndrome (SARS) virus (Section 22.5.1)	Subgenomic mRNAs
Plant viruses	<i>Togaviridae</i>	Rubella virus	Polyprotein plus subgenomic mRNA
		Chikungunya virus (Section 22.3.2)	
	<i>Polyviridae</i>	Potato virus Y	Polyprotein
	<i>Alphaflexiviridae</i>	Potato virus X	Subgenomic mRNAs
	<i>Secoviridae</i>	Cowpea mosaic virus	Polyprotein plus segmented genome

The only ssRNA phages that are known are classified in the family Leviviridae (Section 20.2). Viruses in this family, such as phage MS2, have genomes with characteristics of bacterial mRNA; the genomes are polycistronic, with all of the ORFs translated by internal initiation (Section 6.6).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- give examples of picornaviruses and explain their importance;
- describe the picornavirus virion;
- describe the picornavirus replication cycle;
- discuss picornavirus recombination;
- describe experimental systems used for picornavirus studies.

SOURCES OF FURTHER INFORMATION

Book

Chapters 24-27 in Fields Virology, 5th edition (Knipe, D. M. and Howley, P. M., editors-in-chief), Lippincott, Williams and Wilkins

Historical papers

Dulbecco, R. and Vogt, M. (1954) Plaque formation and isolation of pure lines with poliomyelitis viruses. *Journal of Experimental Medicine*, 99, 167-182

Enders, J. F., Weller, T. H., and Robbins, F. C. (1949) Cultivation of the Lansing strain of poliomyelitis virus in cultures of various human embryonic tissues. *Science*, 109, 85-87

Recent papers

Arden, K. E. and Mackay, I. M. (2010) Newly identified human rhinoviruses: molecular methods heat up the cold viruses. *Reviews In Medical Virology*, 20, 156-176

Bostina, M. et al. (2011) Poliovirus RNA is released from the capsid near a twofold symmetry axis. *Journal of Virology*, 85, 776-783

den Boon, J. A. and Ahlquist, P. (2010) Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annual Review of Microbiology*, 64, 241-256

Fuchs, R. and Blaas, D. (2010) Uncoating of human rhinoviruses. *Reviews In Medical Virology*, 20, 281-297

Gavala, M. L., Bertics, P. J., and Gern, J. E. (2011) Rhinoviruses, allergic inflammation, and asthma. *Immunological Reviews*, 242, 69-90

Lin, J.-Y. et al. (2009) Viral and host proteins involved in picornavirus life cycle. *Journal of Biomedical Science*, 16, 103

Nagy, P. D. and Pogany, J. (2012) The dependence of viral RNA replication on co-opted host factors. *Nature Reviews Microbiology*, 10, 137-149

Ogram, S. A. and Flanagan, J. B. (2011) Non-template functions of viral RNA in picornavirus replication. *Current Opinion in Virology*, 1, 339-346

Racaniello, V. R. (2006) One hundred years of poliovirus pathogenesis. *Virology*, 344, 9-16

Sztuba-Solin

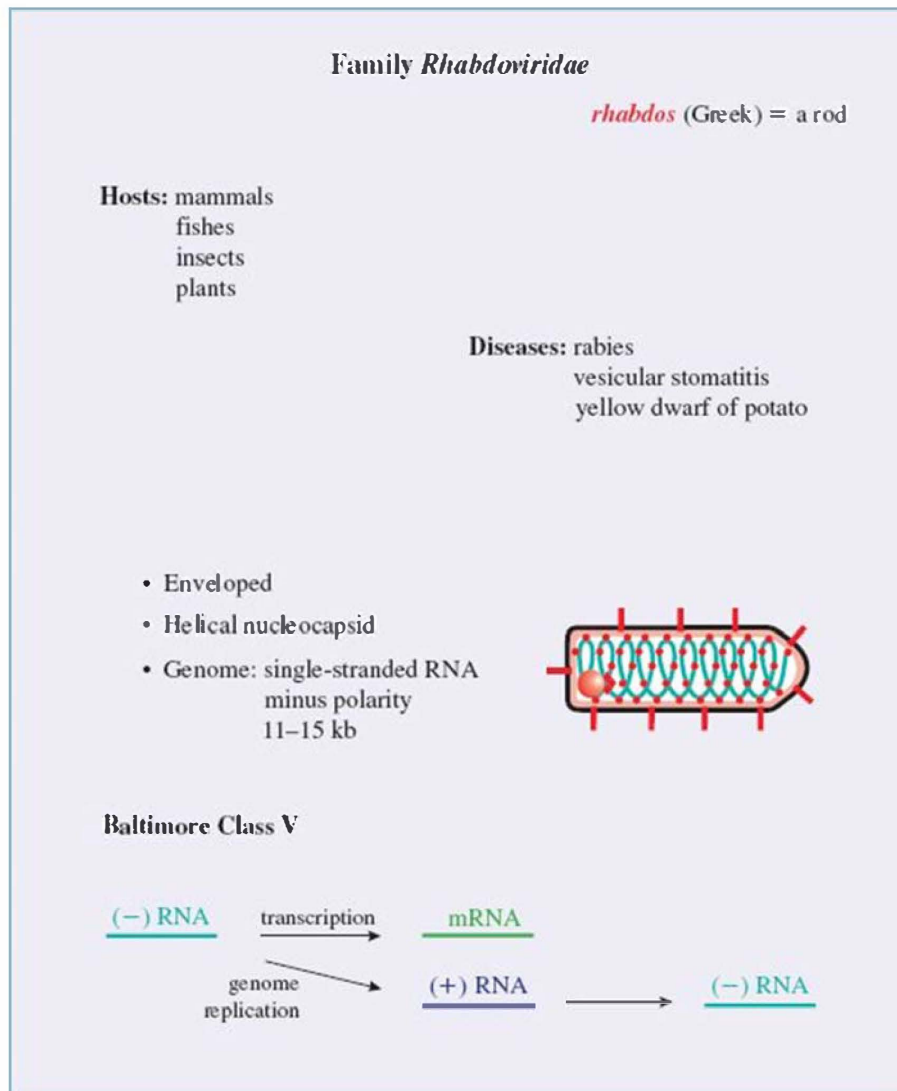
ska, J., Stollar, V., and Bujarski, J. J. (2011) Subgenomic messenger RNAs: mastering regulation of (+)-strand RNA virus life cycle. *Virology*, 412, 245-255

Wimmer, E. (2006) The test-tube synthesis of a chemical called poliovirus: the simple synthesis of a virus has far-reaching societal implications. EMBO Reports, 7 (SI), S3B5-S9

CHAPTER 15

Rhabdoviruses (and Other Minus-Strand RNA Viruses)

CHAPTER 15 AT A GLANCE

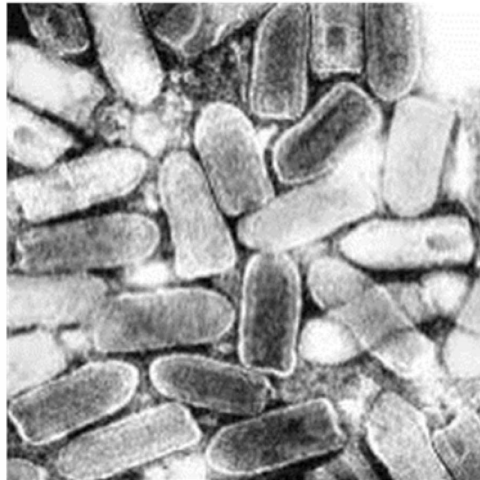


15.1 INTRODUCTION TO RHABDOVIRUSES

The rhabdoviruses have minus-strand RNA genomes in the size range 11–15 kb. The name of these viruses is derived from the Greek *rhabdos*, which means a rod. The virions of some rhabdoviruses, especially those infecting plants, are in the shape of rods with rounded ends, while others, especially those infecting animals, are bullet shaped (Figure 15.1).

Figure 15.1 Negatively stained virions of vesicular stomatitis virus.

Source: Courtesy of Prof. Frederick A. Murphy, The University of Texas Medical Branch.



Rhabdoviruses are found in a wide range of hosts, including mammals, fish, plants, and insects, and many are important pathogens of animals and plants. The rhabdoviruses constitute the family Rhabdoviridae, which contains a number of genera, some of which are listed in Table 15.1.

Table 15.1 Examples of rhabdoviruses

Genus	Name derivation	Hosts	Example
<i>Vesiculovirus</i>	Vesicle = blister	Mammals, fish	Vesicular stomatitis virus
<i>Lyssavirus</i>	<i>Lyssa</i> (Greek) = rage, fury, canine madness	Mammals	Rabies virus
<i>Novirhabdovirus</i>	A non- <i>virion</i> protein is encoded.	Fish	Infectious hematopoietic necrosis virus
<i>Nucleorhabdovirus</i>	Replication cycle includes a nuclear phase.	Plants, insects	Potato yellow dwarf virus

Many rhabdoviruses have very wide host ranges and replicate in diverse types of host, especially the so-called “plant” rhabdoviruses, which replicate in their insect vectors as well as in their plant hosts (Chapter 4).

Before looking at rhabdovirus structure and replication, we consider two important rhabdoviruses, rabies virus and vesicular stomatitis virus (VSV).

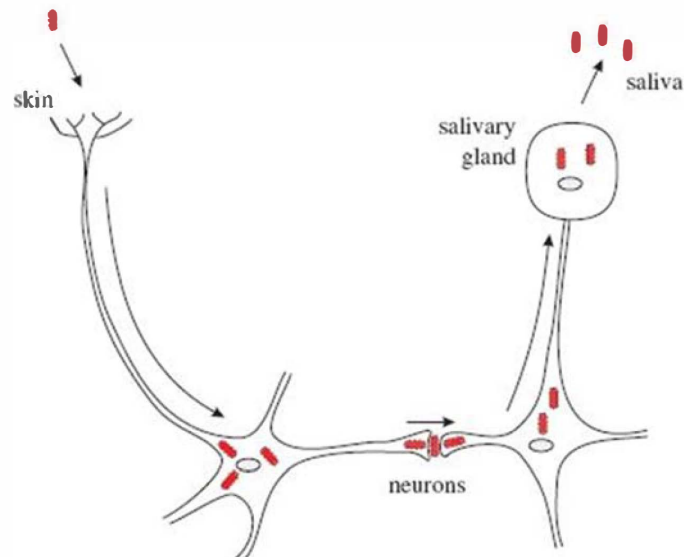
15.2 SOME IMPORTANT RHABDOVIRUSES

15.2.1 Rabies virus

Rabies virus, like many rhabdoviruses, has an exceptionally wide host range. In the wild it has been found infecting many mammalian species, while in the laboratory it has been found that birds can be infected, as well as cell cultures from mammals, birds, reptiles, and insects.

Infection with rabies virus normally occurs as a result of virus in saliva gaining access to neurons through damaged skin. The infection spreads to other neurons in the central nervous system, then to cells in the salivary glands, where infectious virus is shed into the saliva (Figure 15.2).

Figure 15.2 Rabies virus infection of the animal body. After entering the body through damaged skin, a virion infects a neuron via the nerve endings and is transported to the cell body, where virus replication takes place. The infection spreads to other neurons and to salivary gland cells, which shed virions into the saliva.

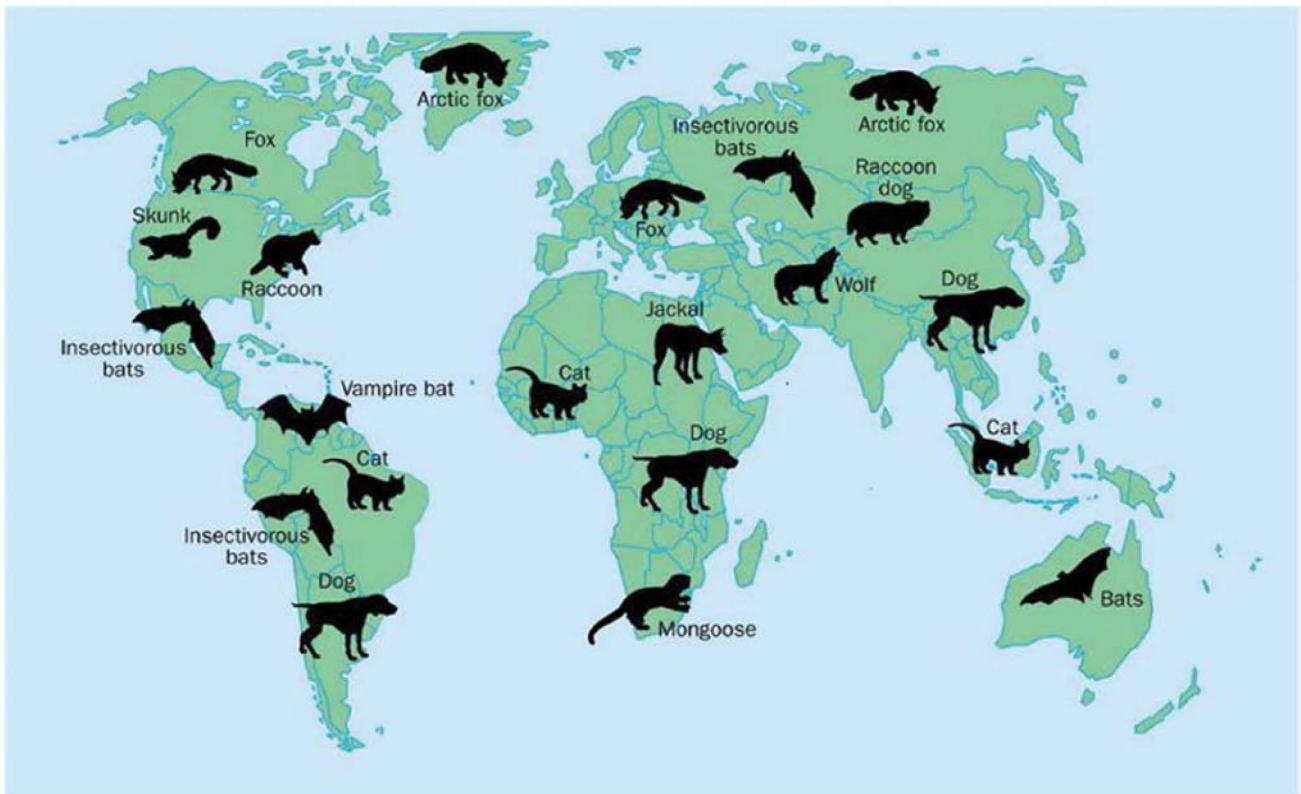


Each year rabies kills large numbers of humans, dogs, cattle, and other animal species; precise numbers are not known, but for humans it is estimated that rabies causes more than 55 000 deaths annually. Most rabies infections of humans are acquired via bites from rabid dogs, though a few people have become infected after receiving an organ transplant from a rabies-infected individual. Rabies in humans is nearly always fatal, having a higher mortality than any other infectious disease.

Rabies is endemic in wild animals in many parts of the world, often one animal species serving as the major reservoir (Figure 15.3). In Western Europe the major reservoir is the red fox.

Figure 15.3 Rabies virus reservoirs.

Source: Rupprecht et al. (2002) *The Lancet Infectious Diseases*, 2, 327. Reproduced by permission of Elsevier Limited and the authors.



Vaccines have been developed to provide protection to humans (e.g. veterinary surgeons), domestic animals (especially dogs), and wild animals (e.g. foxes) at risk from rabies virus infection. Rabies vaccines have been incorporated into food baits (Figure 15.4) attractive to wild mammals, and dropped from aircraft over fox-inhabited regions in Europe and coyote- and raccoon-inhabited regions in the US. The first vaccine to be used was an attenuated vaccine, but more recent vaccines have contained a recombinant vaccinia virus that expresses the rabies virus G protein. Vaccination of wild mammals has been very successful in bringing rabies under control in a number of countries.

Figure 15.4 Wild mammal bait containing rabies vaccine.

Source: Courtesy of Michael Roiland, Pinellas County, Florida, US.



Rabies is normally absent from the UK. In the past, this status was maintained through the requirement for a quarantine period for certain animal species, including dogs, on entry to the country. That policy has been largely replaced with

a “pet passport scheme,” which involves giving rabies vaccine to animals prior to entry, and implanting an identifying microchip in each vaccinated animal.

Many viruses related to rabies virus have been found in bats around the world, and have been classified in the genus *Lyssavirus* along with the original rabies strains. Infection in the bat can have a number of possible outcomes: there may be subclinical infection, or there may be disease from which the bat may or may not recover. There are occasional cases of human rabies resulting from bites from infected bats. One such victim was David McRae, a licensed bat handler in Scotland, who died in 2002 after being bitten by an insectivorous bat.

Vesicle = blister

Stomatitis = inflammation of mucous membrane in the mouth

15.2.2 Vesicular stomatitis virus

VSV causes disease in a variety of animals, including cattle, horses, sheep, and pigs, affected animals developing lesions on the feet and in the mouth similar to those in foot and mouth disease (Section 14.2.5). The disease can result in significant economic damage due to decreased milk and meat production, and the imposition of quarantines and trade barriers. Vesicular stomatitis is endemic in the tropics and there are epidemics in some temperate areas, but it has never been found in the UK.

VSV has a very wide host range. As well as infecting domestic livestock, there is evidence of infection in wild animals including bats, deer, and monkeys. This evidence is the presence in these animals of neutralizing antibodies to the virus. VSV has been isolated from a number of insect species, including mosquitoes, sand flies, and black flies. Its natural cycle is unknown, but it is possible that it is transmitted between mammals by one or more of these types of insect.

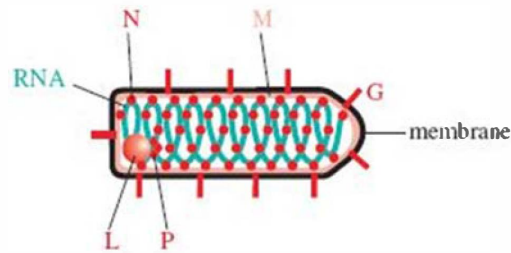
In the laboratory, VSV can replicate in cell cultures derived from mammals, birds, fish, insects, and nematode worms. Much of our understanding of rhabdovirus structure and replication comes from studies with VSV, which is much safer than rabies virus to work with. Three species of VSV are recognized.

15.3 THE RHABDOVIRUS VIRION AND GENOME ORGANIZATION

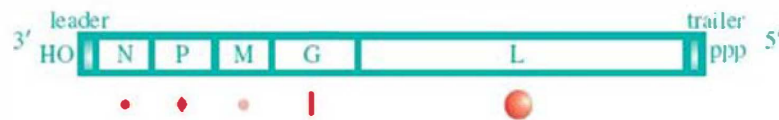
The rhabdovirus virion is an enveloped, rod- or bullet-shaped structure containing five protein species (Figures 15.1 and 15.5).

Figure 15.5 Rhabdovirus virion and genome organization. Between each of the genes for the structural proteins is a short intergenic sequence. At the ends of the genome are non-coding sequences: a “leader” at the 3’ end and a “trailer” at the 5’ end.

Virion components



Genome organization and gene products



The nucleoprotein (N) coats the RNA at the rate of one protein molecule to nine nucleotides, forming a nucleocapsid with helical symmetry. Associated with this ribonucleoprotein are the minor virion proteins P (phosphoprotein) and L (large). The L protein is well named, its gene taking up about half of the genome (Figure 15.5). Its large size points to the fact that it is a multifunctional protein, as will be described later. The helical ribonucleoprotein is coated with a helical layer of M (matrix) protein, forming a layer between the nucleocapsid and the envelope. Trimers of G (glycoprotein) form spikes that protrude from the envelope.

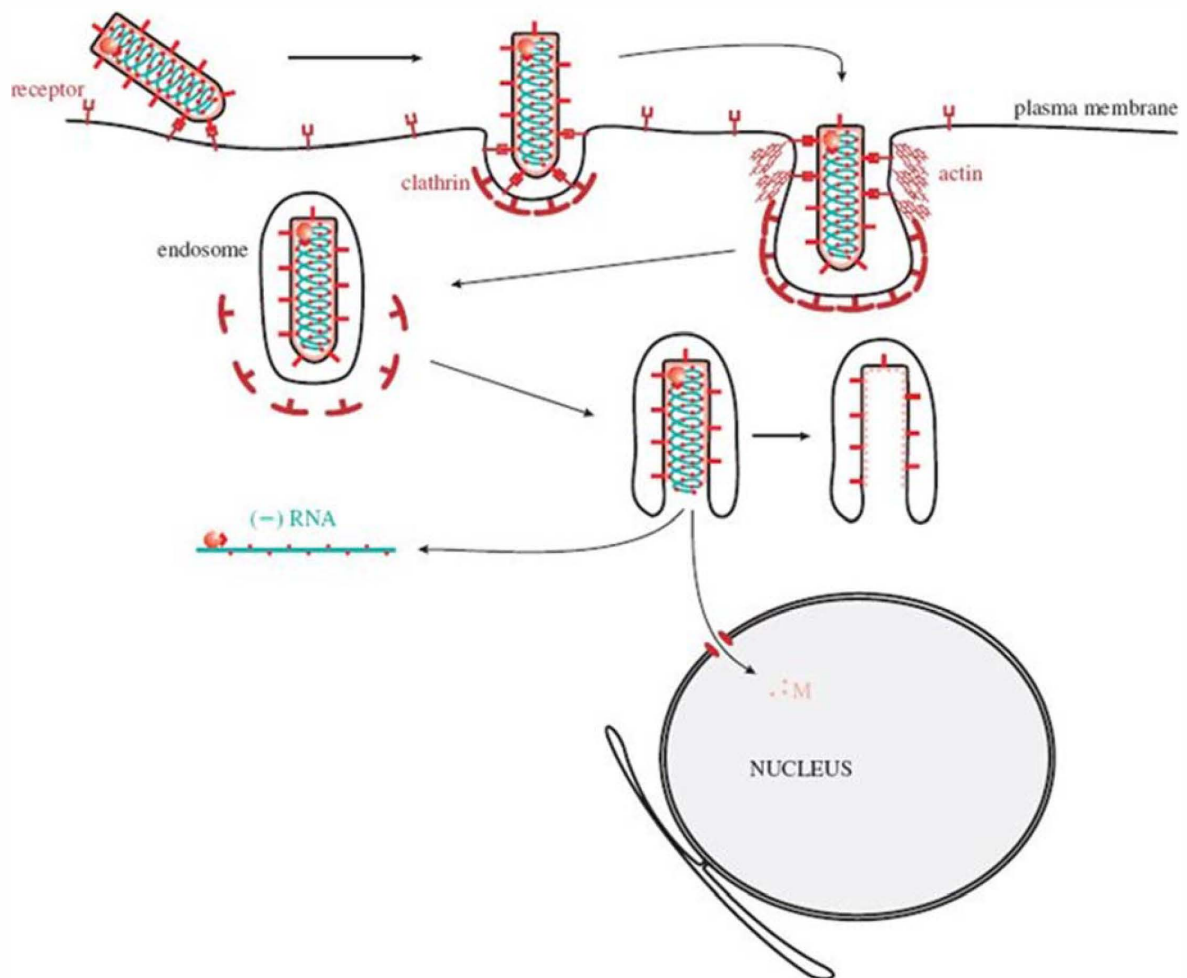
The genomes of all rhabdoviruses encode these five proteins, and many rhabdoviruses encode one or more additional proteins.

15.4 RHABDOVIRUS REPLICATION

15.4.1 Attachment and entry

A rhabdovirus virion attaches to receptors at the cell surface and is then taken into the cell by endocytosis (Figure 15.6). At first, the invagination of the plasma membrane forms a clathrin-coated pit, but as the pit deepens the coating of the plasma membrane is formed from the protein actin, rather than clathrin. The G protein spikes of the virus envelope are involved in attachment to cell receptors and in fusion of the virion and endosome membranes. This fusion takes place after G molecules have undergone a conformational change induced by a drop in pH within the endosome. The nucleocapsid is released into the cytoplasm after the membranes of the virion and the endosome have fused. Most M protein remains associated with the endosome, but some is released and transported to the nucleus, where it acts to shut down expression of cell genes (Section 15.4.6).

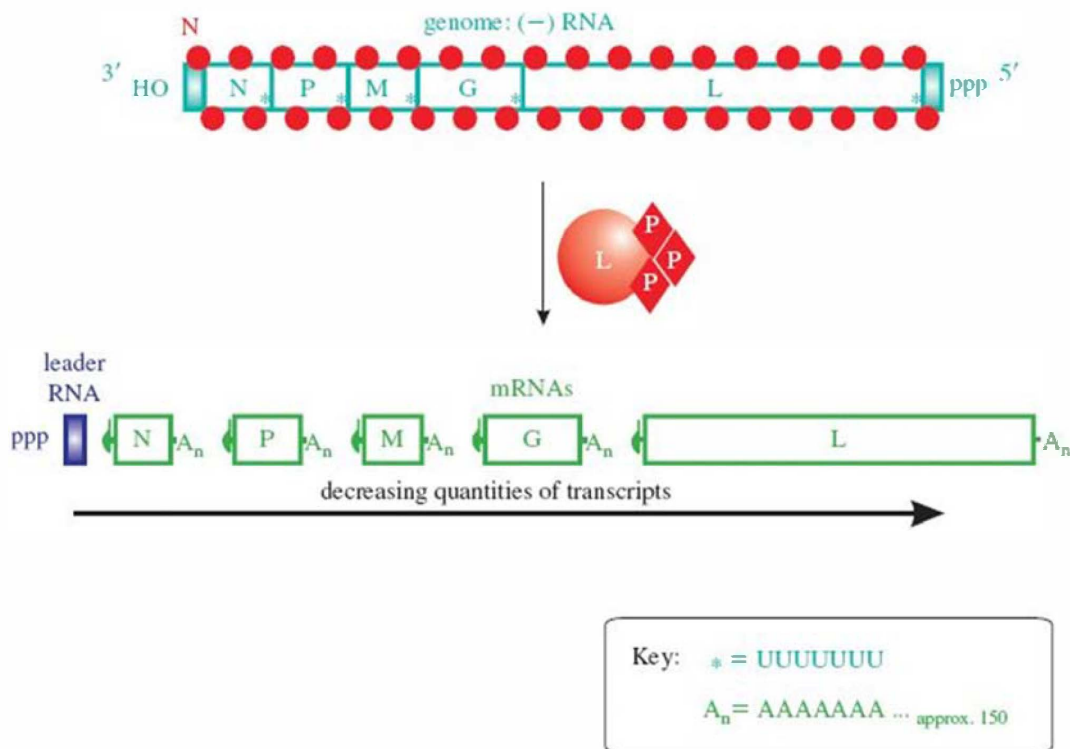
Figure 15.6 Attachment and entry of a rhabdovirus. After endocytosis the membranes of the virion and the endosome fuse, releasing the nucleocapsid into the cytoplasm. Some M protein is transported to the nucleus.



15.4.2 Transcription

Once the RNA and its associated proteins (N, P, and L) are free in the cytoplasm, transcription of the virus genome can begin (Figure 15.7). A plus-strand leader RNA, the function of which is uncertain, and five mRNAs are synthesized.

Figure 15.7 Rhabdovirus transcription. The transcriptase is a complex of L protein with three copies of P protein. The minus-strand genome is transcribed to six plus-strand RNAs: a leader RNA and five mRNAs. Each mRNA is capped at the 5' end, and at the 3' end has about 150 adenylates transcribed by multiple copying from UUUUUUU at the end of each gene.



Transcription is carried out by an RNA-dependent RNA polymerase activity that resides, along with four other enzyme activities, in the L protein (Table 15.2). The polymerase is active only when L is complexed with P protein in the ratio 1L:3P. The requirement for P was demonstrated in experiments in which the P and L proteins of VSV were purified. The individual purified proteins were found to be lacking in polymerase activity, which was restored when the two proteins were mixed.

Table 15.2 Enzyme activities associated with the rhabdovirus L protein

Enzyme	Role
RNA-dependent RNA polymerase	RNA replication
Polyribonucleotidyl transferase	Capping mRNAs
Methyl transferase	Methylation of mRNA caps
Poly(A) polymerase	Polyadenylation of mRNAs
Kinase	Phosphorylation of P protein precursor

Also associated with the L protein are enzyme activities that cap and polyadenylate the mRNAs (Table 15.2). The virus supplies these activities, as the cell enzymes are present only in the nucleus. Capping and polyadenylation of the virus mRNAs proceed by mechanisms different to those carried out by the cell, though the end results are the same: each mRNA is capped at the 5' end and polyadenylated at the 3' end.

As the L-P complex moves along the template RNA from the 3' end to the 5' end a newly synthesized RNA molecule is released at each intergenic sequence. The first

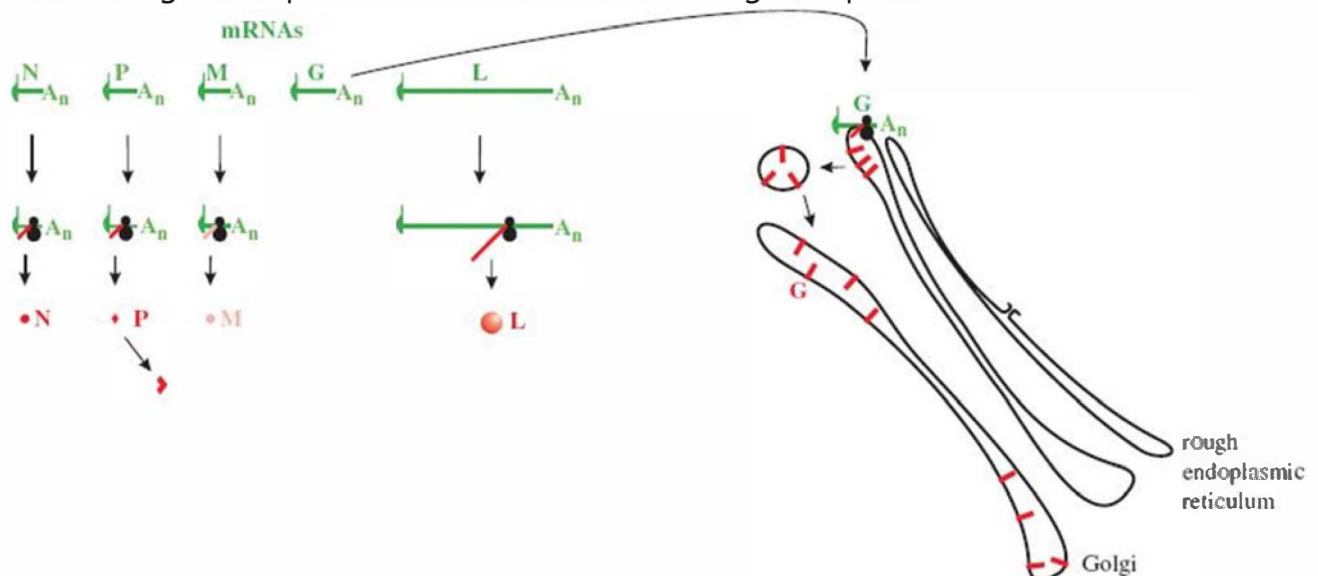
RNA synthesized is the leader RNA and the remainder are mRNAs. Before release, each mRNA is polyadenylated by the poly(A) polymerase activity of L. It is thought that the L-P complex “stutters” at the 5’ end of each gene, where the sequence UUUUUUU is transcribed as about 150 adenylates (Figure 15.7). The enzyme resumes transcription when it recognizes the start of the next gene.

The virus does not need equal amounts of all the gene products; for example, it needs many copies of N protein to coat new RNA, but relatively few copies of L protein. It controls the expression of its genes by controlling the relative quantities of transcripts synthesized. As the enzyme complexes move along the (-) RNA approximately 30% of them detach at the end of each gene, so that fewer and fewer remain associated with the template as they progress toward the 5’ end. Thus, many copies of N are transcribed, but relatively few copies of L (Figure 15.7).

15.4.3 Translation

The virus proteins are translated on free ribosomes, except for the G protein, which is translated in the rough endoplasmic reticulum (Figure 15.8).

Figure 15.8 Rhabdovirus translation and post-translational modifications of proteins. Trimers of P protein are formed after phosphorylation. The G protein is glycosylated in the rough endoplasmic reticulum and the Golgi complex.



As their names imply, the phosphoprotein (P) and the glycoprotein (G) undergo co- and post-translational modifications. One-sixth of the residues in VSV P protein are serine and threonine, and many of these are phosphorylated. The phosphorylation takes place in two steps, the first performed by a cell kinase and the second by the kinase activity of the L protein. After phosphorylation, trimers of P are formed. During translation of the G protein in the rough endoplasmic reticulum oligosaccharides are added to asparagine residues, and further processing of the oligosaccharides takes place in the Golgi complex.

15.4.4 Genome replication and secondary transcription

Once protein synthesis is under way accumulations of N, P, and L proteins form inclusions in which virus RNA synthesis takes place. The minus-strand virus genome is replicated via the synthesis of complementary (+) RNA molecules. A rhabdovirus-infected cell synthesizes about 4 to 10 times more copies of (-) RNA than genome-length (+) RNA. Some copies of the (-) RNA act as templates for the synthesis of new copies of (+) RNA, some are used as templates for further transcription (secondary transcription; Figure 15.9), while some become the genomes of progeny virions (Figure 15.10).

Figure 15.9 Rhabdovirus genome replication and secondary transcription. The (-) RNA genome is the template for genome-length (+) RNA synthesis, which in turn is the template for further (-) RNA synthesis. (-) RNAs serve as templates for further (+) RNA synthesis and for secondary transcription. (-) RNAs and genome-length (+) RNAs become coated with N protein shortly after synthesis.

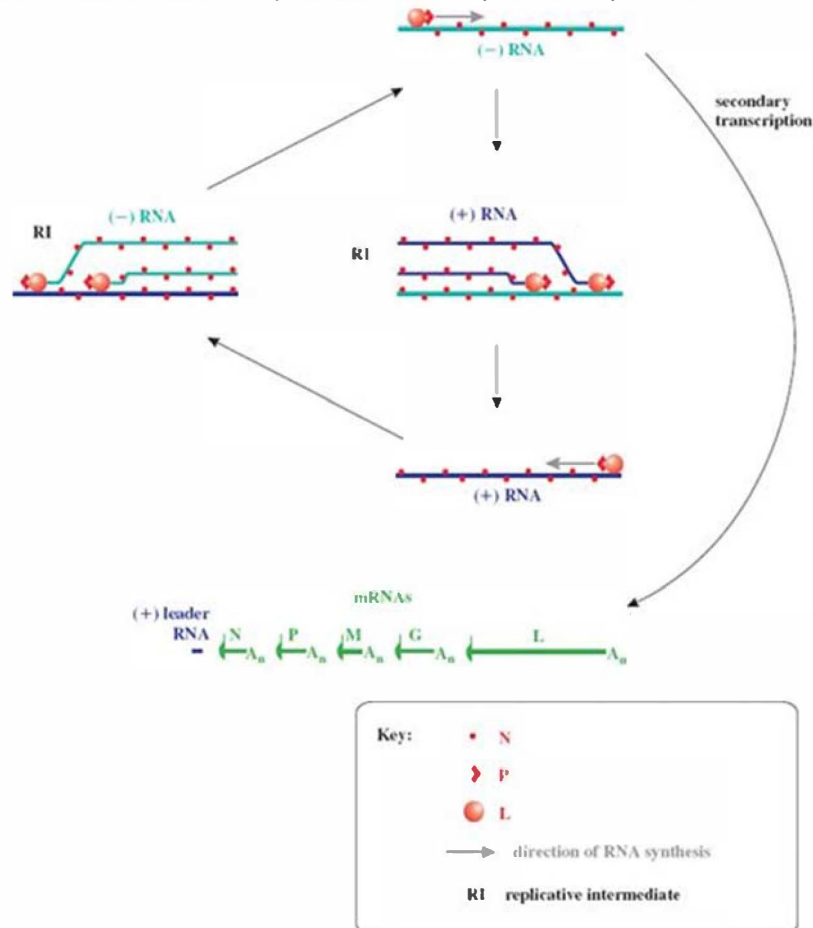
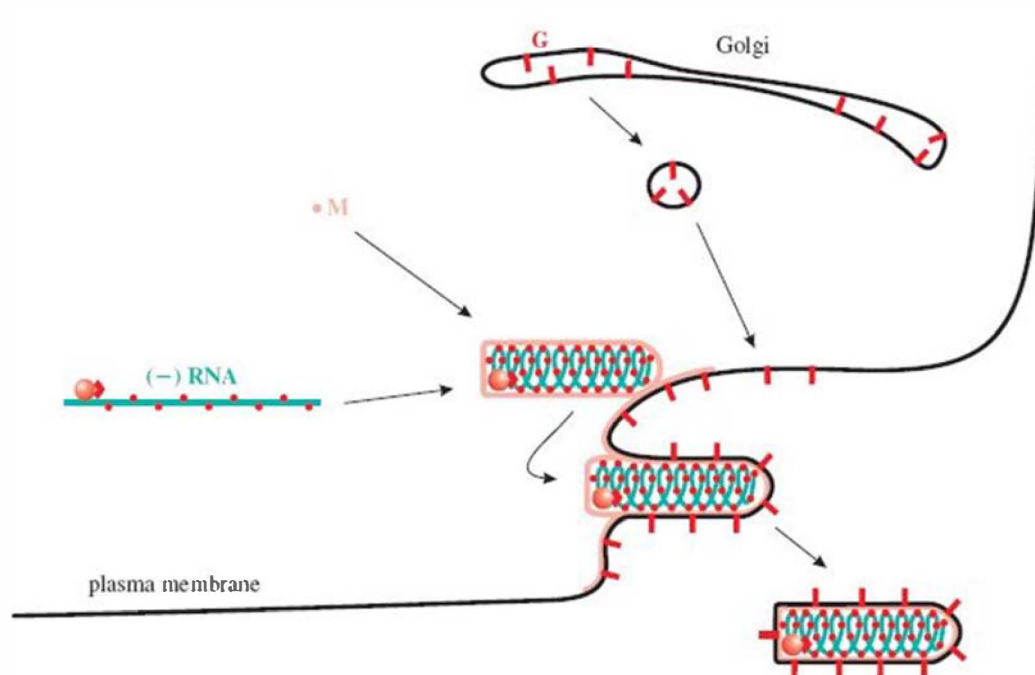


Figure 15.10 Rhabdovirus assembly and exit. M protein coats nucleocapsids, which then bud from regions of the plasma membrane that have been modified by the insertion of G protein.



The initiation of RNA synthesis does not require a primer. Replicative intermediates can be detected in infected cells, as with the plus-strand RNA viruses (Chapter 14). We noted earlier that the leader RNA and the mRNAs are produced as a result of the RNA polymerase recognizing termination signals in the (-) RNA template (Figure 15.7). During genome replication, however, the termination signals are ignored as the enzyme must remain associated with the template to produce genome-length (+) RNA. Another difference between the two processes is that during genome replication the newly synthesized (+) RNA quickly becomes coated with N protein, whereas the mRNAs are not coated. The genome and the genome-length (+) RNA are never present in the cell as naked molecules, but are always associated with N protein, which protects them from ribonucleases. This is true for all minus-strand RNA viruses.

The differences between the processes that result in synthesis of mRNAs and genome-length (+) RNA from the same template are not understood. One suggestion is that there may be differences in the components of the enzyme complexes involved in the two processes, one complex acting as a "transcriptase" and the other acting as a "replicase."

15.4.5 Assembly of virions and exit from the cell

It was noted above that both minus-strands and plus-strands of genome-length RNA are coated with N protein. Only coated minus-strands, however, are selected to form virions, because of the presence of a packaging signal at the 5' end of the minus-strand.

The M protein plays several important roles in the assembly process. It condenses the nucleocapsid into a tightly coiled helix and it links the nucleocapsid with a region of the plasma membrane into which copies of the G protein have been inserted (Figures 8.3(a) and 15.10). Virions bud from these regions of the plasma

membrane, acquiring their envelopes in the process (Figure 15.11). The M protein has a late (L) domain that binds cell proteins involved in the budding process (Section 8.3.1).

Figure 15.11 Electron micrograph of vesicular stomatitis virions budding from the plasma membrane.

Source: Brown et al. (2010) *Viruses*, 2, 995, with permission of the authors and the publisher.



15.4.6 Inhibition of host gene expression

Rhabdovirus infection of a cell results in strong inhibition of host gene expression. The M protein, whose important roles in virion assembly have just been described, appears to play major roles in this inhibition. There is evidence that the M protein inhibits transcription by all three host RNA polymerases, that it blocks intracellular transport of cell RNAs and proteins, and that it inhibits translation of cell proteins. One effect of these activities in animal cells is that the synthesis of interferon (Section 9.2.1.b) is inhibited.

15.4.7 Role of the nucleus

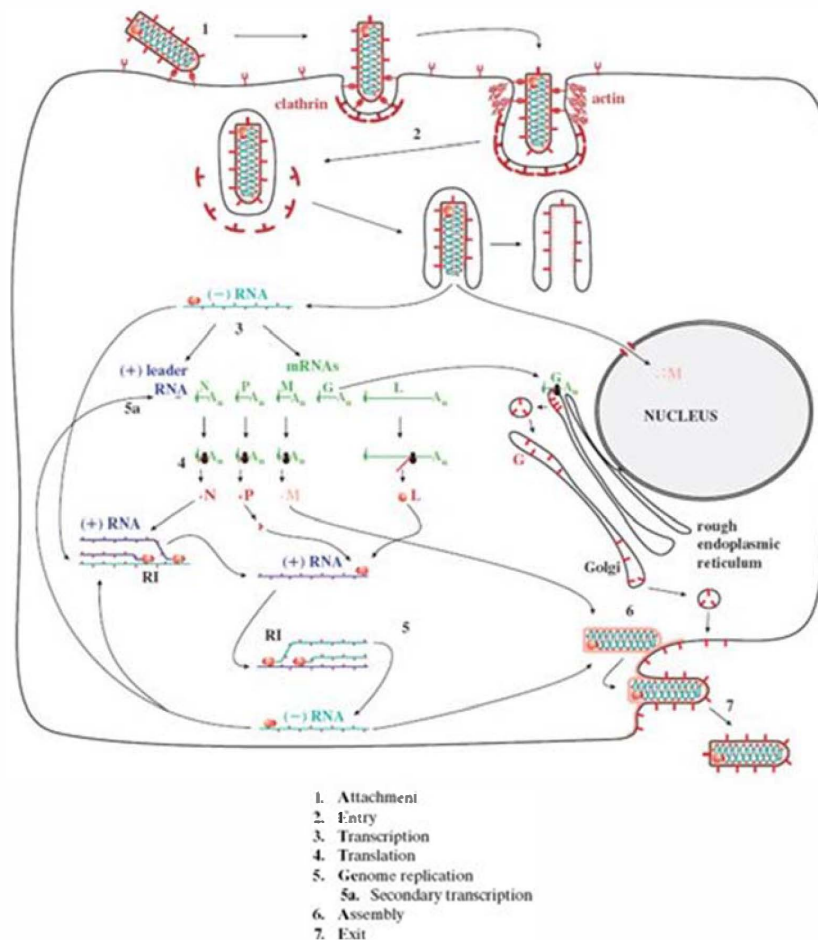
There appears to be no significant role for the nucleus in the replication of most rhabdoviruses. Enucleated cells support VSV replication, with only a small drop in yield compared with normal cells. This is not true for all rhabdoviruses, however, as members of the genus *Nucleorhabdovirus* replicate in the nucleus of their plant and insect hosts, with virions budding through the inner nuclear membrane.

15.4.8 Overview of rhabdovirus replication

The rhabdovirus replication cycle is summarized in Figure 15.12.

Figure 15.12 The rhabdovirus replication cycle.

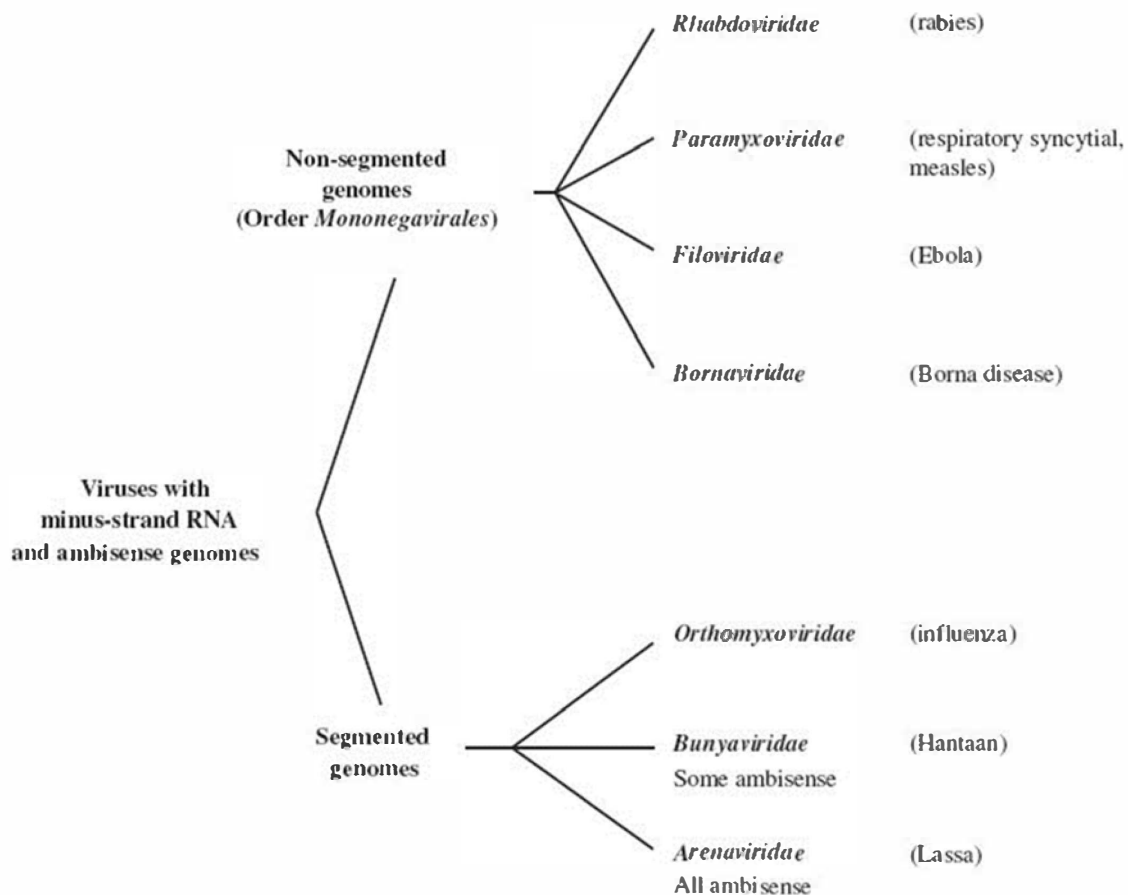
RI: replicative intermediate



15.5 OTHER MINUS-STRAND RNA VIRUSES AND VIRUSES WITH AMBISENSE GENOMES

Figure 15.13 shows the families of viruses with minus-strand RNA genomes. Many of these viruses have segmented genomes, some of which are ambisense, meaning that one or more of the RNA segments each encodes two genes, one in the plus sense and one in the minus sense.

Figure 15.13 Families of viruses with minus-strand RNA genomes and ambisense genomes. Examples of viruses in the families are shown in brackets.



Some rhabdoviruses and some bunyaviruses have plant hosts, but most minus-strand RNA viruses have animal hosts. Amongst them are three of the world's major human pathogens: influenza, measles, and respiratory syncytial viruses, together responsible for millions of cases of serious disease and approximately a million deaths each year.

Measles virus is one of the main causes of childhood death in developing countries and is still responsible for some deaths in industrialized countries (Section 22.7.1). Infection results in immunosuppression, which renders the host more susceptible to secondary infections with a range of bacterial and viral pathogens, and these cause most measles-associated deaths. Measles virus can also cause lethal infections of the central nervous system. The influenza viruses get a chapter to themselves (Chapter 16), while filoviruses and hantaviruses are discussed in Chapter 22.

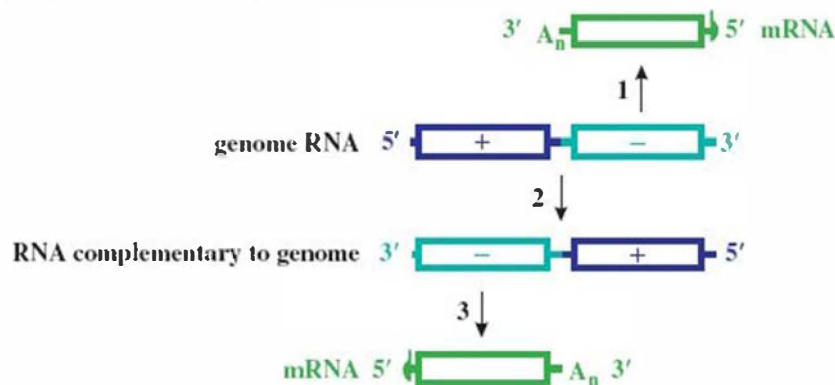
Viruses in the families Paramyxoviridae, Filoviridae, and Bornaviridae, like the Rhabdoviridae, have non-segmented genomes that are transcribed to mRNAs by their RNA-dependent RNA polymerases terminating and reinitiating on (+) RNA templates. These four families constitute the order Mononegavirales (viruses with monopartite genomes composed of negative sense RNA).

Viruses with ambisense genomes include those in the family Arenaviridae (two genome segments, both ambisense) and some members of the family Bunyaviridae, (one or two of the three genome segments are ambisense).

The minus sense gene of an ambisense RNA is expressed by transcription of a mRNA. The plus sense gene is expressed by synthesis of an RNA complementary to the genome, followed by transcription of the mRNA for that gene (Figure 15.14).

Figure 15.14 Expression of genes encoded in an ambisense RNA. The genome RNA encodes one gene in the plus sense and one in the minus sense.

1. The gene encoded in the minus sense is transcribed into mRNA.
2. The genome is transcribed into a complementary RNA.
3. The other gene is transcribed into mRNA.



15.6 REVERSE GENETICS

Reverse genetics have been used to investigate rhabdoviruses and other minus-strand RNA viruses, though the procedures are not as straightforward as with the plus-strand RNA viruses (Section 14.6.1). Minus-strand RNA is not infectious, so techniques had to be devised that not only generate virus genomes from cDNA, but also supply the RNA polymerase and the nucleoprotein that coats the newly synthesized RNA (Section 15.4.4). After much painstaking work, there are now procedures that enable the recovery of infectious minus-strand RNA virus from a cDNA. As with the plus-strand RNA viruses, a defined mutation can be introduced into a gene in the cDNA to investigate its function in the replication cycle and its role, if any, in virus virulence. Reverse genetics are also being explored as a tool to engineer virus strains with reduced virulence for use in vaccines.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- discuss the importance of rabies virus, vesicular stomatitis virus, and other minus-strand RNA viruses;
- describe the rhabdovirus virion;
- outline the main characteristics of the rhabdovirus genome;
- discuss the replication cycle of rhabdoviruses;
- explain the term "ambisense genome";
- discuss the development of reverse genetics procedures for minus-strand RNA viruses.

SOURCES OF FURTHER INFORMATION

Books

Bleck, T. P. and Rupprecht, C. E. (2009) Rhabdoviruses. Chapter 40 in *Clinical Virology*, 3rd edition (Richman, D. D., Whitley, R. J., and Hayden, F. G., editors), ASM Press

Fu, Z. F. (ed.) (2005) *The World of Rhabdoviruses*, Springer

Jackson, A. C. (ed.) (2011) *Advances in Virus Research*, Volume 79: *Research Advances in Rabies*, Elsevier

Samal, S. K. (ed.) (2011) *The Biology of Paramyxoviruses*, Caister Academic Press

Journals

Albertini, A. A. V. et al. (2012) Molecular and cellular aspects of rhabdovirus entry. *Viruses*, 4,117–139

Galloway, S. E. and Wertz, G. W. (2009) A temperature sensitive VSV identifies L protein residues that affect transcription but not replication. *Virology*, 388, 286–293

Ge, P. et al. (2010) Cryo-EM model of the bullet-shaped vesicular stomatitis virus. *Science*, 327, 689–693

Ogino, T. and Banerjee, A. K. (2011) An unconventional pathway of mRNA cap formation by vesiculoviruses. *Virus Research*, 162,100–109

Rupprecht, C. E., Turmelle, A., and Kuzmin, I. V. (2011) A perspective on lyssavirus emergence and perpetuation. *Current Opinion in Virology*, 1, 662–670

CHAPTER 16

Influenza Virus

CHAPTER 16 AT A GLANCE

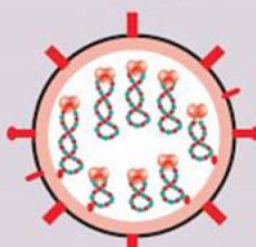
Family *Orthomyxoviridae*

Influenza Viruses

Hosts: birds
mammals

Influenza A Virus

← 80–120 nm →



- Enveloped
- Genome: single-stranded RNA
 - eight segments
 - minus polarity
 - approx. 13.5 kb
- Integral membrane proteins: hemagglutinin (HA)
 - neuraminidase (NA)
 - membrane protein 2 (M2) } drug targets
- Internal proteins: membrane protein 1 (M1)
 - nuclear export protein (NEP)
 - nucleoprotein (NP)
 - polymerase acidic (PA)
 - polymerase basic 1 (PB1)
 - polymerase basic 2 (PB2)
- Non-structural proteins: non-structural protein 1 (NS1)
 - PB1-F2
 - PB1-N40
- Baltimore Class V

16.1 INTRODUCTION TO INFLUENZA VIRUSES

We are devoting a chapter to the influenza viruses because they are responsible for a huge disease burden and for hundreds of thousands of deaths each year. Furthermore, they are very interesting viruses because their replication cycle differs in a number of respects from that of most minus-strand RNA viruses. In particular, they replicate their RNA in the nucleus, whereas most other minus-strand viruses (Chapter 15) carry out all stages of their replication cycle in the cytoplasm.

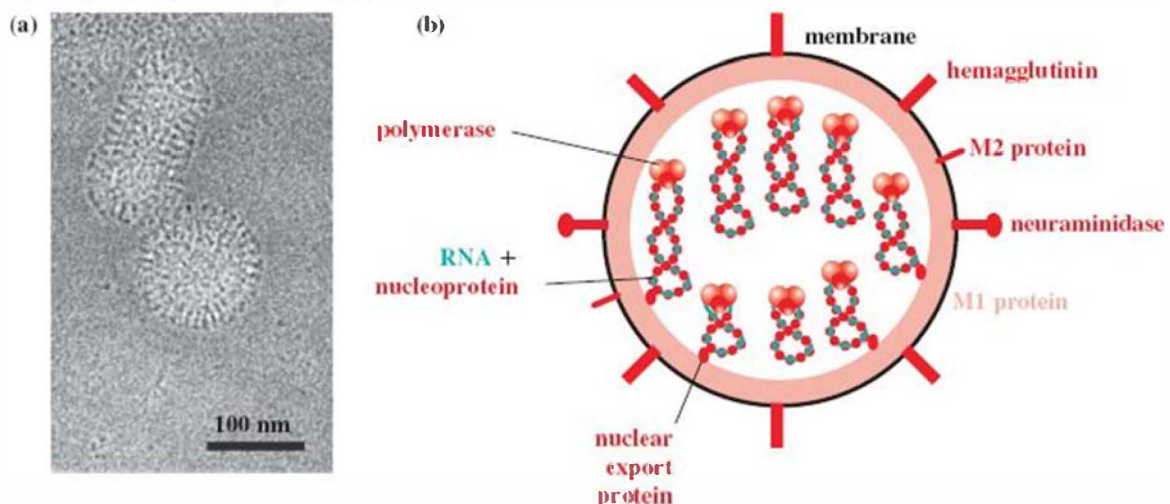
Influenza viruses are classified in the family Orthomyxoviridae. There are three genera (Influenza-virus A, B, and C); the most important of these, and hence the most studied, are the influenza A viruses.

16.2 THE INFLUENZA VIRION

Influenza viruses are enveloped viruses with segmented genomes of ssRNA. The virions are spherical or elongated (Figure 16.1(a)), with diameters 80–120 nm. Isolates that have been passaged many times in the laboratory usually have spherical virions, whereas those that have been passaged only a few times tend to be filamentous.

Figure 16.1 Influenza A virus structure. (a) Cryo-electron micrograph of spherical and elongated virions. (b) The virion components. The envelope contains three integral membrane proteins: hemagglutinin, neuraminidase, and M2 (membrane 2). Underlying the envelope is a layer of M1 (membrane 1) protein, within which is a nuclear export protein and the genome (eight segments of ssRNA coated with nucleoprotein and associated with three polymerase proteins).

Source: (a) Moulès et al. (2011) *Virology*, 414, 51. Reproduced by permission of Elsevier Limited and the authors.



Influenza A and B virions each contain eight segments of ssRNA, while the influenza C virion contains seven segments. There are nine species of virus protein in the influenza A virion, the name of each protein describing a role or property of that protein (Figure 16.1(b)).

Three of the protein species are trans-membrane proteins. In decreasing order of abundance they are:

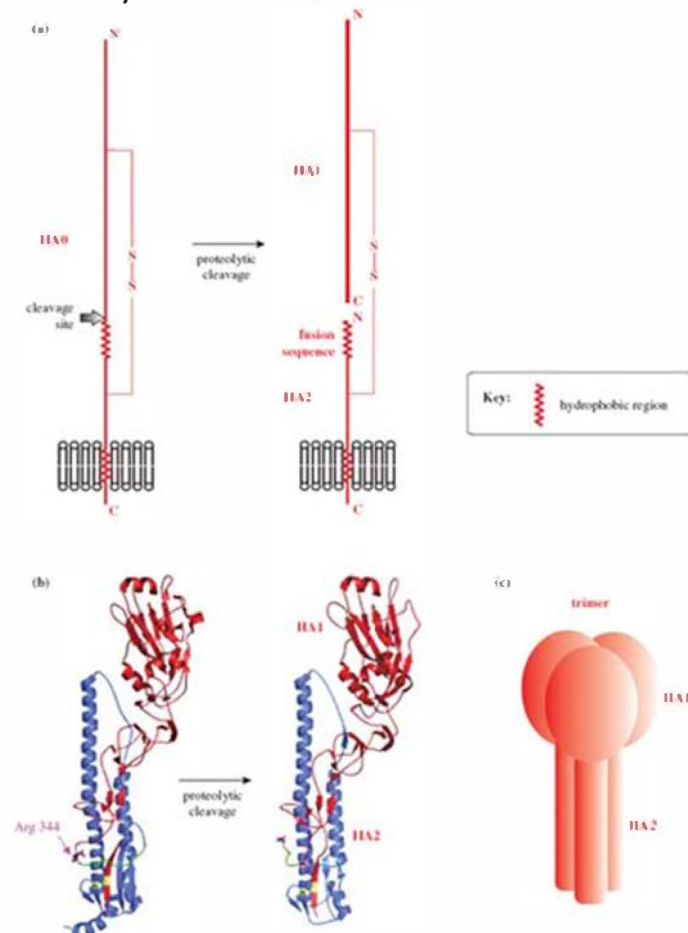
- hemagglutinin (can agglutinate red blood cells of various species);
- neuraminidase (an enzyme, the substrate for which is neuraminic acid);
- M2 (M for membrane; four M2 monomers form an ion channel).

All three trans-membrane proteins have undergone post-translational modifications: for example, the hemagglutinin and neuraminidase are glycosylated, and the hemagglutinin has been cleaved into two subunits, HA1 and HA2, linked by a

disulfide bond (Figure 16.2). HA1 forms a “head” which contains the attachment site for the host cell; HA2 forms a “stem” and is associated with the lipid membrane. Each hemagglutinin spike is a trimer (Figure 16.2 (c)), while each neuraminidase spike is a tetramer. Both the hemagglutinin and the neuraminidase occur as a variety of subtypes, and there is antigenic variation within each subtype. The result is that there are many strains of influenza A virus.

Figure 16.2 Influenza A virus hemagglutinin. (a) The main features of the hemagglutinin. The disulfide-linked subunits (HA1 and HA2) are formed by cleavage of a precursor protein (HA0). (b) Specific features of the 1918 influenza hemagglutinin. The colors in this diagram do not conform to the general color scheme of this book: HA1 (red), HA2 (blue), disulfide bond (yellow), fusion sequence (green), cleavage site of HA0 (Arg 344), C-terminus of HA1, and N-terminus of HA2 (purple). (c) Each hemagglutinin spike consists of a trimer of H1-H2.

Source: (b) Bertram et al. (2010) *Reviews in Medical Virology*, 20, 298. Reproduced by permission of John Wiley & Sons Ltd.



M1, another “membrane” protein, coats the interior surface of the envelope, and within this layer are the genome segments and a nuclear export protein. Each RNA segment is in a ribonucleoprotein (RNP) complex, with the RNA wrapped around nucleoprotein molecules. The 5′ and 3′ ends of each RNA segment are in close proximity and are associated with the RNA polymerase, which is a complex of three

proteins. One of the polymerase proteins is acidic (PA: polymerase acidic), while the other two are basic (PB1 and PB2). The sizes of the RNP complexes vary with the lengths of the RNA molecules within them. Cell proteins, such as tubulin and cyclophilin A, have been reported as components of influenza virus particles.

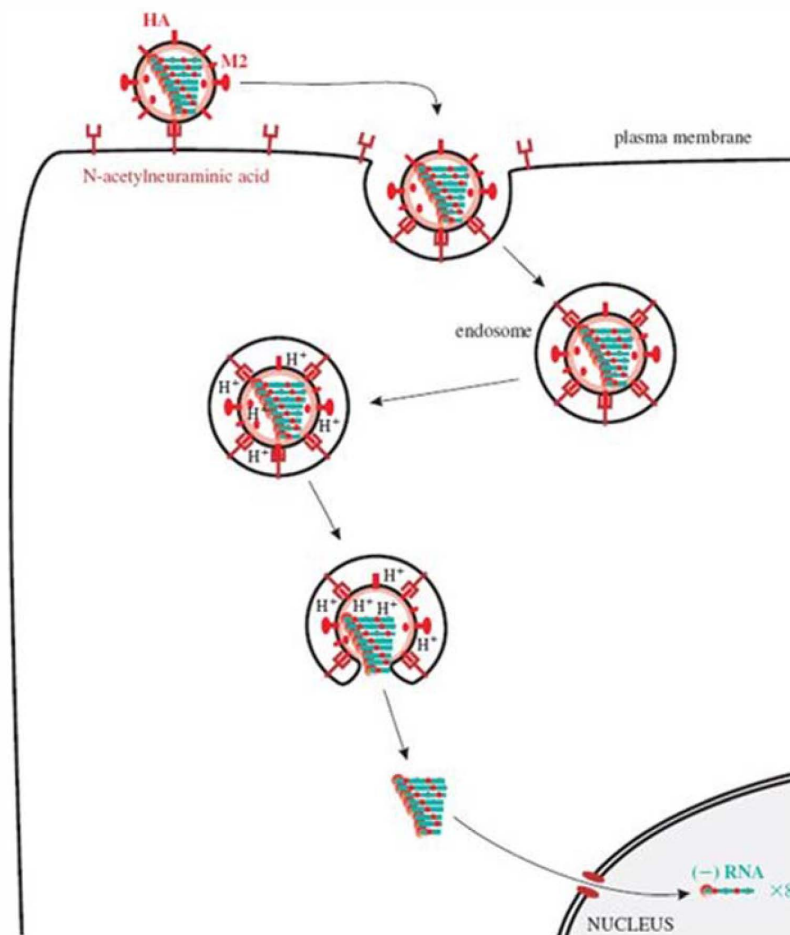
16.3 INFLUENZA A VIRUS REPLICATION

Tissues commonly infected by influenza viruses are respiratory epithelium in humans and other primates, and intestinal epithelium in birds.

16.3.1 Attachment and entry

The virion attaches via a site on HA1 to the cell receptor, which is N-acetylneuraminic acid (neuraminic acid is a nine-carbon monosaccharide). The virion is endocytosed (Figure 16.3) and in the endosome it experiences a fall in pH, which causes a conformational change in the M2 protein ion channels. Hydrogen ions are pumped into the virion, and the drop in pH causes the RNP complexes to dissociate from M1 protein. As the pH in the endosome falls further the hemagglutinin molecules change shape, exposing fusion sequences in their HA2 subunits. The virion membrane fuses with the endosome membrane, releasing the RNP complexes, which are transported to the nucleus where virus RNA synthesis takes place.

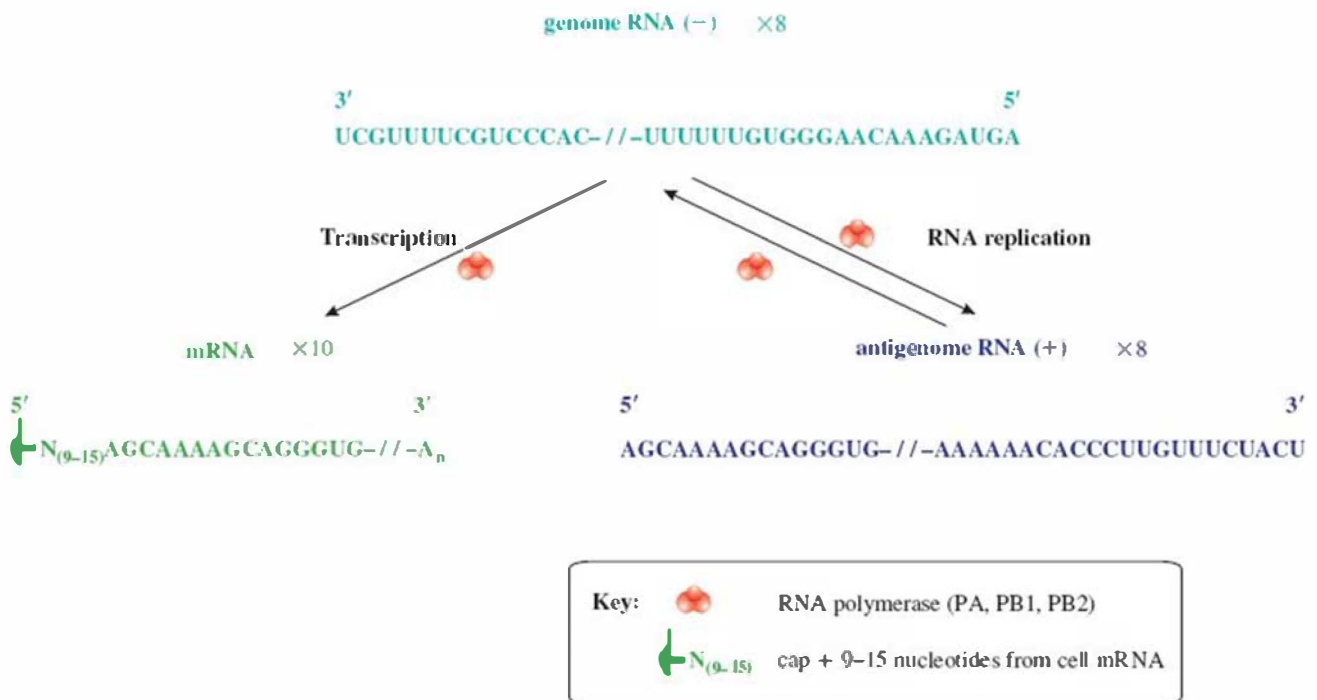
Figure 16.3 Attachment and entry of influenza A virus. After endocytosis the endosome pH falls, changing the shape of M2 and HA molecules. M2 channels pump H⁺ ions into the virion, releasing the RNP complexes from M1 protein; HA fuses the membranes of the virion and the endosome. The RNP complexes are released and transported to the nucleus.



16.3.2 RNA replication

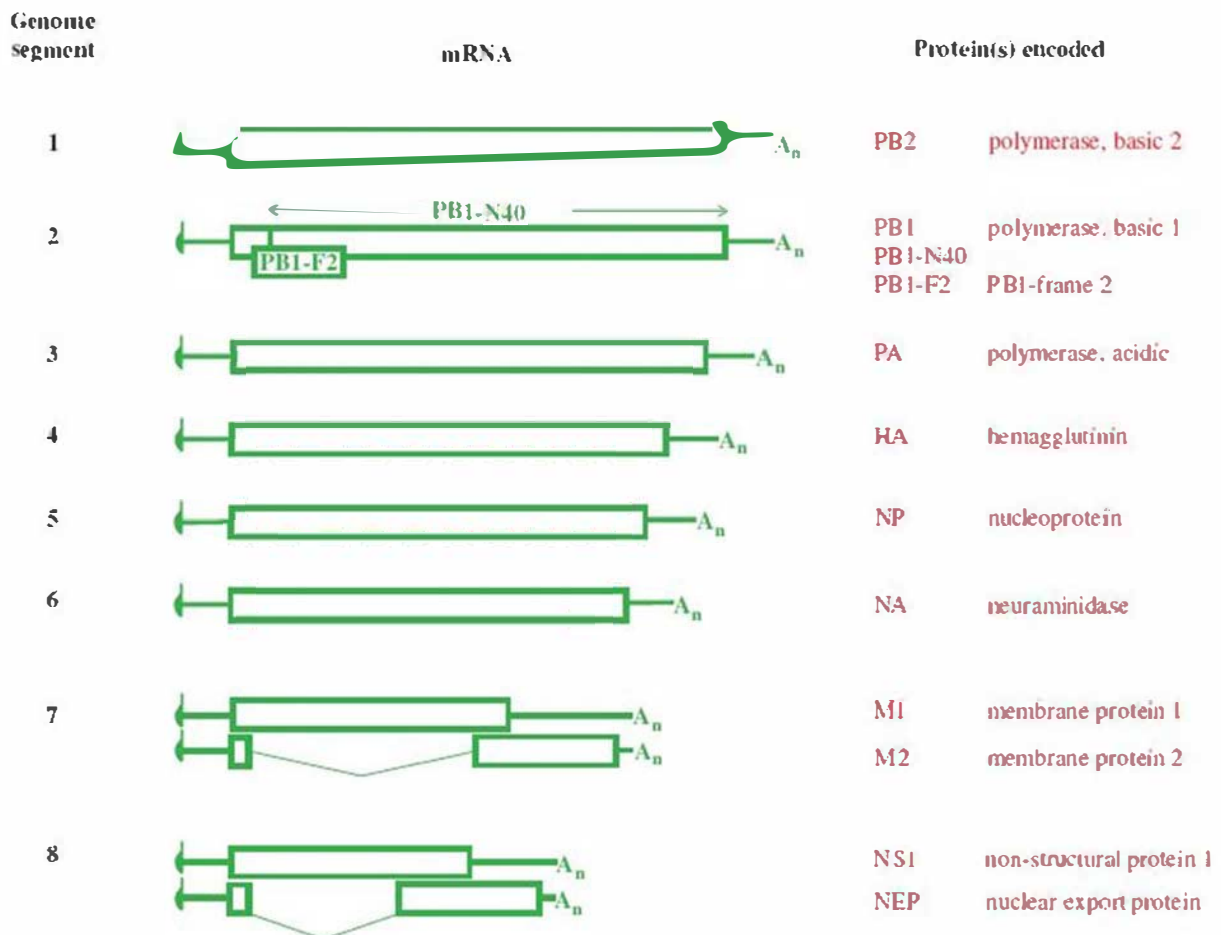
In the nucleus the polymerase (PA, PB1, PB2) associated with each RNP complex transcribes mRNA from the minus-strand genome RNA (Figure 16.4). RNA primers are required for this and these consist of short capped RNAs cut from the 5' ends of cell mRNAs by an endonuclease activity in the polymerase complex. A virus non-structural protein (NS1) is synthesized and is transported to the nucleus where it inhibits the processing and export of cell mRNAs from the nucleus, making them available for cap-snatching.

Figure 16.4 Influenza virus transcription and genome replication. The nucleotide sequences shown are typical ends of influenza A virus RNAs. mRNAs (ten species) and antigenome RNAs are copied from the eight genome RNAs by the viral polymerase. The sequence UUUUUU towards the 5' end of the (-) RNA is part of the polyadenylation signal; the sequence after this is not transcribed to mRNA. The antigenome RNAs act as templates for the synthesis of more (-) RNA.



Ten mRNA species are derived from the eight genome segments, two mRNA species being produced from each of the two smallest segments (Figure 16.5). A proportion of the M1 transcripts are spliced to form the mRNAs for the M2 protein, and approximately 10% of the NS1 transcripts are spliced to form the mRNAs for the nuclear export protein (NEP).

Figure 16.5 Influenza A virus transcripts and the proteins they encode. Codon 40 of the PB1 ORF is the start codon for PB1-N40. The PB1-F2 ORF is within the PB1 ORF, in the +1 reading frame. M2 and NEP are encoded by spliced mRNAs.



The quantities of the viral mRNA species vary during the replication cycle. At first the mRNAs for NP and NS1 are in the majority, while later there is increased production of those encoding HA, NA, and M1.

As well as synthesizing mRNAs the viral polymerase also copies the eight minus-strand genome segments to eight plus-strand antigenome segments to be used as templates for the production of more copies of the genome (Figure 16.4). These plus-strand RNAs, unlike the mRNAs, are not capped and polyadenylated, but they are coated with nucleoprotein, like the genome segments.

16.3.3 Translation and transport of proteins

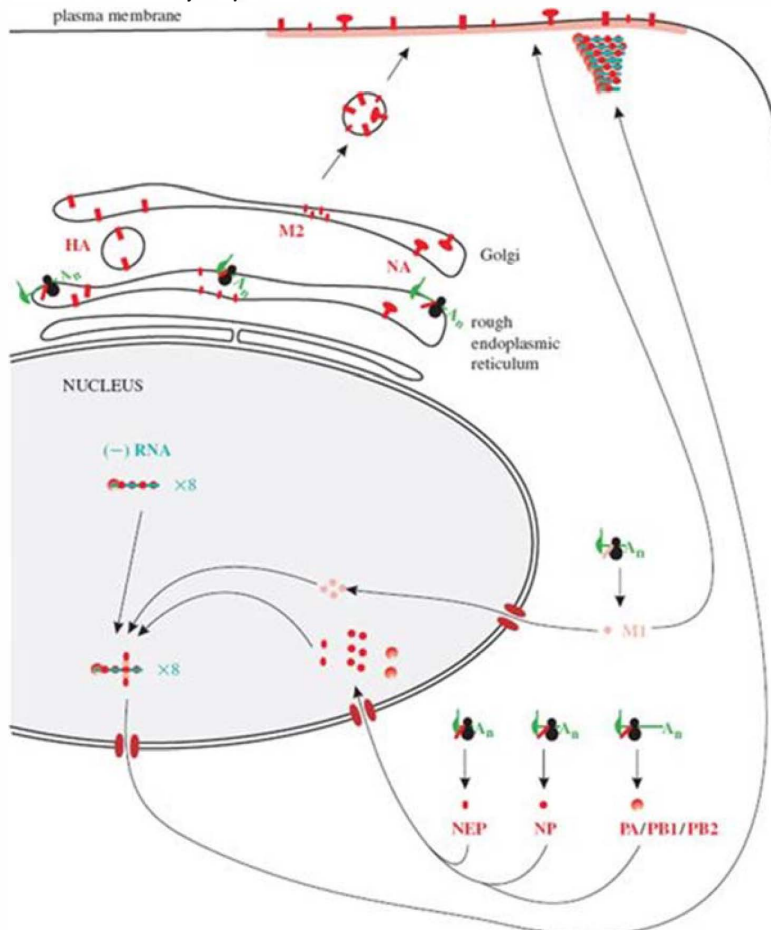
One protein species is translated from each of the ten mRNA species, except for genome segment 2 mRNA from which two or three proteins are translated (Figure 16.5):

- PB1
- PB1-N40 (an N-terminally truncated version of PB1)
- PB1-F2 (encoded by most, but not all, strains of influenza A virus).

The three envelope proteins (HA, NA, and M2) are translated in the rough endoplasmic reticulum and are then transported to the Golgi (Figure 16.6). In these compartments the proteins undergo post-translational modifications and are organized into multimers before transport to the plasma membrane. M2 ion channels play an important role in the transport vesicles, pumping out H⁺ ions to prevent the acid-induced conformational changes in HA molecules. This is in

contrast to the role played by M2 during entry of the virion into the cell, when H⁺ ions were pumped into the endosome to induce conformational changes in HA. The HA0 protein is cleaved to HA1 and HA2 (Figure 16.2) by a host protease.

Figure 16.6 Translation and transport of influenza virus proteins. Envelope proteins accumulate in the plasma membrane, while most other proteins are transported to the nucleus. Newly synthesized genome segments in the nucleus bind viral proteins and are transported to the cytoplasm.



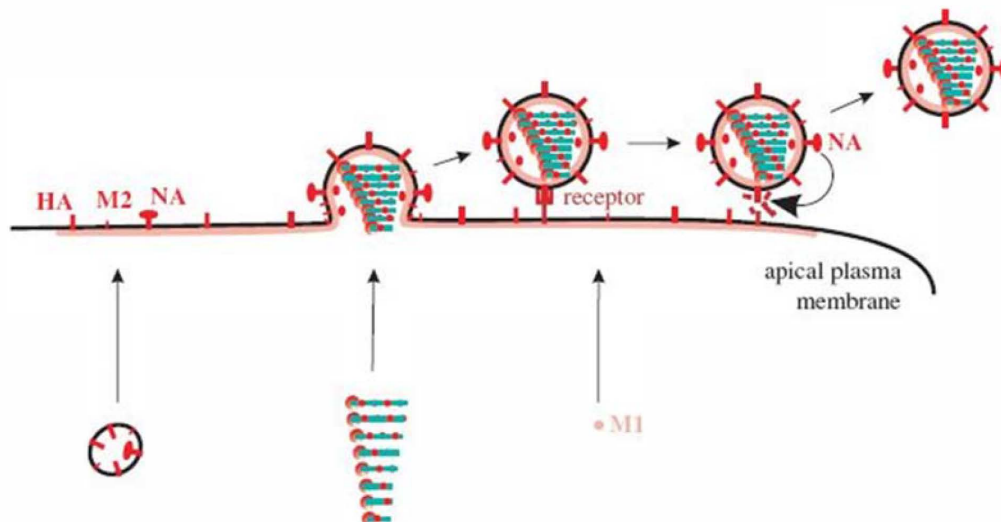
The remaining virus proteins are translated on free ribosomes; some of this protein remains in the cytoplasm, but much is transported to the nucleus. NS1 is a multifunctional protein, with roles in the cytoplasm and the nucleus. It has binding sites for RNA and for many cell and virus proteins. The role of NS1 in making cell mRNAs available to the viral cap-snatching apparatus has already been mentioned; other roles include inhibiting the interferon response of the cell.

After transport into the nucleus, NP, PA, PB1, and PB2 associate with viral RNAs forming RNPs. RNPs containing minus-strand RNA are exported from the nucleus by binding M1 molecules, which in turn bind NEP. A nuclear export signal on NEP is then recognized by the cell nuclear export machinery and the complex is transported through a nuclear pore. The nuclear localization signal on M1 is masked by the bound NEP, thus inhibiting import of the RNP complexes into the nucleus.

16.3.4 Virion assembly and release

The RNP complexes that have been exported from the nucleus accumulate near regions of the plasma membrane at the apical surface of the cell. These regions contain the three integral membrane proteins of the virus, with M1 protein bound to the cytoplasmic side of the membrane; it is from these regions that virions bud (Figure 16.7). Packaging signals near the ends of the virus RNAs ensure that one copy of each genome segment is selected for incorporation into a new virion. There is evidence that the M2 protein plays an important role in the final stages of the budding process. After a virion has budded it is attached to cell receptors via hemagglutinin in the envelope until the neuraminidase destroys the receptors.

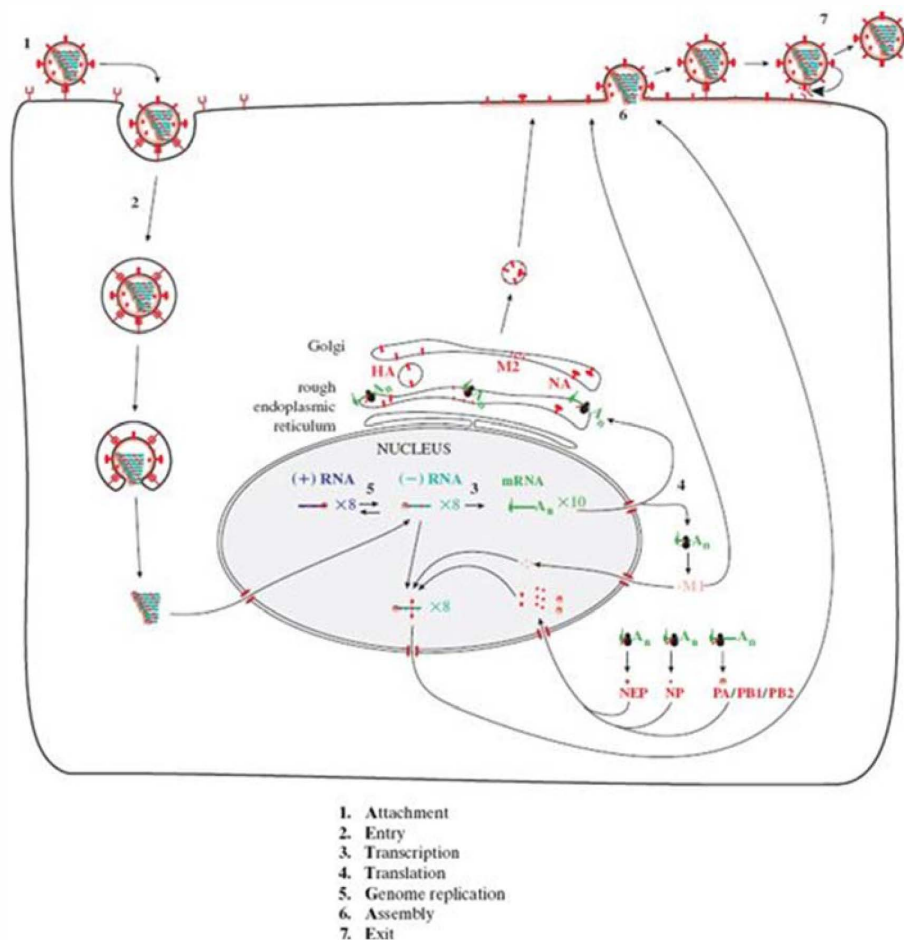
Figure 16.7 Assembly of influenza A virus. Eight genome segments and their associated proteins bud through a region of the plasma membrane coated with M1 and containing the three envelope proteins. The budded virion is bound to cell receptors; it is released when the receptors are degraded by the neuraminidase (NA).



16.3.5 Overview of influenza virus replication

The influenza virus replication cycle is summarized in Figure 16.8.

Figure 16.8 The influenza A virus replication cycle.



16.4 IMPORTANCE OF INFLUENZA VIRUSES

Most cases of influenza are caused by influenza A viruses. Epidemics of influenza B occur at 2–3 year intervals, while influenza C virus causes infrequent outbreaks of mild respiratory disease.

The hosts of influenza A viruses are principally birds that frequent aquatic habitats. The birds (e.g. ducks, geese, gulls) acquire infections by ingestion or inhalation and the viruses infect their intestinal and/or respiratory tracts. Infection of birds with most influenza A virus strains results in few or no signs of disease, but some strains are highly pathogenic and can kill their hosts. The viruses can be spread to new areas when the birds migrate.

Some influenza A viruses infect mammalian species including pigs, horses, and humans; the respiratory tract is the main site of virus replication. Of 16 hemagglutinin and nine neuraminidase subtypes known, only three hemagglutinin (H1, H2, H3) and two neuraminidase (N1, N2) subtypes commonly infect humans, causing abrupt onset of fever, headache, and muscular pain; there may be a sore throat and a cough. Patients are commonly very ill and some die, either as a direct result of the virus infection, especially if the lungs become infected (influenza pneumonia), or indirectly from bacterial pathogens, which are able to infect as a result of damage to the respiratory epithelium. One of these bacterial species is

Haemophilus influenzae; as its name suggests, it was once thought to be the causative agent of influenza.

The H and N glycoproteins occur in a large variety of antigenic forms and are constantly evolving new variants, a process known as antigenic drift. Individual humans have no immunity or limited immunity to a new variant, depending on which virus strains they have previously been infected with, so in most years there are epidemics of influenza, commonly referred to as seasonal flu. As well as influenza A virus strains, influenza B virus may also be involved. In temperate climates these epidemics usually occur during the winter. It is estimated that in most years influenza epidemics cause serious disease in 3–5 million people and 250 000–500 000 deaths. The most vulnerable are young children, the elderly, and those with health problems such as lung disease or heart disease.

Antigenic drift results in relatively small changes to the surface antigens of the virus and these changes occur frequently. From time to time, however, an antigenic shift occurs, where a new influenza A virus appears with an H and/or N gene derived from an avian or swine virus. The process involved in antigenic shift is called reassortment and is described in Section 21.3.3.c. If all or most humans are susceptible to the new virus, then a pandemic can occur.

In recent history influenza A viruses have caused four pandemics (Table 16.1), the most devastating in 1918–19 when the number of people killed by the virus was greater than the number killed during the First World War. The origins of the viruses responsible for the more recent influenza pandemics have been traced. In most cases the virus was a reassortant between a human virus and an avian virus; the 1957 reassortant derived the H and N genes from an avian virus, while the 1968 reassortant derived the H gene from an avian virus. It is thought that reassortment taking place in infected pigs is important in generating new influenza viruses that infect humans. The “swine flu” virus that appeared in 2009 had H and N genes derived from pig viruses, with other genes from human and avian influenza viruses (Figure 16.9).

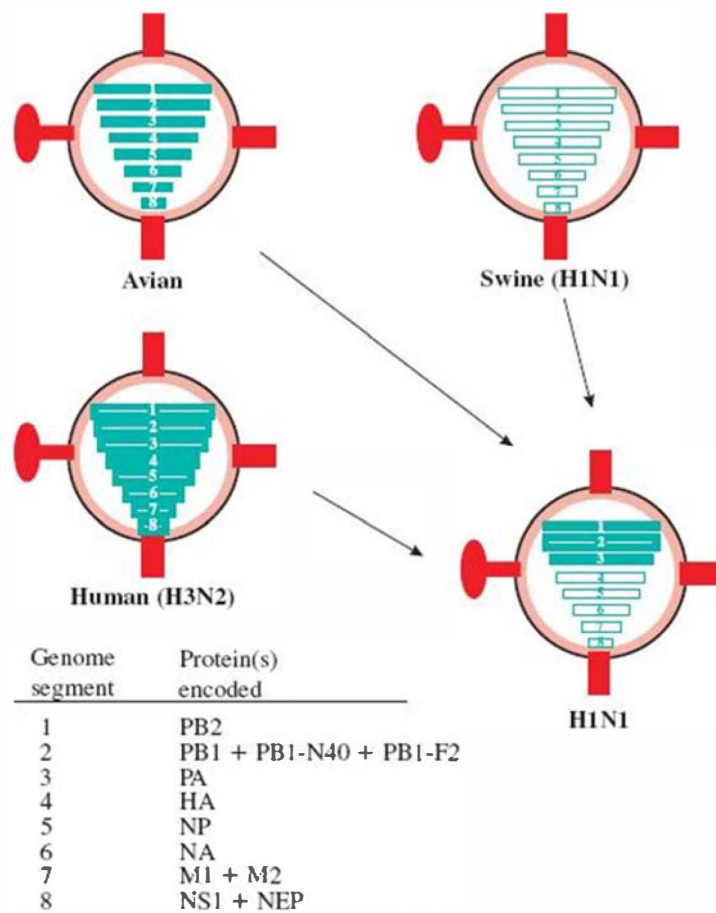
Table 16.1 Influenza A pandemics

Years	Virus subtype ¹	Parental viruses	Common name for pandemic
1918–1919	H1N1	Uncertain	Spanish flu
1957–1958	H2N2	Avian H2N2, human	Asian flu
1968–1969	H3N2	Avian H3, human N2	Hong Kong flu
2009–2010	H1N1	Swine H1N1, human, avian	Swine flu

¹H: hemagglutinin

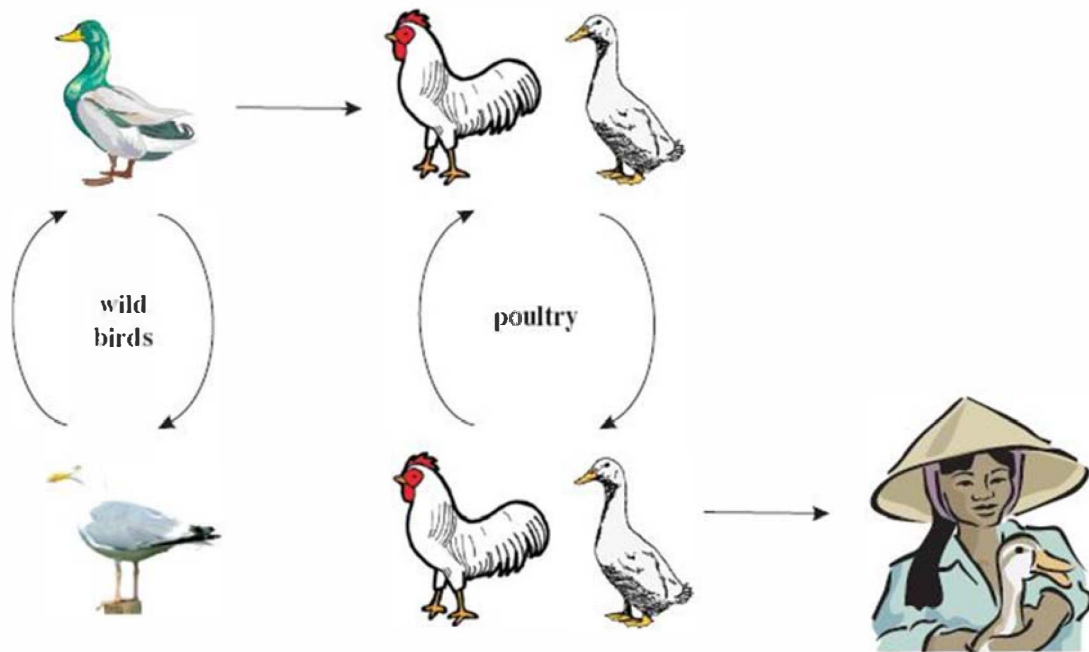
N: neuraminidase

Figure 16.9 Origin of the 2009 “swine flu” virus. The new H1N1 virus was a triple reassortant, with genes derived from swine, human, and avian viruses.



Although most avian strains of influenza A virus have little or no pathogenicity in their avian hosts, some are highly pathogenic in birds; these strains can be transmitted from wild birds to domestic poultry, and can be transmitted to humans (Figure 16.10). This was the situation in Hong Kong in 1997, when an H5N1 virus caused an outbreak of serious disease in poultry. In the early years of the twenty-first century H5N1 viruses spread to other Asian countries and to Europe and Africa. Millions of ducks, chickens, and turkeys died from the disease or were slaughtered. These viruses also infected humans, causing severe respiratory disease and culminating in death in many cases. For those people who became infected the death rate was over 50%, much higher than the death rate for the 1918 virus.

Figure 16.10 Transmission of avian influenza A viruses. The viruses are present in wild birds from which they may be transmitted to poultry. Sometimes there is transmission to humans who have close contact with poultry.



Most, if not all, of the human cases of “bird flu” caused by H5N1 viruses were in people who worked with poultry and who presumably became infected as a result of direct contact with virus on the birds themselves, or in their feces. These avian influenza viruses appeared to have little or no propensity for human-to-human transmission, but there was great concern that mutation or reassortment might give rise to an H5N1 virus that would cause the next pandemic. In the event, the next pandemic was caused by an H1N1 virus of swine origin! H5N1 viruses have also infected other mammalian species including domestic cats, and tigers and leopards in a zoo.

Because of the importance of influenza viruses, vaccines have been developed for use in vulnerable groups (Chapter 25). Most influenza vaccines are produced in chick embryos and are specific for particular virus strains. The adoption of cell culture production methods and the development of vaccines with broader specificity would be significant advances. Much effort has also gone into the development of anti-influenza virus drugs (Section 26.3.6). So far the M2 ion channel and the neuraminidase have been successful drug targets.

16.5 REVERSE GENETICS

Influenza viruses are among the minus-strand RNA viruses to which the techniques of reverse genetics have been applied (Section 15.6). For an influenza A virus this means transfecting cells with 12 different plasmids: one for each of the eight genome segments plus plasmids encoding PA, PB1, PB2, and NP. These procedures have been used in studies to characterize viral proteins and to determine pathogenicity factors.

The techniques of reverse genetics have been applied to the 1918 virus. Using RT-PCR, influenza virus RNA sequences have been recovered from lung tissues of some victims. In one project, tissue that had been fixed in formalin and embedded

in paraffin was the source, while another project used tissue from a body that had been buried in permafrost in Alaska. The virus sequences that were recovered differ from any H1N1 viruses known today and the origins of the 1918 virus remain obscure.

Infectious virus of the 1918 strain was generated. This work was controversial, but it enabled study of the roles of individual virus genes in pathogenicity for laboratory animals such as mice. It was shown that the HA, NA, and PB1 genes of the 1918 virus conferred pathogenicity to influenza A strains that were non-pathogenic in mice.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- describe, with the aid of a labeled diagram, the influenza A virion;
- describe the influenza A virus genome;
- write an illustrated account of the influenza A virus replication cycle;
- explain the roles of the influenza A virus proteins;
- explain the terms “antigenic drift” and “antigenic shift” in the context of influenza viruses;
- explain the importance of influenza in humans and birds.

SOURCES OF FURTHER INFORMATION

Books

Digard, P., Tiley L., and Elton, D. (2009) Orthomyxovirus genome transcription and replication. Chapter 9 in *Viral Genome Replication* (Cameron, C. E., Götte, M., and Raney, K. D., editors), Springer

Rappuoli, R. and Del Giudice, G., editors (2011) *Influenza Vaccines for the Future*, 2nd edition, Springer-Verlag

Van-Tam, J. and Sellwood, C. (eds) (2010) *Introduction to Pandemic Influenza*, CABI

Wang, Q. and Tao, Y. J. (2010) *Influenza: Molecular Virology*, Caister Academic Press

Journals

Chase, G. P. et al. (2011) Influenza virus ribonucleoprotein complexes gain preferential access to cellular export machinery through chromatin targeting. *PLoS Pathogens*, 7 (9), e1002187

Fukuyama, S. & Kawaoka, Y. (2011) The pathogenesis of influenza virus infections: the contributions of virus and host factors. *Current Opinion in Immunology*, 23, 481–486

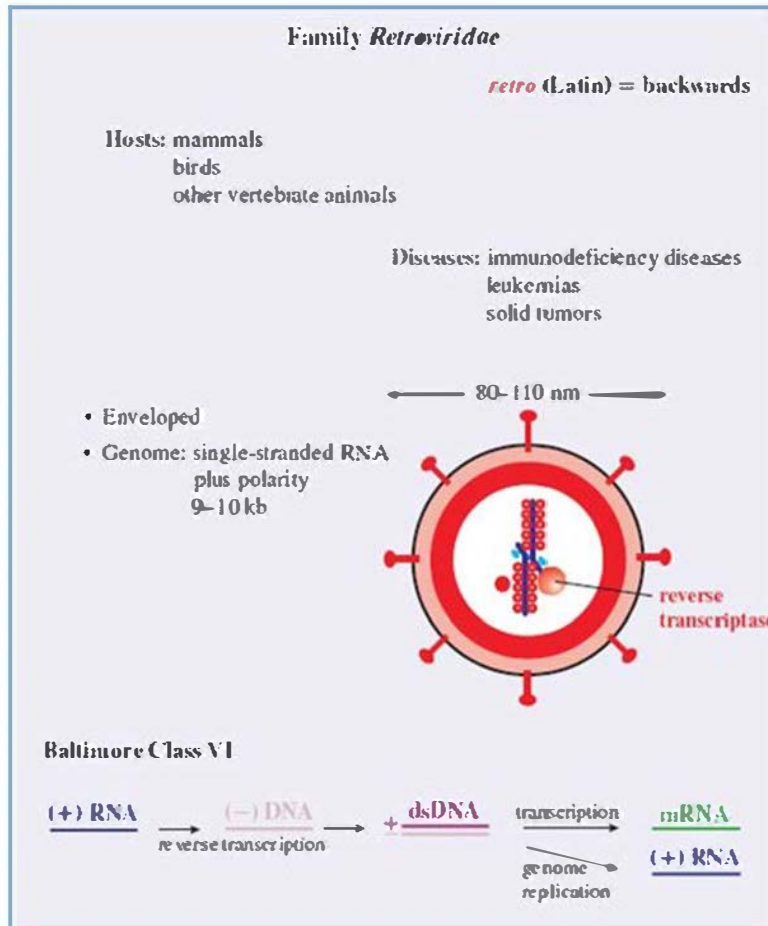
Hutchinson, E. C. et al. (2010) Genome packaging in influenza A virus. *Journal of General Virology*, 91, 313–328

- Klenk, H. D., Garten, W., and Matrosovich, M. (2011) Molecular mechanisms of interspecies transmission and pathogenicity of influenza viruses: lessons from the 2009 pandemic. *Bioessays*, 33, 180–188
- Pielak, R. M. and Chou, J. J. (2010) Influenza M2 proton channels. *Biochimica et Biophysica Acta*, 1808, 522–529
- Robertson, J. S. and Inglis, S. C. (2011) Prospects for controlling future pandemics of influenza. *Virus Research*, 162, 39–46
- Rossman, J. S. and Lamb, R. A. (2011) Influenza virus assembly and budding. *Virology*, 411, 229–236
- Taubenberger, J. K. and Kash, J. C. (2011) Insights on influenza pathogenesis from the grave. *Virus Research*, 162, 2–7
- Yángüez, R. and Nieto, A. (2011) So similar, yet so different: selective translation of capped and polyadenylated viral mRNAs in the influenza virus infected cell. *Virus Research*, 156, 1–12

CHAPTER 17

Retroviruses

CHAPTER 17 AT A GLANCE



17.1 INTRODUCTION TO RETROVIRUSES

The retroviruses are RNA viruses that copy their genomes into DNA during their replication. Until the discovery of these viruses it had been dogma that the transfer of genetic information always occurs in the direction of DNA to RNA, so finding that some viruses carry out “transcription backwards” (reverse transcription) caused something of a revolution. We now know that reverse transcription is carried out not only by these RNA viruses but also by some DNA viruses (see Chapter 19) and by uninfected cells.

REVERSE TRANSCRIPTASES ARE USED IN MOLECULAR BIOLOGY

RTs have the potential to copy any RNA into DNA, even in the absence of specific tRNA primers. This has made them indispensable tools in molecular biology, where they have a number of applications, including the production of cDNA libraries and the reverse transcription–polymerase chain reaction (RT–PCR).

Two commonly used RTs are those from avian myeloblastosis virus and Moloney murine leukemia virus.

The aim of the current chapter is to provide a general introduction to the retroviruses, which have been found in all classes of vertebrate animal, including fish, amphibians, birds, and mammals. The human immunodeficiency viruses (HIV-1 and HIV-2) are retroviruses and the next chapter is devoted entirely to these viruses. Many retroviruses can cause cancer in their hosts, and some aspects of this are discussed in Chapter 23.

17.2 RETROVIRUS VIRION

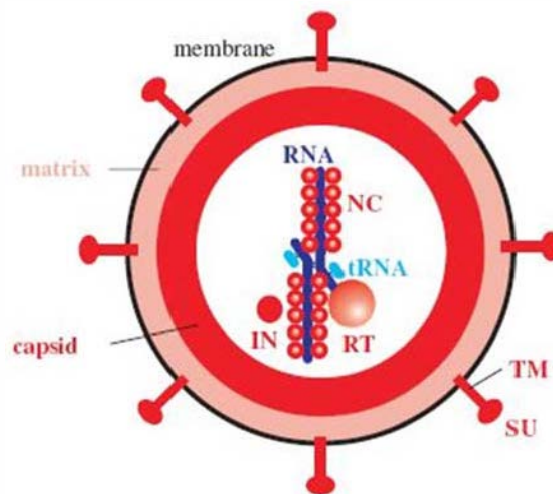
Retroviruses are unusual in that the virion contains two copies of the genome (Figure 17.1). The two RNA molecules are present as a dimer, formed by base pairing between complementary sequences. As well as the virus RNA, the virion also contains molecules of host cell RNA that were packaged during assembly. This host RNA includes a molecule of transfer RNA (tRNA) bound to each copy of the virus RNA through base pairing. The sequence in the virus RNA that binds a tRNA is known as the primer binding site (PBS). Each retrovirus binds a specific tRNA (Table 17.1).

Figure 17.1 Retrovirus virion and genome organization

Genes	gag	group-specific antigen
	pol	polymerase
	env	envelope
Proteins	CA	capsid
	IN	integrase
	MA	matrix
	NC	nucleocapsid
	PR	protease
	RH	ribonuclease H
	RT	reverse transcriptase
	SU	surface glycoprotein
	TM	transmembrane glycoprotein
Sequences with roles in replication	PBS	primer binding site
	PPT	polypurine tract
	R	repeat sequence
	U3	unique sequence at 3' end of genome

	U5	unique sequence at 5' end of genome
	ψ	Greek letter psi, from packaging signal

Virion components



Genome organization and gene products

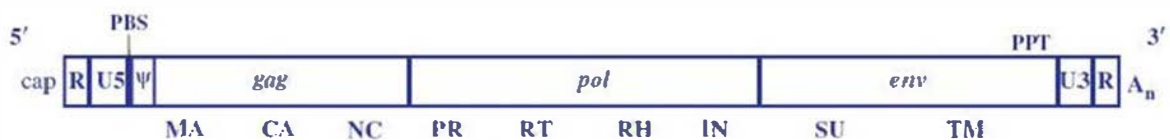


Table 17.1 Examples of tRNAs used by retroviruses as primers

tRNA	Retrovirus
tRNA ^{pro}	Human T-lymphotropic viruses 1 and 2 Murine leukemia virus
tRNA ^{lys3}	HIV-1 and 2 Mouse mammary tumor virus

A number of protein species are associated with the RNA. The most abundant protein is the nucleocapsid (NC) protein, which coats the RNA, while other proteins, present in much smaller amounts, have enzyme activities (Table 17.2).

Table 17.2 Enzyme activities present in the retrovirus virion

RNA-dependent DNA polymerase (reverse transcriptase; RT)
DNA-dependent DNA polymerase
Ribonuclease H (RNase H)
Integrase
Protease

Encasing the RNA and its associated proteins is the capsid, constructed from a lattice of capsid (CA) protein. The capsid has the shape of a sphere, a cone, or a

cylinder, depending on the virus. A layer of matrix (MA) protein lies between the capsid and the envelope. Associated with the envelope are two protein species: a transmembrane (TM) protein bound to a heavily glycosylated surface (SU) protein. In most retroviruses the bonds between the TM and SU proteins are non-covalent. The TM–SU proteins are present in the envelope in trimers (groups of three).

The genes encoding the virus proteins are organized in three major regions of the genome (Figure 17.1):

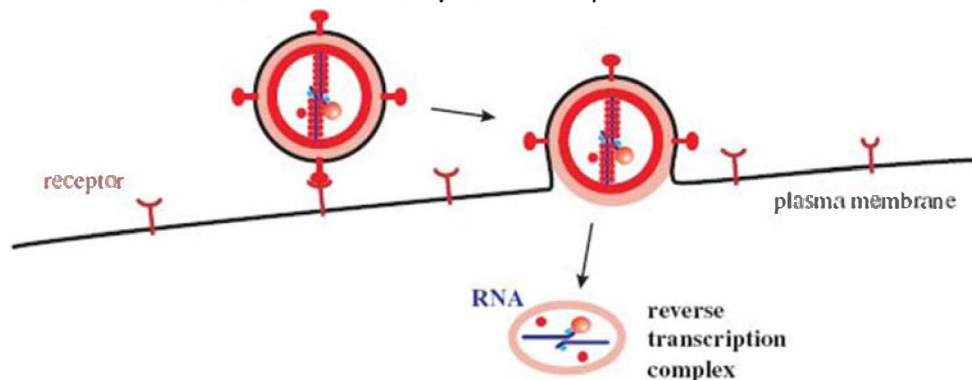
- gag (group-specific antigen)—internal structural proteins;
- pol (polymerase)—enzymes;
- env (envelope)—envelope proteins.

17.3 RETROVIRUS REPLICATION

17.3.1 Attachment and entry

A site on the envelope SU protein binds to a cell receptor. This interaction causes a conformational change in the TM protein that allows a hydrophobic fusion sequence to fuse the virion membrane and a cell membrane (Section 5.2.4.b). Most retroviruses fuse their membrane with the plasma membrane of the cell (Figure 17.2), though some are endocytosed and fuse their membrane with an endosome membrane. The structure that is released into the cytoplasm loses some proteins and a reverse transcription complex is formed.

Figure 17.2 Retrovirus attachment and entry. Fusion of the virion membrane with the plasma membrane of the cell releases the virion contents, which undergo modification to form a reverse transcription complex.

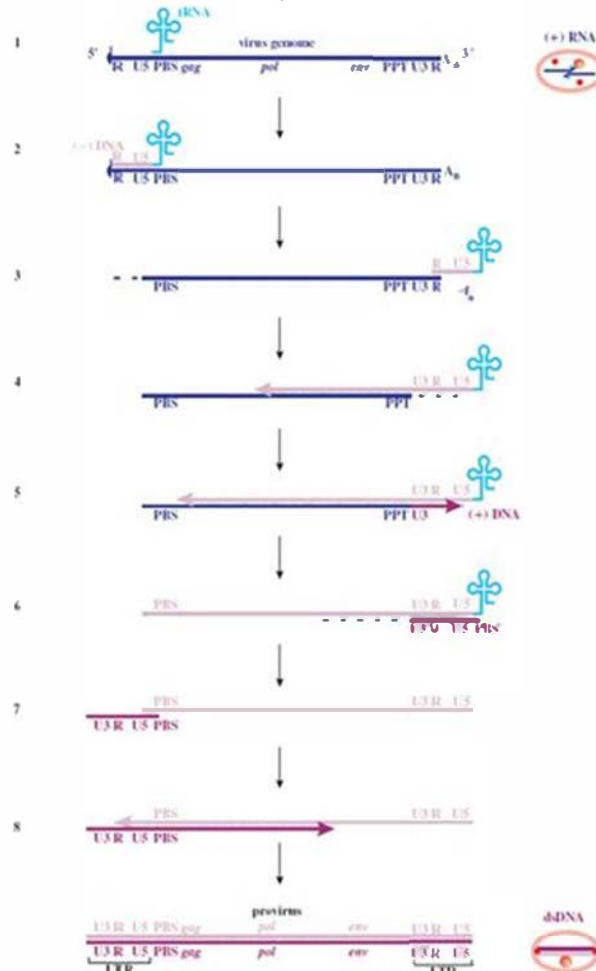


17.3.2 Reverse transcription

Reverse transcription takes place within the reverse transcription complex; some of the detail is indicated in Figure 17.3. Synthesis of both the (–) DNA and the (+) DNA begins at the 3′—OH of a primer RNA. The primer for synthesis of the (–) DNA is the tRNA bound to the genome, while the primer for synthesis of the (+) DNA is a polypurine tract (PPT) in the virus genome. The latter becomes accessible as a result of hydrolysis of the genome RNA from the 3′ end by the RNase H, which is an enzyme that specifically digests RNA in RNA–DNA duplexes.

Figure 17.3 Retroviral reverse transcription. LTR: long terminal repeat. PBS: primer binding site. PPT: polypurine tract (a sequence made up entirely, or almost entirely, of purine residues). R: repeat sequence. U3: unique sequence at 3' end of genome. U5: unique sequence at 5' end of genome.

1. A copy of the virus genome with a tRNA bound at the PBS.
2. The reverse transcriptase begins (–) DNA synthesis at the 3' end of the tRNA.
3. The RNase H digests the RNA from the RNA–DNA duplex. The (–) DNA attaches at the 3' end of either the same RNA strand or the second copy of the genome.
4. Elongation of the (–) DNA continues, while the RNase H degrades the template RNA from the 3' end as far as the PPT.
5. Synthesis of (+) DNA begins.
6. The remaining RNA is degraded.
7. The (+) DNA detaches from the 5' end of the (–) DNA template and attaches at the 3' end.
8. Synthesis of both DNA strands is completed.

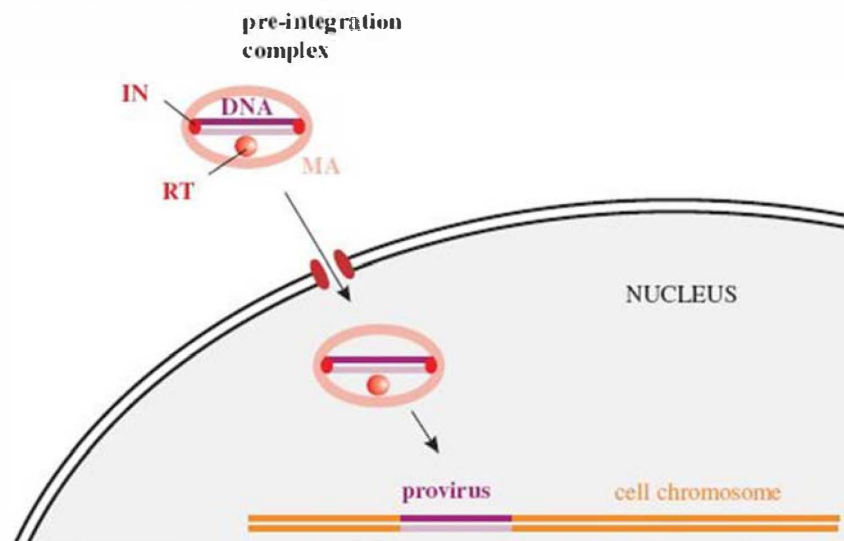


During synthesis of the two DNA strands, each detaches from its template and re-attaches at the other end of the template through base pairing. The DNA that results from reverse transcription (the provirus) is longer than the RNA genome. Each of the termini has the sequence U3–R–U5, known as a long terminal repeat (LTR), one terminus having acquired a U3 sequence and the other a U5 sequence.

17.3.3 Integration of the provirus

The provirus, still associated with some virion protein, is transported to the nucleus as a pre-integration complex. The proteins of the pre-integration complex include integrase (bound to each end of the DNA), reverse transcriptase, and matrix protein (Figure 17.4). The pre-integration complex of some retroviruses, including HIV, is transported toward the nucleus via the microtubule network. For most retroviruses, entry into the nucleus requires breakdown of the nuclear envelope, an event that occurs during mitosis; this means that there can be a productive infection only in dividing cells. HIV and related viruses, however, can productively infect resting cells, as the pre-integration complexes of these viruses are able to enter intact nuclei (Section 18.4.2).

Figure 17.4 Transport of pre-integration complex to the nucleus and integration of the provirus into a cell chromosome.

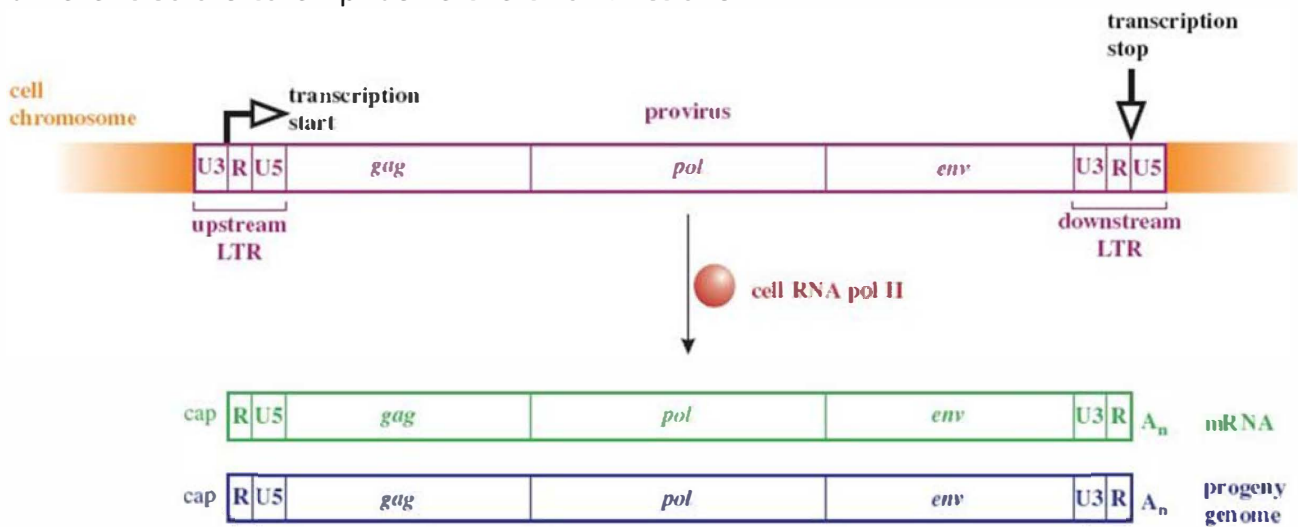


The integrase cuts the DNA of a cell chromosome and inserts the provirus into the gap; integration is completed by the DNA repair machinery of the cell. The integrated provirus genes may be expressed immediately, or there may be little or no expression of viral genes, in which case a latent infection has been initiated. If a latently infected cell divides, the provirus is copied along with the cell genome and each of the daughter cells has a copy of the provirus.

17.3.4 Transcription and genome replication

The two LTRs of the provirus have identical sequences, but are functionally different; transcription is initiated in one and terminated in the other. Transcription factors bind to a promoter in the upstream LTR, then the cell RNA polymerase II starts transcription at the U3–R junction (Figure 17.5). Transcription continues into the downstream LTR. There is a polyadenylation signal in the R region and transcription terminates at the R–U5 junction. Each transcript is capped and polyadenylated. Some transcripts will function as mRNA and a proportion of these become spliced; others will become the genomes of progeny virions.

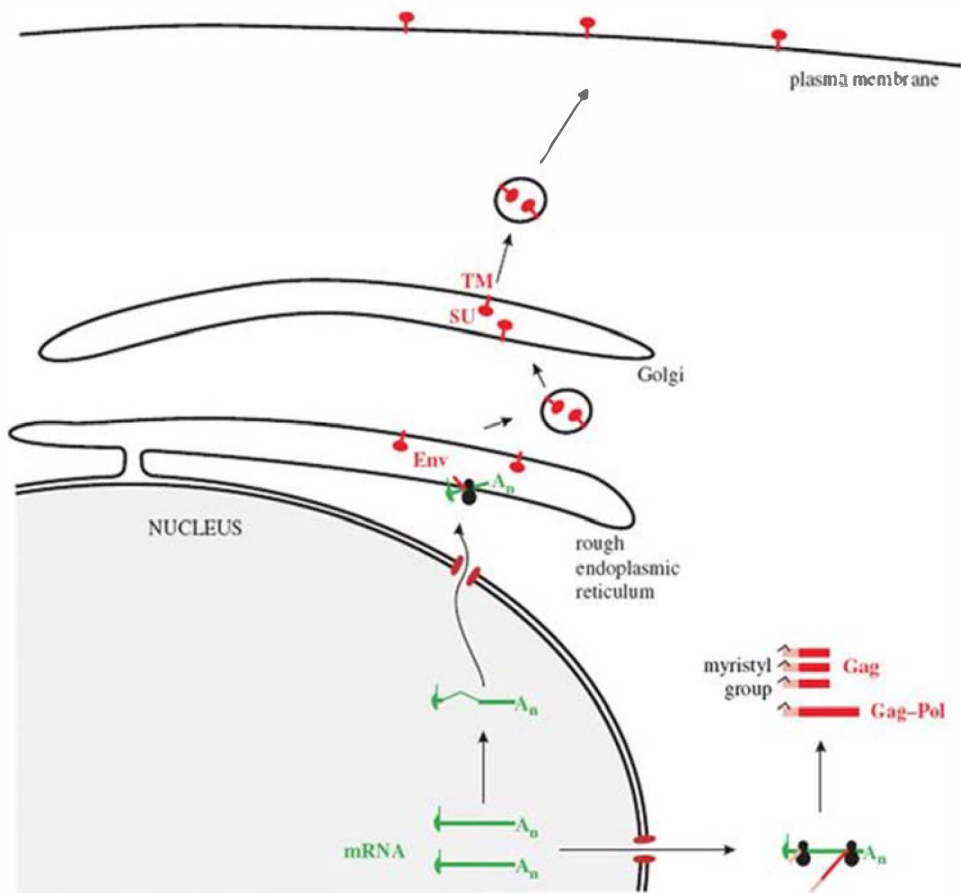
Figure 17.5 Retrovirus transcription. Genome-length RNAs are synthesized. Some will function as mRNAs (shown in green) while some will become the genomes of progeny virions (shown in blue). These RNAs are identical; they are shown in different colors to emphasize the two functions.



17.3.5 Translation and post-translational modifications

The *env* gene is translated from spliced mRNAs in the rough endoplasmic reticulum, where glycosylation commences (Figure 17.6). The Env protein molecules are transported to the Golgi complex, where they are cleaved by a host protease into SU and TM molecules. The two cleavage products remain in close association, and after further glycosylation they are transported to the plasma membrane.

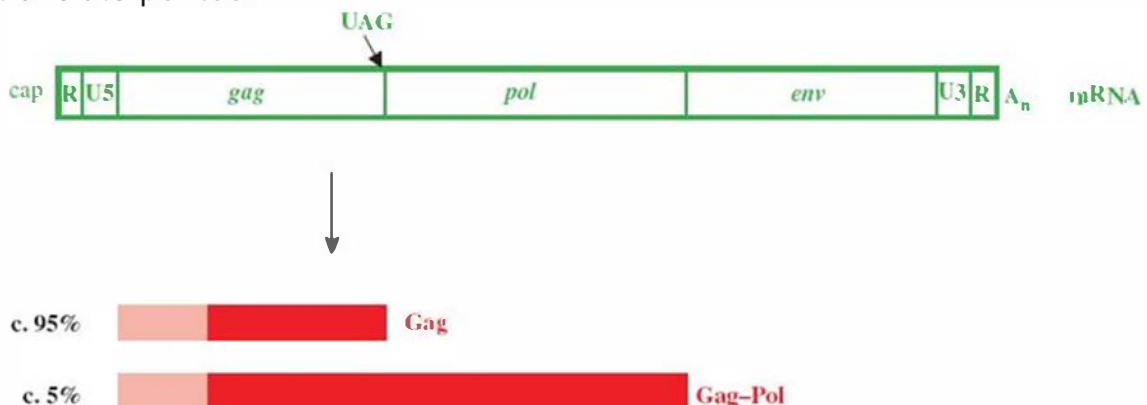
Figure 17.6 Retrovirus translation and post-translational modifications. Gag and Gag–Pol are translated on free ribosomes, and then myristylated (most retroviruses). Env is translated in the rough endoplasmic reticulum and transported to the plasma membrane via the Golgi complex; en route it is glycosylated and cleaved to SU and TM.



The proteins encoded by the gag and pol genes are translated from genome-length mRNAs into Gag and Gag-Pol polyproteins (Figure 17.6). Retroviruses require much greater quantities of Gag than Pol, and have evolved mechanisms to synthesize the required amount of each. These mechanisms involve approximately 95% of ribosomes terminating translation after the synthesis of Gag, while the other ribosomes continue translation to synthesize Gag-Pol.

One mechanism, used by murine leukemia virus, involves reading through a stop codon at the end of gag (Figure 17.7). A “suppressor tRNA” incorporates an amino acid at the stop codon and translation continues into pol, which is in the same reading frame as gag.

Figure 17.7 Translation of Gag-Pol by reading through a stop codon (UAG). Approximately 5% of ribosomes incorporate an amino acid at the gag stop codon and translate pol too.



The majority of retroviruses, however, ensure the correct proportions of the Gag and Pol proteins by a ribosomal frameshifting mechanism (Figure 17.8). Here gag and pol are in different reading frames and there is a -1 shift in reading frame before the gag-pol junction in about 5% of translations. This mechanism is used by HIV-1.

Figure 17.8 Translation of Gag-Pol by ribosomal frameshifting. Approximately 5% of ribosomes shift into a different reading frame before the gag stop codon and translate pol too.

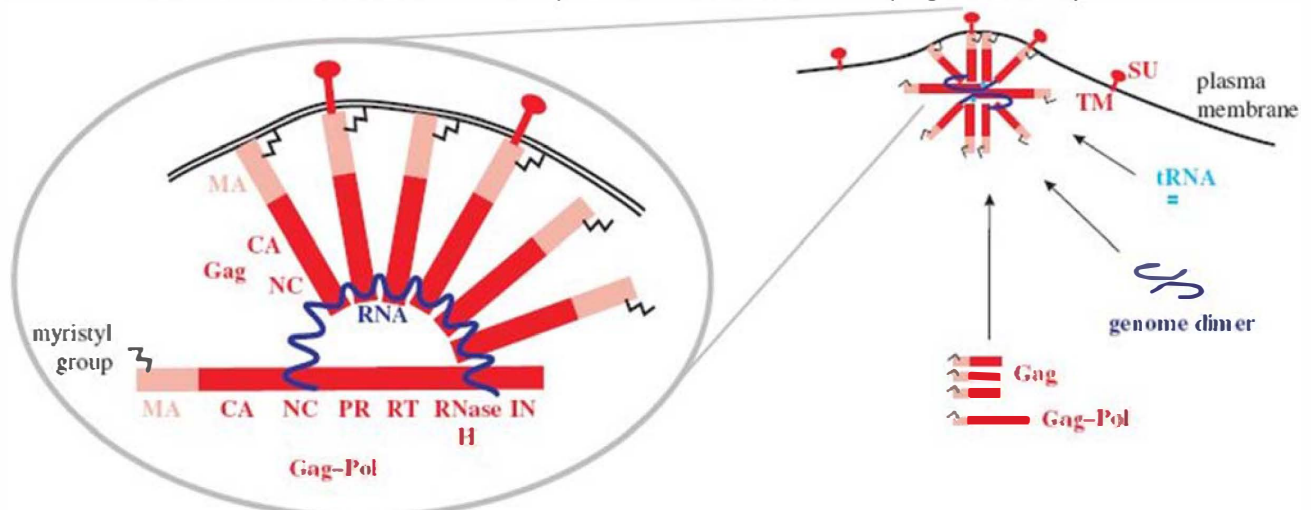


The Gag and Gag-Pol proteins of most retroviruses are myristylated at their N termini.

17.3.6 Assembly and release of virions

Some retroviruses form immature nucleocapsids that are then transported to the plasma membrane. Most retroviruses, however, assemble their components on the inner surface of the plasma membrane (Figure 17.9).

Figure 17.9 Retrovirus assembly – early stages. A genome dimer associates with cell tRNAs and with Gag and Gag-Pol proteins. The domains of Gag and Gag-Pol are indicated in the inset. The order of the Gag domains MA-CA-NC is the same as the exterior-to-interior order of the proteins in the virion (Figure 17.1).



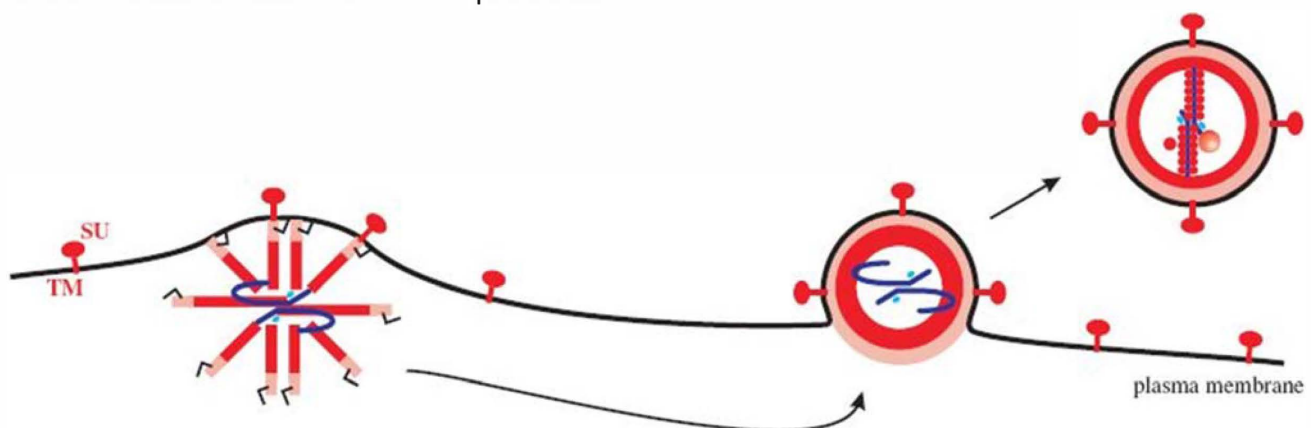
Two copies of genome RNA form a dimer by base pairing between complementary sequences. The region of interaction between the RNA molecules is said to form a "kissing-loop complex." A tRNA molecule binds to the PBS of each copy of the genome (Figure 17.1). Copies of the Gag protein bind via their NC domain to the

RNA; the proteins bind first to the packaging signal (ψ) near the 5' end of the RNA, which then becomes coated with many copies of Gag and a few copies of Gag–Pol (Figure 17.9).

Gag and Gag–Pol proteins become anchored to the plasma membrane by the myristyl groups at their N termini. The associations are stabilized through electrostatic interactions between positive charges in the MA domains and negatively charged phosphate groups in the membrane. The MA domains also bind to the cytoplasmic tails of TM proteins in the membrane.

The immature virion acquires its envelope by budding from the cell surface (Figure 17.10). At this stage multiple copies of the Gag and Gag–Pol proteins are arranged radially with their N termini facing outward and their C termini inward. Gag has one or two late (L) domains that bind host cell factors involved in the budding process (Section 8.3.1). During and/or after the budding process the Gag and Gag–Pol polyproteins are cleaved by the virus protease. The cleavage products of Gag form the matrix, the capsid and the protein component of the nucleocapsid, while the Pol cleavage products are the virion enzymes.

Figure 17.10 Retrovirus assembly – late stages. The envelope is acquired by budding from the plasma membrane. During and after budding Gag and Gag–Pol are cleaved to form the virion proteins.

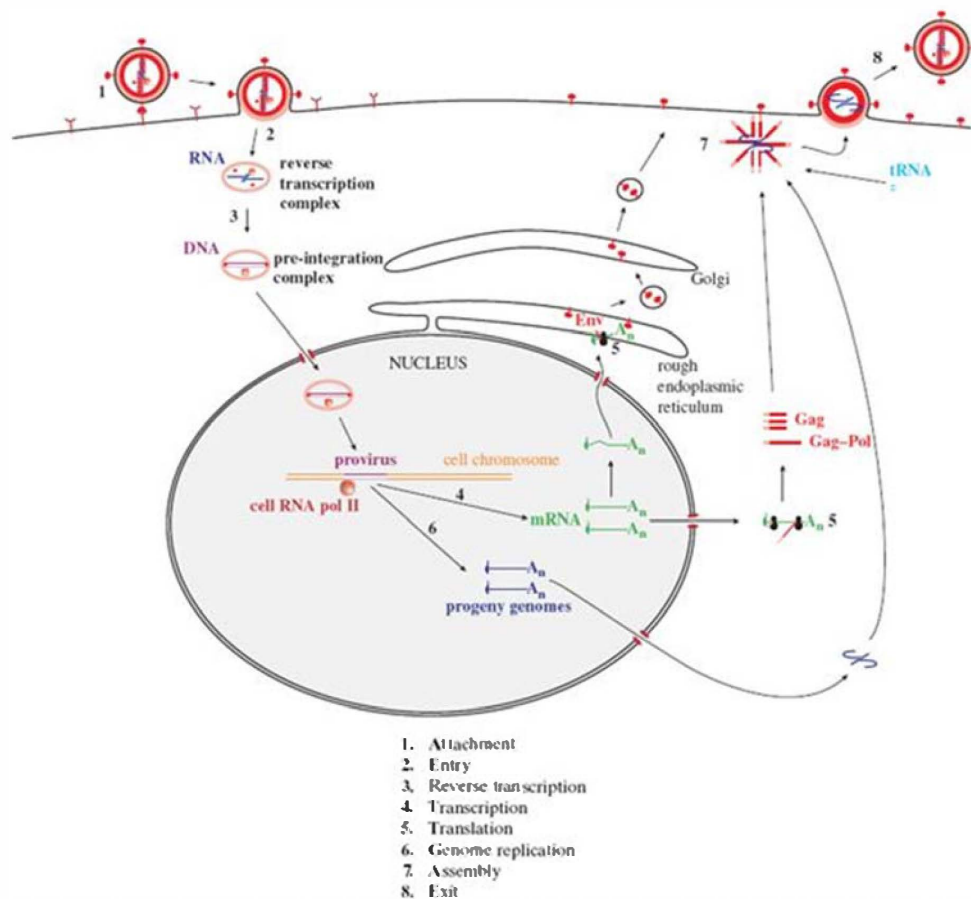


If the retrovirus-infected cell is close to a susceptible cell then the two cells may form a close connection, called a virological synapse, across which infectious virus may be transmitted, thus avoiding exposure to extracellular host defenses such as antibodies. HIV can be transmitted from cell to cell in this way, and for human T-lymphotropic virus 1 it is the main mode of transmission.

17.3.7 Overview of retrovirus replication

The retrovirus replication cycle is summarized in Figure 17.11.

Figure 17.11 The retrovirus replication cycle. Note that there is an additional step in retrovirus replication: reverse transcription, which takes place between entry and transcription.



17.4 EXAMPLES OF RETROVIRUSES

Retroviruses can be classed as either simple or complex, depending on the complexity of their genomes. The simple retroviruses are those that have only the three standard retrovirus genes (*gag*, *pol*, *env*), or in some cases one additional gene, called an oncogene because its expression might result in its host cell developing into a tumor cell. An example of an oncogene is *src* in the genome of Rous sarcoma virus, which infects chickens (Figure 17.12).

Figure 17.12 Rous sarcoma virus genome. There is an oncogene (*src*) in addition to the three standard retrovirus genes.



The complex retroviruses have additional genes, the products of which have a variety of functions in the replication cycle. The human immunodeficiency viruses are complex retroviruses; because of their importance the next chapter is devoted entirely to them.

The genera of the family Retroviridae and some representative viruses are listed in Table 17.3. As the names of the viruses imply, many of them are causative agents of disease in mammals (feline leukemia virus), birds (Rous sarcoma virus), and fish (walleye dermal sarcoma virus).

Table 17.3 Examples of retroviruses

Simple retroviruses		Complex retroviruses	
Genus	Virus examples	Genus	Virus examples
<i>Alpharetrovirus</i>	Rous sarcoma virus	<i>Deltaretrovirus</i>	Human T-lymphotropic viruses 1 & 2
<i>Betaretrovirus</i>	Mouse mammary tumor virus	<i>Epsilonretrovirus</i>	Walleye dermal sarcoma virus
		<i>Lentivirus</i>	Human immunodeficiency viruses 1 & 2
<i>Gammaretrovirus</i>	Murine leukemia virus Feline leukemia virus	<i>Spumavirus</i>	Simian foamy virus

17.5 RETROVIRUSES AS GENE VECTORS

Some genetically modified retroviruses are used as gene vectors. They can introduce genes into a variety of cell types, where the genes are expressed at high levels after integration into the cell genomes. Technologies have been developed for expression of genes in cell cultures and for clinical treatments of genetic disorders and cancers. One of the most commonly used viruses in these applications is murine leukemia virus. Lentiviral vectors have also been developed; these have the advantage that they can deliver genes into non-dividing cells and tissues.

Patients with the genetic disorder X-linked severe combined immunodeficiency (X-SCID) have been successfully treated with retroviral vectors. The procedure involves taking stem cells from the patient and infecting them with a recombinant retrovirus, the genome of which contains the functional gene that is defective in the patient's genome. If the vector provirus is integrated into the cell genome, and if the functional gene is expressed, then the patient is able to develop a normal immune system. The successes, however, have been tempered by the development of cancer in a few treated patients.

17.6 ENDOGENOUS RETROVIRUSES

It has been known for some time that the genomes of vertebrate animals contain retroviral sequences. The genomes of most of these endogenous retroviruses (ERVs) are defective. Sequencing the human genome has revealed the presence of almost 100 000 human ERV (HERV) sequences, and ERVs have been found in the genomes of other animal species as they have been sequenced.

Some ERVs are closely related to normal retroviruses (exogenous retroviruses); for example, there are ERVs in mice that have very similar sequences to the genome of mouse mammary tumor virus. It is highly likely that ERVs originated as a result of exogenous retroviruses infecting germ line cells (sperm and/or egg). If one of these cells with an integrated provirus survived to be involved in the reproductive process, then each cell in the body of the offspring would contain a copy of the provirus. Over time ERVs have copied themselves to other locations in the genome, giving rise to families of related ERVs. The copying process is known as

retrotransposition, and is carried out by other genome elements as well as ERVs. Collectively these elements are known as retrotransposons (Section 21.3.5.a).

Most ERVs are inactive for most of the time, but some ERVs can be stimulated into activity by, for example, infection with other agents. Stimulation of an ERV may result in transcription of ERV sequences and synthesis of ERV proteins. There is evidence that host animals have hijacked some ERV proteins for their own benefit; for example, in some mammalian species the ability of ERV envelope proteins to fuse membranes is exploited during the development of the placenta.

If the ERV has a full complement of retroviral genes, then ERV particles may be produced. If the ERV lacks a gene essential for the complete replication cycle the missing function may be supplied by another ERV or by an exogenous retrovirus. Some ERVs do not replicate in cells of the species in which they occur, but are able to replicate in the cells of other species; for example, some mouse ERVs and some pig ERVs can replicate in human cells.

Because of these findings, concern has been expressed that there may be a risk of transmitting retroviruses from pigs into humans when pigs are used as sources of cells and tissues because of shortages of their human counterparts for transplant purposes.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- describe the retrovirus virion;
- describe the main features of the retrovirus genome;
- explain the main features of the retrovirus replication cycle;
- give examples of retroviruses and explain their importance;
- discuss endogenous retroviruses.

SOURCES OF FURTHER INFORMATION

Books

Galetto, R. and Negroni, M. (2009) Retroviruses. Chapter 6 in *Viral Genome Replication* (Cameron, C. E., Götte, M., and Raney, K. D., editors), Springer

Kurth, R. and Bannert, N. (ed.) (2010) *Retroviruses: Molecular Biology, Genomics and Pathogenesis*, Caister Academic Press

Lisch, D. (ed.) (2009) Transposable elements and retroviruses. Chapter 17 in *Lewin's Genes X* (Krebs, J. E., Goldstein, E. S., and Kilpatrick, S. T., editors), Jones & Bartlett

Journals

Cherepanov, P., Maertens, G. N., and Hare, S. (2011) Structural insights into the retroviral DNA integration apparatus. *Current Opinion in Structural Biology*, 21, 249–256

- Desfarges, S. and Ciuffi, A. (2010) Retroviral integration site selection. *Viruses*, 2, 111–130
- Goff, S. P. (2007) Host factors exploited by retroviruses. *Nature Reviews Microbiology*, 5, 253–263
- Johnson, S. F. and Telesnitsky, A. (2010) Retroviral RNA dimerization and packaging: the what, how, when, where, and why. *PLoS Pathogens*, 6 (10): e1001007
- Martin-Serrano, J. and Neil, S. J. D. (2011) Host factors involved in retroviral budding and release. *Nature Reviews Microbiology*, 9, 519–531
- Sattentau, Q. J. (2010) Cell-to-cell spread of retroviruses. *Viruses*, 2, 1306–1321
- Suzuki, Y. and Craigie, R. (2007) The road to chromatin – nuclear entry of retroviruses. *Nature Reviews Microbiology*, 5, 187–196
- Waheed, A. A. and Freed, E. O. (2010) The role of lipids in retrovirus replication. *Viruses*, 2, 1146–1180

Wolf, D. and Goff, S. P. (2008) Host restriction factors blocking retroviral replication. Annual Review of Genetics, 42, 143–163

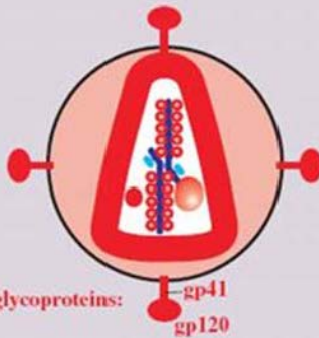
CHAPTER 18

Human Immunodeficiency Viruses

CHAPTER 18 AT A GLANCE

Two major types: HIV-1 and HIV-2
Complex retroviruses
Causative agents of acquired immune deficiency syndrome (AIDS)

- Enveloped
- Genome: single-stranded RNA plus polarity
9.3 kb
- Baltimore Class VI



HIV-1 glycoproteins: gp120, gp41

HIV-1 Auxiliary Proteins	Functions include:
Tat	transcription factor
Rev	late gene expression
Nef	protection from immune surveillance
Vif	protection from APOBEC proteins
Vpr	targeting pre-integration complex to nucleus
Vpu	virion budding

18.1 INTRODUCTION TO HIV

There are two types of human immunodeficiency virus (HIV-1 and HIV-2), which each evolved from a different simian immunodeficiency virus (SIV). Both HIVs emerged in the late twentieth century. HIV infection damages the immune system, leaving the body susceptible to infection with a wide range of bacteria, viruses, fungi, and protozoa. This condition is called acquired immune deficiency syndrome (AIDS). It should be noted that the word "virus" in the phrase "HIV virus" is superfluous and that scientists use the abbreviation "AIDS", not "Aids"!

HIV continues to spread, with millions of new infections each year. AIDS continues to kill large numbers of people, and has become the fourth biggest cause of mortality in the world. The magnitude of this problem has resulted in the allocation of huge resources to the study of these viruses, major objectives being the development of anti-viral drugs and a vaccine. So far there has been some success in achieving the first of these objectives.

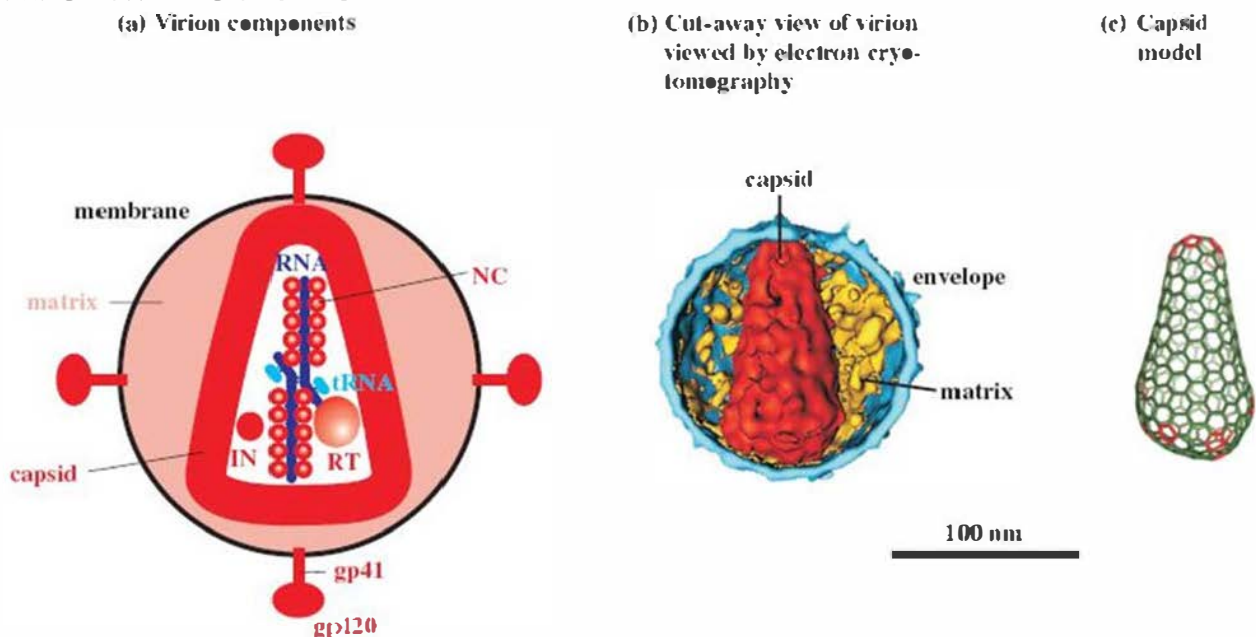
HIV-1 is largely responsible for the AIDS pandemic, while HIV-2 is mainly restricted to West Africa. Consequently the emphasis of this chapter is on HIV-1, as it has been studied more intensively than HIV-2.

18.2 HIV VIRION

The virion has the general characteristics of retroviruses described in the previous chapter but, in contrast to most retroviruses, the capsid is cone-shaped (Figure 18.1), with a diameter of 40–60 nm at the wide end and about 20 nm at the narrow end. Most of the capsid protein is present as hexamers, with five pentamers at the narrow end of the cone and seven pentamers at the wide end of the cone. Generally, there is one capsid per virion, though virions with two or more capsids have been reported.

Figure 18.1 HIV virion. (a) Virion components. IN: integrase. NC: nucleocapsid protein. RT: reverse transcriptase. The TM and SU glycoproteins indicated are those of HIV-1 (gp41 and gp120). (c) Capsid model, showing protein hexamers in green and pentamers in red.

Sources: (b) Grünewald and Cyrklaff (2006) *Current Opinion in Microbiology*, 9, 437. (c) Ganser-Pornillos, Yeager, and Sundquist (2008) *Current Opinion in Structural Biology*, 18, 203. (b) and (c) reproduced by permission of Elsevier Limited and the authors

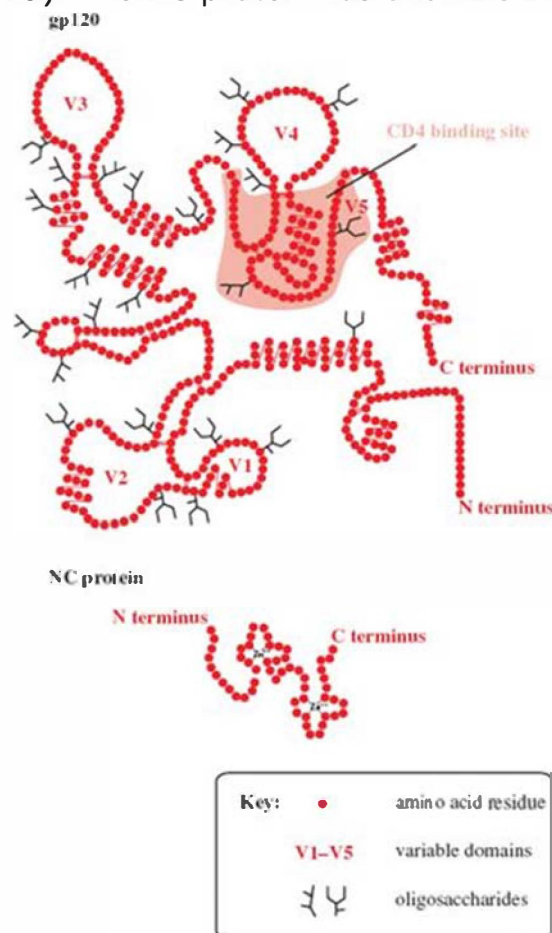


The diameter of the HIV virion measured in negatively stained preparations is in the range 80–110 nm, while results from cryo-electron microscopy are at the upper end of this range or greater.

The TM and SU proteins of HIV-1 have approximate molecular weights of 41 kD and 120 kD, respectively, and are named gp41 and gp120 (gp = glycoprotein). gp120 is heavily glycosylated and has five domains near the surface that are highly variable (V1–V5; Figure 18.2). The C terminus of gp41 is inside the virion, where it is bound to the MA protein. Spikes can be seen at the surface of virions in electron

micrographs, with an average of 14 spikes on each virion. Each spike is a gp41–gp120 trimer. The equivalent glycoproteins in HIV-2 (gp38 and gp130) are unrelated to those of HIV-1, whereas most of the internal proteins of the two viruses are related. The NC proteins are similar to those of many other retroviruses; they have zinc fingers (Figure 18.2) and are highly basic (29% of the amino acid residues of HIV-1 NC protein are basic).

Figure 18.2 Structures of HIV-1 gp120 and NC protein. gp120 has five domains that are highly variable (V1–V5). The NC protein has two zinc fingers.



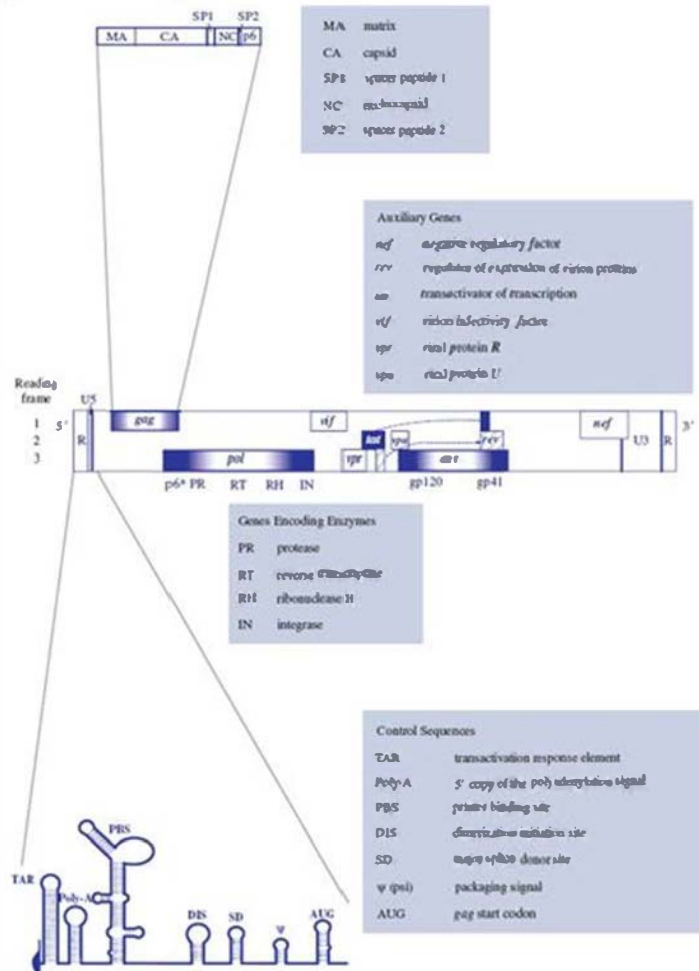
As well as the standard retrovirus proteins, the HIV-1 virion also contains the following virus proteins: Nef, Vpr, Vif, p1, p2, p6, and p6. The presence of host proteins has also been reported, including major histocompatibility complex class II proteins associated with the envelope, and cyclophilin A associated with the capsid.

18.3 HIV GENOME

HIV-1 and HIV-2 have genomes about 9.3 kb in length. In addition to the standard retroviral genes gag, pol, and env (Figure 17.1) there are auxiliary genes, so the viruses are classed as complex retroviruses. The proteins encoded by the auxiliary genes can interact with a wide range of HIV and cell components. Each protein has multiple roles, and together they control virus gene expression, modify the host's immune response, and are involved in transporting virus components within the cell.

The organization of the HIV-1 genome is shown in Figure 18.3. All three reading frames are used and there is extensive overlapping; for example, part of *vpu* in frame 2 overlaps *env* in frame 3. The sequences for *tat* and *rev* are split, the functional sequences being formed when the transcripts are spliced.

Figure 18.3 HIV-1 genome organization



HIV-2 has similar genes to HIV-1, except that it has no *vpu* gene, but it has a *vpx* gene, which is related to *vpr*.

18.4 HIV-1 REPLICATION

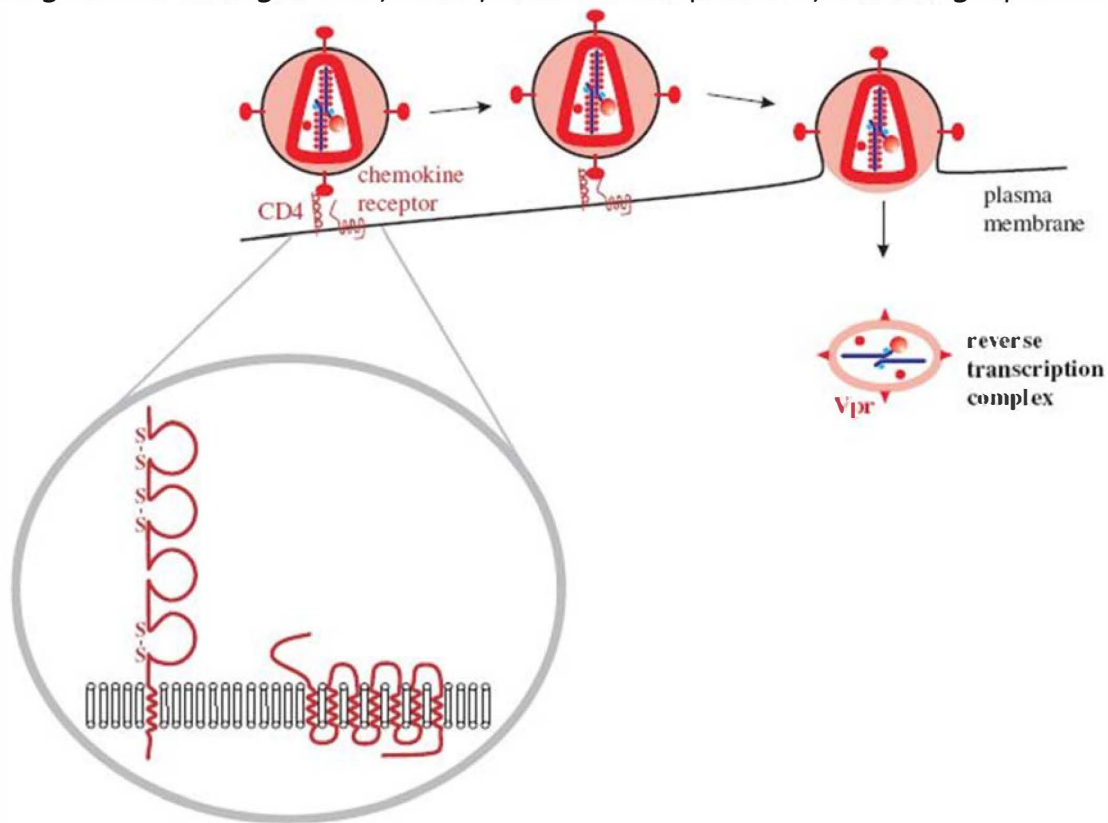
A general description of the retrovirus replication cycle was given in the previous chapter. Here we shall concentrate on details specific to HIV-1.

18.4.1 Attachment and entry

The cell receptor for HIV-1 is CD4 (Figure 18.4), which is a member of the immunoglobulin superfamily of molecules. CD4 is found on several cell types, including helper T cells and some macrophages; CD4 T cells are the main host cells. Attachment of the virion occurs when a site on gp120 (Figure 18.2) recognizes a site on the outer domain of CD4.

Figure 18.4 HIV-1 attachment and entry. The receptor is CD4; each of the loops represents an immunoglobulin-like domain, three of which are stabilized by disulfide bonds. The co-receptor is a chemokine receptor. Fusion of the virion and cell

membranes releases a structure that forms a reverse transcription complex, consisting of the virus genome, tRNA, and several proteins, including Vpr.



As well as binding to receptor molecules gp120 must also attach to a co-receptor on the cell surface. The molecules that act as co-receptors have seven transmembrane domains and are chemokine receptors. During immune responses they bind chemokines and these interactions control leukocyte trafficking and T cell differentiation. Most chemokines fall into one of two major classes, determined by the arrangement of cysteine residues near the N terminus: C—C and C—X—C, where C = cysteine and X = any amino acid. The chemokine receptors are designated CCR and CXCR, respectively. A number of these molecules on T cells act as co-receptors for HIV-1, particularly CCR5 and CXCR4 (Table 18.1).

Table 18.1 Co-receptors for HIV-1 strains and categories of T cell infected.

HIV-1 strain	Main co-receptor(s) used		CD4 T cells infected	
	CCR5	CXCR4	Naive cells	Memory cells
R5	+	-	-	+
R5X4	+	+	+	+
X4	-	+	+	+

Please see Chapter 9 for explanation of the different classes of T cell.

Most HIV-1 strains use CCR5 and are known as R5 strains. It is interesting to note that in some individuals who have had multiple exposures to the virus, but have not become infected, there is a 32-nucleotide deletion in the CCR5 gene. Individuals who are heterozygous for this mutation are less susceptible to infection with HIV-1,

while those who are homozygous express no CCR5 on their cells and are highly resistant. The mutation is found mainly in Europeans.

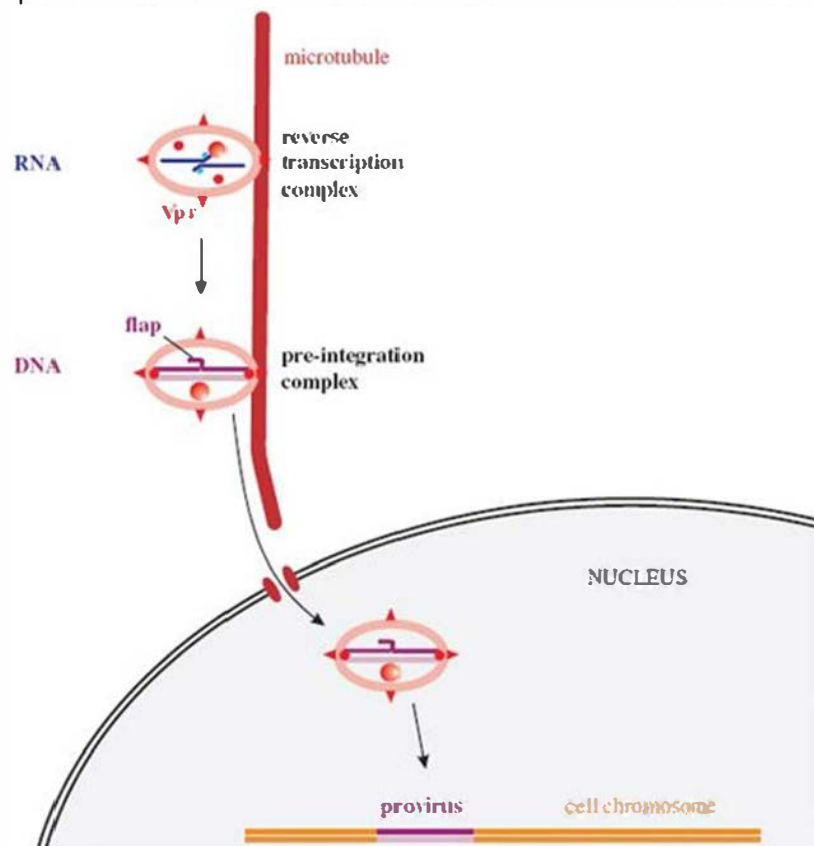
HIV-1 strains that use CXCR4 as a co-receptor are known as X4 strains, and there are some HIV-1 strains (R5X4 strains) that can use either co-receptor. R5 strains do not infect naive T cells, but all three strains infect memory T cells.

The interaction of gp120 with the receptor and co-receptor results in a dramatic rearrangement of gp41, which proceeds to fuse the membranes of the virion and the cell. The matrix and nucleocapsid are released into the cytoplasm and develop into the reverse transcription complex, which contains the MA, Vpr, RT, and IN proteins, as well as the virus genome and tRNA (Figure 18.4).

18.4.2 Reverse transcription and transport to the nucleus

The reverse transcription complex associates rapidly with microtubules (Figure 18.5). Reverse transcription is primed by tRNA_{lys3}, and proceeds as outlined in Section 17.3.2.

Figure 18.5 HIV-1 reverse transcription and integration of the provirus. A DNA flap is formed during reverse transcription. There is evidence that the pre-integration complex is transported toward the nucleus via the microtubule network.



Whereas the proviruses of most retroviruses are entirely dsDNA, those of HIV and other lentiviruses have a short triple-stranded sequence known as a central DNA flap (Figure 18.5). This comes about because there are two initiation sites for (+) DNA synthesis; as well as the polypurine tract (PPT) toward the 3' end of the virus

genome, there is also a central PPT within the pol region. Synthesis of the (+) DNA initiated at the 3' PPT stops soon after reaching the (+) DNA initiated at the central PPT, resulting in a short overlapping DNA sequence. This DNA flap plays a vital role in the early stages of infection.

After reverse transcription has been completed, the pre-integration complex, which contains host proteins as well as virus proteins, is transported into the nucleus. There is evidence that transport toward the nucleus takes place on the microtubule network. As discussed in Chapter 17, most retroviruses can productively infect only if there is breakdown of the nuclear membranes. The pre-integration complex of HIV, however, can enter an intact nucleus, such as that of a resting T cell or a macrophage, and is presumably transported through a nuclear pore.

The viral integrase and cell enzymes are involved in the integration of the provirus into a cell chromosome; cell enzymes remove the DNA flap and repair the gap. There is evidence that integration of the provirus in a resting memory CD4 T cell may result in a latent infection. Latently infected cells can provide a reservoir of infection that is significant for the survival of the virus in individuals receiving anti-retroviral drug therapy. In many cells, though, provirus integration is the prelude to a productive infection in which two phases of gene expression can be distinguished.

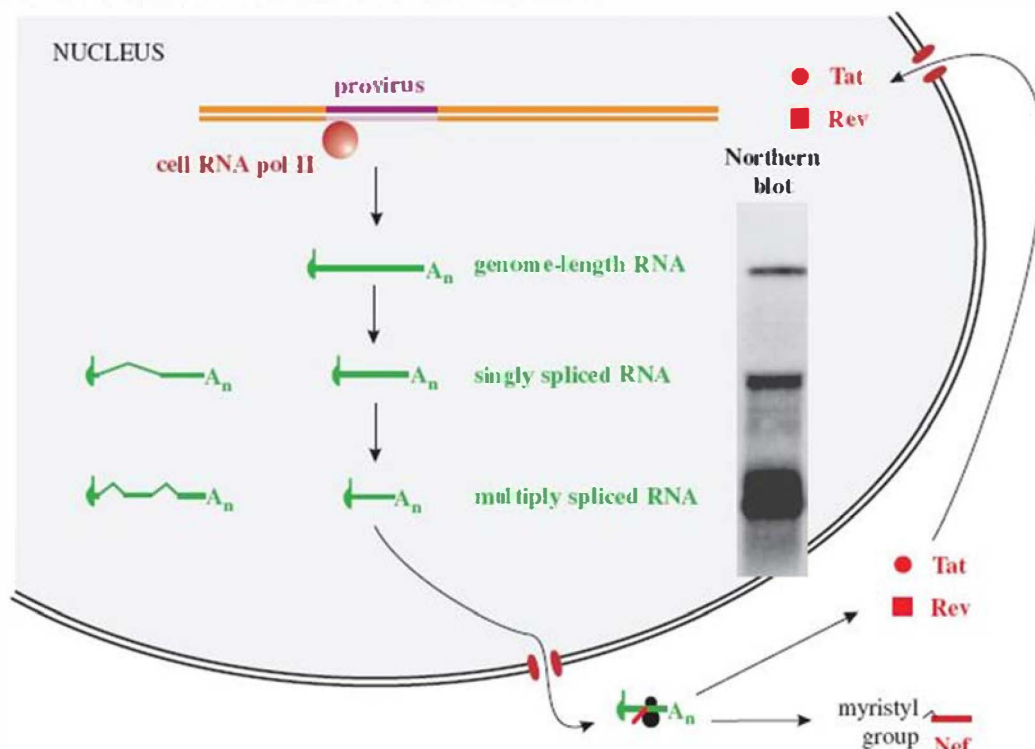
18.4.3 Early gene expression

Transcription is initiated after cell transcription factors bind to promoter and enhancer sequences in the U3 region of the upstream LTR. This region has binding sites for a number of cell transcription factors, including AP-1, NF- κ B, and Sp-1 (Figure 6.4). Transcription is terminated in the downstream LTR; the polyadenylation signal AATAAA is in the R region and transcripts are polyadenylated at the R-U5 junction.

Many of the transcripts are spliced, resulting in three size classes of virus transcript. These can be detected in infected cells using northern blotting (Figure 18.6). The largest RNAs are genome length (about 9.3 kb), while the other two size classes are each made up of a number of mRNA species that have undergone splicing; mRNAs that have been spliced once are around 4.5 kb, while mRNAs that have undergone two or more splicing events ("multiply spliced" transcripts) are around 2 kb. The virus genome has a number of splice donor sites (one is indicated in Figure 18.3) and acceptor sites; these enable splicing events that result in more than 40 mRNA species.

Figure 18.6 HIV-1 early gene expression. Genome-length RNA is transcribed then much of it is spliced, giving rise to two further size classes of RNA that can be detected in northern blots of RNA from infected cells. Early in infection most of the RNA is multiply spliced and is transported to the cytoplasm, where the Tat, Rev, and Nef proteins are translated. Nef is myristylated and performs a number of roles in the cytoplasm, while Tat and Rev are transported to the nucleus.

Source: Northern blot from Malim et al. (1990) Cell, 60, 675; reproduced by permission of Elsevier Limited and the authors.



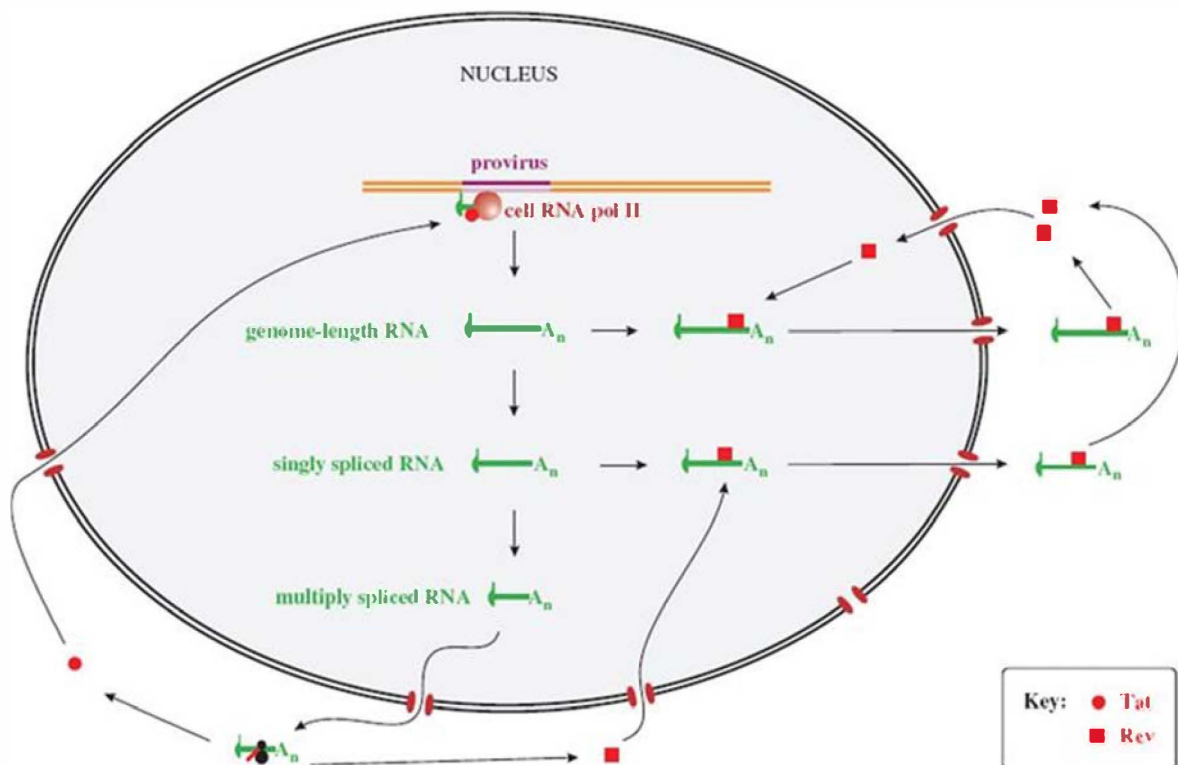
Early in infection most of the primary transcripts are multiply spliced and these RNAs are translated into the Nef, Tat, and Rev proteins (Figure 18.6).

18.4.3.a Roles of Nef, Tat, and Rev

The Nef (Negative regulatory factor) protein acquired its name because it was originally thought to have an inhibitory effect on HIV replication, though later work showed that this protein enhances replication in a number of ways! Nef reduces the quantity of CD4 in the plasma membrane, making it less likely that another HIV virion will bind to the cell and initiate a superinfection. Reduced CD4 at the cell surface also means reduced interference with the release of progeny virions. Nef can also bind to MHC proteins, reducing their expression at the cell surface and helping to shield HIV-infected cells from immune surveillance.

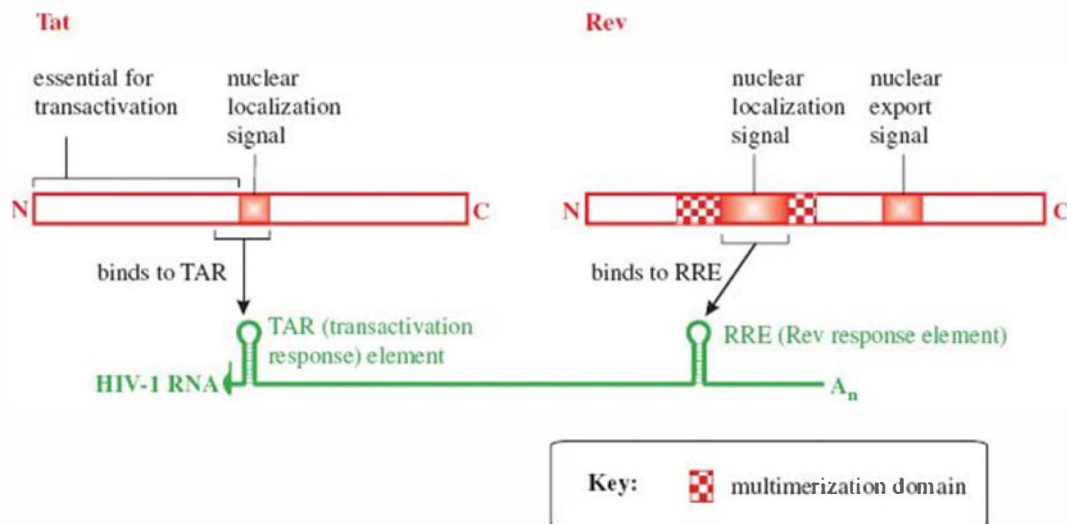
The roles of the Tat and Rev proteins are summarized in Figure 18.7.

Figure 18.7 Roles of Tat and Rev. Tat binds to nascent transcripts and helps to ensure that the entire virus genome is transcribed. Rev binds to genome-length RNA and singly spliced RNA and aids their transport to the cytoplasm, where the late proteins are translated. Rev is recycled to the nucleus.



The Tat (Transactivator of transcription) protein plays an important role in enhancing transcription. A nuclear localization signal (Figure 18.8) directs Tat to the nucleus, where it binds to a sequence at the 5' end of nascent virus transcripts: this RNA sequence is known as the transactivation response (TAR) element (Figures 18.3 and 18.8). Cell proteins also bind to TAR, and among these proteins is a kinase, which phosphorylates components of the RNA polymerase complex. Phosphorylation increases the processivity of the enzyme along the proviral template. Tat therefore functions as a transcription factor, but an unusual one in that it binds to RNA, not to DNA. In the absence of Tat most transcripts are incomplete, though early in infection sufficient are completed to allow the synthesis of a small amount of Tat, which then significantly boosts the synthesis of genome-length RNA.

Figure 18.8 HIV-1 Tat and Rev proteins and their binding sites in the virus RNA. The TAR and RRE regions of the RNA have complex secondary structures. The RRE is present in genome-length RNA and the singly spliced RNAs, but it is absent from the multiply spliced RNAs.



The other early protein, Rev (Regulator of expression of virion proteins; Figure 18.8), is responsible for a shift from early to late protein synthesis. The cell protein importin β binds to a nuclear localization signal on Rev and transports it into the nucleus where it binds to the Rev response element (RRE) in the virus RNA. The RRE is present in the unspliced and singly spliced transcripts, but is absent from the multiply spliced transcripts. More Rev molecules bind and form multimers through protein–protein and protein–RNA interactions. After binding, Rev undergoes a conformational change that masks the nuclear localization signal and exposes the nuclear export signal. Cell proteins associate with the bound Rev. The RNA–Rev–cell protein complexes are transported through nuclear pores to the cytoplasm, where the complexes dissociate. Rev is recycled to the nucleus (Figure 18.7) while the late proteins are translated from the genome-length and singly spliced mRNAs.

18.4.4 Late gene expression

Translation of the late proteins is shown in Figure 18.9. Gag and Gag–Pol are translated from unspliced transcripts, with Gag–Pol translated when a ribosomal frameshift takes place (Section 17.3.5). This occurs on roughly 5% of occasions when a ribosome crosses the frameshift signal UUUUUUA at the junction of the NC and SP2 domains of gag (Figure 18.3). This sequence (known as a slippery sequence), together with a downstream secondary structure, causes the ribosome to slip from reading frame 1 to reading frame 3 (Figure 18.10). The slippery sequence is reminiscent of the 7-U sequence, over which the RNA polymerase “stutters,” at the end of each rhabdovirus gene (Section 15.4.2). After the frameshift has taken place translation continues through the pol region, yielding the Gag–Pol polyprotein.

Figure 18.9 HIV-1 late gene expression. Vpu and Env are translated from singly spliced RNAs in the rough endoplasmic reticulum. The inset shows translation of Vpu and Env from a bicistronic mRNA. Env is synthesized when the vpu start codon is bypassed during leaky scanning. The remaining proteins are translated on free

ribosomes: Vif and Vpr from singly spliced RNAs, and Gag and Gag–Pol from genome-length RNAs.

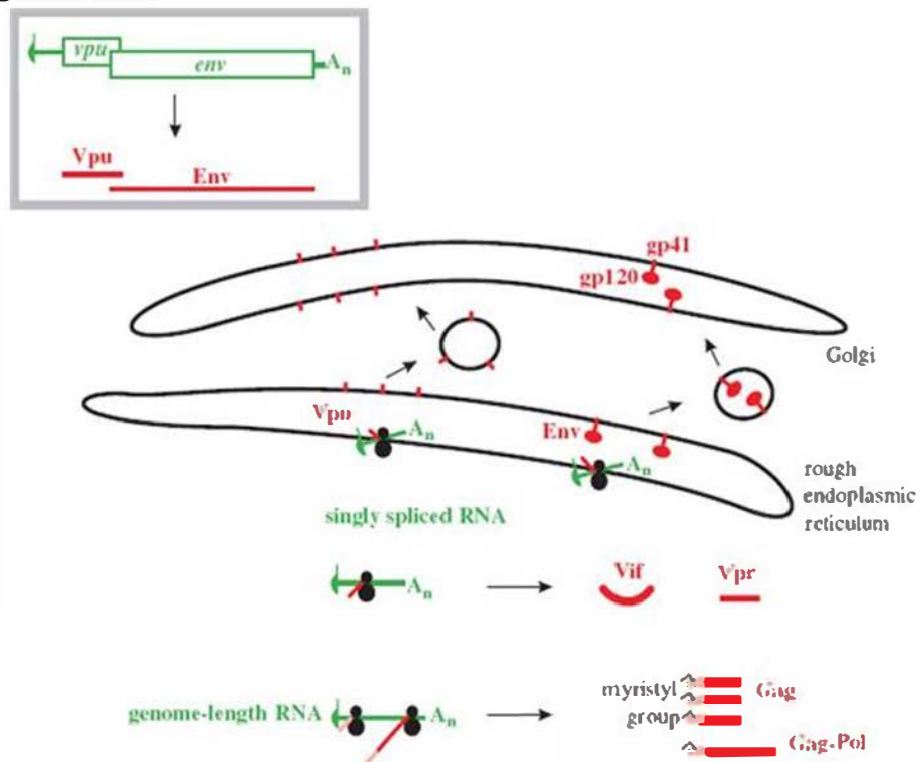
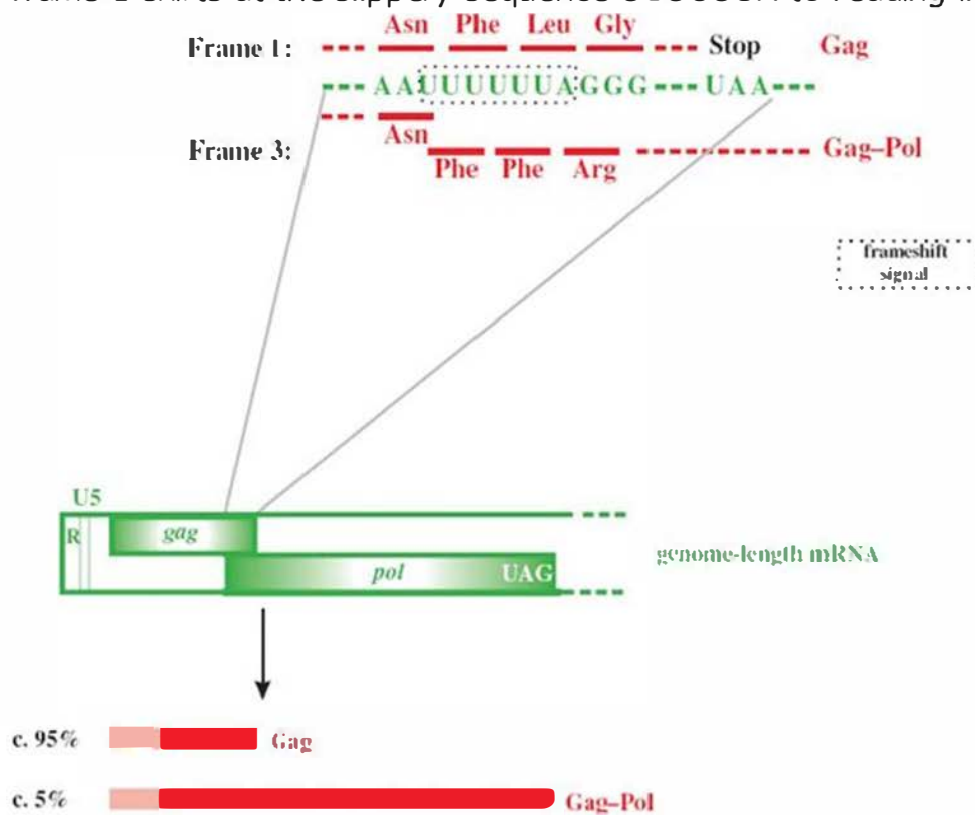


Figure 18.10 Expression of HIV-1 Gag–Pol by ribosomal frameshifting. A ribosome reading in frame 1 shifts at the slippery sequence UUUUUUA to reading in frame 3.



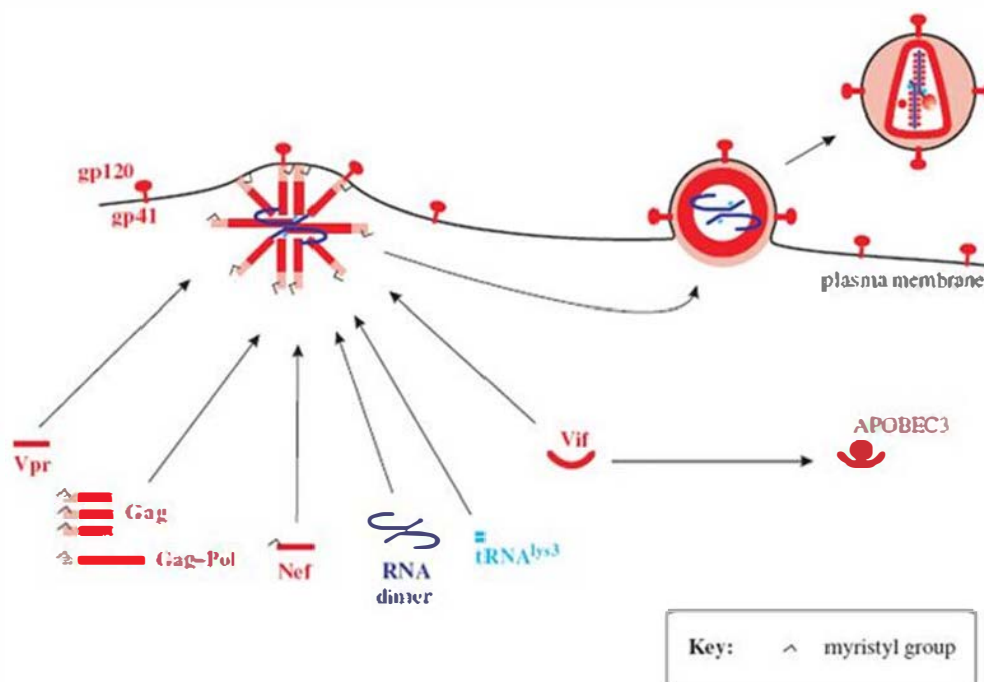
The remaining virus proteins (Vif, Vpr, Vpu, and Env) are translated from singly spliced transcripts. Vpu and Env are translated in the rough endoplasmic reticulum from a bicistronic transcript (Figure 18.9). Env becomes heavily glycosylated, acquiring an apparent molecular weight of 160 kD. Trimers of Env are formed

before the molecules are cleaved to form the envelope proteins gp120 and gp41; the cleavage is carried out by furin, a host protease located in the Golgi complex. Some of the virus proteins are phosphorylated, including Vpu and the MA region of some Gag molecules.

18.4.5 Assembly and exit of virions

General aspects of retrovirus virion assembly were described in Section 17.3.6. Here we focus on aspects specific to the assembly of HIV-1 virions; these are summarized in Figure 18.11.

Figure 18.11 Assembly of HIV-1 virions. All of the virus proteins in the diagram are incorporated into new virions, along with two copies of the virus genome and cell tRNA_{Lys3}. Vif binds to APOBEC3 proteins and prevents their incorporation into virions.



Formation of the RNA dimer that will constitute the genome of a new virion commences by base pairing between complementary sequences in the loop of the dimerization initiation site near the 5' end of each RNA (Figure 18.3). Gag and Gag-Pol molecules bind to the RNA dimer and form an orderly arrangement. The proteins bind via the basic regions and zinc fingers of their NC domains (Figure 18.2); initial binding is to the packaging signal ψ in the virus genome (Figure 18.3). Other domains of Gag and Gag-Pol bind a number of virus and cell proteins, some of them becoming incorporated into the virion. CA domains bind the host protein cyclophilin A, while p6 domains bind the virus protein Vpr. Nef and Vif are also incorporated into virions. p6 domains also bind cell proteins thought to be important for the release of the budding virion from the host cell. These proteins bind to p6 sequences known as late (L) domains (Section 8.3.1); an example is the sequence proline-threonine-alanine-proline (PTAP) which binds tumor suppressor gene 101 protein.

As well as becoming a virion component, the Vif protein also plays an important role in ensuring the exclusion from progeny virions of cell enzymes (APOBEC3F and APOBEC3G) that could interfere with replication in the next host cell (Section 9.2.1.d). If the vif gene is mutated these enzymes can be incorporated into virions; when reverse transcription is under way in the next host cell the enzymes can deaminate deoxycytidine to deoxyuridine, inducing lethal mutations. Vif prevents the incorporation of these enzymes into virions by binding to them and forming a complex of cell proteins that target the APOBEC3 proteins for degradation.

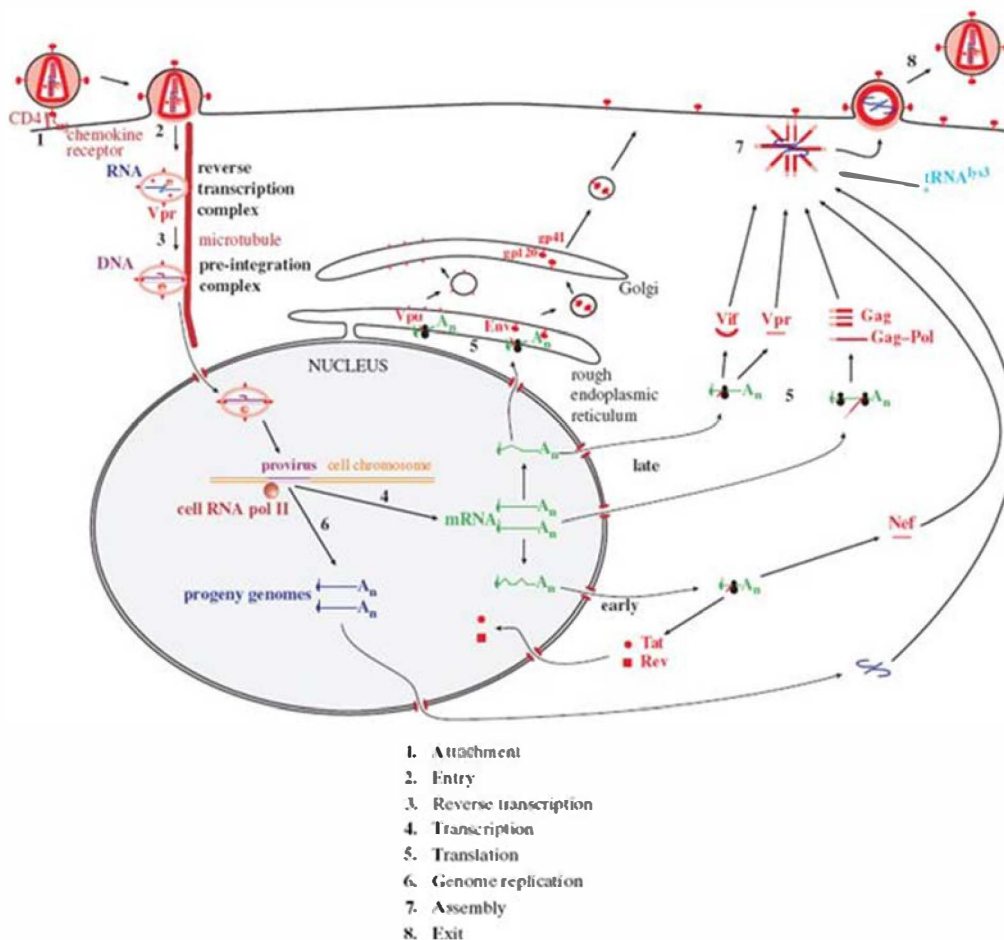
Budding occurs at cholesterol-rich regions of the plasma membrane known as lipid rafts. Vpu is an integral membrane protein and is required for efficient budding of HIV-1. It counteracts the effect of the cell protein tetherin, which would otherwise "tether" budding virions to the plasma membrane and prevent their release. Vpu also reduces the expression of CD4 at the cell surface, supplementing the activity of the Nef protein (Section 18.4.3.a).

During or after budding the virus protease is activated. The protease is a dimer, derived from Gag–Pol dimers that undergo self-cleavage; it cleaves the Gag and Gag–Pol polyproteins into the constituents of the mature virion. The CA molecules are arranged into a lattice of 12 pentamers plus many hexamers, to form the conical capsid.

18.4.6 Overview of HIV-1 replication

The HIV-1 replication cycle is summarized in Figure 18.12.

Figure 18.12 The HIV-1 replication cycle. Note the additional step in retrovirus replication: reverse transcription, which takes place between entry and transcription.



18.5 HIV-1 VARIABILITY

During the replication of retrovirus genomes there is a high error rate as a result of the lack of a proofreading mechanism (Section 21.3.3.a) and further variability may be generated through the process of recombination (Section 21.3.3.b). HIV-1 is a particularly variable virus, having evolved into a number of groups and subtypes; its variability is manifest in characters such as its antigens, host cell range, and resistance to drugs.

18.5.1 Antigens

HIV-1 antigens show a high degree of variability. The surface protein gp120 is one of the most variable, in spite of the constraint imposed by the overlap between the env and vpu genes (Figure 18.3). There are five domains of gp120 that are especially variable (Figure 18.2). The exceptionally high variability of gp120 presumably results from evolutionary pressure exerted by the immune response of the host. Interestingly, the Nef protein, which is not on the virion surface, is also highly variable.

18.5.2 Host cell range

The existence of HIV-1 strains with different co-receptor preferences was mentioned in Section 18.4.1. Transmission of HIV-1 to a new host is almost always with R5 strains, and these predominate during the acute and asymptomatic phases

of infection. In about 50% of people that develop AIDS the virus evolves to give rise to X4 and R5X4 strains.

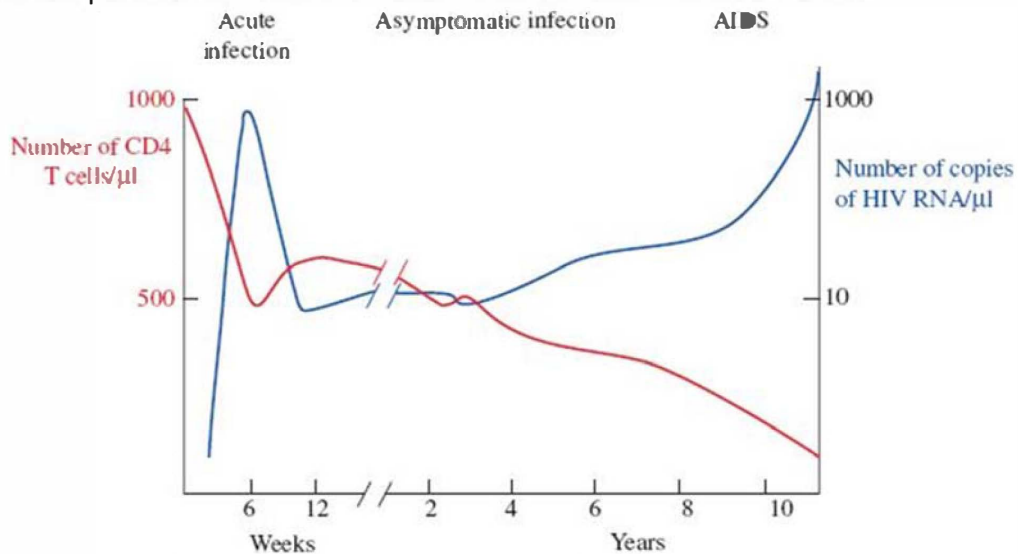
18.5.3 Resistance to drugs

The presence of anti-retroviral drugs in the body of an infected host exerts an evolutionary pressure on the virus and drug-resistant variants can quickly emerge. HIV-1 drug resistance is discussed further in Chapter 26.

18.6 PROGRESSION OF HIV INFECTION

Shortly after a person becomes infected with HIV there is a huge rise in viremia (the presence of virus in the blood; Figure 18.13), and in some people there is an illness resembling glandular fever or influenza. The host's immune responses then control virus replication to some extent and there is a period of asymptomatic infection. In the absence of intervention with drugs this period typically lasts for 8–10 years, but it may be significantly shorter or longer than this, depending on characteristics of both the host and the virus.

Figure 18.13 Levels of CD4 T cells and HIV RNA in blood during HIV infection. The concentration of HIV RNA is a measure of viremia. Shortly after infection viremia rises, then it falls off and relatively low levels are detectable throughout the asymptomatic period. A rise in viremia heralds the onset of AIDS.



Extensive virus replication continues throughout the asymptomatic period, with estimates of more than 10^{10} HIV-1 virions produced each day. The infection persists in spite of immune responses against the virus and there are several reasons for this.

- Antibodies against gp120 and gp41 do not bind readily to HIV-1 virions because of the small number of spikes in the virion envelope. There is an average of 14 spikes per virion.
- Cells that play crucial roles in immune responses (CD4 T cells and macrophages) are killed by HIV infection.

- There is evidence that HIV-1 components induce apoptosis of non-infected CD4 T cells.
- As the infection proceeds the virus evolves, producing new antigenic variants that are not recognized by the antibodies and T cells produced earlier.
- In latently infected cells the virus is shielded from the immune system.

In about 50% of infected individuals X4 and R5X4 strains of HIV-1 emerge during the asymptomatic period and the co-receptor preference shifts from CCR5 to CXCR4 (Section 18.5.2). Individuals in whom X4 strains emerge at an early stage are likely to progress more rapidly to AIDS.

For a while, the body is able to tolerate the onslaught on the immune system by rapidly replacing most of the cells that have been destroyed, but the concentration of CD4 T cells in the blood steadily declines until a point is reached where the level of viremia rises sharply and AIDS develops (Figure 18.13). AIDS is characterized by infections with pathogenic micro-organisms. Brain disease and/or cancer may also develop; virus-associated cancers, such as Kaposi's sarcoma and non-Hodgkin's lymphoma (Chapter 23) are the most common types of cancer in AIDS.

In comparison with HIV-1, HIV-2 infections are associated with longer asymptomatic periods, slower progression of disease, and lower rates of transmission.

18.7 PREVENTION OF HIV TRANSMISSION

As well as the measures that individuals can take to prevent transmission of HIV, there are also measures that society can take using a variety of tools and procedures, some of them developed by virologists. It is important that potential donors of blood, organs, and semen are screened for HIV infection; this can be done by testing their blood for HIV-specific antibodies (Chapter 2). The preparation of blood products for hemophiliacs can include treatment with lipid solvents and detergents to destroy the virions of HIV (and other enveloped viruses, such as hepatitis B virus).

There is a risk of transmitting HIV, and several other viruses (e.g. hepatitis B and C viruses), if syringes and needles are used to inject more than one person. The use of "auto-disable" syringes ensures that this cannot happen. This simple, but ingenious, invention results in the syringe plunger breaking if an attempt is made to use the syringe more than once.

The risk of HIV transmission from a mother to her child is between 15 and 40%, the higher levels of risk being associated with breast-feeding. This risk can be greatly reduced by anti-retroviral drug treatment of the woman before and after birth. In many countries, including the UK, treatment of HIV-infected women has resulted in a decline in the number of HIV-positive children. Although drug

treatment reduces the risk of virus transmission to children, it does not cure infected individuals.

Anti-retroviral drugs are given as post-exposure prophylaxis to protect individuals from HIV infection following needlestick injuries and risky sexual activities. The use of drugs against HIV is discussed in detail in Chapter 26.

A major goal of HIV research is the development of an effective vaccine. The deployment of such a vaccine could dramatically reduce transmission rates of the virus, but after the expenditure of much effort there is no sign on the horizon of an HIV vaccine suitable for mass immunization. There are a number of reasons for this, including the ability of the virus to rapidly evolve antigenic variants as a result of its high mutation rate (Section 18.5).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- explain the importance of HIV;
- describe, with the aid of a labeled diagram, the HIV-1 virion;
- describe the HIV-1 genome;
- write an illustrated account of the replication cycle of HIV-1;
- explain some of the roles of the HIV-1 auxiliary gene products;
- discuss the variability of HIV-1;
- discuss the effects of HIV infection on the host;
- evaluate approaches to the prevention of HIV transmission.

SOURCES OF FURTHER INFORMATION

Books

Guatelli, J. C. et al. (2009) Human immunodeficiency virus. Chapter 33 in *Clinical Virology*, 3rd edition (Richman, D. D., Whitley, R. J., and Hayden, F. G., editors), ASM Press

Harrington, P. R. and Swanstrom, R. (2008) The biology of HIV, SIV and other lentiviruses. Chapter 20 in *Sexually Transmitted Diseases*, 4th edition (Holmes K. K. et al., editors), McGraw-Hill

Jeang, K.-T. (ed.) (2007) *HIV-1: Molecular Biology and Pathogenesis*, Academic Press

Levinson, W. (2008) Chapter 45 in *Review of Medical Microbiology and Immunology*, 10th edition, McGraw-Hill

Levy, J. A. (2007) *HIV and the Pathogenesis of AIDS*, 3rd edition, ASM Press

Spearman, P. and Freed, E. O. (ed.) (2009) *HIV Interactions with Host Cell Proteins*, Current Topics in Microbiology and Immunology series, Volume 339, Springer

Journals

Carlson, L.-A. et al. (2010) Cryo electron tomography of native HIV-1 budding sites. *PLoS Pathogens*, 6 (11): e1001173

Das, A. T., Harwig, A., and Berkhout, B. (2011) The HIV-1 Tat Protein has a versatile role in activating viral transcription. *Journal of Virology*, 85, 9506–9516

Grewe, B. and Überla, K. (2010) The human immunodeficiency virus type 1 Rev protein: ménage à trois during the early phase of the lentiviral replication cycle. *Journal of General Virology*, 91, 1893–1897

Groom, H. C. T., Anderson, E. C., and Lever, A. M. L. (2009) Rev: beyond nuclear export. *Journal of General Virology*, 90, 1303–1318

Hatzioannou, T. and Bieniasz, P. D. (2011) Antiretroviral restriction factors. *Current Opinion in Virology*, 1, 526–532

Killian, M. S. and Levy, J. A. (2011) HIV/AIDS: 30 years of progress and future challenges. *European Journal of Immunology*, 41, 3401–3411

Moir, S., Chun, T.-W., and Fauci, A. S. (2011) Pathogenic mechanisms of HIV disease. *Annual Review of Pathology: Mechanisms of Disease*, 6, 223–248

Murakami, T. (2008) Roles of the interactions between Env and Gag proteins in the HIV-1 replication cycle. *Microbiology and Immunology*, 52, 287–295

Permanyer, M., Ballana, E., and Esté, J. A. (2010) Endocytosis of HIV: anything goes. *Trends in Microbiology*, 18, 543–551

Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011) Atomic-level modelling of the HIV capsid. *Nature*, 469, 424–427

Poropatich, K. and Sullivan, D. J. (2011) Human immunodeficiency virus type 1 long-term non-progressors: the viral, genetic and immunological basis for disease non-progression. *Journal of General Virology*, 92, 247–268

Stolp, B. and Fackler, O. T. (2011) How HIV takes advantage of the cytoskeleton in entry and replication. *Viruses*, 3, 293–311

Warrillow, D., Tachedjian, G., and Harrich, D. (2009) Maturation of the HIV reverse transcription complex: putting the jigsaw together. *Reviews in Medical Virology*, 19, 324–337

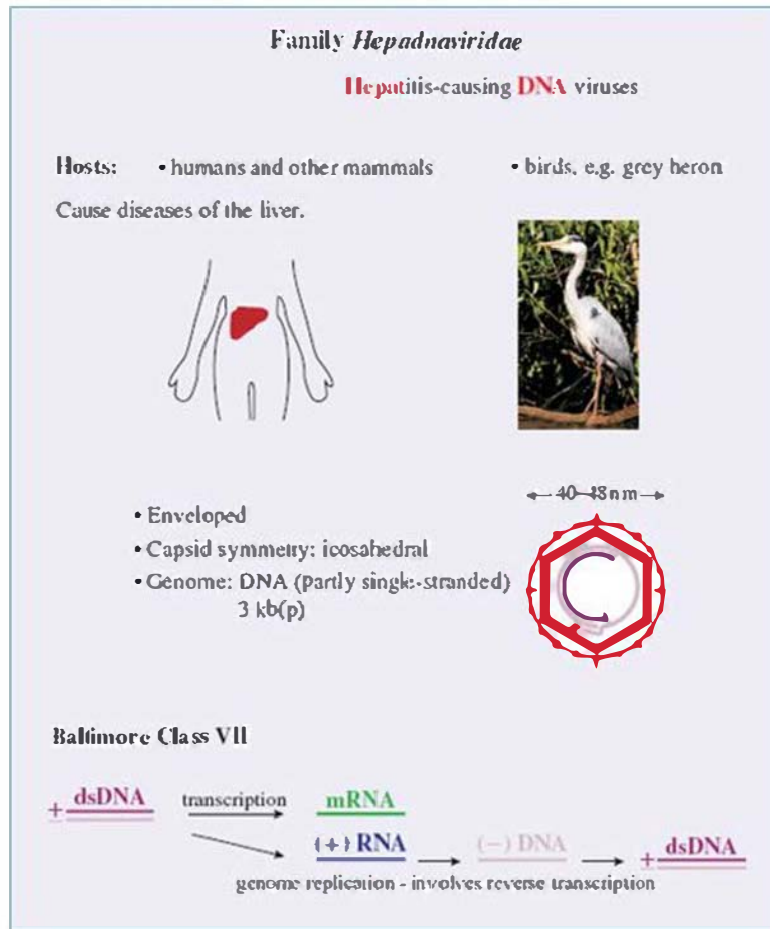
Watts, J. M. et al. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, 460, 711–716

CHAPTER 19

Hepadnaviruses (and Other Reverse-Transcribing DNA Viruses)

CHAPTER 19 AT A GLANCE

Photograph of grey heron courtesy of Robert Carter.



19.1 INTRODUCTION TO HEPADNAVIRUSES

The hepadnaviruses got their name because they cause hepatitis and they have DNA genomes. They are known as hepatitis B viruses (HBVs) and are classified in the family Hepadnaviridae. Some members infect mammals and some infect birds; examples include woodchuck hepatitis virus and heron HBV. The best known hepadnavirus is that which infects humans; it is commonly referred to as HBV, and is of major importance as an agent of disease and death. Duck HBV, on the other hand, is non-pathogenic in its natural host.

The hepadnaviruses are especially fascinating for two reasons. First they have very small genomes, which are used with great economy to encode the virus proteins and to control expression of the virus genes. Second, their DNA genomes are

replicated via an RNA intermediate. In other words, their replication involves reverse transcription, so they are very different from DNA viruses that replicate their DNA directly to DNA. The discovery of the mode of replication of the hepadnavirus genome led to the creation of Baltimore Class VII.

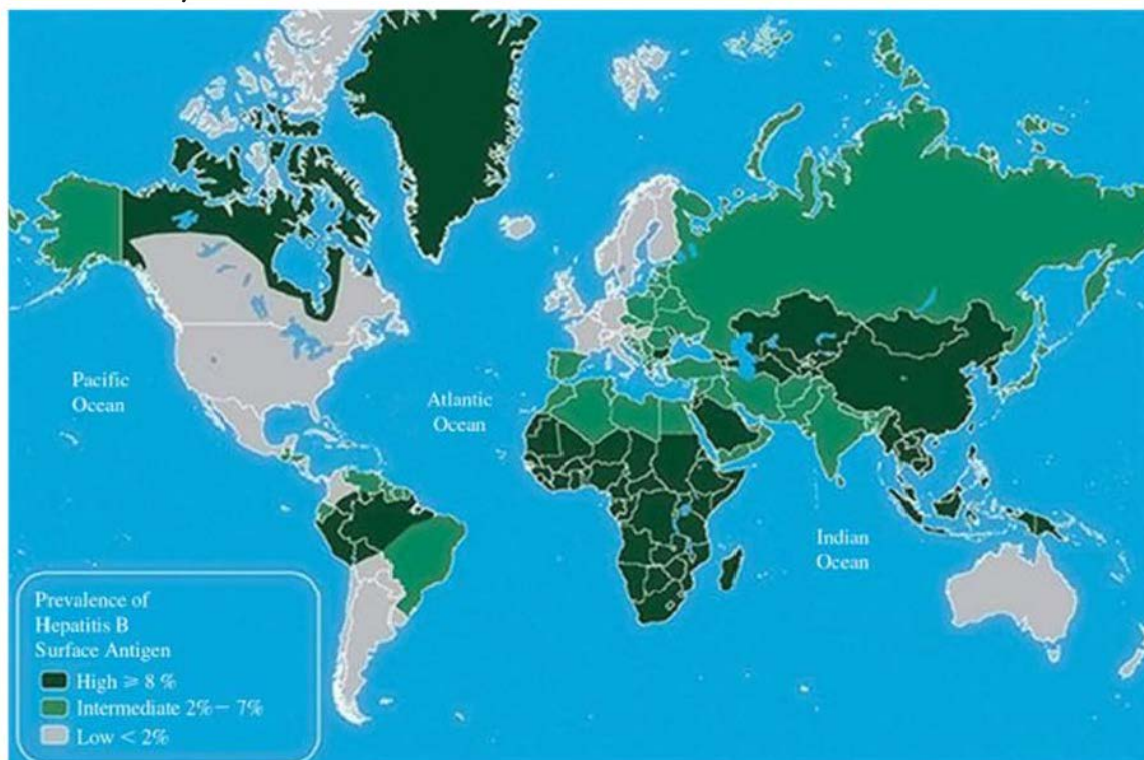
DNA viruses that replicate via RNA have also been found in plants; an example is cauliflower mosaic virus. These viruses, together with the hepadnaviruses, have been termed pararetroviruses. This chapter will concentrate on HBV.

19.2 IMPORTANCE OF HBV

No one knows how many people are infected with HBV, but the figure is estimated to be around 350 million. The majority are in Asia and many are in Africa (Figure 19.1). There are high percentages of people infected in the far north of North America and Greenland, but because of the low populations of these regions the numbers infected are relatively small.

Figure 19.1 World distribution of HBV infection.

Source: Courtesy of the US Centers for Disease Control and Prevention.



Virus is present in the blood and semen of infected individuals and the modes of transmission generally parallel those for HIV transmission. There are over 50 million new HBV infections each year, the majority in babies who acquire the infection from their mothers. Many infections are thought to result from the reuse of syringes and needles for injections, mainly in the developing world.

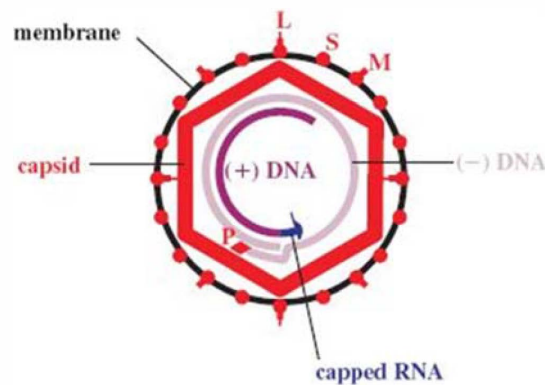
Many HBV infections result in mild symptoms or are asymptomatic, especially in children. It is in children, however, that HBV infection is most likely to become persistent, with 90%–95% of those infected as newborn infants becoming long-term carriers, compared with 1%–10% of those becoming infected as adults.

Individuals who are persistently infected with HBV may remain healthy for much of the time, but some develop severe hepatitis, which may lead to cirrhosis and eventually to liver cancer (Chapter 23). These diseases resulting from HBV infection cause over half a million deaths each year.

19.3 HBV VIRION

The virion is roughly spherical, with a diameter of about 42 nm. The main virion components are an envelope enclosing a capsid, inside which is the DNA and P (polymerase) protein (Figure 19.2).

Figure 19.2 The HBV virion. L, M, S: large, medium, and small envelope proteins. P: polymerase.



19.3.1 DNA

The genome is made up of two strands of DNA, one of which is incomplete; hence the DNA is partly single-stranded and partly double-stranded. A short sequence is triple-stranded as a result of a complementary sequence at the 5' ends, and this results in the DNA having a circular conformation (Figure 19.2). The genome is very small, with a length of about 3.2 kb(p).

At the 5' end of each of the DNA strands there is a covalently linked molecule: a capped RNA on the short strand and a protein (P) on the long strand.

19.3.2 P (polymerase) protein

The virion contains at least one molecule of P (polymerase) protein. The N terminus of P constitutes a terminal protein domain; this is separated by a spacer from a domain with reverse transcriptase activity (Figure 19.3). At the C terminus there is a domain with ribonuclease H (RNase H) activity. P also has DNA-dependent DNA polymerase activity.

Figure 19.3 Domains of the HBV P (polymerase) protein.



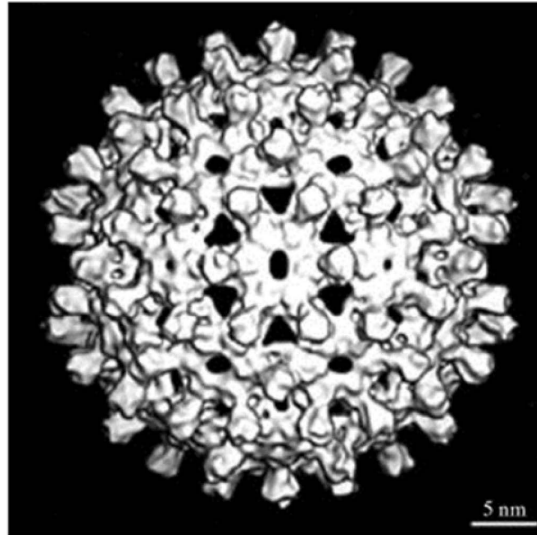
19.3.3 Capsid

The capsid, which has icosahedral symmetry, has holes in it and short spikes protrude from its surface (Figure 19.4). It is constructed from dimers of the C (core) protein, which is largely α -helical, unlike the capsid proteins of many other

viruses. The C terminus of the C protein is highly basic due to the presence of a large number of arginine residues; this region binds the virus genome.

Figure 19.4 HBV capsid. Derived from cryo-electron microscopy images of capsids assembled in *E. coli* cells expressing HBV C protein.

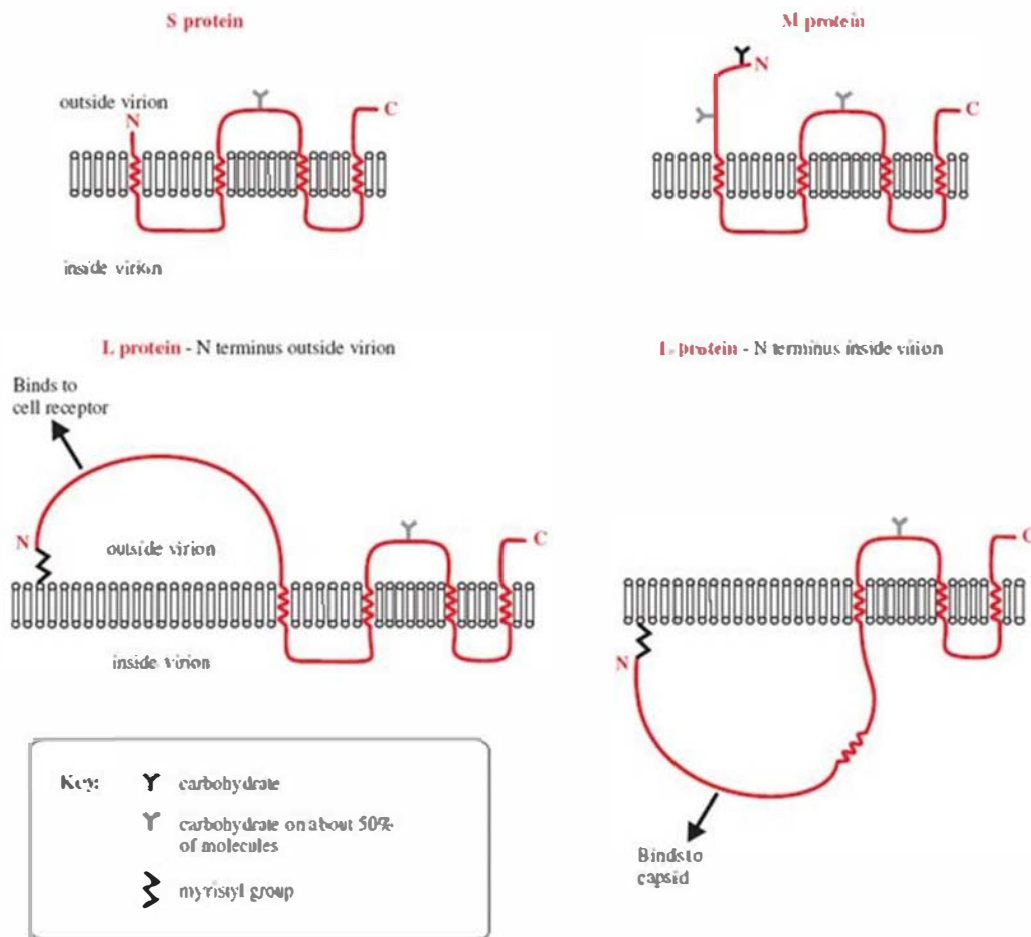
Source: Watts et al. (2002) *The EMBO Journal*, 21, 876. Reproduced by permission of Nature Publishing Group and the authors.



19.3.4 Envelope

The virion envelope contains three protein species designated as small (S), medium (M) and large (L). The M and L proteins are longer versions of the S protein (Figure 19.5), which is the most abundant of the three. Each protein has one or more glycosylation sites, though not all molecules are glycosylated. The surface regions of the envelope proteins constitute an antigen known as hepatitis B surface antigen (HBsAg).

Figure 19.5 Arrangement of HBV small (S), medium (M), and large (L) proteins in the virion envelope. The N termini of about 50% of L protein molecules are outside the membrane and about 50% are inside.



The binding site for cell receptors is near the N terminus of the L protein, but only about 50% of the L molecules have the N terminus on the outside of the virion; the N termini of the remaining L molecules are on the inside bound to the capsid.

19.4 NON-INFECTIOUS PARTICLES

An unusual and intriguing feature of HBV infection is the presence in the blood of not only virions but also large quantities of non-infectious particles that have been released from infected liver cells (Figure 19.6). These particles are composed of lipid and virus envelope proteins, but they do not contain nucleocapsids. Some of the particles are spheres and some are filaments; both have diameters of 22 nm and the filaments have variable lengths up to 200 nm.

Figure 19.6 HBV structures present in the blood of an infected person.

Structure	Diameter (nm)	Typical concentration ($\mu\text{g/ml}$)
Virion	42	0.1
Non-infectious particles		
Sphere	22	100
Filament	22	1

All the particles (virions and non-infectious particles) are much more abundant in the blood than in the liver, and the non-infectious particles, especially the spheres, vastly outnumber the virions. Presumably the virus has a good reason for investing in the production of such large amounts of non-infectious material, but that reason is not clear. The non-infectious particles may be a decoy for antibodies, thereby providing virions with some protection from the host's immune system.

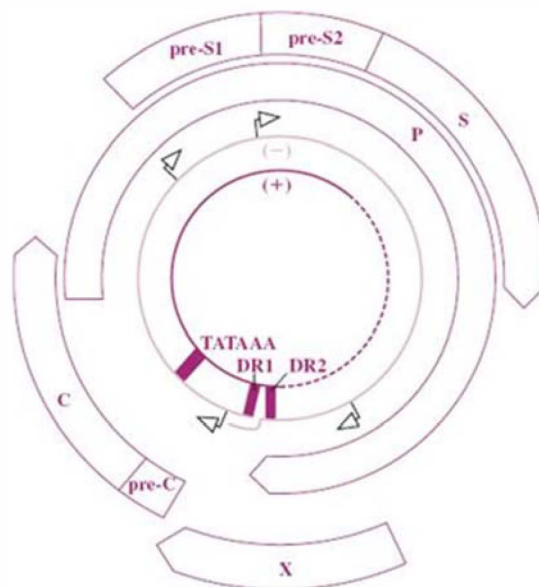
19.5 SOLUBLE VIRUS PROTEIN

In addition to the virions and non-infectious particles a soluble virus protein is found in the blood of some infected individuals. This protein is known as hepatitis B e antigen (HBeAg). It is similar to the C protein, but has 10 extra amino acid residues at the N terminus and lacks 34 amino acid residues at the C terminus. The function of HBeAg, like that of the non-infectious particles, may be to protect the virus from the immune response of the host.

19.6 HBV GENOME

At 3.2 kb, the HBV genome is very small, though there are viruses with smaller genomes (Section 3.2.1). There are four ORFs, from which seven proteins are translated (Figure 19.7), so a large amount of coding information is packed into the small genome. The virus cleverly achieves this by using every nucleotide in the genome for protein coding and by reading more than half of the genome in two reading frames. The P ORF, which occupies about 80% of the genome, overlaps the C and X ORFs and the S ORF is entirely within the P ORF. A further way in which the virus maximizes its coding capacity is by expressing the L protein in two different conformations that have different functions. In one conformation the L protein displays a site which can bind to a cell receptor, while in the other it binds the virion envelope to the capsid (Figure 19.5).

Figure 19.7 HBV genome organization. The incomplete (+) strand and the complete (–) strand of the DNA are shown surrounded by the four ORFs.



Key:	ORF's	C:	core protein
		P:	polymerase
		S:	small protein
		pre-S2 + S:	M protein
		pre-S1 + pre-S2 + S:	L protein
		△	promoter
		TATAAA	polyadenylation signal
		DR1 DR2	direct repeats

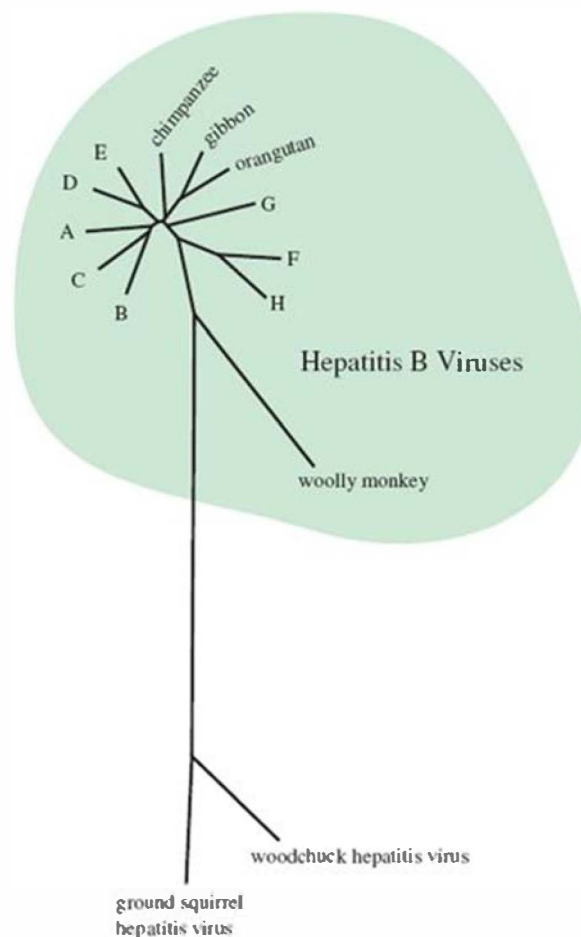
As the entire genome is involved in coding for protein, it follows that all the regulatory sequences, such as promoters, are within protein coding sequences. The genome contains direct repeats of 11 nucleotides known as DR1 and DR2.

Translation of the S region produces the S protein, translation of pre-S2 \rightarrow S produces the M protein and translation of the complete pre-S1 \rightarrow pre-S2 \rightarrow S ORF produces the L protein. Similarly, expression of the pre-C \rightarrow C region gives rise to two proteins. Even though the virus maximizes the use of its small genome it encodes only seven proteins, so it is heavily dependent on host cell functions. Host proteins that have been demonstrated to play roles in the replication of HBV include enzymes, transcription factors, and chaperones.

19.7 HBV GENETIC GROUPS

Sequencing of many HBV isolates has revealed eight genetic groups (genotypes A \rightarrow H), and has demonstrated that these human viruses are related to similar viruses in other primate species (Figure 19.8).

Figure 19.8 Phylogenetic tree showing relationships between hepatitis B viruses from humans (genotypes A \rightarrow H) and other primate species. Relationships with hepatitis viruses from two other mammalian species (woodchuck and ground squirrel) are also shown.



The genetic groups are fairly restricted geographically; genotype A predominates in Northern Europe while genotypes B and C are mainly in Asia.

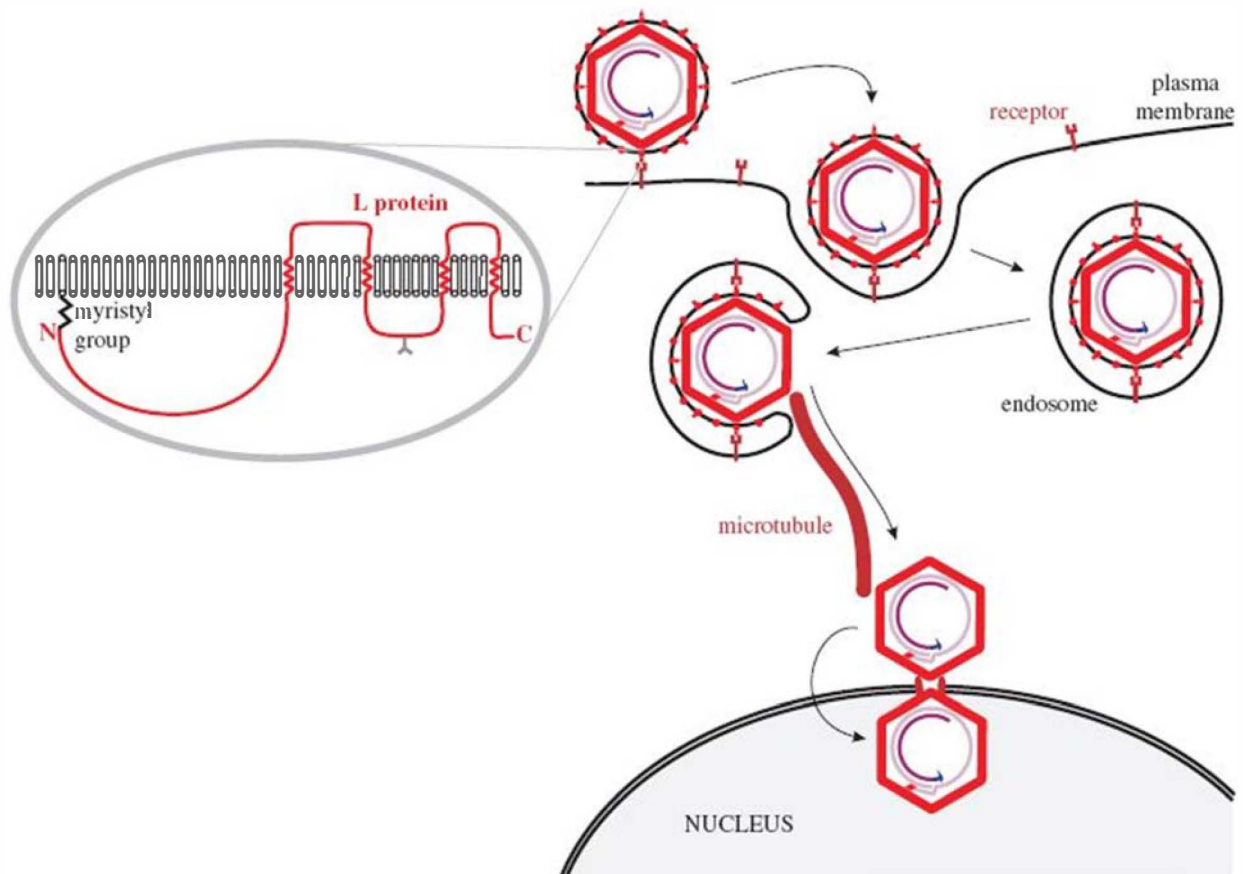
19.8 HBV REPLICATION

Hepatocytes (liver cells) are the host cells for HBV in the body. For a long time studies on HBV replication were limited because primary cultures of human hepatocytes were the only cells available that would support the complete replication cycle of HBV. Some hepatoma cell lines could be infected using HBV DNA (a procedure known as transfection), but this did not permit study of the processes of entry and uncoating. New cell culture systems, including a hepatoma cell line, are available that support the complete replication cycle of HBV.

19.8.1 Attachment

The identity of the hepatocyte receptor to which HBV attaches is not yet known, though various candidates have been proposed, such as polymeric IgA and annexin V. The virus binds to the receptor via a site on its L protein (Figure 19.9).

Figure 19.9 Attachment and entry of HBV. The binding site on the virus is near the N terminus of the L protein. Endocytosis is followed by release of the nucleocapsid and transport to the nucleus.

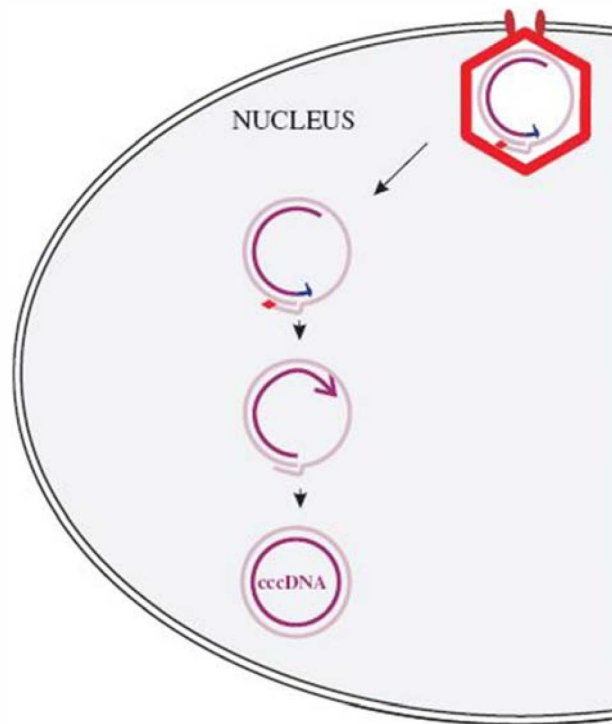


19.8.2 Entry

The virion is endocytosed then the nucleocapsid is released from the endosome by fusion of the virion and endosome membranes. The nucleocapsid is transported to the nucleus on a microtubule. In Figure 19.9 the nucleocapsid is shown entering the nucleus, as evidence has been published that HBV capsids can pass through nuclear pores. It is not certain, however, whether this happens, or whether uncoating occurs within a nuclear pore, or whether the genome is released into the nucleus with the capsid remaining in the cytoplasm.

Once the virus genome is free in the nucleus it is converted into a circular DNA molecule (Figure 19.10). The covalently bound P protein is removed from the 5' end of the minus strand, which is shortened to remove the third strand of the triple-stranded region. The RNA is removed from the 5' end of the plus strand, while DNA synthesis at the 3' end makes the entire molecule double-stranded. The ends of each strand are then ligated to form covalently closed circular DNA (cccDNA). It is uncertain whether all these modifications to the virus genome are carried out by host enzymes, or whether the virus P protein plays a role in some.

Figure 19.10 Release of HBV genome from the capsid and conversion into cccDNA.



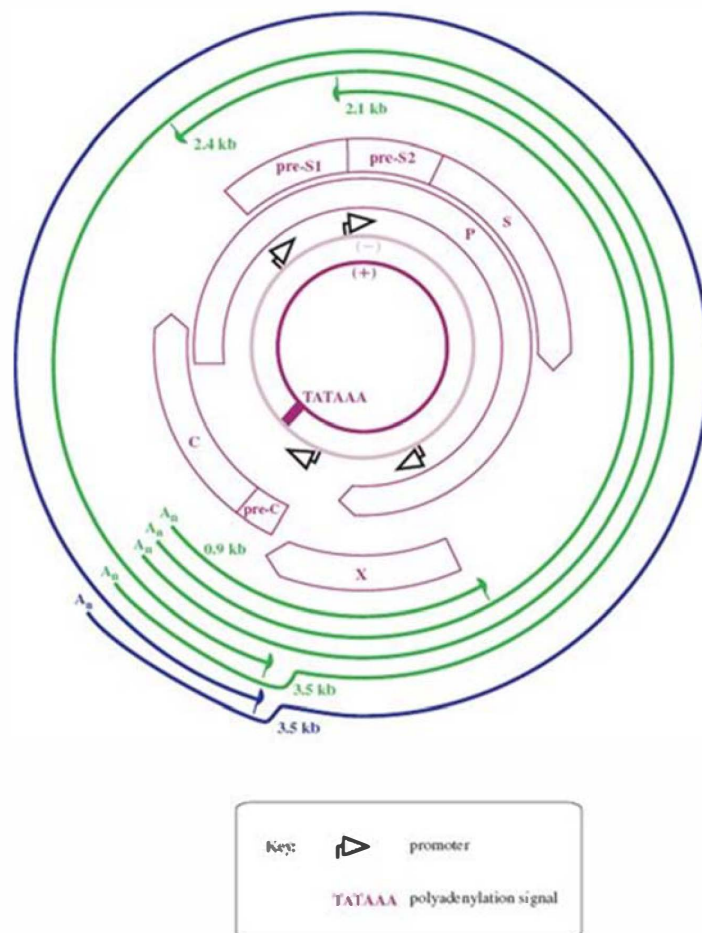
The virus DNA is not replicated in the nucleus, but more copies are brought into the nucleus later in the replication cycle.

19.8.3 Transcription

The cccDNA is the template for transcription. You will recall (Figure 19.7) that the HBV genome has four promoters, sited upstream of the pre-S1, pre-S2, X, and pre-C regions. Only the pre-S2 promoter has a TATA box. At least two of the promoters are highly specific to liver cells and it has been shown that some of the transcription factors involved are liver cell proteins. This is at least part of the explanation why HBV is specific to liver cells. Rates of transcription from all four promoters are controlled by binding of cell transcription factors to two enhancer sequences present in the HBV genome. It is thought that the virus protein X is also a transcription factor, though it appears not to bind to DNA and so is probably not a typical transcription factor.

Transcription is carried out by the cell RNA polymerase II and results in the synthesis of four size classes of RNA (Figure 19.11). All the transcripts are capped at the 5' end and polyadenylated at the 3' end. They all use the same polyadenylation signal (TATAAA), so they have a common 3' end.

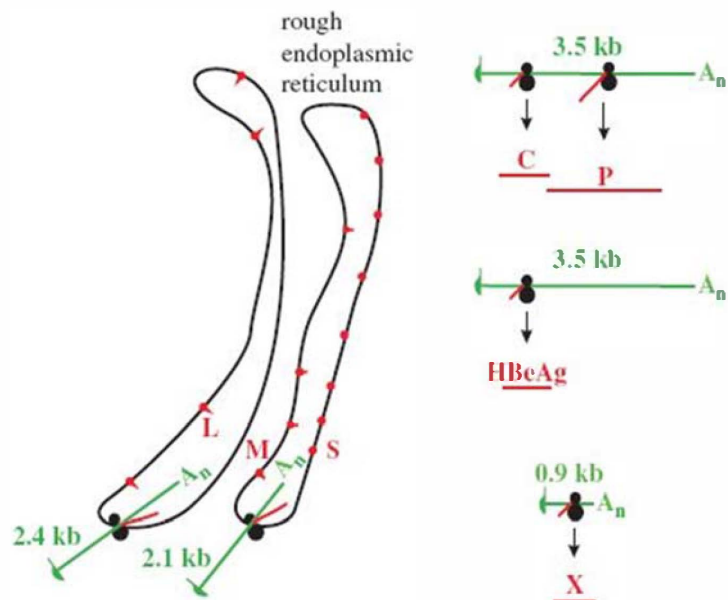
Figure 19.11 HBV transcripts. The four size classes of transcript are shown. Some of the larger RNAs in the 3.5 kb class will function as mRNAs and some will function as pregenomes; these functions are indicated by the green and blue colors respectively.



Note that the 3.5 kb RNAs are longer than the genome, which is 3.2 kb. During their synthesis, part of the genome is transcribed twice, hence these RNAs have direct terminal repeats. In order to synthesize the 3.5 kb RNAs the RNA polymerase must ignore the polyadenylation signal the first time. Studies with duck HBV indicate that this is controlled by a DNA sequence (called PET = positive effector of transcription), which somehow allows the RNA polymerase to continue through the polyadenylation signal the first time. PET suppresses a sequence called NET (negative effector of transcription), which is needed for termination of transcription. The polyadenylation signal is not ignored when the RNA polymerase meets it the second time. If NET is deleted the RNA polymerase makes additional circuits around the DNA, producing very long transcripts.

19.8.4 Translation

Translation of the HBV proteins is depicted in Figure 19.12.
Figure 19.12 Translation of HBV proteins.



Of the RNAs that are approximately 3.5 kb long there are two subsets that differ slightly in size. The shorter subset, which does not include the start codon for the pre-C sequence, acts as mRNA for the C and P proteins. The C ORF is upstream of the P ORF, and most ribosomes that bind to these mRNAs commence translation at the C start codon.

There are at least three further start codons before the P start codon. To translate the P ORF it appears that a ribosome bypasses the upstream start codons by *leaky scanning*. Thus HBV produces much more capsid protein than polymerase, just as retroviruses control the relative amounts of Gag and Pol proteins synthesized (see Section 17.3.5). The shorter subset of 3.5 kb RNAs also acts as pregenomes (see below).

HBeAg is translated from the longer subset of the 3.5 kb mRNAs, which include the start codon for the pre-C sequence. After translation HBeAg is secreted from the cell.

There are also subsets of the 2.1 kb mRNAs. The M protein is translated from the longest subset and the S protein from several shorter subsets.

19.8.4.a Post-translational modifications

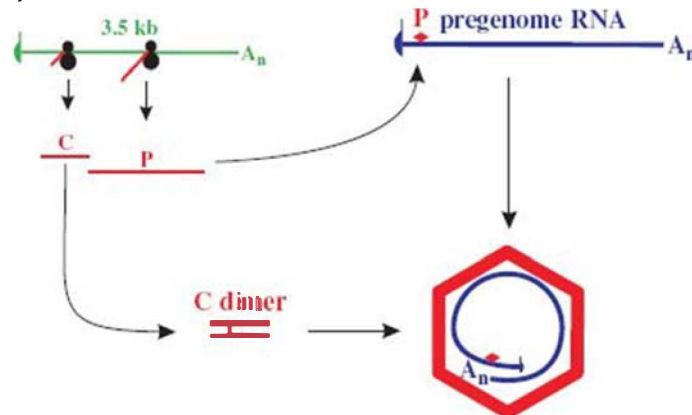
Some of the envelope protein molecules become glycosylated and L is myristylated at the N terminus (Figure 19.5). It has been found that myristylation of L is not required for efficient virion assembly but virions produced in its absence are not infective.

Initially all the L protein N termini are on the cytoplasm side of the endoplasmic reticulum membranes, but after translation about 50% of the N termini are moved through the membranes into the lumen of the endoplasmic reticulum. It is thought that the cell heat shock protein Hsc70 plays a role in this process.

19.8.5 Nucleocapsid assembly

It is known that the C protein can self-assemble, because capsids can be formed in *Escherichia coli* recombinants expressing the HBV C gene (Figure 19.4). The C protein forms dimers, which are assembled into capsids (Figure 19.13). Most of the capsids are composed of 120 dimers, but a minority are built from 90 dimers and are smaller.

Figure 19.13 Early stages of HBV assembly. A capsid is assembled from C protein dimers; it acquires a copy of pregenome RNA bound to P protein and several cell proteins (not shown).



A molecule of P protein, along with several cell proteins, binds to a molecule of 3.5 kb RNA, which will function as pregenome RNA (Figures 19.13 and 19.14). P binds at a site within a sequence known as O_{μ} (Greek letter epsilon) that has a high degree of secondary structure. The O_{μ} sequence is within the terminal repeat and so is present at both ends of the pregenome RNA. For some reason though, P binds only at the 5' end. The O_{μ} structure with its bound P protein acts as the packaging signal for incorporation of the pregenome RNA into a capsid.

Figure 19.14 HBV pregenome with P protein bound near the 5' end. Not to scale. DR1, DR2: direct repeats. The pregenome has terminal repeats, hence there is a copy of DR1 and O_{μ} at each end of the RNA.



19.8.6 Genome synthesis

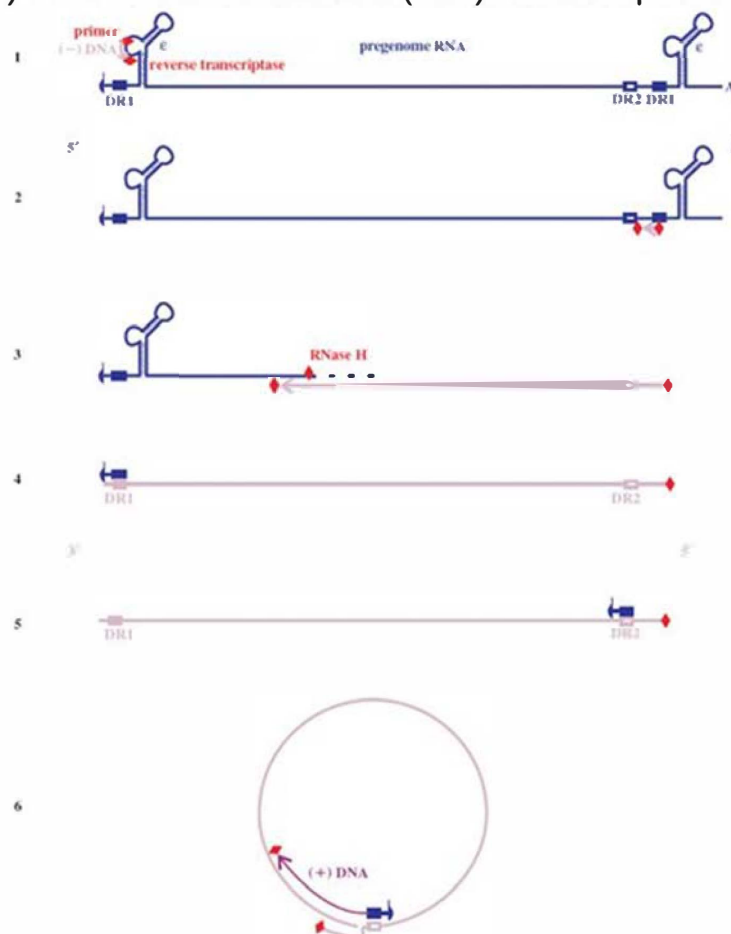
DNA synthesis takes place in the capsid. The pregenome RNA is the template and DNA is synthesized by the reverse transcriptase domain of P (Figure 19.3) from nucleotides that enter through the holes in the capsid. The terminal protein domain of P acts as the primer for the initiation of minus-strand DNA synthesis. A covalent bond is formed between the OH group of a tyrosine residue near the N terminus of P and the first nucleotide. There are conflicting reports as to whether

the terminal protein becomes cleaved from the reverse transcriptase after the initiation of DNA synthesis.

A model for HBV DNA synthesis is outlined in Figure 19.15. Initially a 4-nucleotide sequence of (ϕ⁻) DNA is synthesized. It is transferred to a complementary sequence in DR1 near the 3ϕ⁻I end of the pregenome and DNA synthesis continues to the 5ϕ⁻I end of the pregenome template. The RNase H activity of the P protein degrades the pregenome RNA from the RNAϕ⁻DNA duplex. All the RNA is removed except for a remnant of 15ϕ⁻I18 bases including the cap. The ϕ⁻OH group at the 3ϕ⁻I end of the RNA remnant acts as the primer for the synthesis of (+) DNA.

Figure 19.15 Synthesis of HBV DNA by reverse transcription (model).

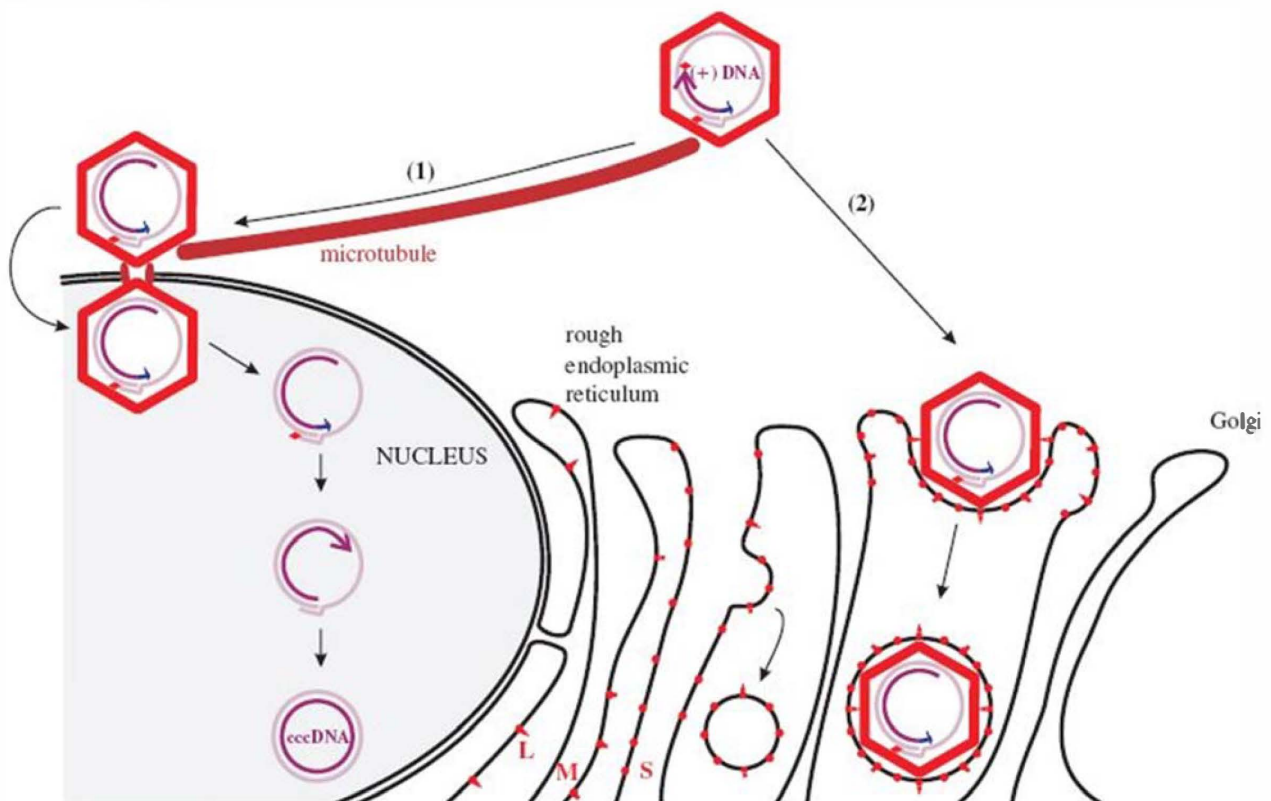
1. Initiation of (ϕ⁻) DNA synthesis near the 5ϕ⁻I end of the pregenome RNA.
2. Transfer of DNA to the same complementary sequence in DR1 near the 3ϕ⁻I end of the pregenome.
3. Continuation of (ϕ⁻) DNA synthesis, and degradation of the pregenome by RNase H.
4. Completion of (ϕ⁻) DNA synthesis. All of the RNA has been degraded, except for a short sequence at the 5ϕ⁻I end.
5. Transfer of RNA to the same complementary sequence near the 5ϕ⁻I end of the (ϕ⁻) DNA.
6. Synthesis of (+) DNA on the circularized (ϕ⁻) DNA template.



19.8.7 Envelope acquisition and exit

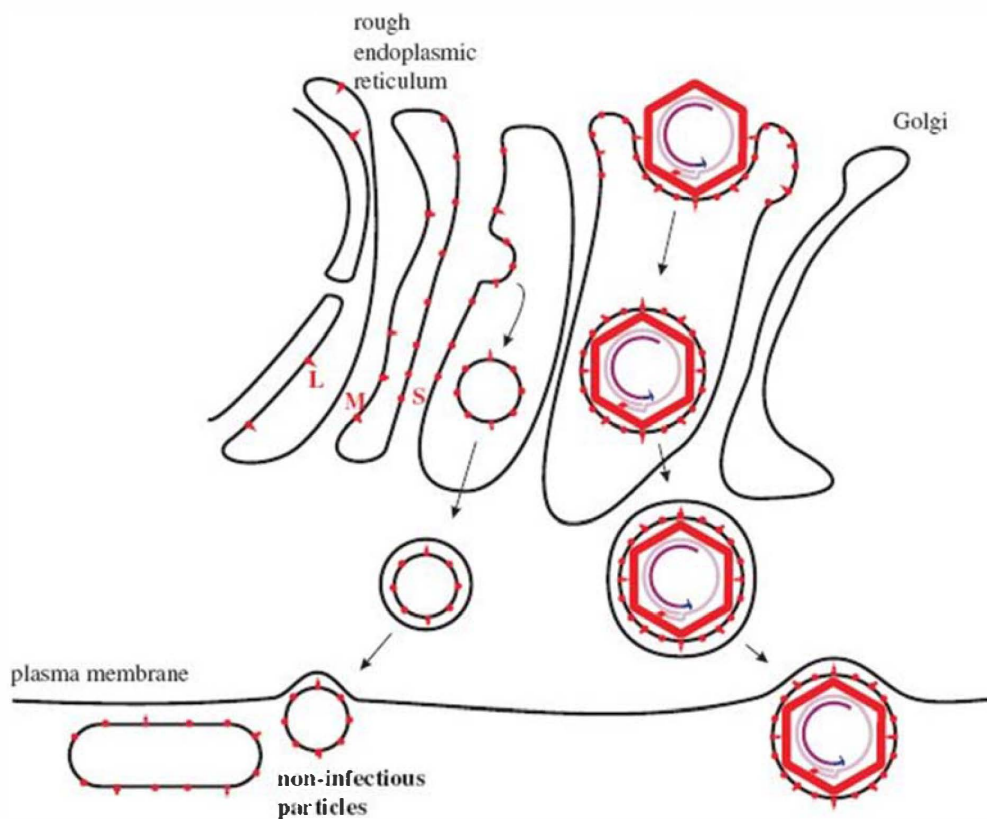
During (+) DNA synthesis a nucleocapsid can either migrate to the nucleus to increase the pool of cccDNA or it can undergo a maturation event that enables it to bud through a membrane containing virus envelope proteins (Figure 19.16). DNA synthesis ceases on budding, as the nucleocapsid is cut off from the pool of nucleotides in the cytoplasm. This explains why the (+) DNA in the virion is incomplete.

Figure 19.16 Roles of progeny HBV nucleocapsids. During DNA synthesis a nucleocapsid may either (1) move to the nucleus where its DNA boosts the pool of virus DNA or (2) bud through a membrane containing virus envelope proteins to form a virion.



The membranes through which budding occurs are part of a compartment between the endoplasmic reticulum and the Golgi complex. As budding proceeds, copies of the L protein with their N termini on the cytoplasmic surfaces of the membranes bind to capsids (Figure 19.5). Virions are transported to the cell surface, where they are released from the cell. It is thought that transport occurs within vesicles, which fuse with the plasma membrane, releasing the virions by exocytosis. Non-infectious spheres and filaments are also released from the cell (Figure 19.17).

Figure 19.17 Exit of HBV virions and non-infectious particles from the infected cell by exocytosis.

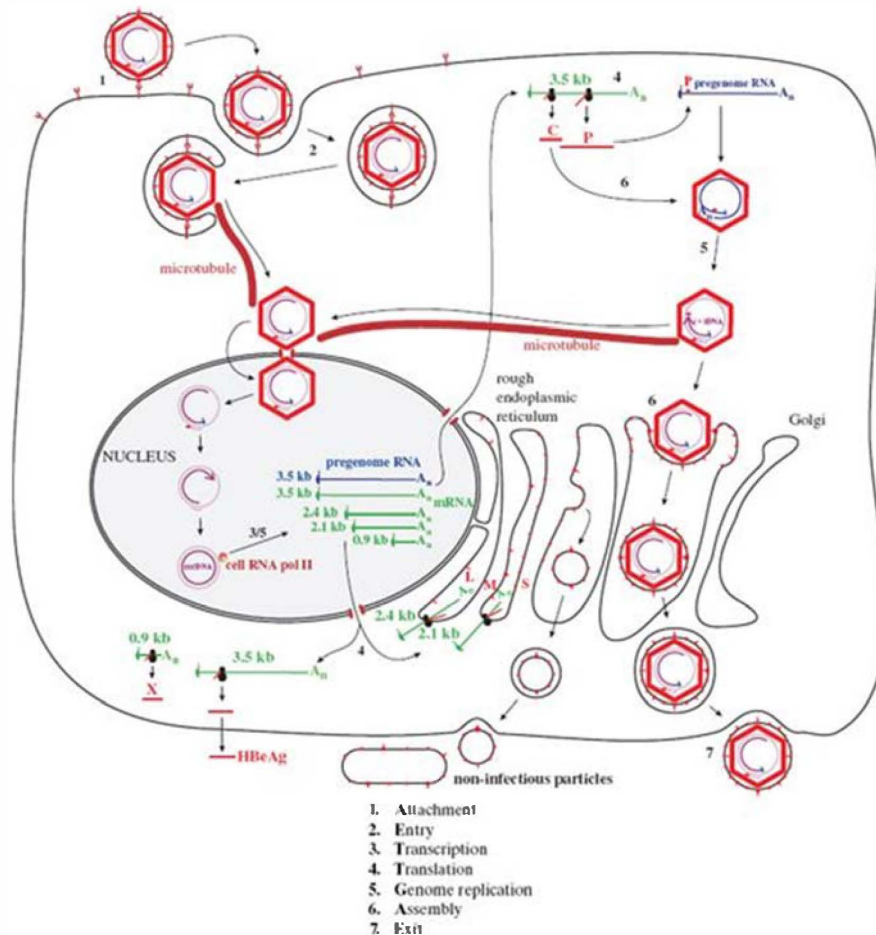


The host hepatocyte is not killed by HBV infection; the cell may survive for many months, releasing large quantities of virions and non-infectious particles. No cytopathic effect is observed in productively infected cell cultures (except for those infected with some HBV mutants). When the liver becomes damaged in an HBV-infected individual it is not as a result of cell damage caused by the virus, but it is due to the killing of HBV-infected cells by the body's immune system.

19.8.8 Overview of HBV replication cycle

A model for HBV replication is outlined in Figure 19.18. Some aspects of this model have yet to be proven unequivocally and much interesting research remains to be done to unearth more details of the replication of this virus.

Figure 19.18 Outline of the HBV replication cycle. Genome replication involves the synthesis of RNA in the nucleus, then copying from RNA to DNA (reverse transcription) in capsids. Stages in virion assembly include capsid construction and acquisition of the envelope by budding. The mode of exit of virions and non-infectious particles from the cell is not shown.



19.9 PREVENTION AND TREATMENT OF HBV INFECTION

The original HBV vaccines consisted of non-infectious spheres and filaments extracted from blood donations from HBV carriers. Now much vaccine is produced using recombinant yeast cells containing the gene for the S protein (Section 25.7). Vaccination programs have been successful in reducing the percentage of HBV carriers in some parts of the world, such as Taiwan and Alaska. The requirements to mass-produce HBV vaccine cheaply and to vaccinate many more of those at risk are major challenges facing the world.

Alpha-interferon (Section 9.2.1.b) has been used to treat HBV-infected persons. This treatment does not eliminate the infection, but it results in a significant reduction in viremia in about 20%–30% of cases. There is a price to pay for interferon treatment in the form of side-effects such as influenza-like symptoms and weight loss, which may necessitate reducing the dosage or even discontinuing the treatment.

The drug lamivudine, a nucleoside analog (Section 26.3.1.c), is used to treat HBV infection. In many ways lamivudine is an improvement over α -interferon as it suppresses virus replication with a low incidence of side-effects, is administered by mouth rather than by injection, and is cheaper. Long-term treatment, however,

results in the appearance of lamivudine-resistant HBV mutants, though they appear to be susceptible to other nucleoside analogs such as adefovir.

19.10 OTHER REVERSE-TRANSCRIBING DNA VIRUSES

There are some plant viruses with dsDNA genomes that replicate by reverse transcription. These viruses are classified in the family Caulimoviridae, which includes viruses with isometric virions, such as cauliflower mosaic virus, and viruses with rod-shaped virions, such as rice tungro bacilliform virus.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- explain the importance of HBV;
- describe the HBV virion and non-infectious particles;
- outline the main features of the HBV genome;
- describe the HBV replication cycle;
- evaluate means of preventing and treating HBV infection.

SOURCES OF FURTHER INFORMATION

Books

Seeger, C., Zoulim, F., and Mason, W. S. (2007) Hepadnaviruses. Chapter 76 in *Fields Virology*, 5th edition (Knipe, D. M. and Howley, P. M., editors-in-chief), Lippincott, Williams and Wilkins

Tavis, J. E. and Badtke, M. P. (2009) Hepadnaviral genomic replication. Chapter 7 in *Viral Genome Replication* (Cameron, C. E., Giffette, M., and Raney, K. D., editors), Springer

Thompson, A. J. V., Bell, S. J., and Locarnini, S. A. (2009) Hepatitis B virus. Chapter 31 in *Clinical Virology*, 3rd edition (Richman, D. D., Whitley, R. J., and Hayden, F. G., editors), ASM Press

Journals

Martin-Vilchez, S. et al. (2011) The molecular and pathophysiological implications of hepatitis B X antigen in chronic hepatitis B virus infection. *Reviews in Medical Virology*, 21, 315-329

Nguyen, D. H., Ludgate, L., and Hu, J. (2008) Hepatitis B virus-cell interactions and pathogenesis. *Journal of Cellular Physiology*, 216, 289-294

Patient, R. et al. (2007) Hepatitis B virus subviral envelope particle morphogenesis and intracellular trafficking. *Journal of Virology*, 81, 3842-3851

Shin, M.-K. et al. (2008) Circularization of an RNA template via long-range base pairing is critical for hepadnaviral reverse transcription. *Virology*, 371, 362-373

Urban, S. et al. (2010) The life cycle of hepatitis B virus. *Journal of Hepatology*, 52, 282B`284

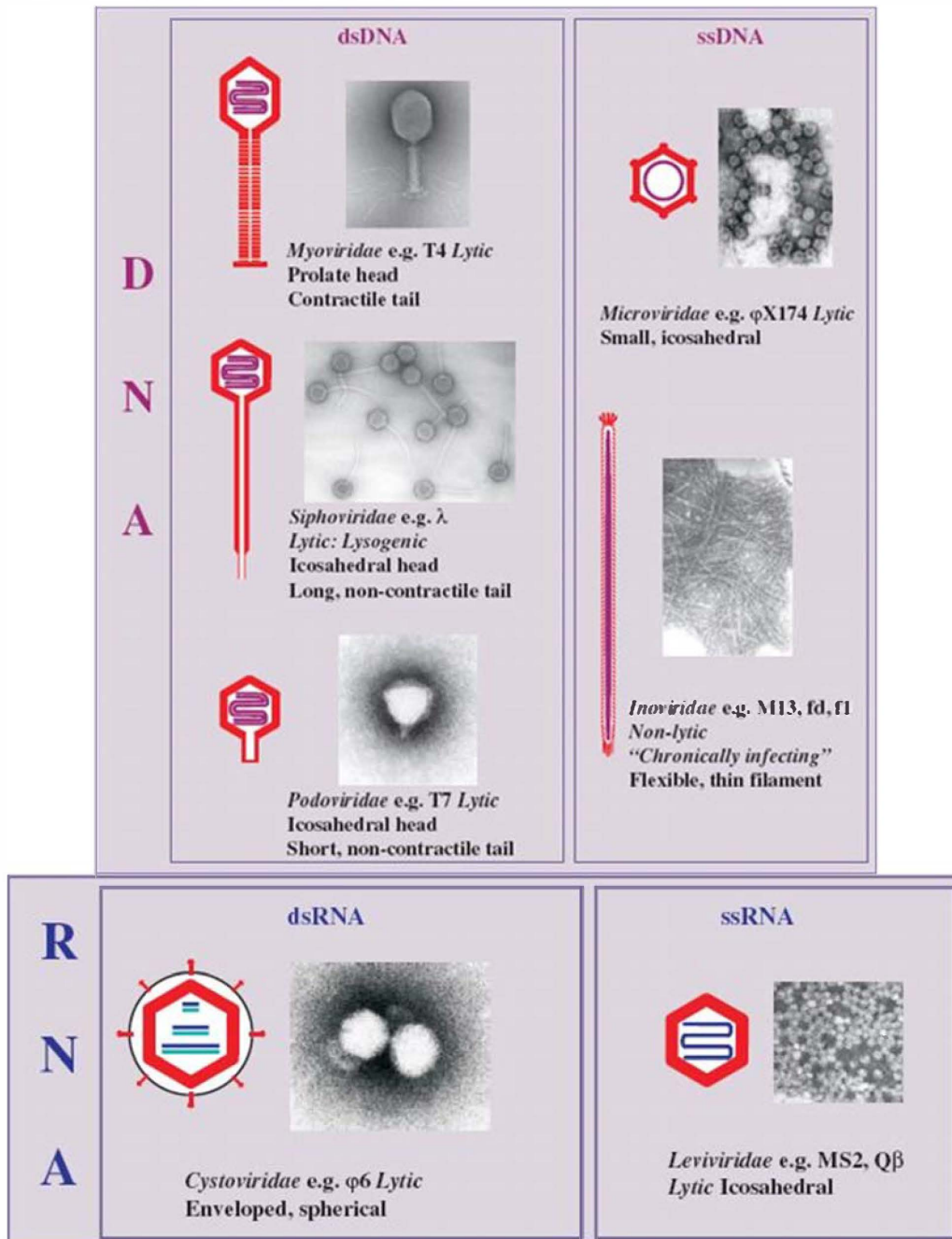
Vartanian, J.-P. et al. (2010) Massive APOBEC3 editing of hepatitis B viral DNA in cirrhosis. PLoS Pathogens, 6 (5): e1000928

CHAPTER 20

Bacterial Viruses

CHAPTER 20 AT A GLANCE

Electron micrographs reproduced by permission of Professor H.-W. Ackermann, Professor R. Duda and Professor R. Inman



20.1 INTRODUCTION TO BACTERIAL VIRUSES (BACTERIOPHAGES)

Bacterial viruses, known as bacteriophages or phages (from the Greek phagein, "to eat"), were discovered independently by Frederick Twort (1915) and Felix d'Herelle (1917). A diversity of phages has subsequently been identified and grouped into a number of families. Phage diversity is reflected in both morphological and genetic characteristics. The genome may be DNA or RNA, single- or double-stranded, circular or linear, and is generally present as a single copy. Morphology varies from simple, icosahedral and filamentous phages to more complex tailed phages with an icosahedral head. The majority of phages are tailed.

Phages are common in most environments where bacteria are found and are important in regulating their abundance and distribution. The host controlled modification and restriction systems of bacteria are presumed to protect against phage infection: restriction is leveled against invading double-stranded phage DNA, whilst self-DNA is protected by the modification system. However, in response, certain phages have evolved anti-restriction mechanisms to avoid degradation of their DNA by restriction systems.

Broadly, phages can be classified as either virulent or temperate. A virulent phage subverts the cellular apparatus of its bacterial host for multiplication, typically culminating in cell lysis (for obligately lytic phages) and release of progeny virions. In rare cases, for example the filamentous ssDNA phage M13, progeny are continuously extruded from the host without cell lysis. Accordingly, M13 has been referred to as a "chronically infecting" phage. Temperate phages have alternative replication cycles: a productive, lytic infection or a reductive infection, in which the phage remains latent in the host, establishing lysogeny. The latter generally occurs when environmental conditions are poor, allowing survival as a prophage in the host (which is referred to as a lysogen). During lysogeny the phage genome is repressed for lytic functions and often integrates into the bacterial chromosome, as is the case for phage lambda (λ), but it can exist extrachromosomally: for example, phage P1. The prophage replicates along with the host and remains dormant until induction of the lytic cycle. This occurs under conditions that result in damage to the host DNA. The phage repressor is inactivated and the lytic process ensues. Such a mechanism allows propagation of phages when host survival is compromised. A resident prophage can protect the host from superinfection by the same or similar strains of phages by repressing the incoming phage genome (a phenomenon known as superinfection immunity).

Some temperate phages contribute "lysogenic conversion genes" (for example, diphtheria or cholera toxin genes) when they establish lysogeny, thereby converting the host to virulence. Phages can also mediate bacterial genome rearrangements and transfer non-viral genes horizontally by transduction. Tailed phages are the

most efficient particles for horizontal (lateral) gene transfer, with the tail effectively guiding injection of DNA into the bacterial cell. Such phage activities generate variability and are a driving force for bacterial evolution.

Bacteriophages have had key roles in developments in molecular biology and biotechnology. They have been used as model systems for animal and plant viruses, and have provided tools for understanding aspects of DNA replication and recombination, transcription, translation, gene regulation, and so on. The first genomes to be sequenced were those of phages. Restriction enzymes were discovered following studies on phage infection of different hosts and laid the basis for the development of gene cloning. Phage-encoded enzymes and other products are exploited in molecular biology. Certain phages have been adapted for use as cloning and sequencing vectors and for phage display. Phages are utilized in the typing of bacteria, in diagnostic systems, as biological tracers, as pollution indicators, and in food and hospital sanitation (see also Section 1.2.2). There is also renewed interest in the therapeutic potential of phages, due largely to the rapid emergence of antibiotic-resistant bacteria and of new infectious diseases. However, realization of this potential will depend on a number of factors, including improved methods of large-scale production and purification of phages, appropriate protocols for administering phages, and modifications to enhance therapeutic properties, in order to remove phage-encoded toxins, avoid clearance of phages by the host defense system, and so on.

This chapter considers the biology of RNA and DNA phages. Properties and applications of a selection of phages that employ different strategies for phage development will be discussed, with emphasis on tailless single-stranded RNA and DNA phages.

RNA PHAGES

20.2 SINGLE-STRANDED RNA PHAGES

Single-stranded RNA phages are small, icosahedral viruses of the family Leviviridae (from the Latin *levis*, "light"), discovered by Timothy Loeb and Norton Zinder in 1961. Phages in this family have high mutation rates and some of the smallest RNA genomes known. They are plus-strand viruses (with the genome acting as mRNA), containing only a few genes, and infect various Gram-negative bacteria, including *E. coli*, *Pseudomonas* spp., and *Caulobacter* spp. Those infecting the enterobacteria do so by way of the sex pilus. Representative ssRNA phages of the genus *Levivirus* serogroup I (e.g. MS2 and f2) and *Allolevivirus* serogroup III (e.g. Qbeta, Q β) are considered here.

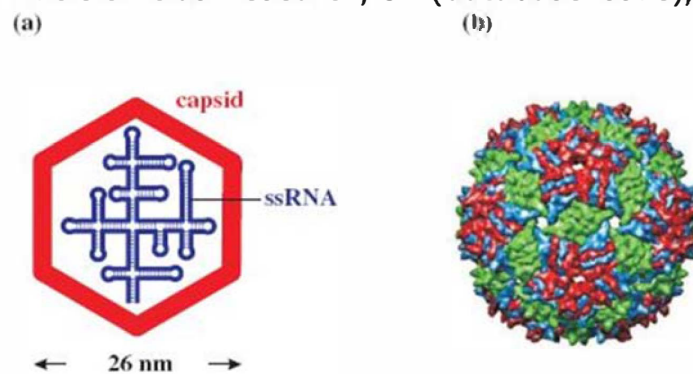
20.2.1 Virion structure of ssRNA phages

RNA phages typically comprise 180 molecules of major coat (capsid) protein (CP) per virion, one molecule of maturation (A) protein, required for infectivity and

maturation, and a linear ssRNA genome of about 3500–4200 nucleotides that displays considerable (up to 80%) secondary structure (Figure 20.1).

Figure 20.1 Structure of ssRNA phage MS2. (a) Virion components. (b) Reconstruction of capsid showing subunit organization. The icosahedral capsid comprises about 180 copies of major coat protein (as dimers). There is also one copy of maturation protein (minor virion protein), for host infection by recognition of the sex pilus, and the ssRNA genome.

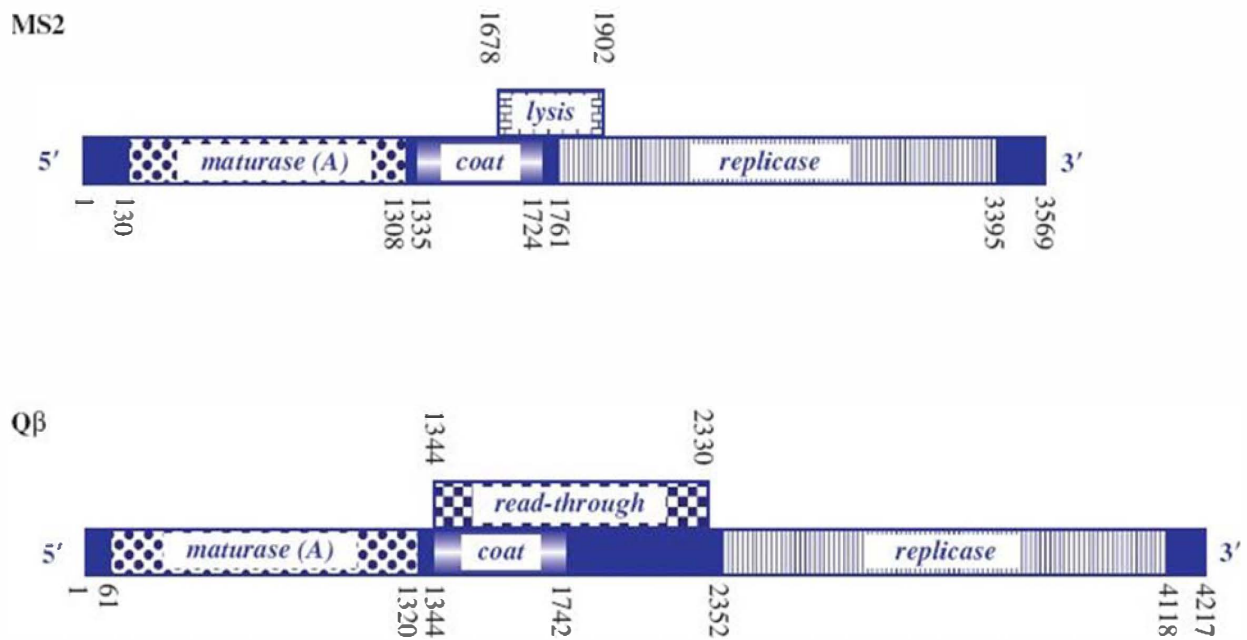
Source: (b) Image rendered by Chimera, from Virus Particle Explorer (VIPER) database (Shepherd C. M. et al. (2006) VIPERdb: a relational database for structural virology, *Nucleic Acids Research*, 34 (database issue), D386–D389).



20.2.2 Genome of ssRNA phages

MS2 was the first virus identified that carried its genome as RNA and the genetic map is shown in Figure 20.2. The genome is 3569 nucleotides, with four open reading frames (ORFs) for major coat protein, maturation (A) protein, replicase (subunit II) and lysis protein, and is characterized by intergenic spaces. It was the first phage RNA genome to be completely sequenced.

Figure 20.2 Genetic maps of ssRNA phages MS2 and Q β . Nucleotide coordinates are shown along the maps. The solid blue boxes are intergenic regions involved with ribosome binding and regulation of translation. MS2: genes for maturase (maturation protein); coat (capsid) protein; replicase protein for RNA-dependent RNA replication; and lysis protein (from overlapping coat and replicase genes). The Min Jou interaction involves nucleotides 1427–1433 and 1738–1744. Q β : genes for maturase; coat protein; replicase protein; and read-through protein. The Q β genome is larger than that of MS2, but there is no separate lysis gene. The maturation protein additionally mediates lysis.



Lysis protein is encoded by an overlapping ORF between the distal part of the coat protein gene, which appears to be required to regulate expression of the lysis gene, and the proximal part of the replicase gene, which encodes the essential residues for functioning of the lysis protein. Occasionally, ribosomes traversing the coat protein gene fail to maintain the correct reading frame so that once translation of coat protein terminates they can reinitiate at the start of the lysis protein ORF by shuffling a short distance. Efficiency of translation of the lysis protein ORF by such reinitiation is low. Thus only small amounts of lysis protein are synthesized, in turn guaranteeing that sufficient copies of coat protein will be available for assembly of virus particles before cell lysis occurs.

The genome of QB, which is about 4000 nucleotides, is larger than that of MS2. It encodes four proteins: coat protein, replicase, maturation (A2) protein, which also mediates cell lysis, and a minor coat protein (read-through protein, A1), which is generated as a result of reading through the coat protein gene into the intergenic region (Figure 20.2). Read-through is due to a leaky coat protein terminator (UGA). This allows a low level of misincorporation of tryptophan at the terminator signal and subsequent ribosome read-through. The read-through protein constitutes 3–7% of the virion protein and has a role in host infection. Such inefficient termination of translation is not found in the group I phages, where there are tandem stop codons (UAA, UAG) at the end of the coat protein gene and no read-through protein is produced.

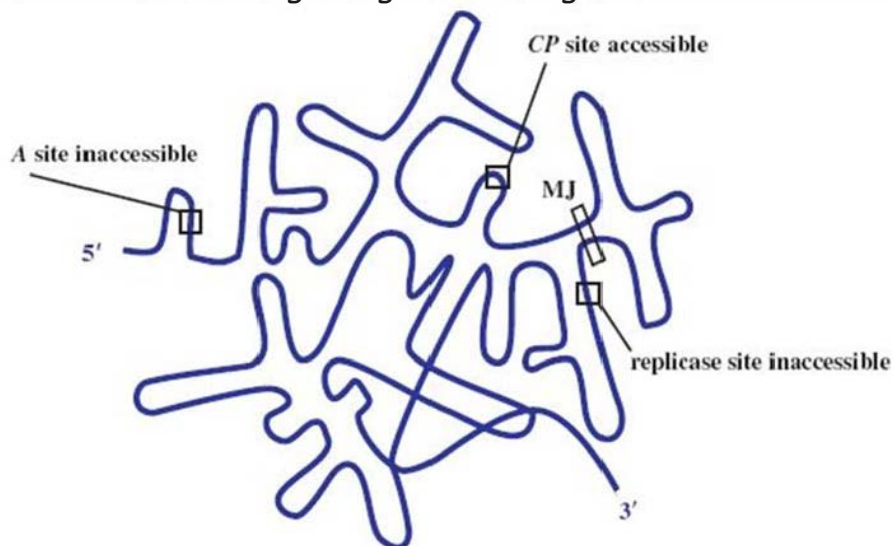
Not only does the ssRNA genome encode the phage proteins, it also determines specific secondary and tertiary structure that regulates translation, replication, and other functions.

20.2.3 Replication cycle of ssRNA phages

For coliphage MS2, infection involves attachment to pilin (the pilus subunit) along the length of a sex (F) pilus of a susceptible host via the A protein. Such binding of phages can block conjugation in *E. coli*.

The plus-strand RNA genome directs protein synthesis immediately upon infection. It adopts a complex secondary structure that affects access of the ribosome binding sites (RBSs) to host ribosomes during translation. Control of gene expression involves transactions between the RNA secondary structure and the initiation of translation. Initially the RBS for the coat protein gene is accessible, occupying the loop position of a hairpin structure, whilst those for the maturation (A) and the replicase gene are not, being embedded in secondary structure (see Figure 20.3). The replicase gene is expressed only after the coat protein gene has been translated. Such translational coupling between coat protein and replicase genes involves the so-called Min Jou (MJ) interaction. This is a long-distance interaction (LDI), in which a sequence just upstream of the replicase gene RBS is base-paired to an internal sequence of the coat protein gene. The MJ interaction represses translation of the replicase gene. However, passage of ribosomes beyond the first half of the coat protein gene temporarily unfolds this LDI by breaking the MJ interaction. This opens up the RBS for translation of the replicase gene. The rate of refolding of the LDI in part determines the level of expression of replicase. Another LDI, called the van Duin (VD) interaction, which borders the MJ, also appears to contribute to the repression of replicase translation.

Figure 20.3 Gene expression in ssRNA phage. Diagrammatic representation of secondary structure of the genome with ribosome binding site of coat protein (CP) gene accessible and those of A and replicase genes blocked. MJ: Min Jou, long distance interaction between sequences upstream of the replicase ribosome binding site and sequences near the beginning of the CP gene.

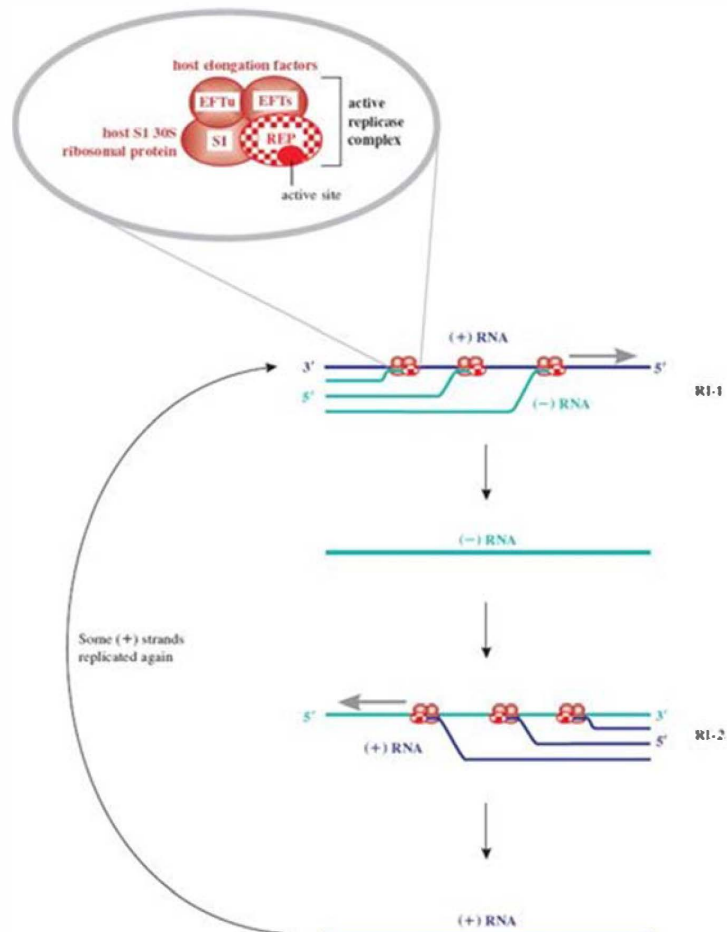


After about 10 to 20 minutes, translation of the replicase gene stops, due to repression by coat protein, which binds to a RNA stem-loop containing the

translation initiation region of replicase. This ensures that only a catalytic amount of replicase is translated. Too much replicase could poison the host. Such binding also appears to serve as an assembly initiation signal. The latter half of the infection cycle is devoted to synthesis of coat protein, which is required in large quantities for virion assembly. A small amount of lysis protein is synthesized during translation of the coat protein gene, due to an occasional reading frame error.

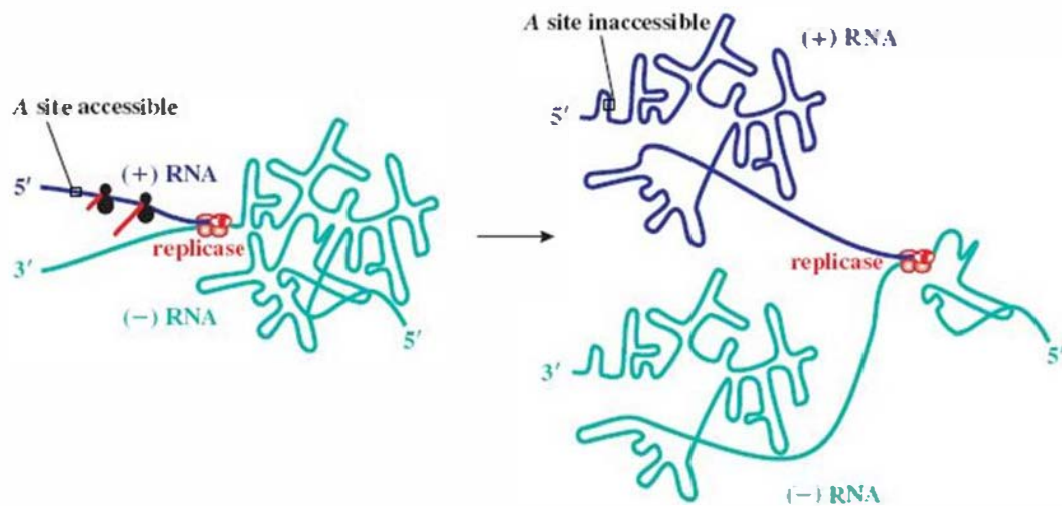
Once replicase is synthesized, replication of the genome can occur; the genome thus switches from a template for translation to a template for replication. There would be a topological problem if replication and translation occurred simultaneously, since they proceed in opposite directions along the template (3' 5' and 5' 3' respectively). However, it appears that translation of coat protein is repressed by replicase, which may block the coat protein ribosome entry site, and then recommences when sufficient (–) RNA has been synthesized. The replicase associates with a number of host proteins to form a phage RNA-specific polymerase (Figure 20.4), which makes both (+) and (–) strands of RNA. Replication occurs in two stages and involves replicative intermediates (RIs) composed of multiple, nascent RNA strands synthesized from a single template, and replicase: RI-1 with (+) RNA as template, and RI-2 with (–) RNA as template. Thus minus strands act as templates for synthesis of plus strands for use as mRNA. Secondary structure, which forms as each new strand is synthesized from the template strand, serves to keep the two strands apart, in turn preventing formation of a double-stranded intermediate. It is noteworthy that the error rate for such RNA replication is quite high compared with that for DNA replication. Replicase is error prone and the lack of a double-stranded intermediate means there is no template for error correction.

Figure 20.4 Replication of ssRNA phage genome. Replication of the single-stranded phage genome occurs through production of two replicative intermediates, RI-1 and RI-2. First, the (–) strand RNA is synthesized 5' to 3' antiparallel and complementary to the (+) sense template strand by replicase, in a multi-branched structure. The (–) strand RNA then serves as template for formation of new (+) strand genomic RNA. This is the only role for the (–) strands. Newly replicated (+) strands can be recycled in replication, translated to yield the capsid proteins, or encapsidated in the formation of progeny phages. The inset shows the activation of the replicase protein (REP) by associating with host proteins: for MS2 two elongation factors (EFTs and EFTu) and the ribosomal S1 protein. Replicase is an RNA-templated RNA polymerase, synthesizing both plus (+) and minus (–) strands of phage RNA 5' to 3' through specificity of the replicase for the 3' end of both template strands. S1 protein is not required for new (+) strand synthesis.



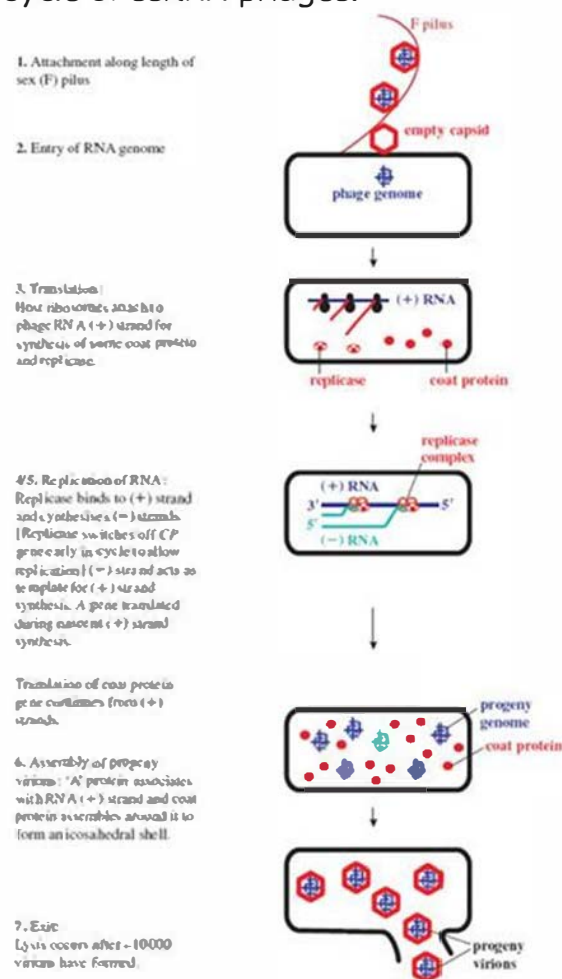
Expression of the A gene occurs independently of the others and is limited to periods of nascent plus strand synthesis. The RBS of the A gene is masked by folding of the RNA genome through an LDI in which the Shine–Dalgarno sequence of A pairs with a complementary sequence upstream in the 5' untranslated leader region. The initiator site of A becomes exposed transiently when plus strand synthesis begins, due to a delay in folding of the newly synthesized strand. This allows ribosomes to translate the A gene until the RBS is sequestered (Figure 20.5). The number of molecules of A protein is thus maintained in line with the number of new RNA plus strands. There appears to be upregulation of the A protein in Q β , consistent with its role in cell lysis as well as infectivity. The LDI is further downstream than in MS2, so that the RBS is exposed for longer, resulting in higher levels of A gene expression.

Figure 20.5 Expression of A gene of ssRNA phage. A gene expression occurs when plus (+) strand RNA is newly synthesized. The start site for A is then exposed, but only transiently, until the ribosome binding site is sequestered by folding of the (+) strand through the long-distance interaction involving the Shine–Dalgarno sequence and its upstream complement.



Phage assembly involves spontaneous aggregation, in which the (+) RNA associates with the A protein and is encapsidated through specific recognition of the phage RNA by coat protein dimers. Progeny virions are normally released by cell lysis (see Figure 20.6 for a summary of the replication cycle), with a burst size of about 104, compared with a few hundred for DNA phages.

Figure 20.6 Replication cycle of ssRNA phages.



Thus the minimalist genome of the ssRNA phages is utilized very efficiently. Coding capacity is expanded through the use of overlapping genes. Host proteins are utilized to activate phage proteins and some phage proteins have multiple functions. The phage genome serves as a model system for investigating such

mechanisms as translational coupling, translational repression, and the role of secondary structure of mRNA in the control of gene expression.

20.3 DOUBLE-STRANDED RNA PHAGES

Phages in the family Cystoviridae (from the Greek *kystis*, “bladder, sack”) contain a dsRNA genome, which is segmented and packaged in a polyhedral inner core with a lipid-containing envelope. Phi6 ($\phi 6$) was the first member of the family to be isolated and has been extensively studied. The genome comprises three linear segments: RNA L (large) of about 6400 nucleotides, RNA M (medium) of about 4000 nucleotides, and RNA S (small) of about 3000 nucleotides.

Transcription of the double-stranded genome (for synthesis of new plus strands) involves a phage RNA-dependent RNA polymerase. The plus-strand transcripts serve as replication templates and mRNAs. Translation of the L segment produces the early proteins that assemble to form the polymerase complex; the M segment produces structural proteins for membranes and spikes in particular, and the S segment produces structural proteins for the capsid, membrane assembly, lysis and entry, and a non-structural protein for envelopment of the capsid.

Phage $\phi 6$ infects its host, *Pseudomonas syringae* (pv. *phaseolicola*), by way of the pilus, which retracts to bring the virion into contact with the cell, and the nucleocapsid enters. Uncoating occurs inside the cell and the polymerase is released. Transcription of the genome is temporally controlled and virions assemble in the cytoplasm, with their envelope deriving from the host. Packaging of the plus-strands occurs in the order S—M—L. These single-stranded precursors are then replicated into mature double-stranded genomes inside the capsid. About 100 virions are released following cell lysis.

DNA PHAGES

20.4 SINGLE-STRANDED DNA PHAGES

There are two groups of ssDNA phages: icosahedral and filamentous. Representative icosahedral phages, $\phi X174$ and S-13, were amongst the first studied. Later, in the 1960s, filamentous phages were isolated by Don Marvin and Hartmut Hoffmann-Berling, and by Norton Zinder and co-workers.

20.4.1 Icosahedral ssDNA phages

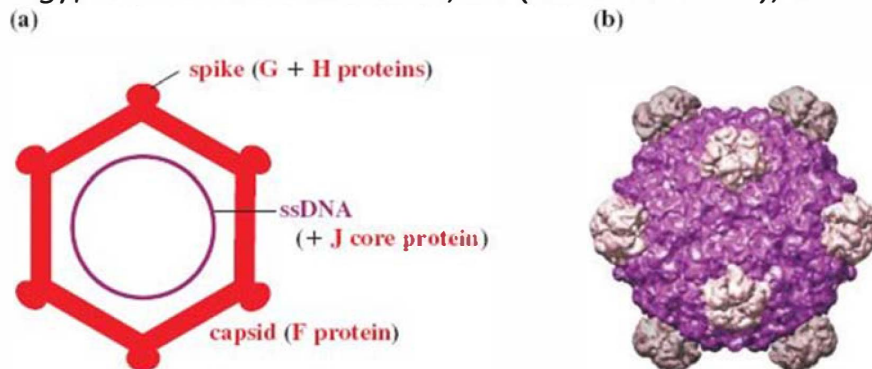
Icosahedral ssDNA phages belong to the family Microviridae (from the Greek *micros*, “small”). Such phages provided the first evidence for overlapping genes and revealed the economy of genetic coding, which is a feature of several small viruses, including hepatitis B virus (Section 19.6) and the ssRNA phages (Section 20.2). Studies on replication of these phages led to the discovery of rolling circle replication and to the identity of various genes encoding proteins for host DNA replication. $\phi X174$ has been most extensively studied.

20.4.1.a Virion structure of phage ϕ X174

The virion of ϕ X174 contains proteins F, G, H, and J; F forms the main shell and the spike protein G protrudes on the icosahedral five-fold axes. A molecule of H is found on each spike and acts as a pilot protein. Part of H may lie outside the shell and part within, spanning the capsid through the channels formed by the G proteins. The projections are involved in attachment to the host and delivery of the genome into the host cell. Highly positively charged DNA binding protein J is associated with the ssDNA genome (Figure 20.7).

Figure 20.7 Structure of phage ϕ X174. (a) Virion components. (b) Reconstruction of capsid showing subunit organization. The virion contains 60 molecules of major coat protein F (48.4 kD). The spike at each vertex of the icosahedron, composed of five molecules of G protein (19.0 kD) and one molecule of H protein (35.8 kD), has a role in host recognition and attachment. Protein J (4.0 kD) binds to the phage genome for condensation of DNA during packaging. This protein subsequently binds to the internal surface of the capsid to displace the internal scaffolding protein B (13.8 kD).

Source: (b) Image rendered by Chimera, from Virus Particle Explorer (VIPER) database (Shepherd C. M. et al. (2006) VIPERdb: a relational database for structural virology, *Nucleic Acids Research*, 34 (database issue), D386–D389).



20.4.1.b Genome of phage ϕ X174

The genome is a circular ssDNA molecule of 5386 nucleotides, coding for eleven proteins. It was the first entire DNA genome to be sequenced (Section 2.9.1). The genes are tightly clustered with little non-coding sequence. The length of the genome is smaller than its coding capacity. However, the different proteins are generated through extensive use of overlapping genes translated in alternative reading frames or employing different start codons. Gene A contains an internal translation initiation site to encode protein A, which corresponds to the C-terminal region of protein A; B is encoded within A in a different reading frame; K is at the end of gene A and extends into gene C and is translated in a different frame to A and C; E is totally within D, but in a different reading frame; the termination codon for D overlaps initiation codon for J (see Figure 20.8). All genes are transcribed in the same direction.

Figure 20.8 Genetic organization of phage ϕ X174. The numbers are coordinates of the genes. The direction of transcription and approximate location of origin (ori) of replication for viral (+) strand are shown. Functions of the gene products (with coordinates of the genes including termination codons) are as follows:

gene A (3981–136) protein, viral strand synthesis and RF replication;

gene A (4497–136) protein, shutting down host DNA synthesis;

gene B (5075–51) protein, capsid morphogenesis;

gene C (133–393) protein, DNA maturation;

gene D (390–848) protein, capsid morphogenesis;

gene E (568–843) protein, host cell lysis;

gene F (1001–2284) protein, capsid morphogenesis—major coat protein;

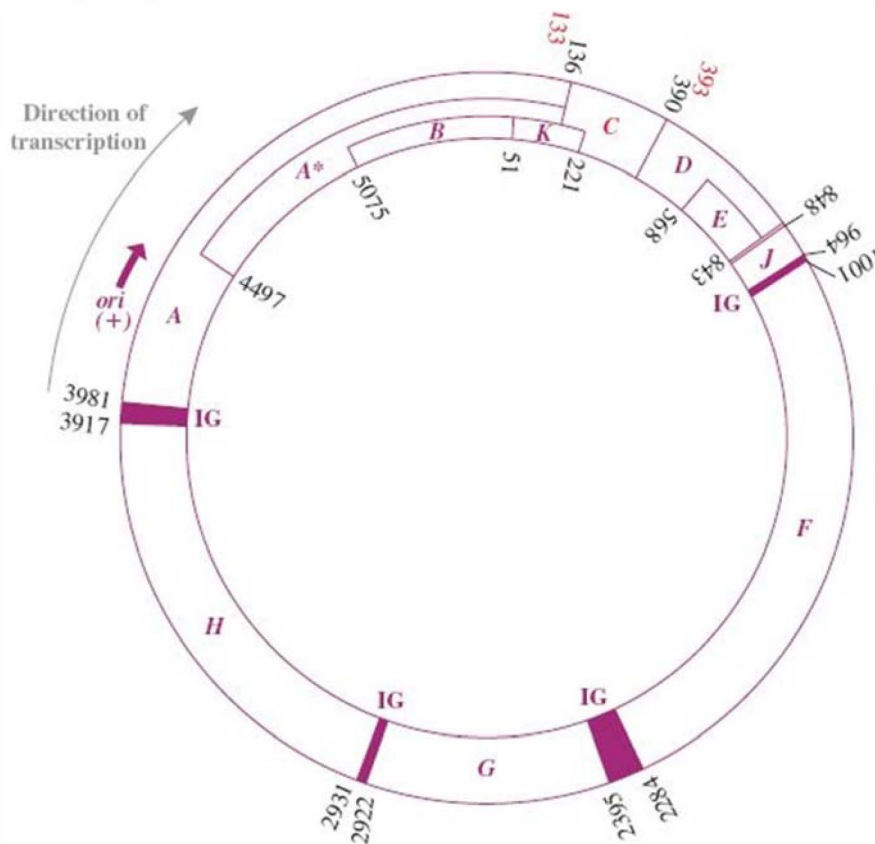
gene G (2395–2922) protein, capsid morphogenesis—major spike protein;

gene H (2931–3917) protein, capsid morphogenesis—minor spike protein;

gene J (848–964) protein, capsid morphogenesis—core protein (DNA condensing protein);

gene K (51–221) protein, function not clear; appears to enhance phage yield (burst size).

IG: intergenic region at borders of genes A, J, F, G, H contains a ribosome binding site and other features.



20.4.1,c Replication cycle of phage ϕ X174

Phage ϕ X174 recognizes the receptor lipopolysaccharide in the outer membrane of rough strains of Enterobacteriaceae, such as *E. coli* and *Salmonella typhimurium*, by

way of protein H. This protein then acts as a pilot involved with delivery of the phage DNA into the host cell for replication.

Like M13 (Section 20.4.2.c), replication of the genome of ϕ X174 occurs in three stages and involves a double-stranded intermediate, the replicative form (RF). The viral plus strand is first converted to the RF by host enzymes. However, this process is more complex than for M13, involving the formation of a primosome that includes a number of proteins to open a hairpin found at the origin (ori) of replication of ϕ X174 and to load the replication apparatus on to the DNA. The minus-strand is transcribed for synthesis of the phage-encoded proteins. Replication of the RF involves rolling circle replication and requires phage-encoded protein A to synthesize new plus-strands. These then serve as templates for minus-strand synthesis to generate the new RFs. The third stage is asymmetric replication of progeny ssDNA plus strands. RF synthesis continues until sufficient structural proteins have been synthesized and assembled into empty precursor particles. This is mediated by scaffolding proteins (D and B), which associate transiently with the coat proteins and induce conformational switches to drive the assembly process. A procapsid forms, with D forming the external shell and B occupying an internal location. Viral DNA plus-strands are then captured, before they can be used as templates for further minus-strands, and are packaged with basic protein J (which neutralizes the negatively charged genome). J protein binds both to the DNA and to the internal surface of the capsid, displacing the B protein, to generate the provirion. Subsequent shedding of the D protein yields the mature infectious virion. Virions accumulate in the cytoplasm and are released by cell lysis, which requires expression of the E gene for endolysin production.

20.4.2 Filamentous ssDNA phages

Filamentous ssDNA phages are in the family Inoviridae (the Greek *ina*, “fiber, filament”). The F-specific filamentous (Ff) phages, notably M13, fd, and f1, have been most extensively studied. They are plus-strand phages and are “male-specific,” infecting *E. coli* strains containing the conjugative plasmid F, by adsorbing to the tip of the F pilus (encoded by the plasmid). Unlike many other DNA phages, filamentous ssDNA phages do not inject their DNA into the host cell; rather, entire phage particles are ingested. Furthermore, these phages do not lyse infected cells, but progeny are continuously extruded through the cell membrane. Despite this, plaque-like zones are formed in a lawn of sensitive bacteria, since infected cells grow more slowly than uninfected cells.

20.4.2.a Genome of Ff phages

The genome of the Ff phages is a circular ssDNA molecule of about 6400 nucleotides. M13, fd, and f1 are 98% identical and generally no distinction will be made between them here. The genes are tightly packed on the genome, with a number of intergenic (IG) regions (Figure 20.9). The major IG region of 507

nucleotides (between genes IV and II) contains a number of regulatory regions: the origin of replication for (+) and (-) DNA synthesis, a strong rho-dependent transcription terminator, after gene IV, and a 78-nucleotide hairpin region, which is the packaging signal (PS) for efficient phage assembly. There is also a strong rho-independent terminator, after gene VIII, and a weaker rho-dependent terminator within gene I.

Figure 20.9 Genetic map of F-specific filamentous phage M13. The main promoters and terminators are shown. ori: origin of replication for viral (+) and complementary (-) DNA strands in main intergenic region. The direction of transcription from promoters (P) is indicated by arrowheads (>). Gene product, approximate molecular size, number of copies, and function are as follows:

pI, 35–40 kD, few copies per cell, membrane protein for assembly;

pII, 46 kD, 103 copies per cell, endonuclease/topoisomerase for replication of RF, viral strand synthesis;

pIII, 42 kD, ~5 copies per virion, minor capsid protein for morphogenesis and adhesion;

pIV, 44 kD, few copies per cell, membrane protein for assembly;

pV, 10 kD, 105 copies per cell, ssDNA binding protein controls switch from RF to viral (+) strand synthesis;

pVI, 12 kD, ~5 copies per virion, minor capsid protein for morphogenesis and attachment;

pVII, 3.5 kD, ~5 copies per virion, minor capsid protein for morphogenesis;

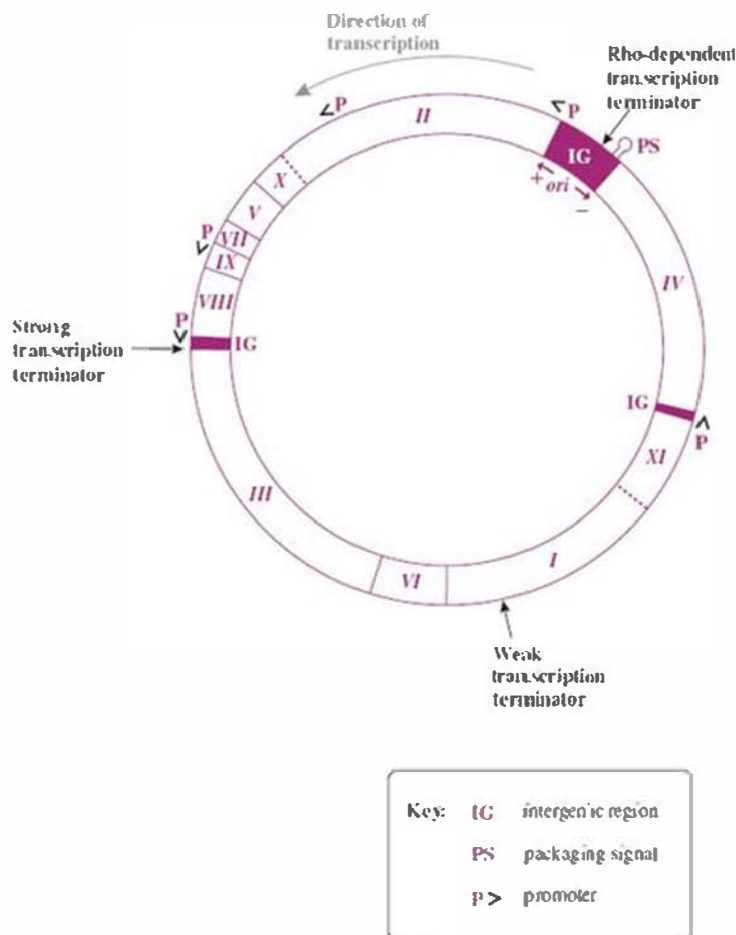
pVIII, 5 kD, 2700–3000 copies per virion, major capsid protein for morphogenesis;

pIX, 3.5 kD, ~5 copies per virion, minor capsid protein for morphogenesis;

pX, 12 kD, 500 copies per cell, replication of RF, viral strand synthesis;

pXI, 12 kD, few copies per cell, membrane protein for assembly.

Source: Modified from Webster (2001) In Phage Display. A Laboratory Manual, Cold Spring Harbor.



The genome encodes 11 proteins. Genes are grouped by function: genes II, V, and X for phage DNA synthesis; genes III, VI, VII, VIII, and IX for capsid structure; and genes I, IV, and XI for assembly.

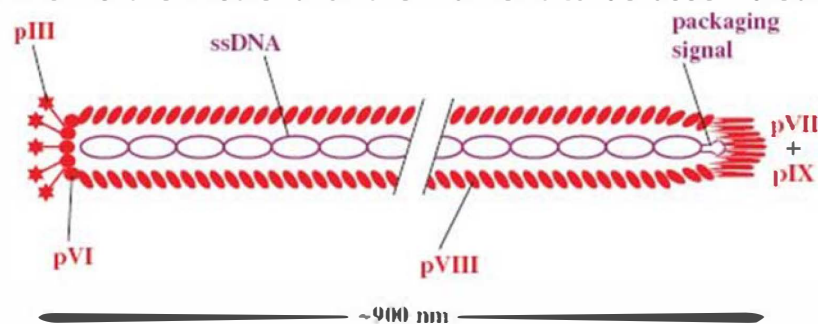
Transcription is complex, involving several strong promoters. The two strong transcription terminators divide the genome into two transcription regions: a frequently transcribed region from gene II to VIII and an infrequently transcribed region from gene III to IV. All the genes are transcribed in the same direction and there is a gradient of transcription with genes near to a terminator being expressed more frequently (e.g., genes VIII and V). The weak terminator in gene I down-regulates expression of the gene, ensuring that large amounts of gene I protein (pI), which would be deleterious to the host, are not synthesized. Gene XI protein (pXI) results from an in-frame translation initiator site at residue 241 of gene I, and corresponds to the C terminus of pI. Likewise gene X protein (pX) is encoded within gene II. Such mechanisms ensure that the quantity of each phage protein synthesized is in line with requirements.

20.4.2.b Virion structure of Ff phages

The capsid is a flexible protein filament of variable length, depending on genome size; the wild-type length is approximately 900 nm. The filament comprises about 2700 copies of the 50 amino acid major coat protein (gene VIII protein, pVIII), arranged in an overlapping array to form a helix (Figure 20.10). The basic positively charged C-terminal domain of pVIII forms the inside wall of the filament and

interacts with the DNA. The N-terminal negatively charged domain is exposed on the outside, giving a negatively charged surface. Hydrophobic interactions hold adjacent coat protein subunits together. The minor coat proteins are present at the ends of the filament. At one terminus (where assembly terminates) are two of these proteins: gene III protein (pIII) and gene VI protein (pVI), that together form an adsorption complex for recognition of the sex pilus. At the other terminus (where assembly initiates) are gene VII protein (pVII) and gene IX protein (pIX). The single-stranded circular supercoiled genome is housed lengthwise in the center of the filament, orientated so that the hairpin packaging loop is located at the pVII/pIX end.

Figure 20.10 Diagrammatic representation of F-specific filamentous ssDNA phage. The capsid comprises five proteins: pVIII, the major coat protein along the length of the filament; pIII and pVI, minor coat proteins that form an adsorption knob at one end of the filament; pVII and pIX, minor coat proteins that form at the other end. Arrangement of these two proteins is not clear: pVII may interact with pVIII and be shielded from the environment, whilst pIX may interact with pVII and be exposed. The single-stranded genome is oriented with the packaging signal at the pVII/pIX end, which is the first end of the filament to be assembled.

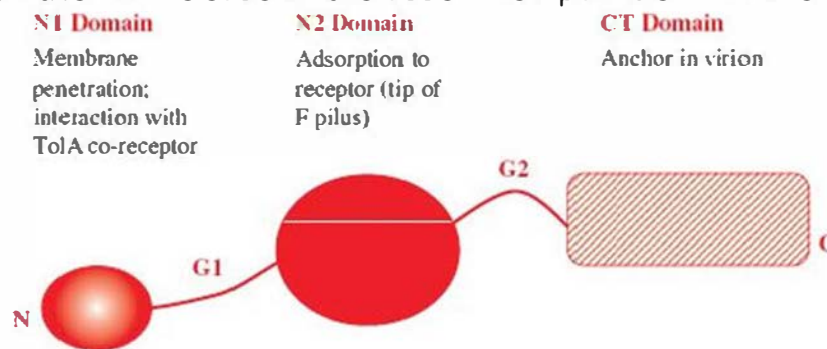


20.4.2.c Replication cycle of Ff phages

The receptor for Ff phages comprises the tip of the F pilus together with TolA, TolQ, and TolR membrane proteins of *E. coli*. The phage binds first to the tip of the F pilus through domain N2 of pIII. The pilus retracts bringing pIII to the periplasm and domain N1 of pIII can then interact with TolA (Figure 20.11).

Figure 20.11 Domains of gene III protein of F-specific filamentous ssDNA phage. The gene III protein comprises three domains: N-terminal domains N1 (or D1) and N2 (or D2) required for host cell infection, and C-terminal CT (or D3) required for phage release from the membrane and structural stability; CT anchors pIII in the phage coat by interacting with pVI. The domains are separated by glycine-rich linker regions: G1 and G2. Domains N1 and N2 (exposed on the surface of phage particles) interact in a horseshoe shaped knob-like structure at one end of the filament. Domain N2 binds to the tip of the F pilus and domain N1 dissociates from N2. N1 binds to the TolA membrane protein. Pilus retraction brings the pIII end of the filament to the periplasm. TolQ, R, and A proteins are required for translocation of phage DNA to the cytoplasm and of coat proteins to the cytoplasmic membrane.

Domain CT contains two sub-domains: one involved in stabilizing the particle once assembled, the other for release of the assembled particle from the membrane.



The phage particle enters the cell and disassembles. TolA, TolQ, and TolR proteins mediate de-polymerization of the coat proteins; the major coat protein and possibly other capsid proteins being deposited in the cytoplasmic membrane to be recycled for new particle formation, as the phage DNA is translocated into the cytoplasm. pIII appears to enter the cell and have a role in the initiation of replication. The circular ssDNA is coated with an E. coli ssDNA binding (ssb) protein to protect against degradation by host nucleases. Phage DNA is plus-strand and DNA replication occurs immediately upon infection in order to provide the (–) DNA for transcription.

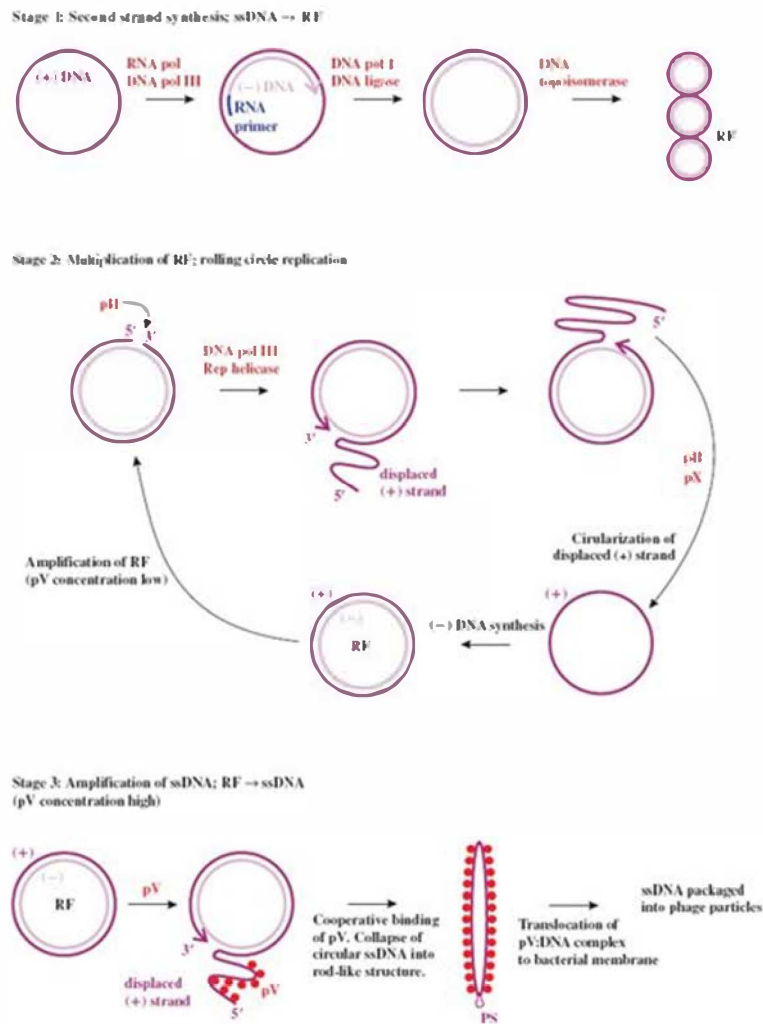
Replication occurs in three stages (Figure 20.12). The first stage is (–) DNA synthesis to generate the double-stranded RF; this depends on host functions. The minus-strand can then be transcribed and translated into phage proteins. A specific endonuclease (gene II protein, pII) synthesized from the first RF is required for the second stage of replication, which involves multiplication of the RF. The plus-strand of the RF is nicked in the major intergenic region, providing a primer (3' end) for synthesis of a new plus-strand, using the minus-strand as template, by rolling circle replication. As DNA polymerase III extends the 3' end, the old plus-strand is displaced, cleaved, and circularized by pII and converted to a dsDNA (RF), which accumulates as a pool in the cell. As DNA replication proceeds there is concomitant synthesis of phage proteins. Gene V protein (pV) accumulates in the cytoplasm and at a critical concentration binds cooperatively to displaced plus strands as they roll off the circle, diverting them from their template role. This initiates stage 3 of replication, in which there is a switch from RF to ssDNA synthesis. The 78-nucleotide hairpin region of the plus-strand remains free of pV and initiates the phage assembly reaction. pV additionally serves as a translational repressor of pII, in turn regulating DNA synthesis. Another protein, pX, also located in the cytoplasm, has a role in synthesis of the single-stranded viral DNA. The nucleoprotein complex of pV dimers and ssDNA forms a filamentous structure and is targeted to the cell membrane, where assembly takes place. Capsid proteins are inserted into the cytoplasmic membrane as they form.

Figure 20.12 Replication of F-specific filamentous phage DNA. Replication occurs in three stages:

1. Second strand synthesis. Plus (+) strand viral genome, introduced into the cell upon infection, is the template for synthesis of the complementary minus (–) strand. Host RNA polymerase (RNA pol) synthesizes an RNA primer in the main intergenic region of the (+) strand. The 3' end of the primer is extended by host DNA polymerase III (DNA pol III) synthesizing the (–) strand. The primer is removed, the gap filled, and a double-stranded, circular, supercoiled replicative form (RF) generated. RF acts as template for production of phage-encoded proteins.

2. Multiplication of RF. RF is replicated by a rolling circle mechanism. Phage-encoded pII, a “nicking–closing” enzyme, nicks RF at a specific site in the main intergenic region of the viral (+) strand. This allows elongation of the (+) strand by DNA pol III, the original (+) strand being simultaneously displaced by Rep helicase. When a round of replication is completed, pII cleaves and circularizes the displaced (+) strand, which is rapidly converted to another RF. Synthesis of the (–) strand requires an RNA primer and host enzymes. During the early stage of infection pV concentration is low and RF amplification occurs rapidly together with expression of phage proteins.

3. Amplification of ssDNA. Conversion of RF to viral (+) strand DNA for packaging is initiated by ssDNA binding protein pV, which represses translation of a number of phage proteins including pII. Inhibition of pII results in a switch from RF synthesis to production of phage particles. Late in the infection process when pV is present in high concentration it sequesters the viral (+) strands as they are displaced and prevents their conversion to RF. pV dimers bind cooperatively to (+) strand and convert the circular ssDNA into a filament-like pV:DNA complex. The DNA is oriented in the complex so that the packaging signal (PS) hairpin protrudes from one end (pV appears to bind weakly to dsDNA). pV:DNA complex is translocated to the inner membrane for assembly of phage particles, with pVII and pIX assembling on to the PS region of the genome. pX has a role in replication, but it remains unclear.

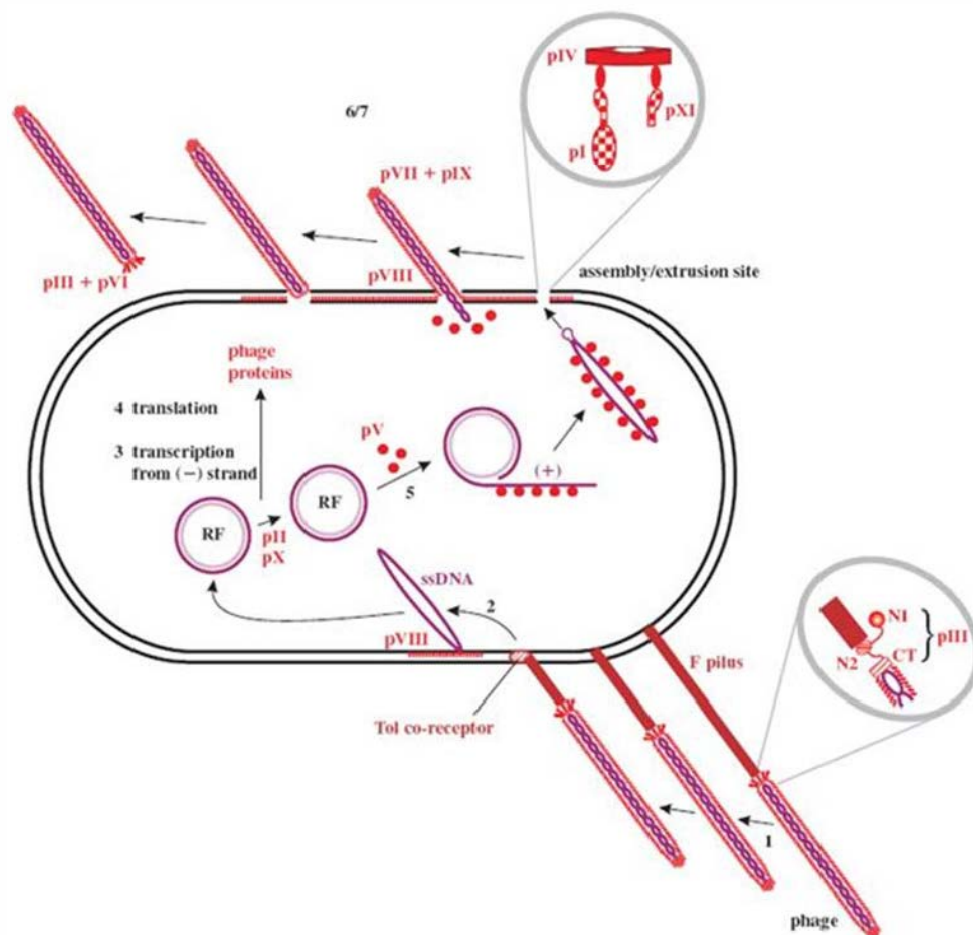


Assembly proteins pIV, pI, and pXI are integral membrane proteins. pIV forms an extrusion pore in the outer membrane and interacts with pI and pXI of the inner membrane to form a channel across both membranes, large enough to accommodate extruding phage particles. Assembly occurs where inner and outer membranes appear in close contact at an extrusion site and requires ATP, proton motive force, and the bacterial protein, thioredoxin. pV is displaced by coat proteins, which are packaged around the DNA as the phage is extruded through the membrane. The process is initiated by interaction of the packaging signal with pVIII, pVII, and pIX, possibly directed by pI. A conformational change induced in pI may allow interaction with pIV to open the gate in the pIV exit pore. The pVII–pIX end of the particle emerges first.

Elongation proceeds as pV is displaced by pVIII, whose positively charged C-terminal region interacts with the negatively charged sugar–phosphate backbone of the DNA. Termination of assembly involves pIII and pVI added to the other end of the filament as it is released, the length of the particle being dictated by the size of the viral DNA. Without pIII, elongation would continue with pVIII encapsidating another viral DNA to generate a polyphage. Thus pIII and pVI are important in maintaining structural integrity and stability of the virion. Particles are continuously extruded without killing the host cell (Figure 20.13).

Figure 20.13 Replication cycle of F-specific filamentous ssDNA phage. Host cell infection is a coordinated, multistep process.

1. Attachment. The phage binds to the tip of the F pilus via pIII, with domain N2 attaching to the tip and dissociating from its normal interaction with domain N1, and CT anchored in the virion with pVI.
2. Entry. Retraction of the pilus brings the phage particle to the periplasm to interact with Tol proteins. Phage coat proteins disassemble into the cytoplasmic membrane; the phage ssDNA is translocated into the cytoplasm.
3. Transcription: using (-) strand of RF as template (from stage (i) of genome replication).
4. Translation: for synthesis of all phage proteins, most of which insert in the cytoplasmic membrane. pII, pV, pX reside in the cytoplasm.
5. Genome replication occurs in three stages:
 - i. ssDNA plus (+) strand is converted to the dsDNA replicative form (RF).
 - ii. Amplification of RF by rolling circle replication involving pII and pX. Amplification of RF continues until pV reaches sufficient concentration to bind to (+) strand DNA and prevent its conversion to RF.
 - iii. pV switches most of the DNA replication from dsDNA RF to ssDNA (+) strand.
6. Assembly. The pV:DNA complex migrates to the cell membrane. pV is removed and replaced by the capsid proteins.
7. Exit. The assembly proteins pIV, pI, and pXI interact, forming a channel through both cell membranes for extrusion of the progeny phages. The pVII/pIX end of the particle emerges first. The phage is then elongated by replacement of pV by pVIII. pVI and pIII are added to the other end of the filament, the C terminus of pIII being anchored in the particle itself. Particles are extruded continuously without cell lysis.



This assembly process, occurring in the cell membrane and involving polymerization of coat proteins around the Ff phage genome, is in marked contrast to that for dsDNA phages, such as T4, which occurs in the cytoplasm and uses pre-formed heads for packaging DNA (Section 20.5.1).

The filamentous ssDNA phage CTX ϕ , which encodes the cholera toxin (the main virulence factor of *Vibrio cholerae*), and a phage associated with invasive meningococci are related to M13. However, unlike the coliphage, these phages can establish lysogeny by site-specific integration into the host chromosome in a similar manner to that of the lambdoid phages (Section 20.5.3).

20.4.2.d Cloning vectors and phage display

The relatively small size and non-lytic infection cycle make the filamentous ssDNA phages easy to manipulate and use as cloning vectors. Furthermore, since filament size is governed by size of the genomic DNA, insert DNA of variable length can be cloned. Vectors are based on the RF, which is easily obtained from infected cells. Cloned ssDNA can be recovered from the phage particles for use in, for example, DNA sequencing, designing DNA probes, and site-directed mutagenesis.

The intergenic (IG) region has been manipulated for the development of cloning vectors. Joachim Messing and co-workers designed a set of vectors, the M13 mp series, by inserting the α -fragment of the *E. coli lacZ* gene containing different polylinkers (multiple cloning sites) into the IG region. Foreign DNA inserted into one of the cloning sites can be identified by insertional inactivation of *lacZ*. Other

vectors derived from filamentous ssDNA phages include phagemids, which combine useful features of plasmids and ssDNA phages. Phagemids enable larger inserts than those used in the M13 mp series to be cloned, double-stranded as plasmids, and amplified and recovered as ssDNA in virions. A phagemid typically contains a selectable marker, an origin of plasmid replication, a cloning region (including polylinker and site for insertional inactivation), and the IG of an ssDNA phage for initiation and termination of DNA replication. Superinfection of the host with an ssDNA helper phage provides phage proteins in trans for single-stranded phagemid DNA replication and those for phage capsid and assembly. Defective helper phages can be used to reduce helper phage titer. Various phagemids have been developed, including those for phage display.

In phage display a peptide/protein is fused to one of the phage coat proteins so that it is displayed on the surface of the virion. The DNA sequence encoding the peptide is cloned in the phage/phagemid vector. Display libraries are produced by cloning large numbers of DNA sequences. Displayed peptides of interest can then be selected on the basis of their ability to bind to specific targets by a process of biopanning. This allows high throughput analysis of protein–protein interactions. After several rounds of screening, phages displaying peptides with the tightest binding capacity will be enriched. There is a direct link between the nucleic acid insert (genotype) and the displayed peptide (phenotype).

Early display systems used phage vectors with peptides fused to pVIII or pIII. Only small peptides with six to eight residues could be efficiently displayed; larger peptides compromised phage infectivity. The problem of fusion protein size could be overcome by using phagemid vectors with a single copy of the chosen coat protein gene for fusion, the wild-type copy of that gene being on the helper phage. Both wild-type and fusion coat proteins contribute to the phage particles so that the surface is phenotypically mixed, and the phages still infective.

A notable application of phage display technology has been in the engineering of antibodies, in which antibody variable domains have been used to construct combinatorial antibody libraries in phagemid vectors. Affinity binding is employed for rapid selection of antibody fragments binding targets of interest.

Phage display was first developed with filamentous phage M13. However, alternative phage systems, such as those of T phages, lambda (λ), and ssRNA phages, have also been employed. The technology promises to revolutionize the study of protein–protein interactions.

20.5 DOUBLE-STRANDED DNA PHAGES

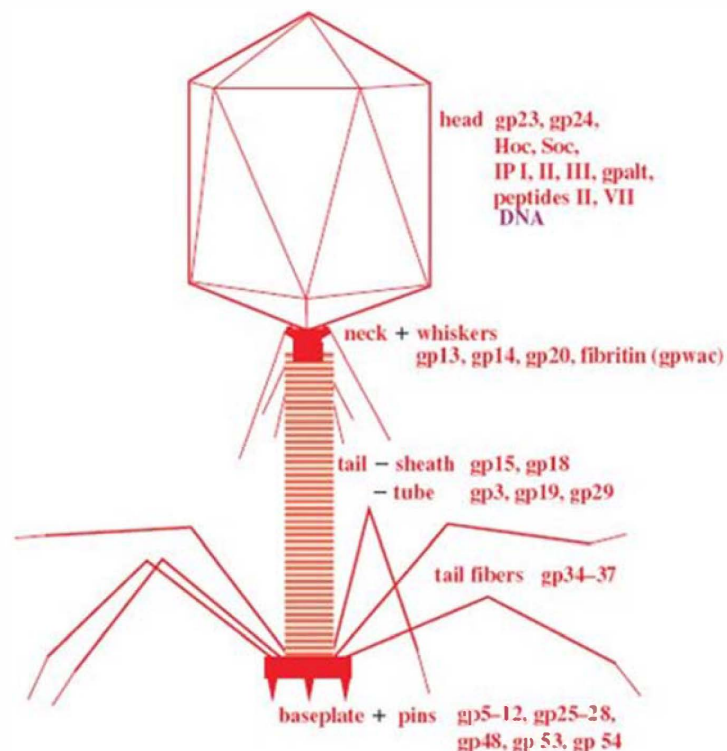
There is a large variety of dsDNA phages, of which the T phages and λ have been particularly well characterized. T2, T4, and T6 belong to the family Myoviridae (from the Greek mys, myos, “muscle,” referring to phages with contractile tails), T1 and T5 together with λ are members of the Siphoviridae (from the Greek siphon,

“tube,” referring to phages with long, flexible, non-contractile tails), and T3 and T7 are in the Podoviridae (from the Greek pous, podos, “foot,” referring to phages with short, non-contractile tails). A representative of each family is considered.

20.5.1 Phage T4

T4, a virulent T-even phage, is one of the largest phages and was the first prokaryotic “organism” providing evidence of gene splicing through the presence of introns in the genome. It has a complex morphology, consisting of an elongated icosahedral head with a number of different types of protein, including two accessory proteins, Hoc (highly antigenic outer capsid protein) and Soc (small outer capsid protein) decorating the external surface, a contractile tail, baseplate, tail fibers, and pins (see Figure 20.14).

Figure 20.14 Diagrammatic representation of T4 phage. The capsid comprises a multiple-protein, prolate icosahedral head separated from the tail by the neck. The tail consists of two helical arrangements of protein: one forming the sheath, the other the tube, with a collar and a baseplate with pins and tail fibers. The pins at the baseplate vertices anchor the phage to the host cell. Proteins are labeled with their corresponding gene name or number. Virus assembly requires a protein scaffold. The dsDNA is densely packed as concentric layers in the head. See Leiman et al. (2003).

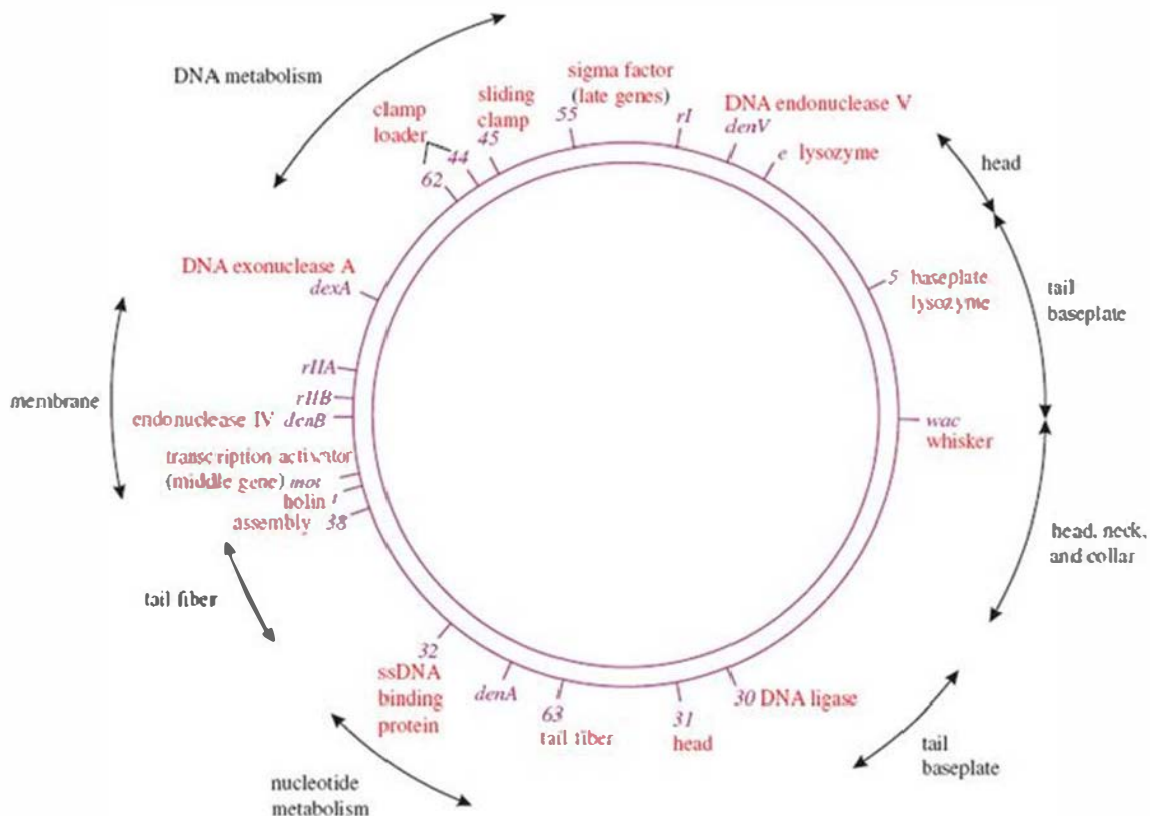


Binding of Soc stabilizes the capsid and may account for the osmotic shock resistance phenotype. The two accessory proteins have been utilized in display fusions (see also Section 20.4.2.d) with, for example, HIV capsid proteins to provide vaccine vectors. In addition, the tail fibers can be exploited in phage display and may be used as protein struts for nanodevices.

T4 tail fibers, pins, and baseplate are involved in binding to the lipopolysaccharide receptor of the *E. coli* host. The tail sheath contracts, driving the internal tail tube into the cell; tail lysozyme (gene product 5, gp5) facilitates digestion of the peptidoglycan layer to reach the inner membrane. The trans-membrane electrochemical potential is required for transfer of the T4 DNA to the cytoplasm. This occurs rapidly, in contrast to the situation with T7 infection (Section 20.5.2). The empty capsid remains extracellular. The phage shuts down the cell; host RNA polymerase stops recognizing host promoters and instead uses phage promoters for T4 gene expression, and the bacterial genome is degraded and recycled in phage DNA synthesis.

The T4 genome is linear dsDNA of about 169 kbp, which is circularly permuted and terminally redundant (a sequence of about 1.6 kbp at one end is directly repeated at the other) due to the mode of replication and packaging by the head-full mechanism. The genome is AT-rich and contains modified bases in the form of 5-hydroxy-methyl-cytosine, rather than cytosine, which protect the phage DNA from many host restriction systems and from phage-encoded nucleases that degrade cytosine-containing host DNA during phage infection. Furthermore, the majority of the hydroxy-methyl groups are glycosylated after DNA synthesis to counteract the host Mcr (modified cytosine restriction) systems, to which the hydroxy-methyl-cytosine residues confer susceptibility. The genome comprises about 300 probable genes assigned by the timing of their expression, with three different types of promoter: early (Pe), middle (Pm), and late (Pl). The early and middle genes encode functions for DNA replication and for regulating expression of the late genes, which encode head and tail components for the phage particles, and functions for cell lysis (Figure 20.15).

Figure 20.15 T4 genetic map showing approximate location of some genes and main functional groups. The map is presented as a circle. For nucleotide positions and functions of genes see Miller et al. (2003).



The regulatory circuit is complex and involves transcriptional termination and anti-termination, transcriptional activators, and replication-coupled transcription. Early genes rely on the transcription apparatus of the host, being transcribed from normal $\sigma 70$ promoters. Early gene products alter the host RNA polymerase in two ways for expression of middle genes. The enzyme is thereby able to read through a transcription terminator by an anti-termination mechanism (see Section 20.5.3 for an explanation) to express genes downstream of early genes and is modified to recognize middle promoters. These promoters differ from $\sigma 70$ promoters in the -35 sequence, and a transcriptional activator, encoded by the T4 *motA* early gene, allows the host RNA polymerase to transcribe from such promoters.

Transcription of the late genes is coupled to replication. Promoters for these genes again differ from $\sigma 70$ promoters and require an alternative sigma factor, encoded by T4 regulatory gene 55, to activate transcription. Binding of gene 55 product to host RNA polymerase allows the enzyme to recognize specifically T4 late gene promoters. Other T4 gene products, including those of genes 44, 45, and 62, which are involved in replication, are additionally required for late gene transcription. RNA polymerase complexed with gp55 seems to need an actively replicating T4 DNA for effective transcription of the late genes. Such coupling between replication and transcription ensures that phage DNA becomes available for packaging as the capsid proteins are synthesized from the late genes. The DNA is packaged into the pre-made heads. These require a protein scaffold to assemble. Packaging is size specific, depending on the size of the head, and is accompanied by a structural change in the procapsid. Tail and baseplate assemble separately. Heads attach to tails and tail fibers are added last to produce mature progeny phages. T4 holin

(gene t protein) creates a hole in the inner membrane so that lysozyme (gene e protein for endolysin) can degrade the peptidoglycan for cell lysis and release of phages.

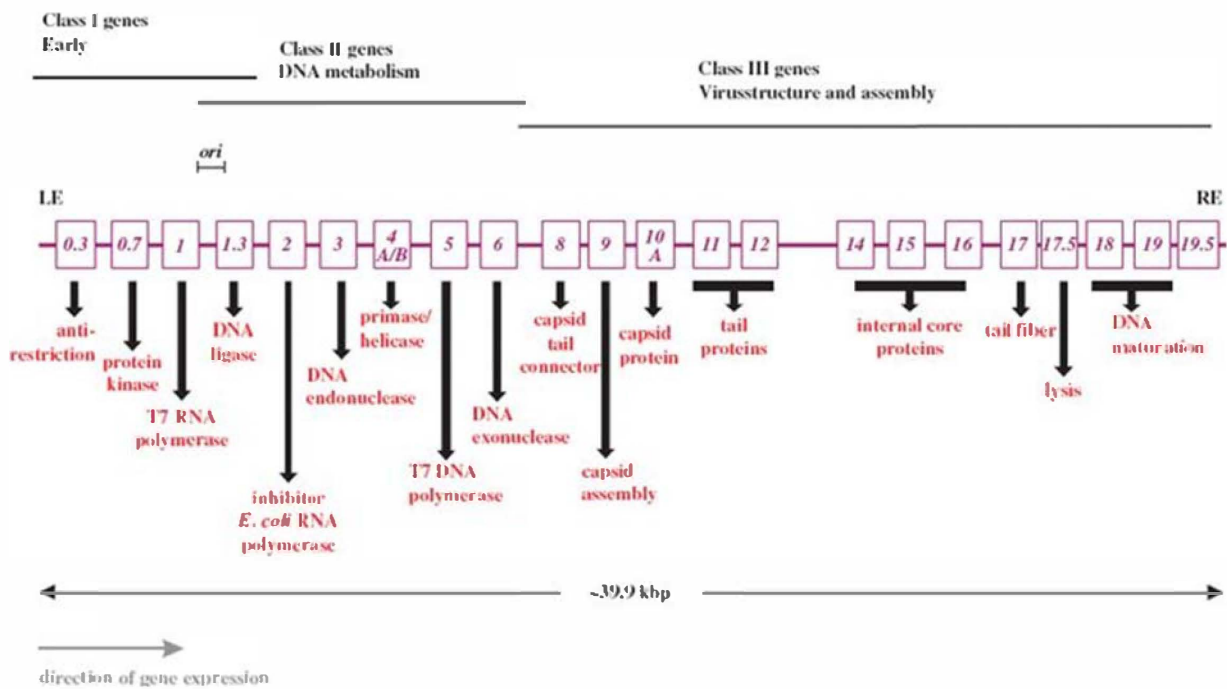
Upon superinfection of T4-infected bacteria, the phage exhibits the phenomenon of "lysis inhibition," whereby the latent period is prolonged, lysis is delayed, and an increased burst size (about 1000) results. It is noteworthy that T4-like phages have been suggested as possible therapeutic agents in, for example, the treatment of *E. coli* diarrhea. Lysis of the majority of diarrhea-causing *E. coli* strains would probably require a cocktail of several phages and/or genetic manipulation to extend host range sufficiently.

20.5.2 Phage T7

T7 is a virulent T-odd phage, with icosahedral head and short, non-contractile tail (Figure 3.18). The linear genome (about 39 kbp) is smaller than that of T4. It is terminally redundant though not circularly permuted and comprises about 50 genes, which are closely packed, divided into three groups and numbered from the genetic left end, the first end to enter the host cell (Figure 20.16). Integral numbered genes are essential; most non-integral numbered are non-essential.

Figure 20.16 Simplified genetic map of phage T7. The T7 genome contains about 50 genes, of which a number are shown. Genes are transcribed from the left end (LE) to the right end (RE), with gene 0.3 expressed first and gene 19.5 last. Three classes (I, II, and III) of genes are identified by function and timing of their expression. Gene 0.3 encodes an anti-restriction protein to avoid restriction of T7 DNA by type I restriction systems of the host immediately upon infection, allowing the phage to propagate. *E. coli* RNA polymerase transcribes class I genes, one of which encodes T7 RNA polymerase, which subsequently transcribes class II (middle) and class III (late) genes from T7-specific promoters. T7 proteins encoded by genes 0.7 and 2 inactivate the host RNA polymerase, switching transcription to the T7 rather than the bacterial genome. T7 proteins encoded by genes 3 and 6 degrade host DNA. T7 DNA is replicated by T7 DNA polymerase encoded by gene 5. *ori* is the primary origin of DNA replication.

Sources: Functions derived from Studier and Dunn (1983) *Cold Spring Harbor Quantitative Biology*, 47, 999–1007, and Steven and Trus (1986) In *Electron Microscopy of Proteins Volume 5: Viral Structure* (Harris, J. R. and Horne, R. W., editors) pp. 1–35, Academic Press.



T7 early genes (up to 1.3) are transcribed by the bacterial RNA polymerase, which is afterwards inactivated. One of these early gene products is T7 RNA polymerase, a very active enzyme that takes over transcription of T7 genes. Late genes are transcribed from strong T7 promoters. T7 transcription is regulated both by the slow delivery of the DNA into the host and by the synthesis of the T7 RNA polymerase, which recognizes only T7 promoters. This specificity of T7 RNA polymerase has been exploited in the development of a number of expression vectors that incorporate these promoters. Degradation of the bacterial chromosome during infection provides the nucleotides for T7 replication, which uses T7 DNA polymerase. This polymerase has been modified and marketed commercially as the DNA sequencing enzyme, Sequenase™.

20.5.3 Phage lambda (λ)

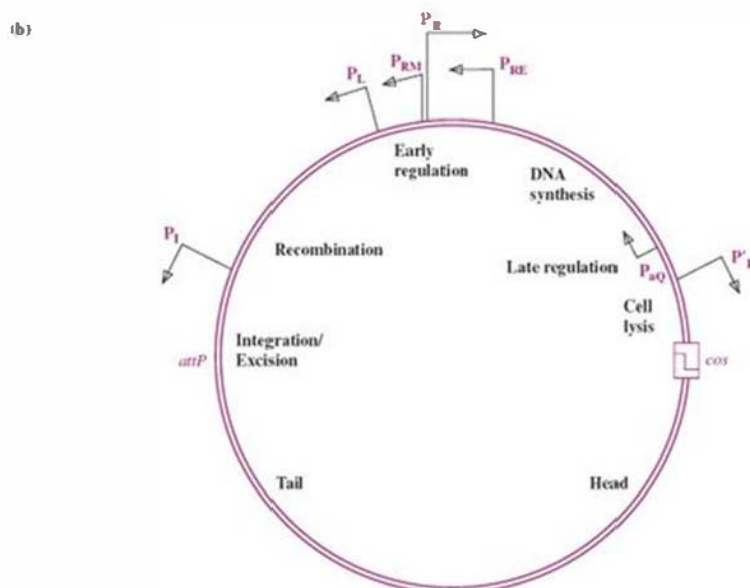
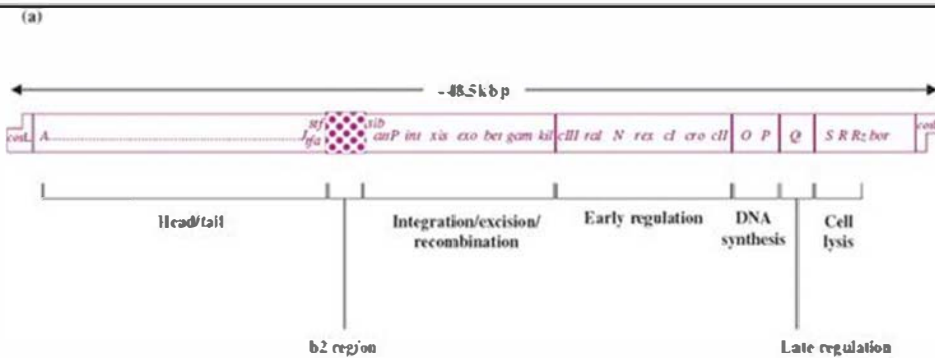
Morphologically, phage λ comprises an icosahedral head containing a dsDNA genome of about 48.5 kbp and a long, flexible tail with central tail tip fiber and side tail fibers. Genes are clustered according to function on the genome (see Figure 20.17), with many encoding head and tail components of the phage, but only two (O and P) for replication, unlike phages T4 and T7, where there are many phage-encoded replication proteins. The genome in the virion is linear, with 12 b single-stranded complementary 5' ends (cohesive ends) that mediate circularization of the DNA after infection of the host. Circularization of the λ genome protects it from degradation. The cohesive ends are generated due to the mode of packaging of the DNA from concatemers (containing multiple λ genomes), which are cut asymmetrically at λ *cos* sites. The specificity of this packaging mechanism has been exploited for efficient introduction of insert DNA into recombinant hosts using λ cloning vectors. To subvert bacterial restriction–modification activities upon infection, phage λ encodes the *Ral* (restriction alleviation) protein. This enhances

methylation of the viral DNA, thereby alleviating restriction by type I restriction enzymes.

Figure 20.17 Genetic organization in phage lambda. (a) Simplified genetic map (non-integrated). Some of the essential λ genes are shown. Gene products/functions are as follows: A to J, head and tail proteins and assembly; stf and tfa, side tail fibers; b2 region, non-essential genes, part of the region replaced in λ cloning vectors; sib, regulatory sequence for int and xis expression—int can be transcribed from PI and PL. Expression from PL is inhibited when sib is present. PL transcript terminates beyond sib and the site forms a secondary structure (hairpin) that is recognized for cleavage by RNase III, removing the int portion (= a form of retroregulation). PI transcript, which terminates at tI (upstream of sib), is stable and int expressed. PL transcript from integrated prophage DNA is also stable, since sib is removed, and int and xis expressed. attP (P.P' or a.a'), λ attachment site—location of site-specific recombination (15 b homology) with host chromosome integration site attB (B.B' or b.b') between gal and bio operons; int, integrase for integration and with xis λ prophage excision; xis, excisionase for λ prophage excision; exo, 5' to 3' exonuclease for recombination; bet, promotes annealing of complementary ssDNA strands for recombination; gam, inhibits host RecBCD nuclease; kil, host killing by blocking cell division; cIII, required for establishment of lysogeny, CIII blocks degradation of CII by host proteases (Hfl A and Hfl B); ral (restriction alleviation), for partial alleviation of host B and K restriction systems; N, anti-terminator protein promotes expression of delayed early genes, causes RNA polymerase to overcome transcription terminators and not stop transcription, requires nut (N utilization) site and host factors; rex [rexB/rexA] (restriction of exogenous phage), blocks superinfection by phage such as P22, expressed in the lysogen; cI, λ repressor—binds to OL and OR to repress transcription from PL and PR, activates transcription from PRM when bound to OR2, cI expressed from PRE to establish lysogeny and from PRM to maintain lysogeny; cro, anti-repressor—binds to OL and OR to repress early transcription, block activation of PRM and promote the lytic cycle; cII, regulatory protein for establishment of lysogeny, activates transcription from PRE for CI synthesis, PI for integrase synthesis, PaQ for antisense Q resulting in reduced expression of Q. Stability of CII controlled by host proteases and CIII. O, P, proteins for phage DNA replication; Q, anti-terminator protein promotes expression of late genes from P'R, requires qut site; S, holin protein, creates holes in cytoplasmic membrane permitting endolysin (R product) to gain access to periplasm for cell lysis—two forms of S protein are translated (from two in-frame translation start sites) and the ratio determines the timing of lysis to ensure progeny phages are ready for release before lysis occurs. R, an endolysin—a transglycosylase that digests peptidoglycan for cell lysis; Rz, for cell lysis; bor, lipoprotein in the membrane of λ lysogen conferring serum resistance on host. cosL, cosR, cohesive ends (left and right respectively)—12 b single-stranded

complementary 5' extensions, anneal to form circular genome in the host. Note that the integrated prophage DNA is a linear permutation of the virion DNA, such that *sib* is no longer downstream of *att*. (b) Simplified circular map. Circularization of genome by annealing of cohesive ends to form the *cos* site brings all late genes (cell lysis, head, tail genes) together. Main promoters are shown:

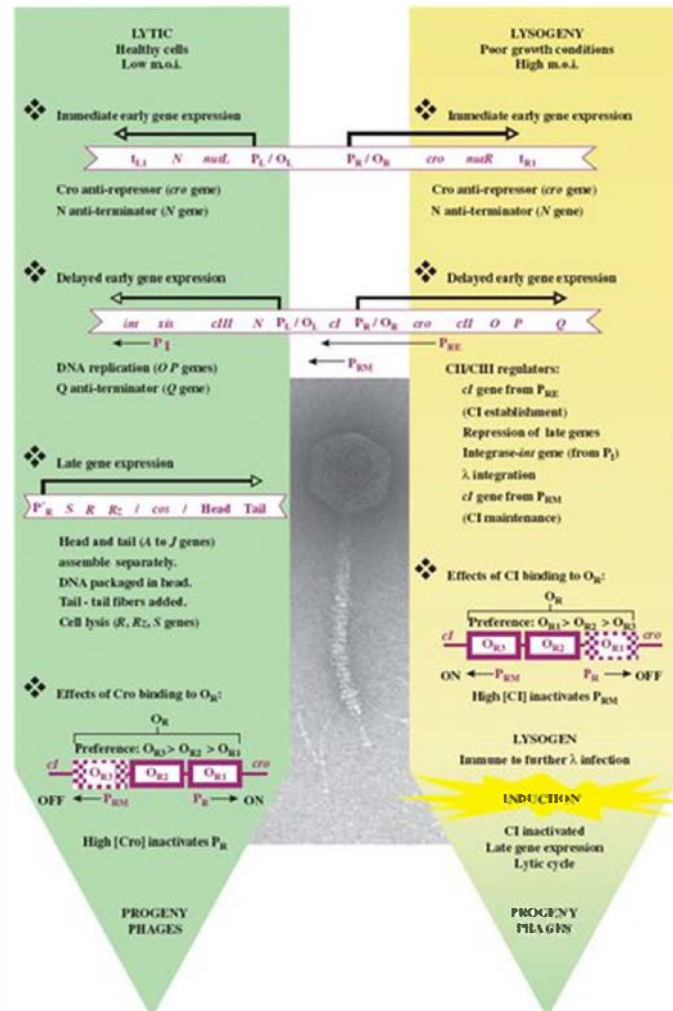
PL	promoter leftward
PR	promoter rightward
P'R	late promoter
PRE	promoter for λ repressor establishment
PRM	promoter for λ repressor maintenance
PI	promoter for <i>int</i>
PaQ	promoter for anti-Q.



As a temperate phage, λ can either establish lysogeny, being maintained as a prophage integrated into the bacterial chromosome between the biotin and galactose operons, or multiply lytically to produce progeny virions upon infection (Figure 20.18).

Figure 20.18 Requirements for lytic cycle and lysogeny in phage lambda. Important genes, gene products/functions required for the lytic cycle and lysogeny are shown. For description of gene functions see Figure 20.17.

Source: Electron micrograph of phage lambda reproduced by permission of Professor R. Duda.



The decision between the lytic cycle and lysogeny depends on a number of genetic, environmental and physiological factors. Gene expression is negatively regulated by repressor-operator interactions, with termination and anti-termination of transcription playing important roles in temporal control.

Phage λ infection involves the E. coli lamB protein receptor, which is recognized by the tip structure of the phage tail. Upon introduction into the host the phage genome circularizes and the bacterial RNA polymerase binds to leftward and rightward promoters (P_L and P_R), transcribing cro and N genes. N protein, together with host proteins, mediates anti-termination at transcription terminators $tL1$, $tR1$, and $tR2$, so that delayed early genes including cII and $cIII$ are expressed. Anti-terminator protein, N , acts by modifying the RNA polymerase, in the presence of nut (N utilization) sequences ($nutL$ and $nutR$), so that it ignores termination. N may additionally impede release of the transcript by the terminator hairpin or by the activity of ρ helicase. Where the concentration of CII (transcription activator

protein) remains below a critical level, which is the case in the majority of cells, transcription from PL and PR continues until Q anti-terminator switches on late genes governing synthesis of phage components for heads and tails, and R, Rz, and S for lysis. The S gene protein allows transport of λ endolysin, encoded by R and Rz, to the periplasm. The R gene product degrades the peptidoglycan of the cell wall, with the Rz product having a minor role in cell lysis. The effectiveness and specificity of phage lytic enzymes make them promising candidates for use as antibacterial agents (“enzymiotics”) in the control of pathogenic bacteria, including antibiotic-resistant bacteria.

Under conditions where CII protein prevails, it activates transcription of cI from PRE (promoter for repressor establishment) and of int from PI (promoter for integrase). This results in integration of the λ genome into the bacterial chromosome and repression of lytic functions, through CI repressor binding to OL and OR, for establishment of lysogeny. CII also activates PaQ for transcription of antisense RNA that reduces expression of Q. Lysogeny is maintained by continued synthesis of CI repressor from PRM (promoter for repressor maintenance). Bacterial proteins, such as HflA protease, and phage-encoded CIII protein affect stability and hence concentration of CII. These proteins are themselves affected by environmental stimuli (e.g., temperature and nutrients).

Physiological stress, effected for instance by exposure of a λ lysogen to UV irradiation, results in damage to DNA and induces the SOS response, in turn activating RecA co-protease. This stimulates a latent protease activity of CI with subsequent cleavage of the repressor, so that it no longer binds to OR and OL. Anti-repressor Cro in turn binds to OR and OL to prevent transcription of cI and stimulate its own transcription. The phage is thereby induced to enter the lytic cycle. Genes xis (for excisionase) and int (for integrase) are expressed and the λ genome excises to direct the lytic cycle.

The λ lysis–lysogeny decision represents a paradigm for control of gene expression in bacteria and lysogenic phages. Relatives of phage λ may carry verocytotoxin (VT) genes to convert their hosts into pathogens upon lysogenization, as is the case with E. coli O157. Phage λ , which itself carries bor, conferring serum resistance on the lysogen, provides a useful model system for studying such lambdoid phages.

Phages of the lambda family and certain filamentous phages (Section 20.4.2.c) are important vectors for spread of virulence genes. Thus, despite potential lethality to their hosts, such phages may confer selective advantages, through, for example, lysogenic conversion and superinfection immunity, to contribute to bacterial fitness in the environment.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- discuss the replication cycle and control of gene expression in ssRNA coliphages;
- outline the infection process of dsRNA phages;
- review the biology of the Ff and icosahedral ssDNA phages;
- describe the structure and replication cycle of dsDNA phages, T4 and T7;
- discuss the molecular events that govern the decision between the lytic cycle and lysogeny in phage λ ;
- explain some uses of phages in modern biotechnology;
- appreciate the role of phages in the evolution of bacterial diversity.

SOURCES OF FURTHER INFORMATION

General

Calendar, R. (ed.) (2006) *The Bacteriophages*, 2nd edition, Oxford University Press

Clokie, M. R. J. and Kropinski, A. (2009) *Methods in Molecular Biology 501: Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterisation and Interactions*, Humana Press

Kutter, E. and Sulakvelidze, A. (eds) (2005) *Bacteriophages—Biology and Applications*, CRC Press

McGrath, S. and van Sinderen, D. (2007) *Bacteriophage Genetics and Molecular Biology*. Caister Academic Press

Wackett, L. P. (2005) Bacteriophage. An annotated selection of World Wide Web sites relevant to the topics in environmental microbiology. *Environmental Microbiology*, 17, 299–300

Single-stranded RNA phages

van Duin, J. and Tsareva, N. (2006) Single-stranded RNA bacteriophages. In *The Bacteriophages*, 2nd edition, pp. 175–196 (Calendar R., editor), Oxford University Press

Double-stranded RNA phages

Mindich, L. (1999) Precise packaging of the three genomic segments of the double-stranded-RNA bacteriophage $\phi 6$. *Microbiology and Molecular Biology Reviews*, 63, 149–160.

Poranen, M. M., Tuma, R., and Bamford, D. H. (2008) Dissecting the assembly pathway of bacterial dsRNA viruses: Infectious nucleocapsids produced by self-assembly. In *Segmented Double-stranded RNA Viruses: Structure and Molecular Biology*, pp. 115–132 (Patton, J. T., editor), Caister Academic Press

Phage $\phi X174$

LeClerc, J. E. (2002) Single-stranded DNA phages. In *Modern Microbial Genetics*, 2nd edition, pp. 145–176 (Streips, U. N. and Yasbin, R. E., editors) Wiley-Liss

Filamentous single-stranded DNA phages

Russel, M., Lowman, H. B., and Clackson, T. (2004) Introduction to phage biology and phage display. In *Phage Display: A Practical Approach*, pp. 1–26 (Clackson, T. and Lowman, H. B., editors), Oxford University Press

Webster, R. (2001) Filamentous phage biology. In *Phage Display. A Laboratory Manual*, pp. 1.1–1.37 (Barbas, C. F. III, Burton, D. R., and Scott, J. K., editors), Cold Spring Harbor Laboratory Press

Phage T4

Karam, J. D. and Miller, E. S. (eds) (2010) Bacteriophage T4 and its relatives (review series) *Virology Journal*, 7, 288–360

Leiman, P. G. et al. (2003) Structure and morphogenesis of bacteriophage T4. *Cellular and Molecular Life Sciences*, 60, 2356–2370

Miller, E. S. et al. (2003) Bacteriophage T4 genome. *Microbiology and Molecular Biology Reviews*, 67, 86–156

Phage T7

Dunn, J. J. and Studier, F. W. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and locations of T7 genetic elements. *Journal of Molecular Biology*, 166, 477–535

Molineux, I. J. (2001) No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Molecular Microbiology*, 40, 1–8

Phage lambda

Gottesman, M. E. and Weisberg, R. A. (2004) Little lambda, who made thee? *Microbiology and Molecular Biology Reviews*, 68, 796–813

Oppenheim, A. B. et al. (2005) Switches in bacteriophage lambda development. *Annual Review of Genetics*, 39, 409–429

CHAPTER 21

Origins and Evolution of Viruses

CHAPTER 21 AT A GLANCE

Possible virus origins:

- RNA molecules that existed before cells
- cell components
- micro-organisms.

Viruses evolve as a result of:

- errors during nucleic acid replication;
- recombination between virus strains;
- reassortment between virus strains;
- acquisition of cell genes.

Evolution of viruses can be monitored by sequencing their genomes and creating phylogenetic trees:

```
graph LR; Root --- Node1; Node1 --- A; Node1 --- Node2; Node2 --- B; Node2 --- Node3; Node3 --- C; Node3 --- D; Node3 --- E;
```

New viruses may evolve after viruses infect new host species, e.g. HIV-1 and HIV-2.

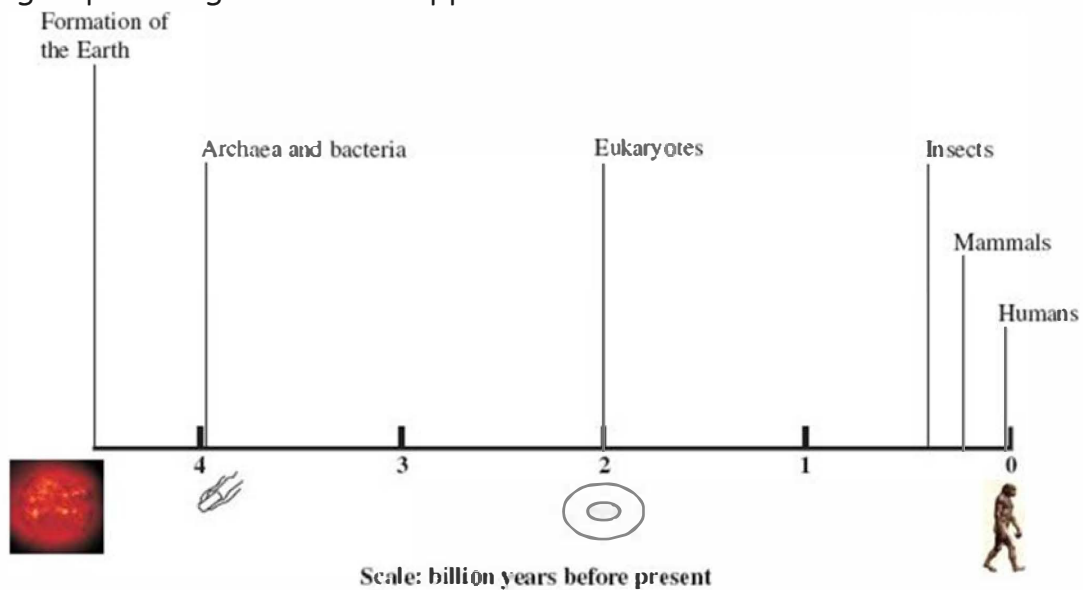
21.1 INTRODUCTION TO ORIGINS AND EVOLUTION OF VIRUSES

In this chapter we shall speculate on how viruses may have originated and consider the mechanisms by which viruses evolved in the past and continue to evolve today. The origins and evolution of many cellular organisms can be inferred from fossils, but the fossil record of viruses is limited. The discussion of the possible origins of viruses will be highly speculative, though when we come to mechanisms of virus evolution we will be on firmer ground as sequencing of virus genomes enables the construction of phylogenetic trees and the monitoring of virus evolution as it occurs. Virus evolution is of great importance in a number of areas of applied virology, including the emergence of new pathogenic viruses (Chapter 22) and the development by viruses of resistance to drugs (Chapter 26).

21.2 ORIGINS OF VIRUSES

Viruses by definition are parasites of cells, so there could be no viruses until cells evolved. There is evidence for the presence of living cells about 3.9 billion years ago (Figure 21.1), though the first cells probably came into existence several hundred million years earlier. It is believed that, prior to the development of these primitive prokaryotes, there was a phase of evolution involving organic molecules. These molecules probably included proteins and RNAs, and some of the latter may have evolved capacities for self-replication.

Figure 21.1 Timeline of the Earth's history. The approximate times when various groups of organisms first appeared are indicated.



The archaea and the bacteria that inhabit the earth today are the prokaryotic descendants of the early cells. It is likely that viruses developed at an early stage in these primitive prokaryotes, but the extent to which the viruses of the modern prokaryotes (Chapter 20) resemble these early viruses is unknown. Eukaryotic cells appeared much later in evolution (Figure 21.1), so on the evolutionary timescale the viruses that infect eukaryotes are much more recent than those of the prokaryotes. The viruses of modern mammals and birds are presumably descended from those that infected their dinosaur ancestors.

The answer to the question "Where did viruses come from?" is "Basically, we do not know!" We can, however, speculate about possible ancestors of viruses; these include molecular precursors of cellular organisms when the earth was young, components of cells, and intracellular micro-organisms. We shall consider each of these in turn.

21.2.1 Molecular precursors of cellular organisms

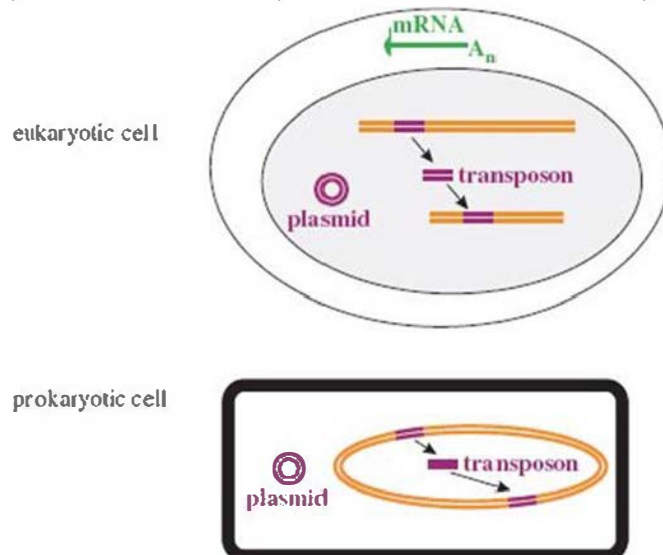
It is thought that on the early Earth, before cellular organisms had come into existence, RNA molecules underwent evolution, developing enzyme activities (ribozymes) and the ability to replicate. Once cells had evolved perhaps some were

parasitized by some of these RNA molecules, which somehow acquired capsid protein genes. If this happened then these were the first viruses.

21.2.2 Components of cells

Perhaps some cellular components evolved abilities to replicate themselves, independent of host cell control, and thus became parasites of those cells. Potential candidates for precursors of viruses include mRNA molecules and DNA molecules such as plasmids and transposons (Figure 21.2).

Figure 21.2 Cell components that are potential candidates as precursors of viruses.



The genomes of many plus-strand RNA viruses of eukaryotic hosts have characteristics of eukaryotic mRNA. For example, the 5' end of the genome is capped in tobamoviruses, the 3' end is polyadenylated in picornaviruses, while the genomes of coronaviruses are both capped and polyadenylated. The plus strands in the dsRNA genomes of reoviruses are also capped and polyadenylated. Could some RNA viruses be descendants of cell mRNAs?

Covalently closed circular DNA molecules known as plasmids are found in prokaryotic and eukaryotic cells. Could some DNA viruses be descendants of such molecules? Some bacterial plasmids carry genes that specify the production of a protein tube (a sex pilus) that can attach to other bacterial cells and allow the transfer of a copy of the plasmid from one cell to the other. Could some phages, such as the filamentous phages or some of the tailed phages (Chapter 20), be descended from ancient plasmids and their sex pili? On the other hand it has been suggested that the sex pili of modern bacteria may have originated from filamentous phages.

Transposons are sequences of mobile DNA in the genomes of prokaryotes and eukaryotes. They are described as mobile because they can move from one part of the genome to another, usually by a cut and paste mechanism. Perhaps some DNA viruses are descendants of transposons.

For any of these DNA or RNA molecules to evolve into a virus the putative virus genome would somehow have to acquire a range of genes, including those for

capsid protein(s) and, in many cases, for a polymerase to replicate the virus genome.

21.2.3 Intracellular micro-organisms

There is strong evidence that the mitochondria and chloroplasts of eukaryotic cells are derived from prokaryotic cells that adopted new modes of life within the cells of host organisms. It is assumed that the ancestors of these organelles adopted parasitic or mutually beneficial modes of life in host cells, and that over time they became increasingly dependent on their hosts, losing the ability to perform various functions, and losing the genes that encode those functions. Perhaps a similar evolutionary process continued further, leading to greater degeneracy and loss of functions such as protein synthesis, until the intracellular intruder was no longer a cell or an organelle, but had become a virus.

A virus that may have arisen in this way is the mimivirus (Section 1.3.1). The 1.2 Mb genome of this virus encodes a wide variety of proteins, including enzymes for polysaccharide synthesis and proteins involved in translation; it also encodes six transfer RNAs.

21.2.4 How did viruses originate?

The answer to this question is that we do not know! The great variety of virion structures, and of virus genome types and replication strategies, indicates that viruses had multiple origins. Small simple viruses, such as parvoviruses and picornaviruses, perhaps evolved from molecular precursors, while some large complex viruses, such as the mimivirus, perhaps evolved from cellular precursors.

We noted in Chapter 3 that the viruses found in certain categories of organism have a predominance of certain types of genome: dsDNA genomes in prokaryote viruses, ssRNA genomes in plant viruses, and dsRNA genomes in fungal viruses. We can now note that certain structural types of virus are restricted to particular categories of host: naked, rod-shaped ssRNA viruses are restricted to plant hosts, while viruses with a head-tail structure are rarely found outside the prokaryotes. The reasons for these distributions presumably concern diverse origins of the viruses in these very different hosts.

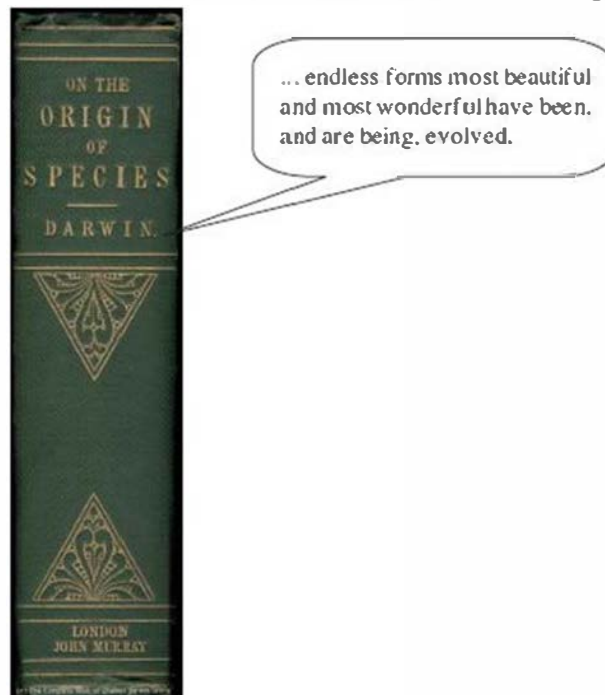
There is one feature, however, that is universal throughout the virus world: icosahedral symmetry. There are icosahedral viruses with dsDNA, ssDNA, dsRNA, and ssRNA genomes, and icosahedral viruses infect bacterial, archeal, fungal, plant, and animal hosts. These viruses have evolved from a variety of origins, but evolution has chosen the icosahedron for capsid design because it is a particularly stable structure.

One final thought on virus origins: we must entertain the possibility that today new viruses are continuing to evolve from molecular and cellular precursors.

21.3 EVOLUTION OF VIRUSES

In his book *On the Origin of Species* (Figure 21.3), Charles Darwin presented concepts of over-reproduction and survival of the fittest in terms of animals and plants. These same concepts can also be applied to viruses. (It can also be argued that the myriad exquisite viral structures entitles viruses, as well as plants and animals, to be included in the *ВЪНЪ*. . . endless forms most beautiful and most wonderful . . .*ВЪНЪ*) Like a cellular organism, a virus has genes that strive to perpetuate themselves. The expression of the combination of genes that constitutes the genome permits the virus to replicate itself. Changes to individual genes and new gene combinations continually throw up new genotypes, most of which are less successful than the parental genotypes and do not survive. Occasionally, however, a new genotype is more successful than the parental genotype and might even supplant it. Sometimes a new virus genotype enables the infection of a new host species.

Figure 21.3 Quotation from Charles Darwin's *On the Origin of Species* (1859).



21.3.1 Viruses from the past

The reverse transcription polymerase chain reaction (RT-PCR) has been used to recover and amplify genomes of at least two RNA viruses from the past: influenza A virus from the 1918-19 pandemic (see Sections 16.4 and 16.5) and tomato mosaic virus. The latter virus is very stable and it can be detected in the atmosphere. Its presence in Arctic ice was predicted, and when cores of ice between 500 and 140 000 years old were examined, the RNAs of 15 strains of tomato mosaic virus were detected.

Virus fossils can also be found as DNA sequences in the genomes of plants, animals and bacteria. These sequences originated from integrations of viral

sequences into the genomes of the ancestors of these organisms, and in many cases only fragments of the original sequences remain.

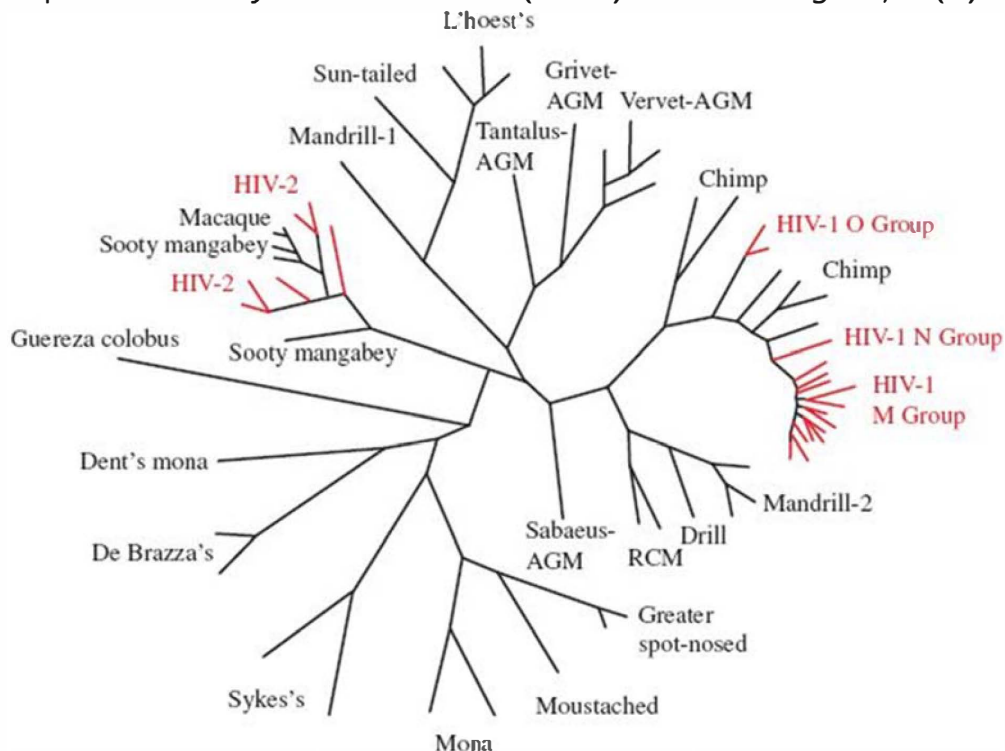
Examples of virus sequences in cell genomes include those of caulimoviruses in plants, flaviviruses in mosquitoes, and hepadnaviruses in birds. Mammalian genome sequencing has uncovered sequences originating from a wide range of DNA and RNA viruses, including retroviruses. The retroviral sequences are known as endogenous retroviruses (Section 17.6) and are the likely sources of retrotransposons (Section 21.3.5.a).

21.3.2 Monitoring virus evolution

Evolutionary relationships between viruses can be found by comparing their genome sequences. The relationships can be portrayed in phylogenetic trees (Section 10.2.1; Figure 21.4). Genome sequencing can also be used to follow changes in rapidly evolving viruses, such as influenza A virus and HIV-1.

Figure 21.4 Phylogenetic tree showing relationships between primate lentiviruses. The tree is based on pol sequences. Simian immunodeficiency viruses (SIVs) are denoted by the name of the host species in black. Human immunodeficiency viruses (HIVs) are indicated in red. AGM: African green monkey. RCM: red-capped mangabey.

Source: Adapted from Mjller and De Boer (2006) PLoS Pathogens, 2 (3): e15.



21.3.3 Mechanisms of virus evolution

In many respects, the underlying processes that drive virus evolution are the same as those that drive the evolution of cellular organisms. These processes involve the generation of genome variants, the vast majority of which are deleterious and do not survive, but a few provide an advantage in a particular niche. For a virus the

niche might be a new host species or the presence of an anti-viral drug, and a variant may proliferate in that niche as a new virus strain. Virus genome variants arise as a result of mutations, recombination, reassortment of genome segments, and acquisition of cell genes. Some viruses are able to undergo two or more of these processes. We shall consider each process in turn.

21.3.3.a Mutations

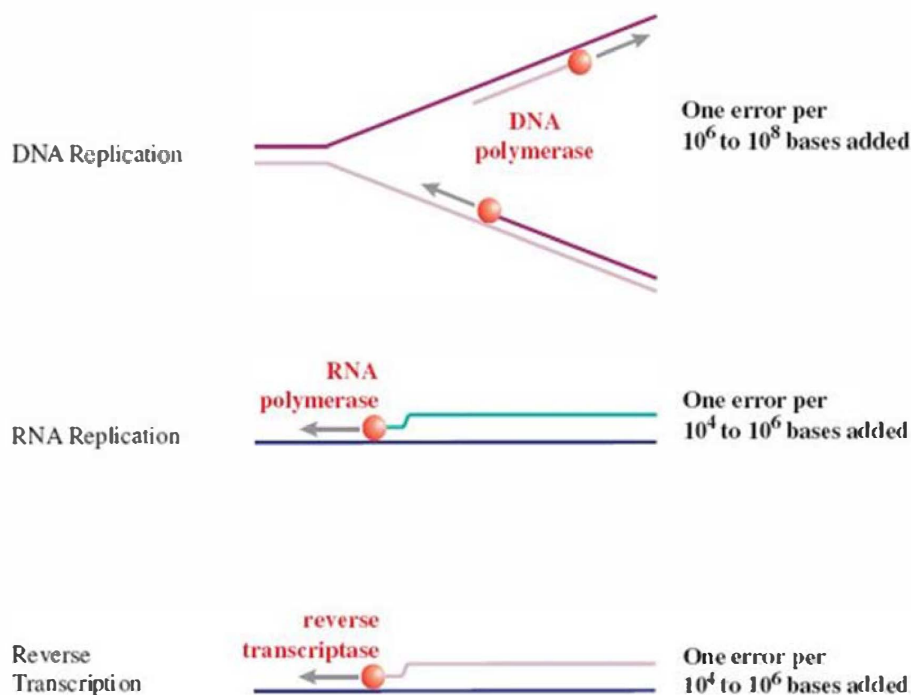
When nucleic acids are copied by polymerases some errors are made. If an error is in a protein-coding sequence and if the error results in a change in the amino acid encoded, then the error results in a mutation. Natural selection operates and the mutations that survive are those that best fit the virus for its continued survival.

There are many selection pressures on viruses. These pressures include the immune response of the host; for example, a new antigenic type of an animal virus that the host's immune system has not previously encountered is at an advantage compared with antigenic types against which the host has acquired immunity. There is, therefore, heavy selection pressure on virus proteins (e.g. HIV-1 gp120) that are targets of the host's immune response, and these proteins tend to be the least conserved. Those regions of the proteins that are targets for neutralizing antibodies are especially variable. There are, however, constraints on the evolutionary process. Attachment proteins must retain the configurations that enable them to bind to cell receptors, and enzymes, such as reverse transcriptases, must retain their catalytic abilities.

For some viruses there are additional constraints. A virus such as potato yellow dwarf virus, which needs to replicate in both a plant host and an insect host, can tolerate a mutation that enhances its replication in the insect host only if the mutation does not compromise its ability to replicate in the plant host, and vice versa.

When virus DNA genomes are replicated, the error rates are much lower than when RNA genomes are replicated (Figure 21.5). This is because DNA-dependent DNA polymerases have a proofreading mechanism that corrects most of the errors, whereas RNA polymerases and reverse transcriptases have no such mechanism. As a result, DNA viruses, such as papillomaviruses, evolve much more slowly than RNA viruses, such as picornaviruses.

Figure 21.5 Error rates of DNA polymerases and RNA polymerases.

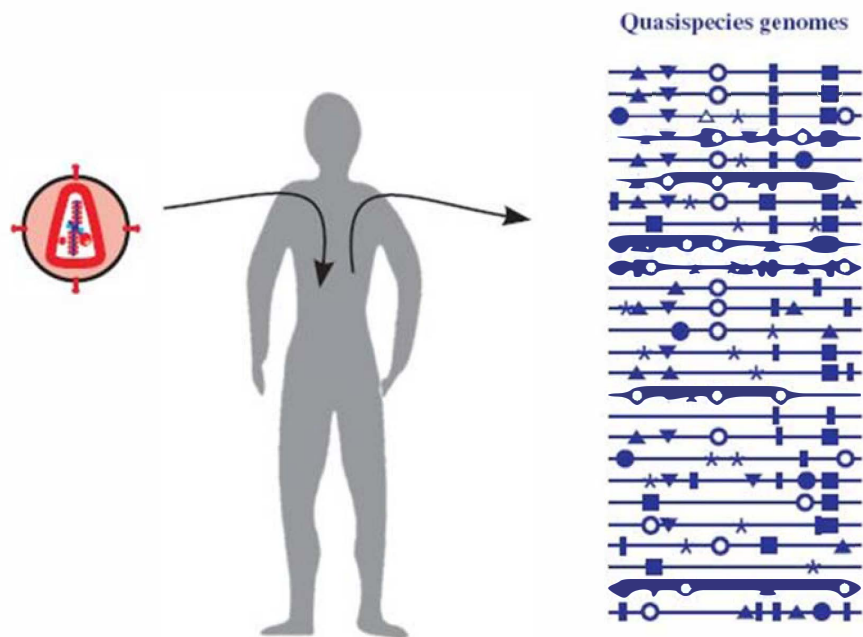


RNA viruses deserve their reputation as Nature's swiftest evolvers. Michael Worobey and Edward Holmes (1999)

The high error rates during RNA replication and reverse transcription mean that for an RNA virus there is no fixed sequence of bases for the genome. Instead, the virus genome exists as a large number of variants; the term *quasispecies* has been coined to describe the group of variants that collectively constitute the genome of an RNA virus (Figure 21.6). Many variants have only a fleeting existence, while those best adapted to a particular niche dominate in that niche. The high mutation rate produces some variants that may enable a virus to adapt to new niches (e.g., an expansion of its host range) should the opportunities arise.

Figure 21.6 Quasispecies genomes of an RNA virus in an infected human. Virus genomes are represented as lines, and mutations as symbols on the lines. Mutations result in a spectrum of related genomes, termed viral quasispecies.

Source: Modified with the permission of the author and the publisher (Springer Verlag) from Domingo et al. (2006) *Current Topics in Microbiology and Immunology*, 299, 51.



The first step in the replication of HIV genomes involves reverse transcription, with an estimated error rate around one error per 10^4 bases added. As the HIV genome contains about 10^4 bases, this means that there is an average of one error each time the genome is reverse transcribed. There is evidence that the human immunodeficiency viruses evolved from viruses that jumped from chimpanzees and sooty mangabeys into humans (Section 21.3.4). Since jumping species, the viruses have evolved rapidly, especially the M (main) group of HIV-1, which has radiated into many subtypes and sub-subtypes.

Rapid evolution of HIV can be detected within the body of an infected individual, in which there may be 10^{10} – 10^{12} new virions produced each day. There is therefore constant production of a huge amount of genetic variability. We mentioned in Chapter 18 that R5 strains of HIV-1 predominate in a newly infected host, but in many individuals as the infection proceeds there is evolution towards X4 and R5X4 strains. The constant production of new variants also means that the virus can evolve rapidly in response to a selection pressure, such as the presence of an anti-retroviral drug.

High diversity is also found in the genomes of other RNA viruses, such as hepatitis C virus, where isolates from a patient can have up to 5% difference in their nucleotide sequences.

The capacity for rapid evolution of RNA viruses creates problems for the development and maintenance of some vaccination programs. Vaccination with attenuated poliovirus vaccine initiates infection in the gut with one or more of the vaccine strains, and the infection normally lasts for 1–2 months. Some immunodeficient vaccinees, however, remain infected for much longer periods of time, and in these individuals the vaccine viruses almost always evolve. There are differences of only a few bases between the genomes of neurovirulent poliovirus strains and those of the vaccine strains (Section 25.2), so reversion to

neurovirulence is common in these chronically infected individuals, who may develop paralytic poliomyelitis and may transmit virulent poliovirus to others.

Another example of RNA virus variability having practical consequences for vaccination is foot and mouth disease virus. Seven distinct serotypes of this virus have evolved (Figure 10.1(a)), and vaccination against one serotype does not provide protection against the others. In addition, there are many subtypes of each serotype and immunity to one subtype confers only partial immunity to the related subtypes.

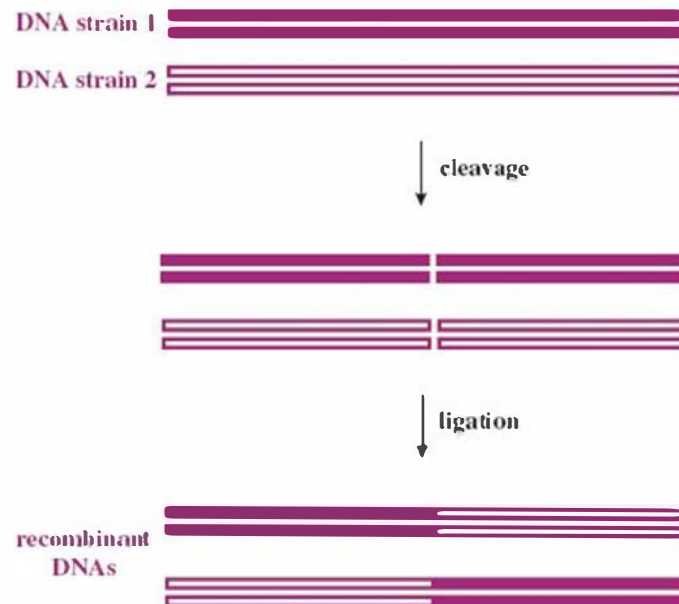
Hepatitis B virus (HBV) is a DNA virus, but, because it replicates its genome by reverse transcription, error rates similar to those of RNA viruses are observed. More than half of the HBV genome is read in two reading frames and regulatory elements are all within protein coding regions (Section 19.6). There are, therefore, severe constraints on the ability of HBV to mutate and remain viable. In spite of these constraints, however, variation is much greater in HBV than in DNA viruses that replicate their genomes solely with DNA-dependent DNA polymerases. HBV mutation rates are lower in overlapping regions of the genome.

21.3.3.b Recombination

Recombination is a process that results in the production of a new genome derived from two parental genomes. For a cellular organism recombination can take place between DNA molecules within a cell. For viruses recombination may occur when a cell is infected with two related viruses; the new virus that is produced is referred to as a recombinant virus. Recombination has been reported for viruses of all four genome types (ssDNA, dsDNA, ssRNA, dsRNA).

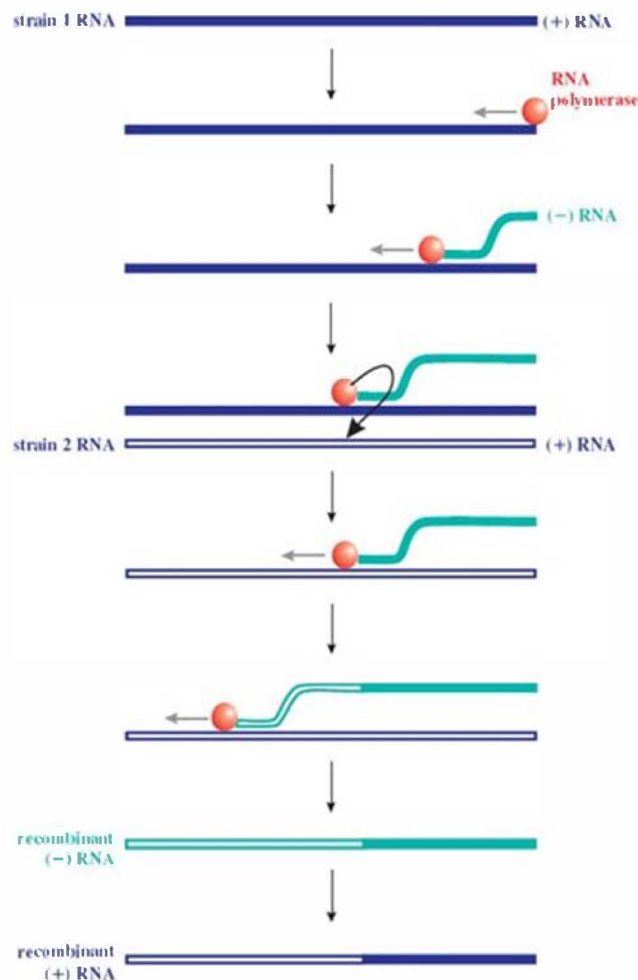
The first evidence for recombination in a virus came from studies with the DNA phage T2, and since then recombination has been documented for other DNA viruses. The formation of recombinants between different baculoviruses (DNA viruses that infect invertebrates such as insects), and between herpes simplex viruses 1 and 2, has been reported. DNA viruses undergo recombination by cleavage of the DNA molecules followed by ligation (Figure 21.7).

Figure 21.7 Recombination between the DNAs of two virus strains.



Recombination is common in ssRNA viruses, and there are reports of recombination in dsRNA viruses, including rotaviruses and bluetongue virus. Some RNA viruses may be able to undergo recombination by mechanisms similar to the cleavage/ligation mechanism just described for DNA viruses. For ssRNA viruses such as picornaviruses, it is thought that recombinant genomes can also be produced by a template switching mechanism. This involves displacement of an RNA polymerase while synthesizing RNA, followed by association of the enzyme with another RNA molecule and resumption of RNA synthesis using the new template (Figure 21.8).

Figure 21.8 Template switching mechanism of recombination in ssRNA viruses. A cell is infected with two strains of a virus. The RNA polymerase starts to replicate strain 1 RNA, but before replication is complete the polymerase and the growing strand of RNA dissociate from the template RNA. The polymerase associates with strain 2 RNA and RNA synthesis is completed on the new template. The result is a recombinant RNA with part of its sequence derived from strain 1 and part from strain 2. When this strand is used as a template to generate new genome strands, the resulting RNA molecules are also recombinant.



Recombination is common in retroviruses and pararetroviruses. It occurs during reverse transcription (Sections 17.3.2 and 19.8.6) by a template switching mechanism similar to that described above, with the reverse transcriptase switching between two RNA templates. For a retrovirus the two RNA templates are in the infecting virion, derived from two virus strains co-infecting the cell in which the virion was assembled. For a pararetrovirus the RNA templates are present in the cytoplasm of a cell co-infected with two virus strains. Some people have been infected with two or even three different strains of HIV-1, and many recombinant strains have evolved; for example, a recombinant derived from HIV-1 subtypes A and E has become prevalent in Thailand.

Virus recombinants with desired combinations of genes can also be generated in the laboratory. Examples are the recombinant vaccinia virus that was produced for vaccination of foxes (Section 15.2.1) and recombinant baculoviruses created for use in gene expression systems.

21.3.3.c Reassortment

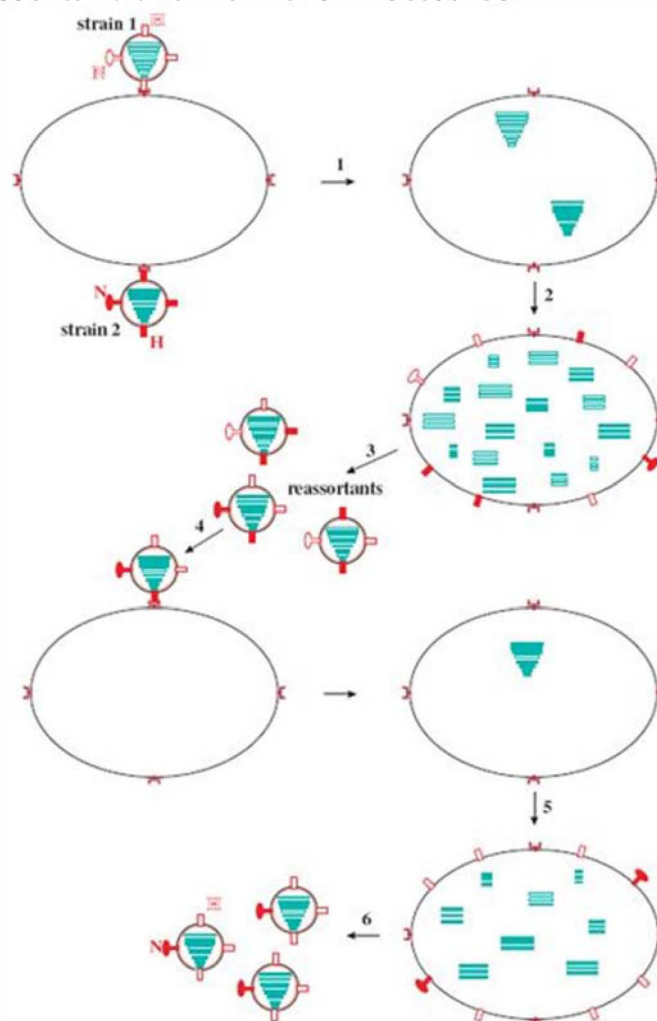
Reassortment is a category of recombination that may occur with segmented genome viruses that have all the segments packaged in one virion; for example, reoviruses (Chapter 13), influenza viruses (Chapter 16), and bunyaviruses (Section 22.2.1). If a cell is co-infected by two strains of one of these viruses the progeny

virions may contain mixtures of genome segments from the two parental strains. These virions are known as reassortants.

The formation of influenza A virus reassortants is illustrated in Figure 21.9. The virus genome is composed of eight segments of minus-strand RNA; two of the segments encode the envelope glycoproteins hemagglutinin (H) and neuraminidase (N). The figure shows how reassortment can give rise to a new virus strain.

Figure 21.9 Influenza virus reassortment

1. A cell is co-infected with two strains of influenza virus. Each strain depicted has a unique hemagglutinin (H) and a unique neuraminidase (N).
2. The RNAs of both strains are replicated.
3. Progeny virions with genomes made up from mixtures of RNA segments from the parental strains are called reassortants. H and N proteins of both virus strains are present in the virion envelopes as the infected cell expresses the genes of both strains.
4. A reassortant infects another cell. This reassortant has the H gene of strain 1 (the fourth genome segment), while the other genes are derived from strain 2.
5. The RNAs of the reassortant are replicated.
6. Virions of the reassortant bud from the infected cell.



21.3.3.d Acquisition of cell genes

The discovery of similarities between some virus proteins and cell proteins has led to the suggestion that some viruses have captured cell genes. The genomes of some retroviruses contain oncogenes that they probably acquired from their host cells, which contain similar genes (proto-oncogenes). An example is the src gene of Rous sarcoma virus (Section 17.4); a very similar gene (c-src) is present in the cells of all vertebrates. A retrovirus might acquire an oncogene by recombination between the proviral DNA and host DNA. Alternatively, during transcription RNA synthesis might continue from the provirus to an adjacent proto-oncogene in the cell chromosome.

Some virus proteins modulate the immune system of the host, often by mimicking cell proteins. For example, some of the human herpesviruses encode proteins similar to cytokines and major histocompatibility complex proteins. Several large DNA viruses encode proteins similar to interleukin-10 and some of these proteins have been shown to suppress cell-mediated immune responses. Are some of the virus genes that encode these proteins derived from cell genes and, if so, how did the viruses acquire the genes?

One possible mechanism is recombination, though in order for this to occur the genomes of the virus and the cell must come into close proximity. This may happen with DNA viruses that replicate in the nucleus (e.g. herpesviruses), but there are many DNA viruses (e.g. poxviruses) that do not enter the nucleus, so cellular and viral genomes are never in close proximity.

A second possible mechanism for viral acquisition of a cell gene might be through synthesis of a DNA copy of a cell mRNA, followed by insertion of the DNA into the virus genome. The DNA synthesis would require a reverse transcriptase, possibly supplied by the cell or by a retrovirus co-infecting the cell.

21.3.4 Evolution of new viruses

From time to time a new virus appears, and mystery surrounds its origin. Often there are suspects for the origin of the virus and these suspects may be viruses in other host species. The genomes of the new virus and the suspects can be sequenced, and if a close match is found between the new virus and one of the suspects then the mystery has probably been solved. It is commonly found that a new virus is one that has infected a new host species in which some further evolution has then taken place.

Crossing the species barrier is one small step to man, one giant leap to mankind.

Mark Klemperer and Daniel Shapiro (2004)

Many virus infections across species barriers are dead ends for the viruses as no transmissions to further individuals occur, but sometimes a virus adapts to enable

transmission from host to host. A virus may have to undergo more than one adaptation in order for it to:

- replicate in the new host;
- evade immune responses of the new host;
- transmit to other individuals of the new host.

In 1978 a new parvovirus appeared in dogs and rapidly spread around the world. The genome sequence of this canine parvovirus was found to be more than 99% identical to that of a parvovirus in cats (feline panleukopenia virus). The new canine virus could infect dogs, but not cats. Further evolution of the virus has given rise to many strains that can infect both dogs and cats.

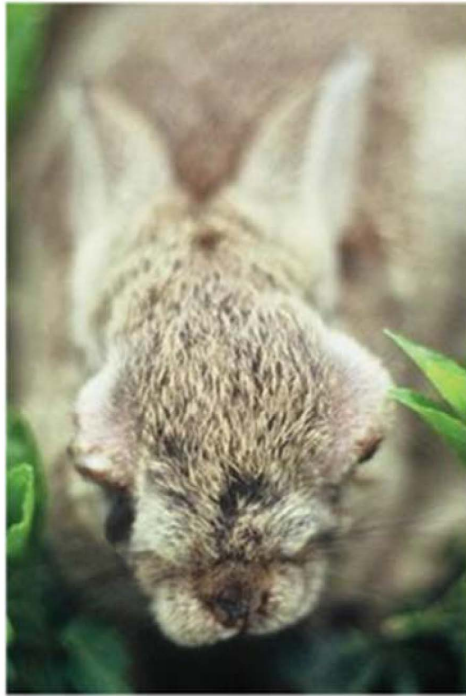
Similarly, the genome sequences of HIV-1 and HIV-2 are very similar to those of simian immunodeficiency viruses found in chimpanzees (SIVcpz) and sooty mangabeys (SIVsm), respectively (Figure 21.4). The human viruses are believed to have originated when these simian viruses jumped species, perhaps when humans came into contact with the blood of infected animals. Other examples of viruses that have emerged in this way are discussed in Chapter 22.

21.3.5 Co-evolution of viruses and their hosts

A virus–host association that has existed for a long period is likely to have evolved a relationship in which the host suffers little or no harm. Examples of such viruses are the dependoviruses (Chapter 12), and some reoviruses, which acquired the “orphans” in their name because they were found to be orphan viruses (not associated with any disease; Chapter 13). The members of these virus groups that infect *Homo sapiens* have probably been with us since we diverged as a separate species.

When a virus extends its host range into a new species it is often much more virulent in the new host, but the virus may then rapidly evolve to a less virulent form. A good example of co-evolution of a virus–host relationship is myxoma virus in the European rabbit. In its original host, a rabbit species in South America, myxoma virus infection results in the development of skin tumors, which eventually heal. In contrast, when the virus infects the European rabbit it causes myxomatosis, characterized by acute conjunctivitis (Figure 21.10), lumps on the skin, and loss of appetite. The disease is almost always fatal, and for this reason the virus was introduced into Australia as a biological control agent against the European rabbit. Within a year of its introduction there was evidence of virus evolution. Myxoma virus strains with attenuated virulence had evolved, apparently selected because they were transmitted more effectively. Subsequently, selection for rabbits with resistance to myxomatosis occurred and both the virus and its new host have continued to co-evolve.

Figure 21.10 Rabbit with myxomatosis.



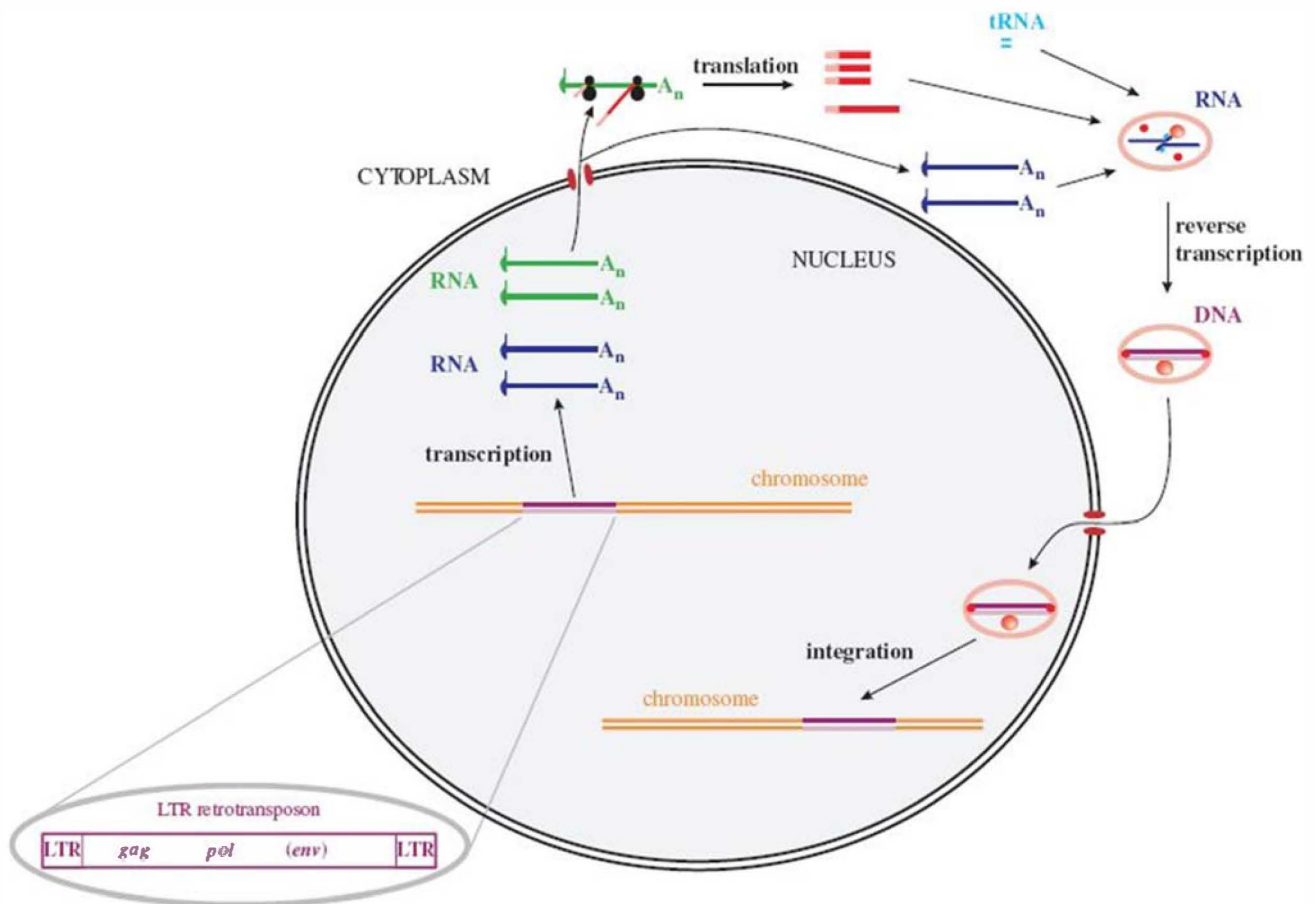
A virulent virus may affect the evolution of its host by causing genes for susceptibility to infection with the virus to become extinct, whilst selecting in favor of genes for resistance. HIV-1 is currently selecting for the 32-nucleotide deletion in the CCR5 gene that confers resistance to infection (Section 18.4.1).

21.3.5.a Integration of virus genomes into host genomes

The ultimate in virus–host co-evolution occurs if a virus genome becomes permanently integrated into the genome of its host. This is probably what happened to create many of the retrotransposons in the genomes of eukaryotes. Retrotransposons are sequences in cell genomes that are copied to new locations in the genome by a process involving reverse transcription, in contrast to transposons, which are DNA sequences that are cut and pasted (Section 21.2.2).

Some retrotransposons consist of two long terminal repeats (LTRs) flanking a gag gene, a pol gene (encodes a reverse transcriptase and an integrase), and sometimes an env gene. These retrotransposons thus resemble the proviruses of retroviruses and they are known as endogenous retroviruses (Section 17.6); they can replicate by transcribing plus-strand RNAs, which are translated into Gag and Gag–Pol proteins. Reverse transcription of the RNA produces a new copy of the retrotransposon, which can be inserted into another site in the cell genome (Figure 21.11).

Figure 21.11 Transposition by a LTR retrotransposon. LTR retrotransposons include endogenous retroviruses, which have gag and pol genes, and some have an env gene. Transcription results in the synthesis of RNA, from which Gag and Gag–Pol proteins are translated. Some of the RNA is reverse transcribed to DNA, which is integrated into a new chromosomal location.



Other types of retrotransposon consist of LTRs flanking only a *gag* gene, or non-retroviral genes, while further types of retrotransposon lack LTRs. It is thought that endogenous retroviruses and probably some other categories of retrotransposon are descended from retroviral proviruses that became inserted into the germ line of their hosts. It is also possible that some retroviruses originated from retrotransposons.

Retrotransposons are found in the genomes of vertebrates, invertebrates, plants, and fungi. The higher up the evolutionary tree, the greater percentage of the genome consists of retrotransposons. In the human genome this figure is over 40%. David Baltimore (one of the discoverers of reverse transcriptase) remarked, "In places, the genome looks like a sea of reverse-transcribed DNA with a small admixture of genes."

Retroviral sequences are not the only viral sequences in eukaryotic genomes. Sequences of a wide range of DNA and RNA viruses are being discovered as the genomes of more animals, plants, and fungi are sequenced.

The genomes of prokaryotes also contain sequences that originated as a result of integration of viral sequences through the process of lysogeny (Section 20.1). Phage sequences (prophages) have been found in many bacterial and archaeal genomes. Some of these prophages are inducible, meaning that they can be activated, virus replication can ensue, and the host can be lysed. Other prophages are defective, and replication cannot be induced. The genes for some

characteristics of bacteria are encoded by prophage sequences (e.g., the cholera toxin gene of *Vibrio cholerae* and toxin genes in *E. coli* O157).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- evaluate theories on the origins of viruses;
- explain how virus evolution occurs through mutation, recombination, reassortment, and acquisition of cell genes;
- assess the value of virus genome sequencing in studies of virus origins and evolution;
- assess the threats posed to man and animals by rapid virus evolution;
- discuss the co-evolution of viruses and their hosts.

SOURCES OF FURTHER INFORMATION

Books

Domingo, E. (ed.) (2006) *Quasispecies: Concept and Implications for Virology*, Current Topics in Microbiology and Immunology series, Volume 299, Springer

Domingo, E., Parrish, C. R., and Holland, J. J. (eds) (2008) *Origin and Evolution of Viruses*, 2nd edition, Elsevier

Holmes, E. C. (2009) *The Evolution and Emergence of RNA Viruses*, Oxford University Press

Kazazian, H. H. (2011) *Mobile DNA: Finding Treasure in Junk*, FT Press

Teycheney, P. Y. and Geering, A. D. (2011) Endogenous viral sequences in plant genomes. In *Recent Advances in Plant Virology*, pp. 343-362 (Caranta, C., Aranda, M. A., and Tepfer, M., editors), Caister Academic Press

Journals

Beauregard, A., Curcio, M. J., and Belfort, M. (2008) The take and give between retrotransposable elements and their hosts. *Annual Review of Genetics*, 42, 587-617

Hoelzer, K. et al. (2008) Phylogenetic analysis reveals the emergence, evolution and dispersal of carnivore parvoviruses. *Journal of General Virology*, 89, 2280-2289

Holmes, E. C. (2011) What does virus evolution tell us about virus origins? *Journal of Virology*, 85, 5247-5251

Johnson, W. E. (2010) Endless forms most viral. *PLoS Genetics*, 6 (11): e1001210

Klempner, M. S. and Shapiro, D. S. (2004) Crossing the species barrier—one small step to man, one giant leap to mankind. *New England Journal of Medicine*, 350, 1171-1172

Lauring, A. S. and Andino, R. (2010) Quasispecies theory and the behavior of RNA viruses. *PLoS Pathogens*, 6 (7): e1001005

Liu, H. et al. (2011) Widespread endogenization of densoviruses and parvoviruses in animal and human genomes. *Journal of Virology*, 85, 9863-9876

Lukashev, A. N. (2010) Recombination among picornaviruses. *Reviews in Medical Virology*, 20, 327-337

Meyerson, N. R. and Sawyer, S. L. (2011) Two-stepping through time: mammals and viruses. *Trends in Microbiology*, 19, 286-294

Nelson, M. I. et al. (2008) Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. *PLoS Pathogens*, 4 (2): e1000012

Sanjuán, R. et al. (2010) Viral mutation rates. *Journal of Virology*, 84, 9733-9748

Shi, B. et al. (2010) Evolution and recombination of genes encoding HIV-1 drug resistance and tropism during antiretroviral therapy. *Virology*, 404, 5-20

Sztuba-Solin

ś
ska, J. et al. (2011) RNA-RNA recombination in plant virus replication and evolution. *Annual Review of Phytopathology*, 49, 415-443

Worobey, M. and Holmes, E. C. (1999) Evolutionary aspects of recombination in RNA viruses. *Journal of General Virology*, 80, 2535-2543

Yamada, T. (2011) Giant viruses in the environment: their origins and evolution. *Current Opinion in Virology*, 1, 58-62

CHAPTER 22

Emerging Viruses

CHAPTER 22 AT A GLANCE

Emerging Viruses	Examples
• Viruses in new hosts	Nipah virus in pigs and humans
• Viruses in new areas	West Nile virus in North America
• Newly evolved viruses	Influenza virus reassortants
• Recently discovered viruses	Human metapneumovirus
• Re-emerging viruses	Measles virus

22.1 INTRODUCTION TO EMERGING VIRUSES

The term “emerging virus” is used in a number of contexts: it may refer to a virus that has recently made its presence felt by infecting a new host species, by appearing in a new area of the world, or by both. Sometimes a virus is described as a “re-emerging virus” if it has started to become more common after it was becoming rare. Foot and mouth disease virus (Section 14.2.5) re-emerges in the UK from time to time.

Human activities may increase the likelihood of virus emergence and re-emergence. Travel and trade provide opportunities for viruses to spread to new areas of the planet, and the introduction of animal species into new areas (e.g., horses into Australia) may provide new hosts for viruses in those areas.

Other activities that may result in virus emergence involve close contact with animals, including the hunting and killing of non-human primates for bushmeat. It has been shown that viruses such as simian immunodeficiency viruses (SIVs) can be present in the meat and there is a risk of acquiring an infection when the meat is handled. Simian-to-human transmission of SIVs is thought to have occurred several times, resulting in the major groups of HIV-1 and HIV-2 (Figure 21.4), and there is concern that further viruses might emerge as a result of contact with bushmeat.

If a virus jumps into a new host species it may undergo some evolutionary changes in the new host, resulting in a new virus. This is how HIV-1 and HIV-2 were derived from simian immunodeficiency viruses, and how a new canine parvovirus was derived from feline panleukopenia virus (Section 21.3.4). Other new viruses emerge when recombination and reassortment result in new viable combinations of genes (Section 21.3.3); new strains of influenza A virus come into this category. Some “new” viruses that are reported are actually old viruses, about which mankind has recently become aware, such as Kaposi’s sarcoma-associated herpesvirus (Chapters 11 and 23).

This chapter considers examples of viruses that have “emerged” and “re-emerged” in the late twentieth century and the twenty-first century. Most of the examples discussed are human viruses.

22.2 VIRUSES IN NEW HOST SPECIES

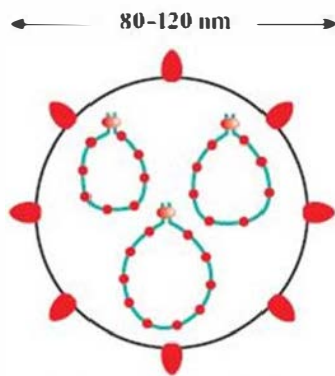
22.2.1 Bunyaviruses

In 1993 in a region of the south-western USA known as Four Corners, some residents became afflicted by an illness resembling influenza; many of them developed severe lung disease and died. The region is normally very dry, but there had been unusually high rainfall and snowfall that had resulted in a burst of plant growth, followed by explosions in the populations of small mammals. One of those mammals was the deer mouse, a species that is attracted to human habitation.

On investigation, it was found that many of the deer mice were persistently infected with a virus that was excreted in their urine, droppings, and saliva. Humans exposed to these materials were becoming infected and developing the respiratory illness. The virus was characterized as a new hantavirus and the disease was called hantavirus pulmonary syndrome.

Hantaviruses are members of the family Bunyaviridae (Figure 22.1). They are named after the River Hantaan in Korea, where the first of these viruses was isolated during the Korean War from soldiers who had developed hemorrhagic fever with renal syndrome. Similar viruses are known elsewhere in Asia and in Europe.

Figure 22.1 Family Bunyaviridae.



- Genome: ssRNA (three segments, which are either all minus strand, or a mixture of minus strand and ambisense). Each segment is circular as a result of base pairing at its ends.
- Nucleocapsid symmetry: helical
- Enveloped, with glycoprotein spikes
- Includes genus *Hantavirus*, e.g. Sin Nombre virus

There was much heated debate about naming the virus that had appeared in Four Corners. Local people did not want it named after a place as they did not want potential tourists to be discouraged by publicity surrounding a virus disease. Eventually the Spanish name Sin Nombre virus (the virus without a name) was agreed.

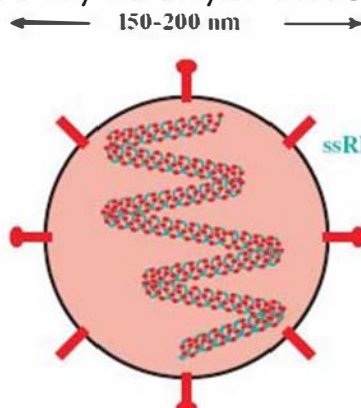
Since 1993 hantavirus pulmonary syndrome has been reported in many parts of North and South America, with mortality rates of around 40%. Some of these cases have been caused by Sin Nombre virus, while many others have been caused by newly described hantaviruses, such as Andes virus. The normal hosts of all these viruses are rodents, in which the viruses can establish persistent infections.

22.2.2 Paramyxoviruses

22.2.2.a Hendra virus

In 1994 at Hendra in south-east Australia, there was an outbreak of pneumonia in horses, then a trainer and a stable hand who worked with the horses developed severe respiratory disease. The horses and the two humans were found to be infected with a virus that had characteristics of the family Paramyxoviridae (Figure 22.2). This family contains well-known viruses such as measles and mumps viruses, but the virus that was isolated was previously unknown. Thirteen of the horses and the trainer died as a result of infection with the virus, which was named Hendra virus.

Figure 22.2 Family Paramyxoviridae.

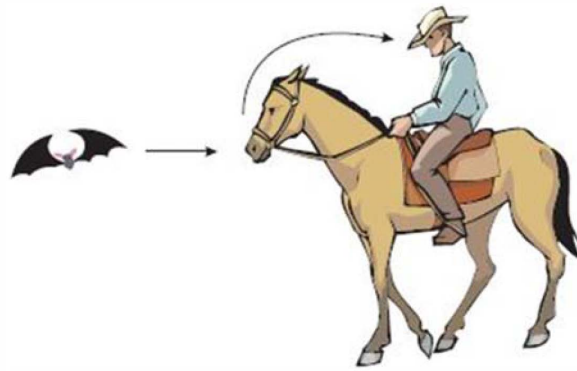


- Nucleocapsid symmetry: helical
- Enveloped, with glycoprotein spikes
- Includes Hendra, Nipah, measles, mumps, and human metapneumoviruses

A couple of years later a man living about 800 km north of Hendra developed seizures and paralysis. He was found to be infected with Hendra virus and he died from meningoencephalitis. It turned out that he had helped with post mortems of

two horses 13 months earlier. Tissue from the horses had been preserved and when it was examined it was found to contain Hendra virus.

It looked as though the infected humans had acquired the virus from horses, but the source of infection for horses was a mystery. An investigation was initiated, and evidence was obtained that the virus is present in populations of fruit-eating bats; Hendra virus-specific antibodies were found in all four local species of the bat genus *Pteropus*, and infectious virus was isolated from one bat. Bats appear to be the normal hosts of Hendra virus, and experimental infection of several bat species has provided no evidence that infection causes disease in these animals. It appears that occasional transmission to horses occurs, with very occasional transmission from horses to humans (Figure 22.3). In horses and humans the outcome of infection is very different, with the virus demonstrating a high degree of virulence. Figure 22.3 Hendra virus transmission routes.



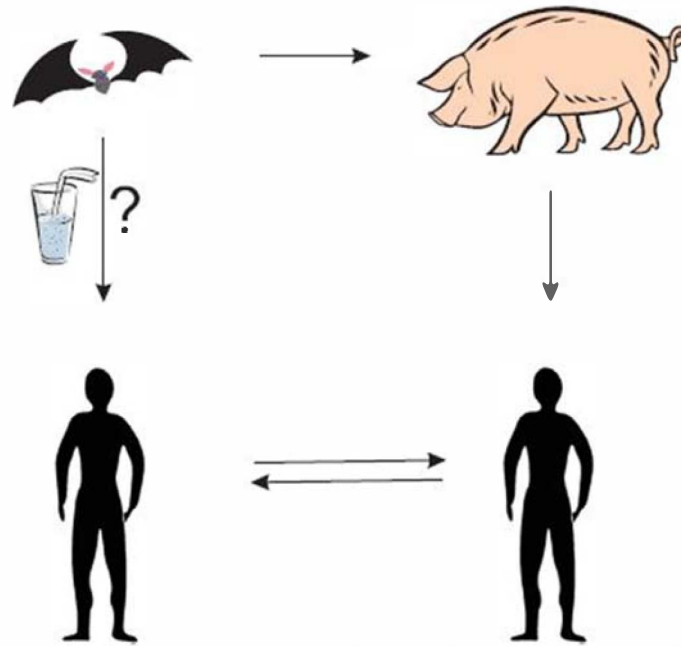
There have been further cases of Hendra virus disease in horses and humans in eastern Australia; many of these cases were fatal.

22.2.2.b Nipah virus

In Malaysia, in 1997, there appeared a new disease of pigs, characterized by respiratory disease and encephalitis. Soon afterwards workers on the affected pig farms began to develop encephalitis; over a two-year period there were several hundred human cases, with over 100 deaths. In 1999 a new virus was isolated from the brain of one of the patients who had died. The virus turned out to be a paramyxovirus with similar characteristics to Hendra virus (Figure 22.2) and it was named Nipah virus. In an attempt to contain the outbreak over one million pigs were slaughtered.

A search was initiated to find the reservoir of infection. As with Hendra virus, bats were investigated and evidence of Nipah virus infection was found in fruit-eating bats of the genus *Pteropus*. Infectious virus was isolated from bats' urine and virus-specific antibodies were found in their blood. Thus the situation parallels that of Hendra virus, with a reservoir of infection in bats and transmission to humans via an intermediate mammalian host (Figure 22.4).

Figure 22.4 Nipah virus transmission routes. The virus can be transmitted from bat to pig, pig to human, and human to human. Bat-to-human transmission may also occur.



More recently, in Bangladesh and India, there have been outbreaks of Nipah virus encephalitis and respiratory disease in humans; many of the cases have been fatal. In these outbreaks there was no evidence of transmission from pigs. It is believed that the virus was transmitted from bats to humans; for example, as a result of humans drinking date palm sap contaminated by bats during its collection. There is also evidence of human-to-human transmission of the virus.

22.3 VIRUSES IN NEW AREAS

22.3.1 West Nile virus

In 1999 some people in New York became ill with apparent viral encephalitis; several of them died. At about the same time several birds in the zoo became ill, and deaths of wild birds, especially crows, were reported. Diagnostic tests revealed that the human patients and the birds were infected with West Nile virus (WNV).

This virus has been known since 1937 when it was first isolated from the blood of a woman in the West Nile region of Uganda. In addition to Africa, the virus is also known in Eastern Europe, West Asia, and the Middle East, but until 1999 it had never been reported in North America.

WNV is classified in the family Flaviviridae (Figure 22.5). The virus has a wide host range. It cycles between birds and mosquitoes, in which it replicates. Infected mosquitoes can transmit the virus to mammalian species, including humans and horses (Figure 22.6). Mammalian infections are normally dead-end as far as the virus is concerned (i.e., the virus is not normally transmitted on to new hosts).

Figure 22.5 Families Flaviviridae and Togaviridae. Flaviviruses and togaviruses have many properties in common, but flaviviruses (around 50 nm diameter) are smaller than togaviruses (70 nm diameter).

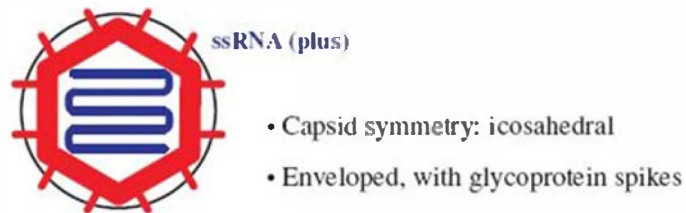
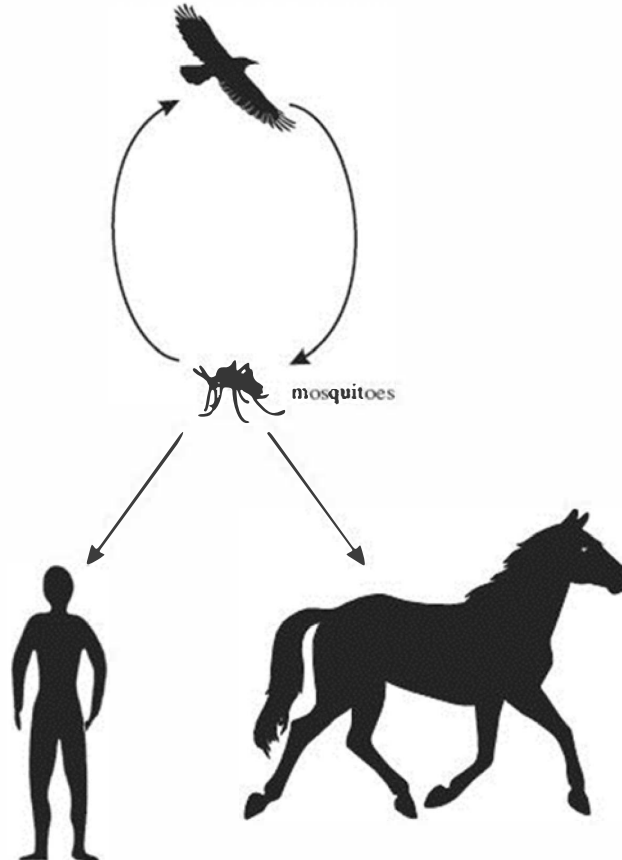
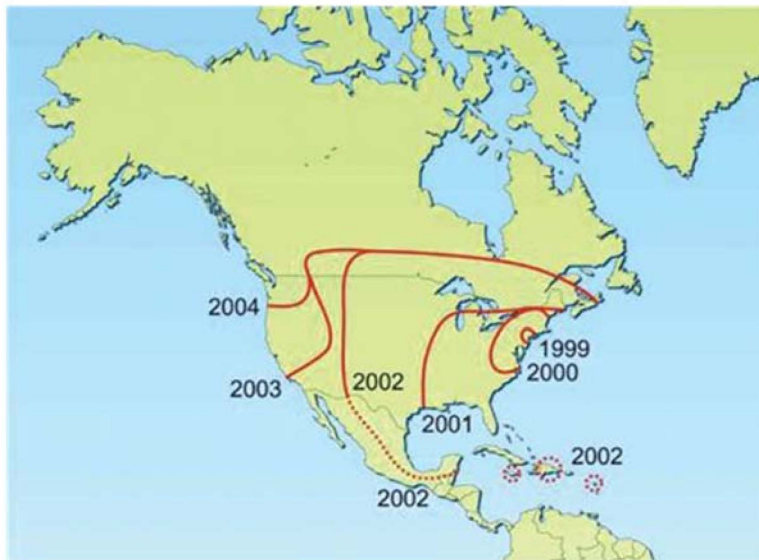


Figure 22.6 West Nile virus transmission routes. The virus cycles between birds and mosquitoes. Sometimes infected mosquitoes transmit the virus to mammals, including humans and horses.



Most infections with WNV are asymptomatic, but about 25% of human infections result in signs and symptoms such as fever and aching muscles. Less than 1% of infections spread to the central nervous system, causing severe illness (such as encephalitis), which may result in death; these cases are mainly in patients over 60 years old. Protective measures for humans include use of insect repellents, wearing “mosquito jackets” and treating mosquito breeding areas with insecticides.

It is not known how WNV was introduced into North America, but once there it spread rapidly throughout the US (Figure 22.7). Within five years it was also present in Canada, Central America, and the Caribbean islands; it is now present in South America. The virus has caused hundreds of human deaths in North America. Figure 22.7 Approximate distribution of West Nile virus in the Americas 1999–2004. Source: Mackenzie et al. (2004) *Nature Medicine*, 10, S98. Reproduced by permission of Nature Publishing Group and the author.



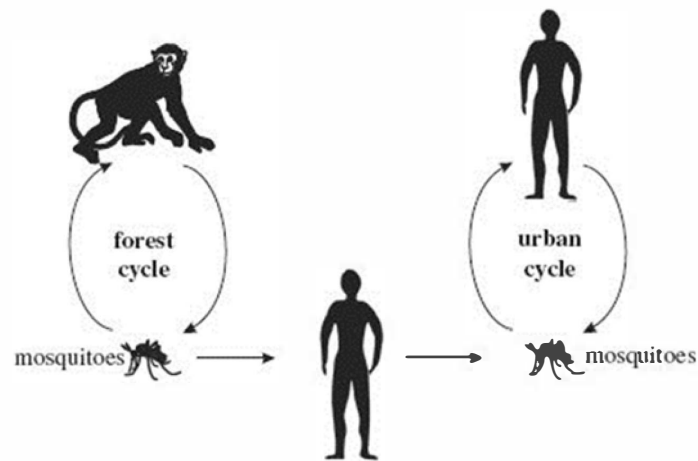
The spread of WNV to North America prompted an investigation to determine whether the virus is present in birds in the UK. WNV-specific antibodies were found in a number of species and, using RT-PCR, WNV RNA was detected in magpies and a blackbird, though all the birds appeared healthy.

22.3.2 Chikungunya virus

Infection of humans with chikungunya virus can cause variable signs and symptoms, but a consistent feature is swelling and pain in many of the body's joints, causing the patient to walk in a stooped position. Chikungunya is a word in one of the African languages, and means "that which bends up." Occasionally the patient dies.

Chikungunya virus is a member of the family *Togaviridae* (Figure 22.5). It was first isolated in Tanganyika (now Tanzania) in the early 1950s, and is now known to occur in many parts of Africa, where it cycles between non-human primates and forest mosquito species, which sometimes transmit the infection to humans. There is also an urban cycle, where the virus is transmitted between humans by one or two mosquito species that are highly adapted to the urban environment (Figure 22.8). The virus replicates in the mosquito vectors.

Figure 22.8 Chikungunya virus transmission routes. In African forests the virus cycles between non-human primates and mosquitoes, which may transmit infection to humans. In urban environments the virus can cycle between humans and mosquitoes.



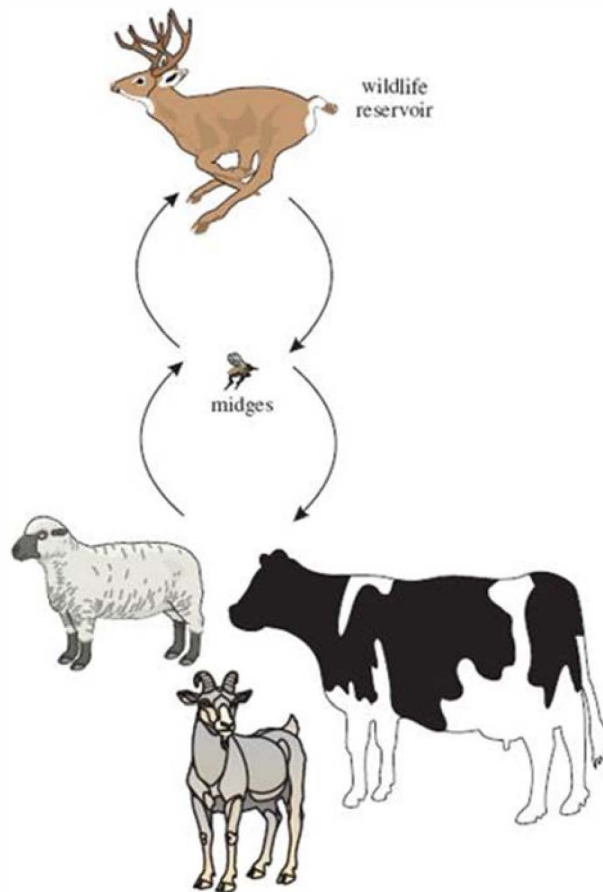
In the twenty-first century there have been urban outbreaks of chikungunya virus in East Africa, India, several parts of Southeast Asia, and on a number of Indian Ocean islands, in particular La Réunion. There have also been cases in Europe and North America, as a result of the virus being introduced by infected travelers. In Italy several hundred people became infected through mosquito transmission of virus originating from a patient who had become infected in India.

22.3.3 Bluetongue virus

Bluetongue virus (BTV), a member of the genus *Orbivirus* in the family *Reoviridae* (Chapter 13), causes disease in ruminant animals, including sheep, goats, and cattle. Like WNV and chikungunya virus, it has a wide host range in vertebrate animals and replicates in insects, which act as vectors.

All ruminant species are susceptible to infection with BTV, but infection results in severe disease and death only in certain hosts, mainly some breeds of sheep and some species of deer. In some affected animals the tongue may become swollen and blue. Infected animals that remain healthy can act as reservoirs of infection. The virus is transmitted between its ruminant hosts by certain species of biting midge in the genus *Culicoides* (Figure 22.9). The virus replicates in the midge vectors and transmission is restricted to those times of the year when the midges are active.

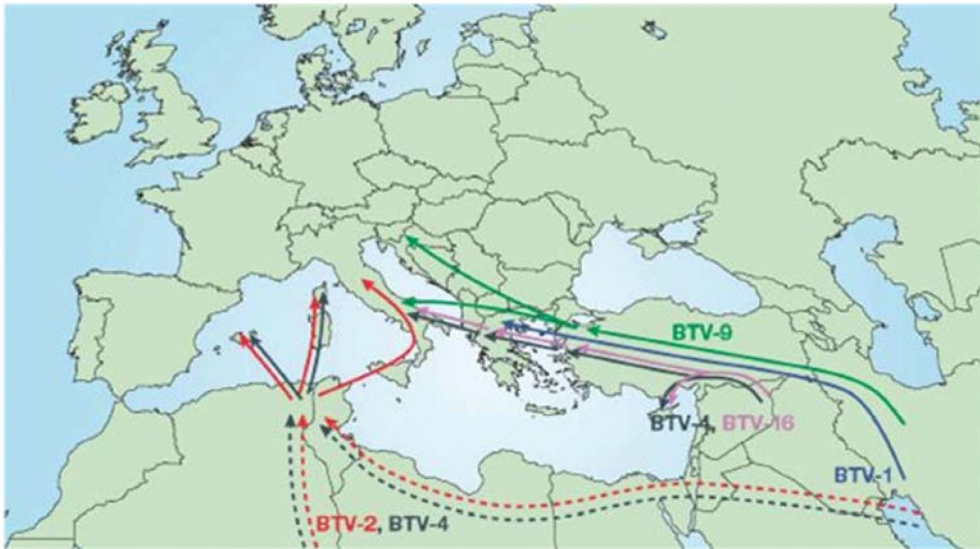
Figure 22.9 Bluetongue virus transmission routes. Midges act as vectors, transmitting the virus between ruminant animals, both wild and domestic.



Bluetongue has been known in Africa and the Middle East for a long time. Occasionally it has made brief incursions into Europe, but in 1998 six serotypes of BTV began to spread north. By 2004 there were bluetongue cases 800 km further north than ever before (Figure 22.10) and over a million sheep had been killed by the disease. One of the BTV serotypes reached Britain in 2007. The midge species that is the main BTV vector had extended its range north, presumably because of climate warming, and this was a major factor responsible for the spread of the virus. A number of measures, including vaccination, eradicated bluetongue from Northern Europe by 2010.

Figure 22.10 Spread of bluetongue virus serotypes in Europe 1998–2004.

Source: Purse et al. (2005) *Nature Reviews Microbiology*, 3, 171. Reproduced by permission of Nature Publishing Group and the authors.



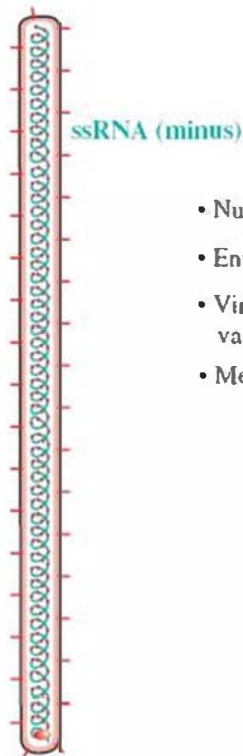
22.4 VIRUSES IN NEW HOST SPECIES AND IN NEW AREAS

22.4.1 Filoviruses

In 1967, in the town of Marburg in Germany, some laboratory workers became ill with a hemorrhagic fever (a terrible disease characterized by diarrhea and vomiting, as well as by hemorrhaging and fever). These people had been in contact with blood, organs, and cell cultures from African green monkeys caught in Uganda. Seven of those affected died, and there were five cases of the illness in hospital staff that had been in contact with patients' blood. Investigations revealed that the monkeys had been infected with a virus that had been transmitted to the laboratory staff, and from them to the hospital staff.

This virus was of a type never previously encountered, with elongated virions, some of which were straight and some were curved. It was named Marburg virus and is now classified in the family Filoviridae, named from the Latin *filum*, meaning a thread (Figure 22.11).

Figure 22.11 Family Filoviridae.



- Nucleocapsid symmetry: helical
- Enveloped, with glycoprotein spikes
- Virion diameter about 80 nm; length variable, up to 1400 nm
- Members: Marburg and Ebola viruses

In 1976 there were outbreaks of a similar disease near the River Ebola in the Democratic Republic of Congo (then Zaire) and in Sudan. A virus similar to Marburg virus was isolated from patients and was named Ebola virus. Since then there have been a number of outbreaks of disease caused by Ebola and Marburg viruses across central Africa, from Côte d'Ivoire in the west to Kenya in the east, and to Angola in the south.

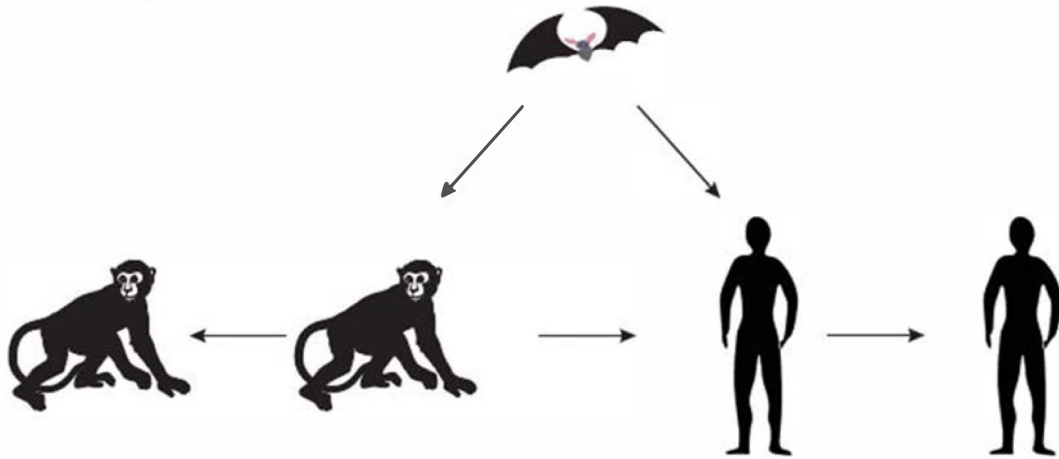
The ways in which outbreaks start have long been a mystery. There is increasing evidence that some outbreaks start when a human becomes infected as a result of contact with the blood of an infected non-human primate. It was known from the original outbreak in Germany that African green monkeys can be infected with Marburg virus. It has since been found that gorillas, chimpanzees, and duikers can be infected with Ebola virus, which may be responsible for significant mortality of these species. Ebola and Marburg viruses are present in the blood of infected hosts and transmission to humans can occur through contact with the flesh of infected animals after they have been hunted and killed. Human-to-human transmission readily occurs through contact with the blood of infected individuals.

There are several species of Ebola virus in Africa. A further species appeared in the US and Italy in monkeys imported from the Philippines, and it appeared in pigs in the Philippines. This Ebola virus species is pathogenic to monkeys and pigs, but there have been no reports of it causing human disease, though there is serological evidence that monkey handlers and pig farmers have been infected.

For a long time it was thought likely that additional animal species act as "reservoirs" of Marburg and Ebola viruses, but it was not known which species these might be. There is now evidence of filovirus infection in bats (Figure 22.12), in which infectious virus, virus RNA, and anti-virus antibodies have been detected.

Some people have become infected with a filovirus after working in mines or visiting caves where bats roost.

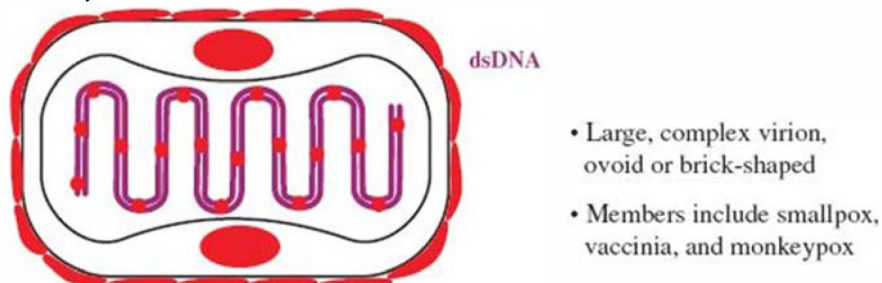
Figure 22.12 Filovirus transmission routes. It is thought that Marburg and Ebola viruses are present in bat populations, from which infection may spread to human and non-human primate populations.



22.4.2 Monkeypox virus

Monkeypox is a disease of monkeys that is similar to smallpox in man. The viruses that cause these diseases are members of the family Poxviridae (Figure 22.13). Since 1970 there have been cases in Africa of human disease caused by infection with monkeypox virus; some cases have been fatal.

Figure 22.13 Family Poxviridae.



In 2003 some people in the US developed a mystery illness that involved a rash and a fever. Investigations revealed that they were infected with monkeypox virus, which had entered the country in small mammals imported from Ghana. The virus had infected pet prairie dogs and had then been transmitted to humans.

Sequencing the genomes of monkeypox virus isolates has revealed the existence of two strains that appear to differ in virulence. The virus that was imported into the US was the less virulent strain.

22.5 NEW VIRUSES

Sometimes new viruses appear, often as a result of infecting new host species. Examples include HIV (Chapter 18), influenza (Chapter 16), and SARS coronavirus.

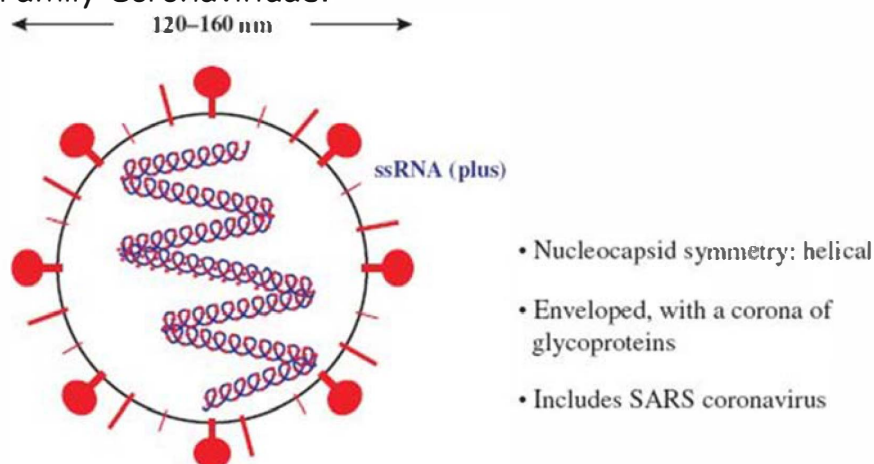
22.5.1 SARS coronavirus

In 2002 a new human respiratory disease emerged in southern China. The following year one of the doctors who had been treating patients traveled to Hong Kong, where he became ill and died. Subsequently, people who had stayed in the same hotel as the doctor traveled to Singapore, Vietnam, Canada, and the US, taking the infectious agent with them. The epidemic of severe acute respiratory syndrome (SARS) was under way.

The signs and symptoms of SARS resemble those of influenza and include fever, aching muscles, sore throat, cough, and shortness of breath. About 90% of patients recovered, but for the remainder the infection proved to be lethal. These were mainly individuals who had an underlying condition such as diabetes, heart disease, or a weakened immune system. On the face of it SARS is a respiratory tract disease, but in many patients the infection spread to other parts of the body. Diarrhea developed in some patients and the virus was shed in the feces and urine for several weeks.

The causative agent was found to be a new coronavirus (Figure 22.14). No natural reservoir for the SARS coronavirus has been found, but a similar virus has been isolated from several bat species. It is likely that coronaviruses have repeatedly crossed into humans from other mammalian species, but in the majority of cases the virus has lacked the ability for efficient human-to-human transmission. It would seem that the SARS coronavirus evolved from a rare virus that had this ability.

Figure 22.14 Family Coronaviridae.



The SARS outbreak was brought under control by quarantine measures, but only after there had been over 8000 cases with nearly 800 deaths.

22.6 RECENTLY DISCOVERED VIRUSES

22.6.1 Human metapneumovirus

In 2001 a new member of the family Paramyxoviridae (Figure 22.2) was isolated from young children in the Netherlands. The virus was named human metapneumovirus. Infection can cause mild respiratory problems, but can also

cause bronchiolitis and pneumonia, with symptoms similar to those caused by infection with respiratory syncytial virus, another paramyxovirus. There is serological evidence that the metapneumovirus had been present in humans for at least 50 years, and that in the Netherlands most children have become infected with the virus by the age of five.

Infants and children with acute respiratory tract infection due to human metapneumovirus have since been reported in a number of countries, including the UK.

22.7 RE-EMERGING VIRUSES

22.7.1 Measles and mumps viruses

Further members of the family Paramyxoviridae (Figure 22.2) are measles and mumps viruses. Both are important human pathogens, especially measles, which is a major cause of mortality in developing countries.

Until the late 1990s cases of measles and mumps had been declining in the UK as a result of widespread uptake of the measles, mumps, and rubella (MMR) vaccine, but after fears were raised surrounding the safety of the vaccine fewer children were vaccinated. Measles and mumps began to “re-emerge” with the annual number of cases of each disease at first in the hundreds, and later thousands.

22.7.2 Bioterrorism

There are concerns that the world may witness the re-emergence of a virus as a result of a deliberate release by terrorists. There is a long list of potential agents, including viruses, that terrorists might consider using. High on the list is variola virus, the causative agent of smallpox, and a member of the family Poxviridae (Figure 22.13). There have been no smallpox cases since the 1970s, so smallpox vaccination programs have been discontinued and most humans lack immunity to this virus. Some governments have stockpiled vaccine to be used in the event of a smallpox release.

22.8 VIRUS SURVEILLANCE

In order that the threats posed by emerging viruses and other agents can be dealt with effectively, it is important that worldwide surveillance systems are in place. There need to be effective systems that warn when agents such as the SARS coronavirus emerge.

The scheme for monitoring influenza virus strains provides a good model for a surveillance system. In this scheme, which is coordinated by the United Nations World Health Organization (WHO), isolates of influenza virus from laboratories around the world are sent to Collaborating Centers for Influenza Reference and Research in London, Atlanta, Melbourne, and Tokyo. Antigenic variation of the virus in humans and in animals (especially birds and pigs) is monitored, and twice a year

the WHO recommends the virus strains to be mass-produced for incorporation into influenza vaccines.

22.9 DEALING WITH OUTBREAKS

Vaccines and/or anti-viral drugs can play important roles in attempts to deal with outbreaks of virus disease. Influenza vaccines and drugs have been deployed against pandemic influenza, and vaccines were used to control the spread of bluetongue virus in Europe. For most viruses, however, there are no vaccines or anti-viral drugs available.

It is important that resources for diagnosis and research of infectious diseases are available when the need arises. A useful case study is provided by the measures taken to deal with the emergence of the SARS coronavirus (Section 22.5.1).

There was some delay in alerting the world to SARS but, once it was apparent that this virus posed a major threat, work got under way in a number of virology laboratories. The virus was isolated in February 2003 and three months later its genome had been sequenced. The following year a paper was published reporting compounds that inhibit replication of the virus, while other papers reported the cell receptor of the virus and the structure of its replicase protein.

Diagnostic laboratory methods to detect evidence of SARS coronavirus in samples from patients rapidly became available; tests based on immunofluorescence and RT-PCR were developed. The health of international travelers was monitored (Figure 22.15) and those found to be infected with the virus were nursed in isolation. These measures brought the SARS outbreak under control.

Figure 22.15 Checkpoint for monitoring the health of people crossing the Thailand–Burma border during the SARS outbreak, June 2003.

Source: Courtesy of Professor A. J. S. Whalley, Liverpool John Moores University.



When there is an outbreak of a highly infectious virus, such as the SARS coronavirus or Ebola virus, infected patients and their contacts are quarantined. Control measures for some outbreaks include slaughter of animals, both those that are infected and those that have been in contact with infected animals. This has been a key measure in dealing with outbreaks of foot and mouth disease (cattle and sheep), Nipah virus (pigs), and avian influenza (poultry).

Medical staff, veterinary staff, and other personnel dealing with humans and animals infected with high-risk viruses must take precautions to protect themselves and to avoid spreading infection. Precautions include wearing protective clothing and breathing filtered air.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- evaluate the term “emerging virus”;
- discuss examples of viruses that have recently appeared in new host species;
- discuss examples of viruses that have recently appeared in new parts of the world;
- discuss examples of new viruses;
- discuss examples of re-emerging viruses;
- assess measures that can be taken to prevent and contain outbreaks of infectious disease.

SOURCES OF FURTHER INFORMATION

Books

Bishop, K. A. and Broder, C. C. (2008) Hendra and Nipah viruses: lethal zoonotic paramyxoviruses. Chapter 9 in *Emerging Infections 8* (Scheld, W. M., Hammer, S. M., and Hughes, J. M., editors), ASM Press

Bock, G. and Goode, J. (eds) (2008) *Novel and Re-emerging Respiratory Viral Diseases*, Novartis Foundation Symposium 290, John Wiley & Sons Ltd.

Diamond, M. S. (ed.) (2009) *West Nile Encephalitis Virus Infection*, Springer

Haagmans, B. L. (2011) Vaccines against newly emerging viral diseases: the example of SARS. Chapter 16 in *Vaccine Design: Innovative Approaches and Novel Strategies* (Rappuoli, R. and Bagnoli, F., editors), Caister Academic Press

Hussein, I. T. M. et al. (2011) Recent advances in hantavirus molecular biology and disease. In *Advances in Applied Microbiology*, Volume 74, pp. 35–75, Academic Press

Mellor, P. S., Baylis, M., and Mertens, P. P. C. (2009) *Bluetongue*, Elsevier

Oldstone, M. B. A. (2009) *Viruses, Plagues, and History: Past, Present, and Future*, Oxford University Press

Plyusnin, A. and Elliott, R. M. (2011) *Bunyaviridae: Molecular and Cellular Biology*, Caister Academic Press

Powers, A. M. (2008) Epidemic emergence of Chikungunya virus: options for control of an enzootic virus. Chapter 7 in *Emerging Infections 8* (Scheld, W. M., Hammer, S. M., and Hughes, J. M., editors), ASM Press

Wolfe, N. (2011) *The Viral Storm: The Dawn of a New Pandemic Age*, Allen Lane/Times

Journals

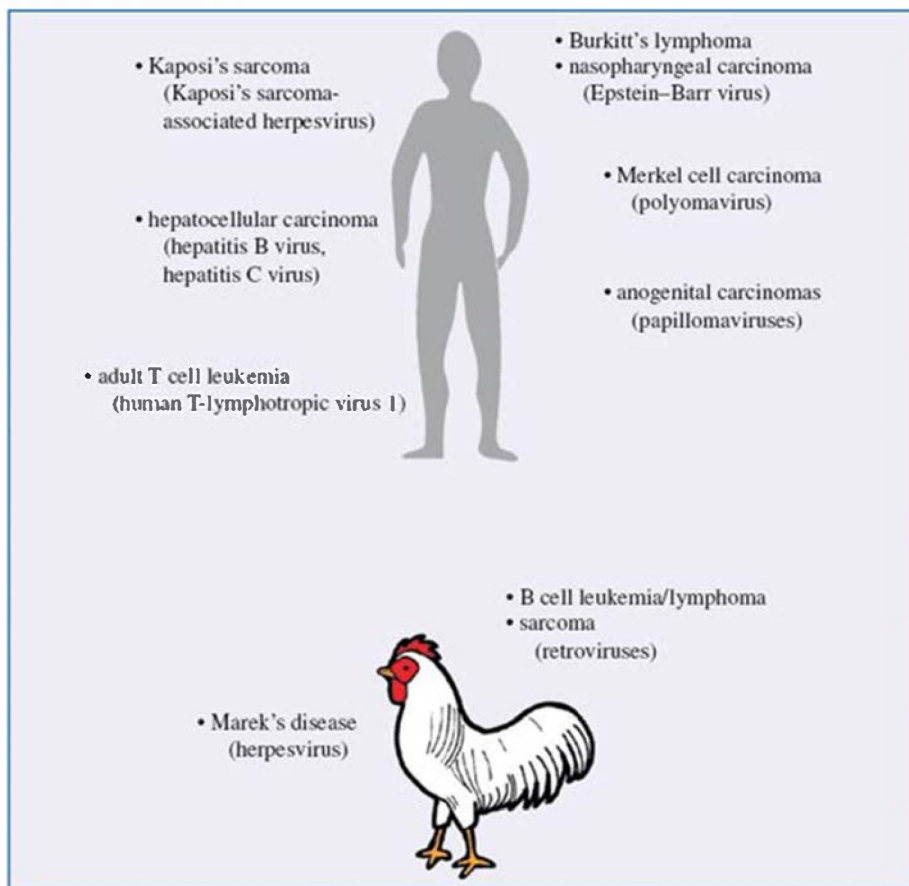
- Bolles, M., Donaldson, E., and Baric, R. (2011) SARS-CoV and emergent coronaviruses: viral determinants of interspecies transmission. *Current Opinion in Virology*, 1, 624–634
- Field, H. and Kung, N. (2011) Henipaviruses—unanswered questions of lethal zoonoses. *Current Opinion in Virology*, 1, 658–661
- Kuzmin, I. V. et al. (2010) Marburg virus in fruit bat, Kenya. *Emerging Infectious Diseases*, 16 : available at <http://www.cdc.gov/eid/content/16/2/352.htm>
- Lim, S. M. et al. (2011) West Nile virus: immunity and pathogenesis *Viruses*, 3, 811–828
- MacNeil, A., Nichol, S. T., and Spiropoulou, C. F. (2011) Hantavirus pulmonary syndrome. *Virus Research*, 162, 138–147
- Morens, D. M. and Taubenberger, J. K. (2011) Pandemic influenza: certain uncertainties. *Reviews in Medical Virology*, 21, 262–284
- Nilsson, A. and Chiodi, F. (2011) Measles outbreak in Africa—Is there a link to the HIV-1 epidemic? *PLoS Pathogens*, 7 (2): e1001241
- Pfeffer, M. and Dobler, G. (2010) Emergence of zoonotic arboviruses by animal trade and migration *Parasites and Vectors*, 3, 35
- Tsetsarkin, K. A. et al. (2011) Chikungunya virus: evolution and genetic determinants of emergence *Current Opinion in Virology*, 1, 310–317

Vasilakis, N. et al. (2011) Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health *Nature Reviews Microbiology*, 9, 532–541

CHAPTER 23

Viruses and Cancer

CHAPTER 23 AT A GLANCE



23.1 INTRODUCTION TO VIRUSES AND CANCER

A cancer in a human or an animal is a malignant tumor and involves continuous proliferation of a clone of cells derived from one of the body's normal cells. The cell undergoes changes known as transformation as a result of events that include mutation, activation of oncogenes (tumor genes), and inactivation of tumor suppressors. The types of cancer that will be discussed in this chapter, and the cell types from which they are derived, are listed in Table 23.1.

Table 23.1 Types of cancer and the cell types from which they are derived

Cancer	Cell type
Carcinoma	Epithelium

Leukemia	Blood-forming
Lymphoma	Lymphocyte
Mesothelioma	Mesothelium
Sarcoma	Connective tissue

For most cancers the full sequence of events that turns a normal cell into a tumor cell (the process of oncogenesis) is not fully understood, but it is thought that between four and six steps are involved. Some of the steps in this transformation can be triggered by environmental factors including some chemicals, some forms of irradiation, and some viruses.

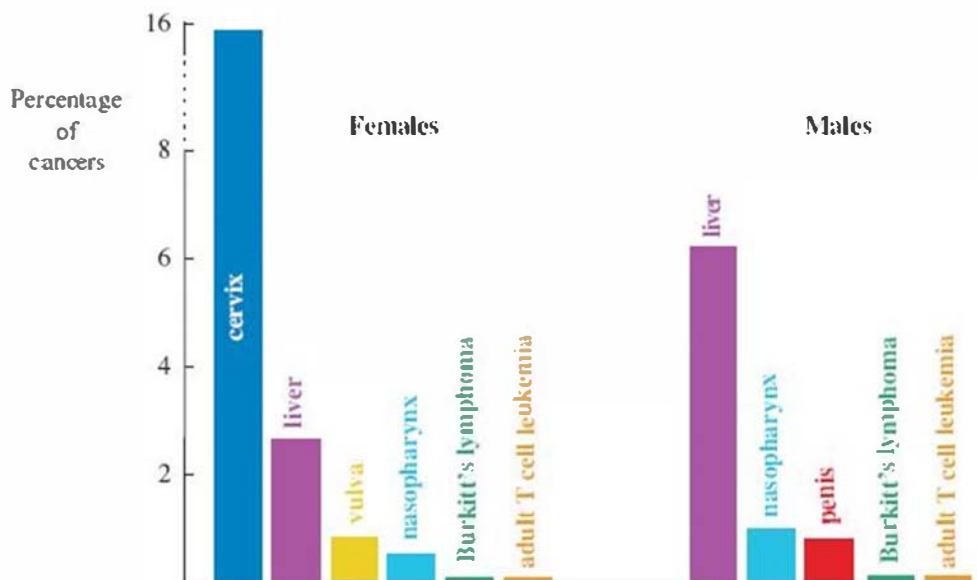
A virus that is able to cause cancer is known as an oncogenic virus. Evidence that a virus is oncogenic includes the regular presence in the tumor cells of virus DNA, which might be all or a part of the virus genome. In some types of tumor the virus DNA is integrated into a cell chromosome, while in other types it is present as multiple copies of covalently closed circular DNA (cccDNA), as discussed in the context of latent virus infections (see Figure 9.7). In many cases one or more of the virus genes are expressed in the tumor cell and virus proteins can be detected.

For some types of cancer there is evidence for involvement of a virus in most, if not all, cases of the cancer. For other types virus DNA and/or proteins are detectable in only a minority of cases and it is possible that the virus is just one of a number of carcinogenic factors that can give rise to these cancers. Some human adenoviruses are oncogenic in that they can transform cells in culture and can cause tumors when inoculated into animals, though there is no evidence that adenoviruses cause cancer in humans. Studies with these viruses have, however, contributed to understanding of oncogenic mechanisms.

This chapter is mainly concerned with viruses that cause human cancers, of which about 20% of cases in females and about 8% in males are thought to be caused by viruses. The relative incidences of the main virus-associated cancers are shown in Figure 23.1. The viruses involved are found worldwide and some are very common, though the prevalence of some varies between regions. We shall also briefly consider some animal cancers caused by viruses.

Figure 23.1 Virus-associated cancers in humans.

Source: Modified from Microbiology Today, August 2005, with the permission of Professor D. H. Crawford (University of Edinburgh) and the Society for General Microbiology.



23.2 PAPILOMAVIRUS-LINKED CANCERS

Cervical carcinoma is the third most common cancer in women, constituting almost 16% of cancers in females (Figure 23.1). In the world each year there are approximately half a million new cases of cervical cancer and 275 000 deaths caused by this cancer. The vast majority of these cases result from infection with a papillomavirus.

The papillomaviruses are small DNA viruses of mammals and birds (Figures 3.16 and 23.2). They enter the body through small abrasions in skin and mucous membranes, and infect basal cells in the epithelium. Sequencing the genomes of human papillomaviruses (HPVs) has differentiated about 200 types. Each HPV type infects a preferred site, such as the hands or the genitals, and infection may result in a benign wart (papilloma) or a carcinoma.

Figure 23.2 Virion characteristics of the families Papillomaviridae and Polyomaviridae. The virions of both families consist of icosahedral capsids containing dsDNA; papillomavirus virions are larger (50–55 nm diameter) than those of polyomaviruses (40–45 nm diameter).

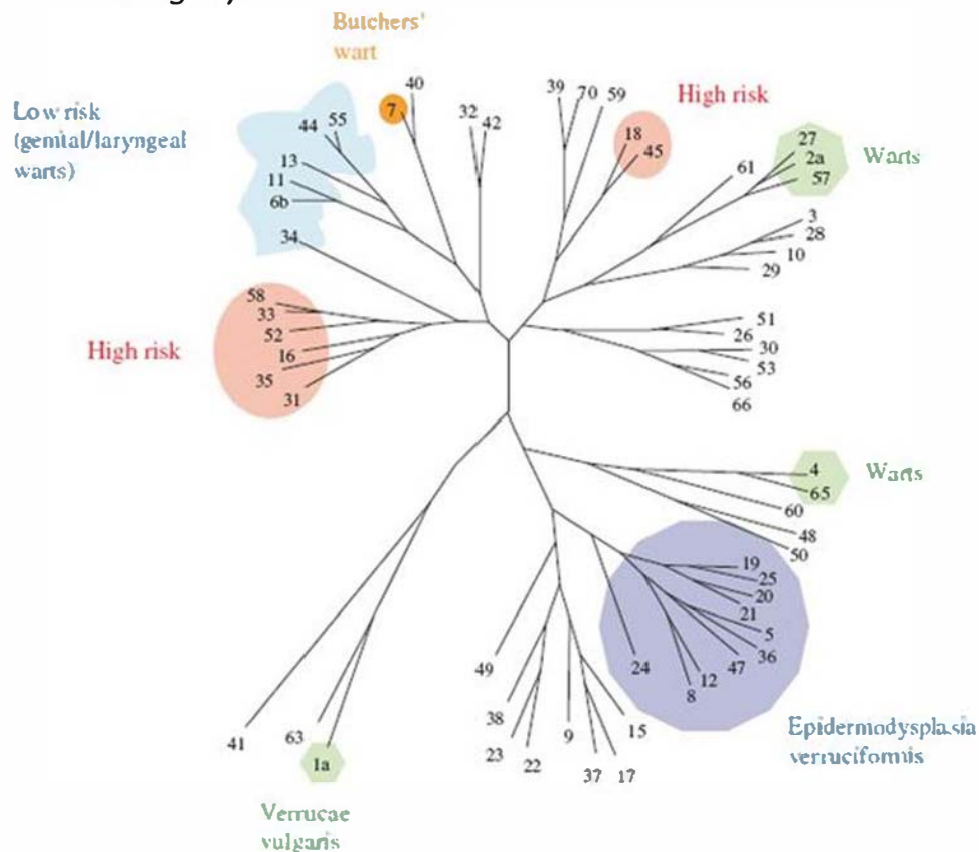


- Genome: dsDNA (covalently-closed circular)
- Capsid symmetry: icosahedral

The papillomaviruses that infect the genitals are transmitted between individuals during sexual contact. Most papillomavirus infections do not become persistent, but in a minority of hosts the infection is not cleared by the host's immune response. In individuals who harbor a persistent infection there is a risk of cancer developing. This risk is associated with about 15 of the HPV types; these "high-risk" types include HPV-16 and 18 (Figure 23.3). Infection with other HPV types that infect the genitals carries little or no risk of cancer; these "low-risk" HPV types include HPV-6 and 11, and are associated almost exclusively with benign genital warts.

Figure 23.3 Phylogenetic tree showing relationships between some HPVs. The high-risk HPVs cluster in two regions of the tree.

Source: Modified from Microbiology Today, August 2005, with the permission of Professor N. J. Maitland (University of York) and the Society for General Microbiology. Data from Los Alamos National Laboratory HPV website (<http://hpv-web.lanl.gov>).



The stages from normal cervical cell to cancer cell involve a number of morphological changes that can be detected by microscopic observation of a sample of cells taken from the cervix. If precancer cells are detected they can be killed or removed to prevent the development of a tumor. This is one approach to cervical cancer screening; an alternative approach involves testing the cells for the DNA of a high-risk HPV. All or part of the genome of one of these viruses can nearly always be found (usually integrated into a chromosome (Figure 9.7)) in precancer cells and in cancer cells. The virus replication cycle is not completed in these cells so no progeny virions are formed.

There is evidence that high-risk HPVs (especially HPV-16) are also causative agents of some carcinomas at other body sites, including the vulva, vagina, penis, anus, mouth, head, and neck. Again the evidence is the presence in the tumor cells of HPV DNA.

HPVs are also involved in a very rare form of skin cancer, which has a genetic basis: epidermodysplasia verruciformis. In this disease the patient is highly susceptible to infection with HPVs, especially a number of types that are rare in the normal population, mainly types 5 and 8 (Figure 23.3). Warts spread over the

entire body during childhood and 25–33% of patients develop cancer (squamous cell carcinoma) in areas of the skin exposed to ultraviolet light (sunlight). HPV DNA can be detected in more than 80% of these cancers but, in contrast to cervical cancers, the DNA is rarely integrated into a cell chromosome.

23.3 POLYOMAVIRUS-LINKED CANCERS

Polyomavirus virions have the same characteristics as those of the papillomaviruses, except that they are slightly smaller (Figure 23.2). They are found in mammals and birds, and most infections are subclinical. Injection of a polyomavirus into certain animal species, however, can induce the development of many types of tumor (“poly-oma”).

Several human polyomaviruses have been discovered and have been found to be extremely common. There is evidence that at least one of them has a role in the development of a cancer: Merkel cell carcinoma, a rare skin cancer that mainly affects elderly and immunosuppressed people. The tumor cells contain the integrated genome of Merkel cell polyoma virus and express the T (= tumor) antigen of this virus.

A monkey polyomavirus, simian virus 40 (SV40), was first isolated from primary cell cultures derived from rhesus monkey kidneys. The cell cultures were being used to produce poliovirus for Salk vaccine. The quality control procedures included injection of vaccine into hamsters, and some of these animals developed tumors. It was subsequently found that injection of SV40 into a newborn hamster could cause a carcinoma, sarcoma, lymphoma, or leukemia, depending on the injection site. Millions of humans became infected with SV40 when they received polio vaccine that was unknowingly contaminated with the virus. There have been suggestions that SV40 is linked with some types of human cancer, including primary brain cancer and malignant mesothelioma, but these suggestions are disputed.

23.4 EPSTEIN–BARR VIRUS-LINKED CANCERS

“Endemic” Burkitt’s lymphoma (BL) is a B cell tumor that occurs with a high frequency in children in central Africa and Papua New Guinea. Shortly after it was first described (by Denis Burkitt) Anthony Epstein established cell lines from the tumor of a patient. The cells were found to be persistently infected with a herpesvirus, which was named Epstein–Barr virus (EBV). Please see Chapter 11 for general characteristics of herpesviruses. Cases of “sporadic” BL and “HIV-associated” BL occur throughout the world, but generally without an association with EBV.

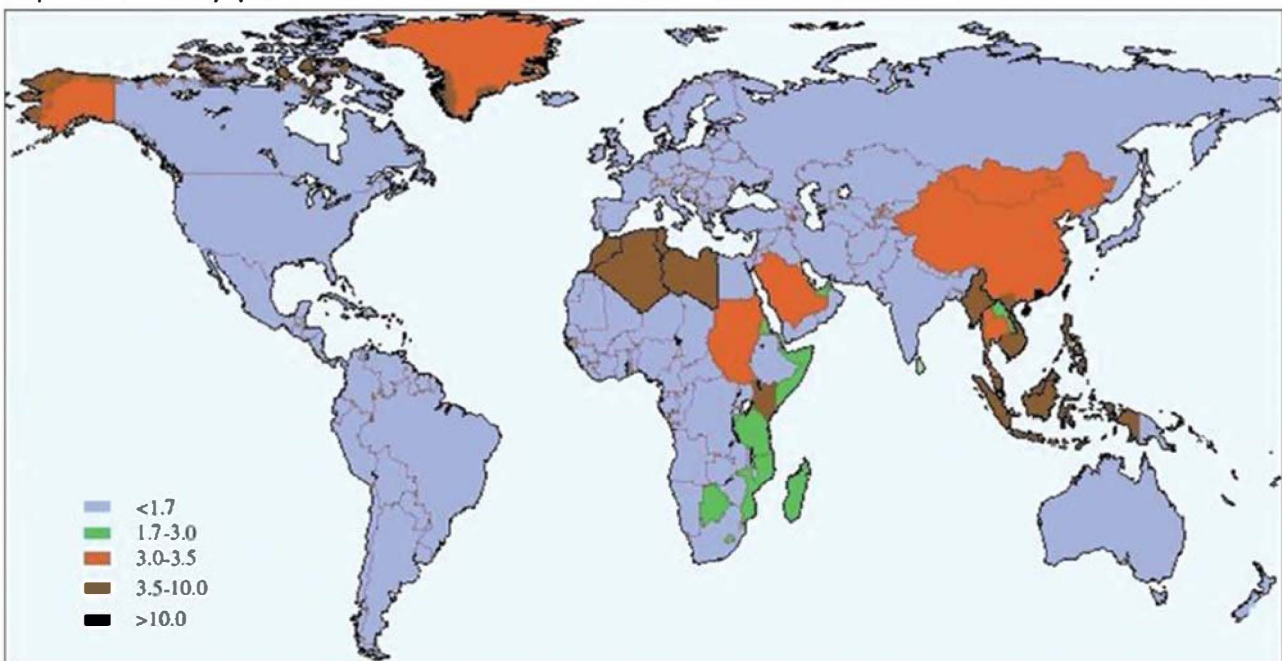
A consistent abnormality in BL tumor cells is a chromosomal rearrangement that results in the c-myc gene being placed next to an enhancer of an immunoglobulin gene. This results in the expression of c-myc at abnormally high levels. The chromosomal rearrangement is present in all cases of BL, irrespective of whether

EBV is present or not. Perhaps EBV is one of several agents that can trigger the rearrangement.

Another tumor that has an association with EBV is nasopharyngeal carcinoma (NPC). This tumor, like BL, is more prevalent in particular regions (Figure 23.4). In both tumors the EBV genome is present in the tumor cells as cccDNA molecules. You will recall that when a herpesvirus infects a cell the linear virus genome becomes circularized (Section 11.5.1).

Figure 23.4 Incidence rates of nasopharyngeal carcinoma in males.

Source: Data (age standardized incidence rates per 100 000) published by Busson et al. (2004) Trends in Microbiology, 12, 356. Map drawn by V. Gaborieau and M. Corbex (Genetic Epidemiology Unit, International Agency for Research on Cancer). Reproduced by permission of Elsevier Limited and the authors.



EBV is found worldwide and the majority of humans are infected with the virus. The factors that restrict high incidences of endemic BL and NPC to particular geographical locations are not understood. The endemic BL zone corresponds closely with the zone of malaria caused by *Plasmodium falciparum*, so this parasite may play a role. In southern China there are clusters of NPC in families, so there may be a human gene that increases the risk. Components of the diet, such as salted fish, are also suspected as co-factors in NPC.

A number of other cancers have been found to have associations with EBV, including:

- some cases of Hodgkin's lymphoma;
- non-Hodgkin's lymphoma in AIDS;
- post-transplant lymphoproliferative disorder.

The last two cancers develop in patients who are immunocompromised as a result of AIDS or immunosuppressive treatment, respectively.

The oncogenic potential of EBV can be demonstrated when cultures of B cells are infected with the virus. Infected cells synthesize a number of virus proteins that push the cells permanently into the cell cycle (Figure 4.5), resulting in the formation of lympho blastoid cell lines. In the body this B cell proliferation is normally controlled by T cells, but in individuals who are T cell immunocompromised the control may be inadequate and a B cell tumor may result.

23.5 KAPOSI'S SARCOMA

Kaposi's sarcoma was first described in the nineteenth century as a rare skin cancer that affected elderly men in the Mediterranean region. Since the arrival of AIDS that picture has changed. It is one of the most common cancers in people with AIDS; in these patients the cancer is more aggressive and it disseminates throughout the body.

In 1994 it was discovered that the tumor cells contain the DNA of a herpesvirus. The virus was subsequently isolated and was named Kaposi's sarcoma-associated herpesvirus (KSHV). It can be found in most parts of the world, but serological studies have shown that it is more common in certain regions, such as central Africa, and it is more common in homosexual men. The risk of developing Kaposi's sarcoma parallels the prevalence of the virus in the population.

There is also good evidence linking KSHV with two other human cancers: primary effusion lymphoma and multicentric Castleman's disease.

KSHV initiates latent infections in B cells and in endothelial cells lining blood vessels. If the host becomes immunocompromised one of these cells may be triggered into cycles of uncontrolled growth and division. If the cell is a B cell, then a primary effusion lymphoma or multicentric Castleman's disease may ensue; if the cell is an endothelial cell, then the outcome may be Kaposi's sarcoma. As with EBV-associated tumors, copies of the KSHV genome are present as cccDNA in the nuclei of the tumor cells.

23.6 ADULT T CELL LEUKEMIA

Adult T cell leukemia is associated with human T-lymphotropic virus 1 (HTLV-1) infection. Regions of the world with a high prevalence of the virus, such as southwest Japan, have a high prevalence of this cancer. HTLV-1 is a retrovirus and the tumor cells each have a copy of the proviral DNA integrated into a chromosome (Chapter 17).

23.7 HEPATOCELLULAR CARCINOMA

Hepatocellular carcinoma (liver cancer) accounts for 4–5% of cancer cases in the world. A number of factors are implicated as causative agents, including consumption of mold toxins in food, but most cases are associated with two viruses: hepatitis B virus (HBV) and hepatitis C virus (HCV). Liver cancer is the most common virus-associated cancer in men (Figure 23.1).

HBV is the most significant agent; the prevalence of liver cancer closely parallels the prevalence of persistent HBV infection (Section 19.2), with the highest incidences in Asia and in central and southern Africa. The other virus associated with liver cancer, HCV, is a flavivirus (Figure 22.5). It is the only class IV virus (plus-strand RNA virus) that is known to be oncogenic. Its genome is not found routinely in the cancer cells, in contrast to the other human oncogenic viruses.

Both HBV and HCV elicit immune responses when they infect the body. In some individuals the immune response successfully eliminates the infection, but in many cases the infection persists for life, as discussed in Section 19.2 for HBV. As far as HCV is concerned, it is estimated that about 75% of individuals who become infected are unable to eliminate the infection and that about 175 million people in the world are infected. In some individuals who are persistently infected with HBV or HCV, liver cirrhosis develops. This results in a high rate of cell division to replace the damaged hepatocytes, during which a mutation might occur in one of the cells, and this might eventually lead to cancer.

23.8 VIRUS-ASSOCIATED CANCERS IN ANIMALS

A number of cancers in animal species have been found to be associated with viruses. Outbreaks of some of these diseases can have serious economic consequences in agriculture. A few examples are given here.

It was demonstrated early in the twentieth century that filterable agents cause a leukosis (a B cell leukemia/lymphoma) and a sarcoma in chickens. These agents were subsequently shown to be retroviruses; the sarcoma-causing virus was named Rous sarcoma virus after its discoverer, Peyton Rous (Section 17.4). Other retroviruses were subsequently found to be causative agents of leukemia in cats and cattle, and a retrovirus known as Jaagsiekte sheep retrovirus was found to be responsible for adenocarcinoma of the lung in sheep.

Marek's disease in chickens is a lymphoma, similar to that caused by retroviruses, but the causative agent is a herpesvirus and the transformed lymphocytes contain multiple copies of the viral genome. Papillomaviruses, already discussed as causative agents of human cancers (Section 23.2), cause cancers in cattle, horses, and rabbits.

23.9 CELL LINES DERIVED FROM VIRUS-ASSOCIATED CANCERS

Cancer cells, which by definition multiply continuously in the body, will often continue to multiply continuously if transferred into cell culture, when permanent cell lines may be established from them. Cell lines derived from most virus-induced cancers contain part or all of the virus genome. The most famous of these, the

HeLa cell line, was established from a cervical cancer in the mid-twentieth century and is now used in laboratories around the world. Each HeLa cell contains a copy of part of the genome of HPV-18; this will be discussed further in Section 23.10.5.

Cell lines derived from Burkitt's lymphoma contain copies of the EBV genome and some produce infectious EBV; the virus was first discovered in one of these cell lines. Multiple copies of Marek's disease virus genome are present in a cell line derived from a Marek's tumor. A cell line derived from a liver cancer contains HBV DNA integrated at seven sites, and the cells secrete HBsAg. Cell lines derived from Kaposi's sarcoma, however, do not contain KSHV DNA; the virus DNA is lost from the tumor cells when they grow in culture.

23.10 HOW DO VIRUSES CAUSE CANCER?

Most virus-induced cancers develop after a long period of persistent infection with an oncogenic virus; for adult T cell leukemia this period is exceptionally long (around 60 years). The virus infections persist in their hosts in spite of immune responses, such as the production of virus-specific antibodies. Some persistent infections are latent for much of the time (e.g. EBV and KSHV), with only small numbers of virus genes expressed. Others, including HBV and HCV infections, are productive. Both of the latter viruses are able to evolve rapidly (Chapter 21), and this probably allows them to keep one step ahead of acquired immune responses.

Although many humans are persistently infected by viruses that are potentially oncogenic, only small percentages of those infected develop virus-linked cancers. Relatively few people develop EBV-related tumors, though over 90% of adults worldwide are infected with the virus. About 3% of women persistently infected with one of the high-risk strains of HPV develop cervical carcinoma, and similar percentages of individuals persistently infected with HBV, HCV, or HTLV-1 develop a virus-linked cancer.

As cancer develops in only a small proportion of virus-infected hosts it is clear that the virus infections alone do not cause cancer. Other factors are involved and these include exposure to particular environmental factors, host genetic factors, and immunodeficiency.

Immunodeficiency increases the risk of a virus-associated tumor, the nature of the immunodeficiency influencing the types of tumor that may develop. AIDS patients are much more likely to develop Kaposi's sarcoma than immunosuppressed transplant recipients. Non-Hodgkin's lymphoma in AIDS and post-transplant lymphoproliferative disorder (EBV-related tumors; Section 23.4) each occurs in a distinct class of immunodeficient patient. Immunodeficiency does not increase the risk of all tumors with a viral link; for example, there is no increased incidence in AIDS patients of NPC or of liver cancer.

The probability of cancer developing in a host infected with an oncogenic virus therefore depends on a complex interplay between the state of the host,

environmental factors to which the host is exposed, and cellular changes induced by the virus infection. There is increasing evidence that virus proteins synthesized during persistent infections play roles in the conversion of normal cells to cancer cells. In some cases virus proteins “deliberately” interfere with cell activities in order to provide appropriate conditions for virus replication. In other cases virus proteins may “accidentally” interact with cell components with outcomes that are of no benefit to the virus. The roles of virus proteins in the development of cancers are discussed in the next sections.

23.10.1 “Deliberate” interference with control of cell activities

A virus needs to manipulate the internal environment of its host cell so that all the requirements for virus replication are fulfilled. A requirement for small DNA viruses of eukaryotes (papillomaviruses, polyomaviruses, and adenoviruses) is that the cell’s DNA synthesizing functions are available (i.e., that the cell is in the S phase of the cell cycle; Figure 4.5).

Many cells in the animal body grow and divide either slowly or not at all; the latter are arrested in the G1 phase of the cell cycle and are said to be in the G0 state. The control of the cell cycle is mediated by many proteins; two that play key roles in humans are p53 and retinoblastoma protein (pRb). The latter protein is so named because the absence of the protein in some individuals, due to mutations in both copies of the gene, leads to the development of retinoblastoma (cancer of the retina). It has been demonstrated that several proteins produced by oncogenic viruses can interact with p53, pRb, and other proteins that control cell growth and division, increasing the probability of a cell being pushed into repeated cycles of division.

Virus proteins that can interfere with control of the cell cycle include the HPV early proteins E6 and E7, which localize to the nucleus and bind to cell proteins. E6 binds to p53, promoting its degradation, and E7 binds to pRb–E2F complexes, causing their dissociation (E2F is a cell transcription factor that activates DNA synthesis). The outcome of these interactions is the transition of the cell to the S phase.

Many other proteins of viruses considered in this chapter can interact with p53 and/or pRB. Examples include:

- SV40 large T antigen (binds to p53 and pRB);
- KSHV latency-associated nuclear antigen (binds to p53 and pRB);
- HTLV-1 Tax protein (binds to p53);
- HBV X protein (binds to p53);
- Adenovirus E1A protein (binds to pRB).

As well as affecting the cell cycle, some viruses may interfere with attempts by the infected cell to destroy itself via one of the apoptosis pathways. The cell protein

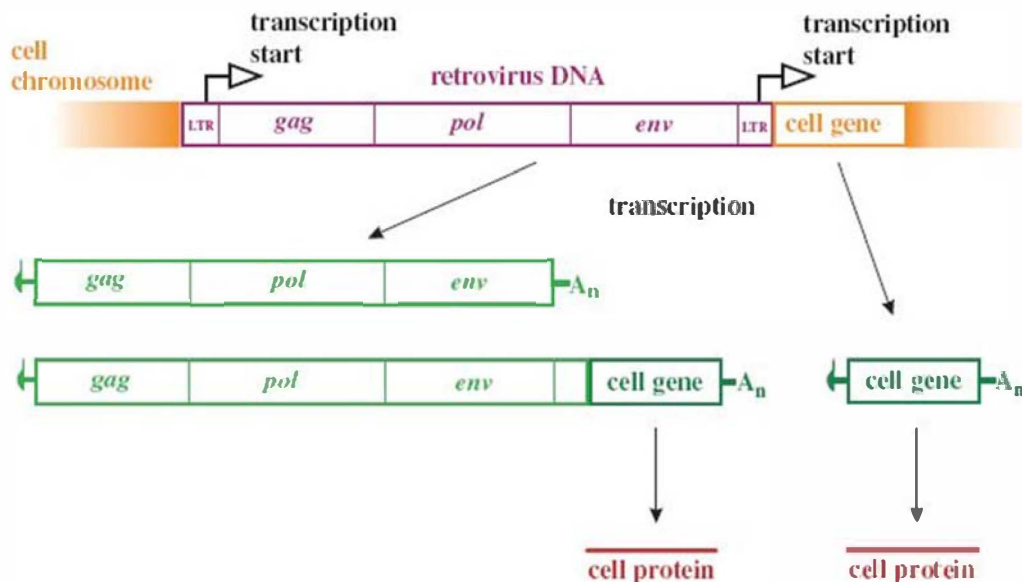
p53 is involved in some of these pathways, which may be blocked as a result of a virus protein binding to p53.

23.10.2 "Accidental" activation of cell genes

Some virus proteins are able to bind to cell proteins that may not be the intended targets and may thus trigger events that are of no value to the virus, but may be harmful to the host. A virus protein might inadvertently push a cell toward a cancerous state by activating a cell gene that is switched off, or by enhancing the transcription of a gene that is normally expressed at a low level.

Some retroviruses can induce tumor formation as a consequence of integrating their provirus into a cell chromosome at a site that puts a cell gene under the control of a virus promoter (Figure 23.5). Some cell genes that can be activated in this way are known as proto-oncogenes because their activation may result in a cancer. The proteins encoded by proto-oncogenes are generally involved in gene regulation or cell cycle control. An example of a proto-oncogene is c-myc. Infection of a chicken with an avian leukosis virus (Section 23.8) can cause a tumor when proviral integration next to c-myc results in greatly enhanced transcription of this gene. There were unfortunate outcomes in some gene therapy trials as a result of proto-oncogene activation by a retrovirus. A retroviral vector was used to introduce a gene into the patients' cells and some of the patients went on to develop leukemia.

Figure 23.5 Activation of a cell gene by insertion of a retroviral provirus. Initiation of transcription at either of the provirus LTRs may lead to expression of the cell gene.



23.10.3 Retroviral oncogenes

Some retroviruses have the ability to cause cancer because of the presence in the viral genome of an oncogene; for example, the Rous sarcoma virus genome contains the v-src gene (Figure 17.12). The retroviral oncogenes were derived from

cell proto-oncogenes and are closely related to them (Table 23.2), hence v-src is closely related to c-src.

Table 23.2 Examples of retroviral oncogenes and the cell proto-oncogenes from which they were derived. The names of some genes are derived from the virus names. Viral oncogenes are prefixed by "v-" while cell proto-oncogenes are prefixed by "c-".

Retrovirus	Viral oncogene	Cell proto-oncogene
Rous sarcoma virus	v-src	c-src
Avian myelocytomatosis virus	v-myc	c-myc
Harvey murine sarcoma virus	v-H-ras	c-H-ras

When a proto-oncogene is mutated or aberrantly expressed it becomes an oncogene. The src gene, which has been implicated in the development of many cancers, encodes a protein kinase; over-phosphorylation of the enzyme's substrates is a key process in the development of these tumors. C-myc encodes a transcription factor and is expressed at abnormally high levels in Burkitt's lymphoma (Section 23.4) and in avian leukosis (Section 23.10.2). Retroviral oncogenes lack the sequences that control the expression of cell proto-oncogenes, so when a retroviral oncogene is expressed in an infected cell the protein is often at a higher concentration than the product of the cell proto-oncogene.

Nearly all of the oncogene-carrying retro viruses are defective as a result of genome deletions; Rous sarcoma virus is an exception. The defective viruses can replicate only with the help of an endogenous retrovirus, or in a cell co-infected with a helper virus; for example, murine sarcoma viruses are defective and a murine leukemia virus may act as a helper. Retroviral oncogenes have the ability to induce rapid formation of tumors, usually 1–6 weeks post-infection, in contrast to tumors induced by other oncogenic viruses, which develop only after years or decades of persistent virus infection.

Oncogene-carrying retroviruses are rarely found in nature, but they have proved to be useful laboratory tools for studying the roles of mutated proto-oncogenes in cancer.

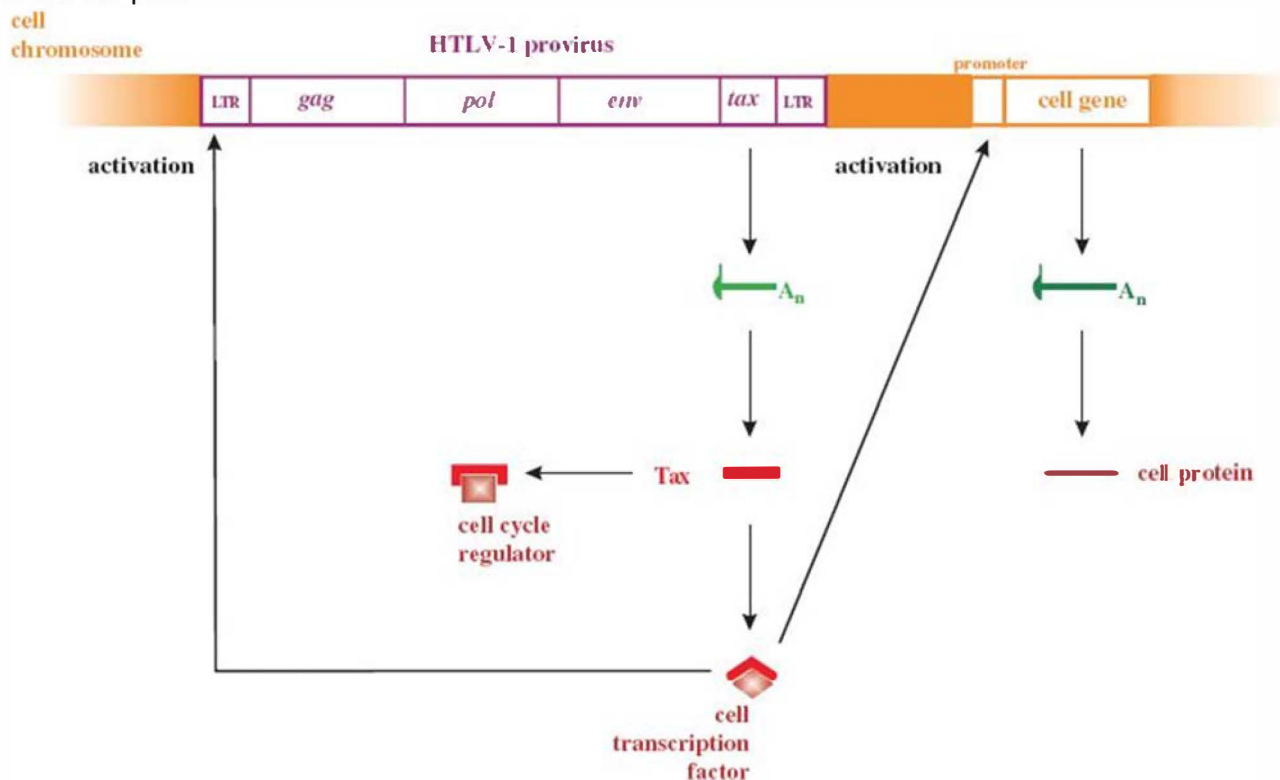
23.10.4 Damage to immune defenses

Interactions between cell proteins and oncogenic virus proteins can lead to breakdown of immune defenses that may allow the development of a cancer. Virus proteins may interfere with apoptosis, and hence prevent the death of virus-infected cells. Some of the proteins produced by EBV and KSHV in latently infected cells can interfere with acquired immune responses.

23.10.5 Overview of cancer induction by viruses

The previous sections have outlined some of the ways in which a virus might be involved in the induction of a cancer. For a number of viruses more than one mechanism might be involved. One such virus is HTLV-1, a complex retrovirus with auxiliary genes (Section 17.4), one of which is called *tax*. The Tax protein, in conjunction with cell transcription factors, functions as a transcription factor for the virus, stimulating the promoter in the LTR (Figure 23.6). The Tax protein, however, can also increase the expression of many cell genes, and it can inactivate cell cycle regulators including p53. Consequently this protein stimulates the cell cycle, inhibits apoptosis, and can induce chromosome instability.

Figure 23.6 HTLV-1 oncogenesis. The virus Tax protein, complexed with cell proteins, activates transcription of the provirus and may also activate transcription of cell genes. Tax can also influence the cell cycle by binding to cell cycle regulators such as p53.

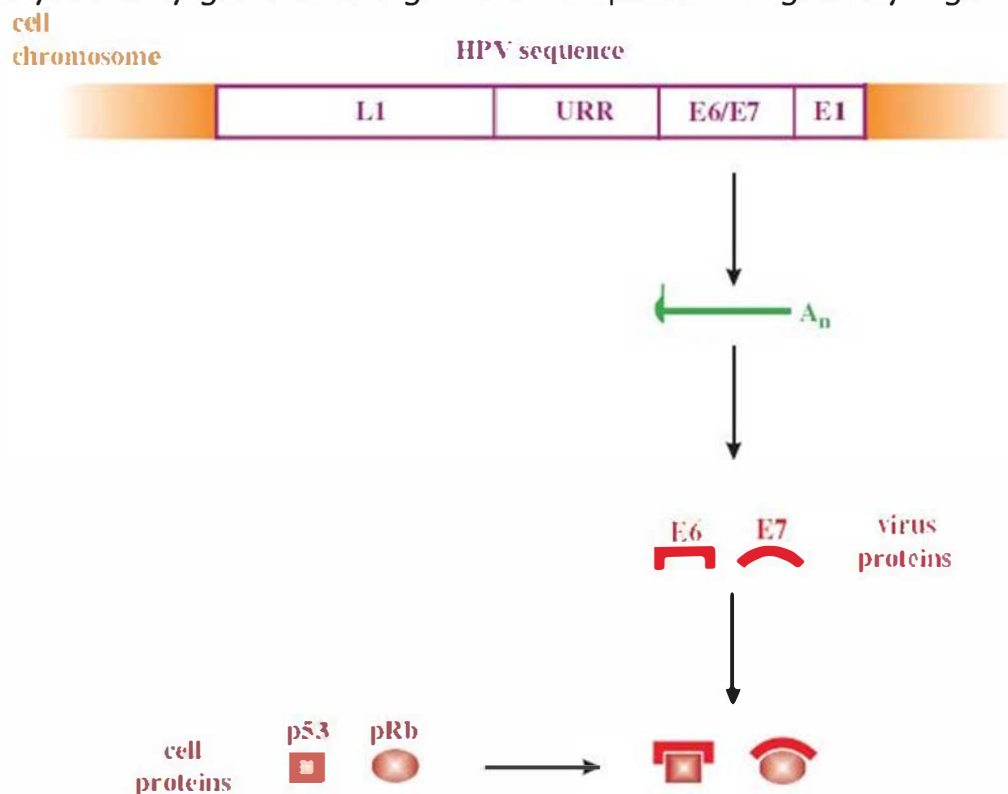


Another virus protein that can have multiple roles in cancer induction is the HBV X protein. This is the only virus protein that is consistently found in HBV tumor cells, and it is found at high levels. Like the HTLV-1 Tax protein, HBV X protein can bind to p53, and it can activate the expression of a wide range of cell genes through interactions with transcription factors. Evidence for the oncogenicity of the X protein includes its ability to induce liver cancer in some strains of transgenic mice. For some cancers, integration of virus DNA into cell DNA may be significant for cancer induction. HBV DNA does not integrate into the cell genome during the normal replication cycle, but in about 80% of HBV tumors virus DNA is integrated. Usually the virus DNA has undergone rearrangements, including deletions that have

destroyed the P and C ORFs, but the S and X ORFs are often intact (see Figure 19.7 for a reminder of the HBV genes). Integration can result in deletions from cell chromosomes.

The high-risk HPVs also have a propensity for integration of all or part of the virus genome into a cell chromosome. In the majority of cervical carcinomas HPV DNA is integrated; often the virus genome is incomplete, but the E6 and E7 genes are always present. The E2 protein normally controls the expression of E6 and E7, but if the E2 gene is absent E6 and E7 may be produced at high levels (Figure 23.7). Expression of these genes in a variety of cell types in culture can transform the cells, and the transformed state can be maintained by the continuous expression of the genes. Part of the genome of HPV-18 is integrated into the genome of HeLa cells, which were derived from a cervical cancer (Section 23.9).

Figure 23.7 Expression of HPV genes from an integrated virus sequence. The HPV-18 sequence that is integrated into the HeLa cell genome is shown. The E6 and E7 proteins are synthesized and bind to the cell proteins p53 and pRb, respectively. E: early gene. L: late gene. URR: upstream regulatory region.



23.11 PREVENTION OF VIRUS-INDUCED CANCERS

Logical approaches to preventing virus-induced cancers include attempting to prevent transmission of the viruses, so knowledge of the modes of transmission of the viruses is important (Table 23.3). This allows the development of strategies aimed at reducing the risks of transmission. Most of the oncogenic human viruses can be transmitted sexually and/or in blood, so measures implemented to prevent

transmission of HIV (Section 18.7) also reduce the risk of transmission of these viruses.

Table 23.3 The main modes of transmission of oncogenic human viruses

Virus	Transmission mode				
	Mother to baby		Via saliva	Sexual	Via blood
	Before or during birth	Breast feeding			
HPV				✓	
EBV			✓		✓
KSHV			✓		
HTLV-1		✓		✓	✓
HBV	✓			✓	✓
HCV	✓			✓	✓

Vaccination provides another approach to reducing rates of virus transmission. Most HBV transmission occurs from mother to baby around the time of birth, so several countries with high incidences of HBV infection have initiated HBV vaccination programs for babies within 2 days of birth. There is evidence from Taiwan of a reduced incidence of liver cancer in children as a result of this program, and in time there should be a reduced incidence in adults too.

Several countries have started vaccination programs against HPV using vaccines containing virus-like particles constructed from the capsid proteins of HPV types 16 and 18, two common high-risk types. The aim is to induce long-term immunity against these viruses, resulting in a reduction in the incidence of cervical and other HPV-associated cancers. Hopefully future vaccines will include a wider spectrum of high-risk types of HPV.

Marek's disease in chickens is economically important and vaccines are used to protect poultry against infection with the causative virus. Some vaccines contain an attenuated strain of the virus, while others contain a related herpesvirus from turkeys.

An alternative approach to the prevention of virus-induced cancers is the attempted elimination from the body of persistent infections with oncogenic viruses. So far this approach has had limited success. HCV infection in some patients has been eliminated by treatment with α -interferon (Section 9.2.1.b), either alone or in combination with ribavirin. Persistent HBV infections may be controlled, though not eliminated, using α -interferon and/or lamivudine. Please see Chapter 26 for more information on these anti-viral drugs.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- outline the characteristics of oncogenic viruses;
- evaluate the evidence for association of viruses with some cancers;
- discuss mechanisms whereby virus infection may induce cancer;

- discuss examples of human and animal cancers caused by viruses;
- suggest how virus-induced cancers may be prevented.

SOURCES OF FURTHER INFORMATION

Books

- Damania, B. and Pipas, J. M. (eds) (2009) *DNA Tumor Viruses*, Springer
- Khalili, K. and Jeang, K. T. (eds) (2010) *Viral Oncology: Basic Science and Clinical Applications*, John Wiley & Sons Ltd.
- Klein, G. and DiMaio, D. (2007) Principles of human tumor virology. Chapter 1 in *The Papillomaviruses* (Garcea, R. L. and DiMaio, D., editors), Springer
- Ou, J.-H. J. and Yen, T. S. B. (2010) *Human Oncogenic Viruses*, World Scientific
- Tschachler, E. (2011) Kaposi's sarcoma. Chapter 31 in *Sexually Transmitted Infections and Sexually Transmitted Diseases* (Gross, G. and Tyring, S. K., editors), Springer
- Weinberg, R. A. (2007) *The Biology of Cancer*, Garland

Journals

- Chemin, I. and Zoulim, F. (2009) Hepatitis B virus induced hepatocellular carcinoma. *Cancer Letters*, 286, 52–59
- de Flora, S. and Bonanni, P. (2011) The prevention of infection-associated cancers. *Carcinogenesis*, 32, 787–795
- Delecluse, H.-J. et al. (2007) Epstein–Barr virus-associated tumours: an update for the attention of the working pathologist. *Journal of Clinical Pathology*, 60, 1358–1364
- Gonçalves, D. U. et al. (2010) Epidemiology, treatment, and prevention of human T-cell leukemia virus type 1-associated diseases. *Clinical Microbiology Reviews*, 23, 577–589
- Houghton, M. (2009) The long and winding road leading to the identification of the hepatitis C virus. *Journal of Hepatology*, 51, 939–948
- McLaughlin-Drubin, M. E. and Munger, K. (2008) Viruses associated with human cancer. *Biochimica et Biophysica Acta*, 1782, 127–150
- Rodríguez, S. M. et al. (2011) Preventive and therapeutic strategies for bovine leukemia virus: lessons for HTLV. *Viruses*, 3, 1210–1248
- Sarid, R. and Gao, S.-J. (2011) Viruses and human cancer: from detection to causality. *Cancer Letters*, 305, 218–227
- Uldrick, T. S. and Whitby, D. (2011) Update on KSHV epidemiology, Kaposi sarcoma pathogenesis, and treatment of Kaposi sarcoma. *Cancer Letters*, 305, 150–162
- Zheng, Z.-M. (2010) Viral oncogenes, noncoding RNAs, and RNA splicing in human tumor viruses. *International Journal of Biological Sciences*, 6, 730–755
- zur Hausen, H. (2009) Papillomaviruses in the causation of human cancers—a brief historical account. *Virology*, 384, 260–265


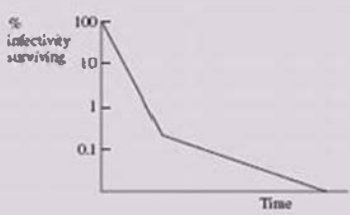
zur Hausen, H. (2009) The search for infectious causes of human cancers: where and why. *Virology*, 392, 1–10

CHAPTER 24

Survival of Infectivity

CHAPTER 24 AT A GLANCE

- Viruses vary greatly in the rate at which they lose infectivity.
- The rate of virus infectivity loss is affected by temperature, radiation, and the chemical environment, including pH.
- There may be two rates of infectivity loss.
- Understanding the kinetics of infectivity loss is important in practical situations, such as devising disinfection and sterilization procedures, storage of virus preparations and vaccine manufacture.
- In vaccine manufacture it is vital to ensure that maximum infectivity is preserved in live vaccines, and that 100% of infectivity is destroyed in inactivated vaccines.



24.1 PRESERVATION OF VIRUS INFECTIVITY

There are several practical situations in which there is a need to preserve virus infectivity. A prime example is during the storage and distribution of live virus vaccines (Sections 25.2 and 25.5), when there is a need to retain the potency of the vaccine. A second example is during the transport and storage of blood, feces, saliva, and other specimens from human and animal patients prior to testing in a diagnostic virology laboratory. Last, but by no means least, virologists need to preserve the infectivity of the viruses that they work with.

24.2 DESTRUCTION OF VIRUS INFECTIVITY

There are many situations in which there is a requirement to destroy virus infectivity. Procedures to minimize the transmission of pathogenic viruses (and other micro-organisms) in hospitals and other health care facilities include the disinfection of surfaces and the sterilization of equipment and materials. A good

example of a situation requiring rigorous action is an outbreak of winter vomiting disease caused by a norovirus; patients' vomit contains significant quantities of virus, which must be inactivated if the outbreak is to be brought under control.

Inactivated (killed) virus vaccines (Section 25.3) are made from preparations of virulent viruses, so it is essential that all infectivity is destroyed. Blood products, such as clotting factors for hemophiliacs and immunoglobulins, are treated to destroy any viruses (e.g. HIV, hepatitis viruses) that might be present in the donated blood. Continuing the theme of human and animal health, water supplies that are free from pathogenic viruses (and other micro-organisms) are of fundamental importance, so water treatment procedures must include processes that destroy these agents. The same applies to the treatment of swimming pool water.

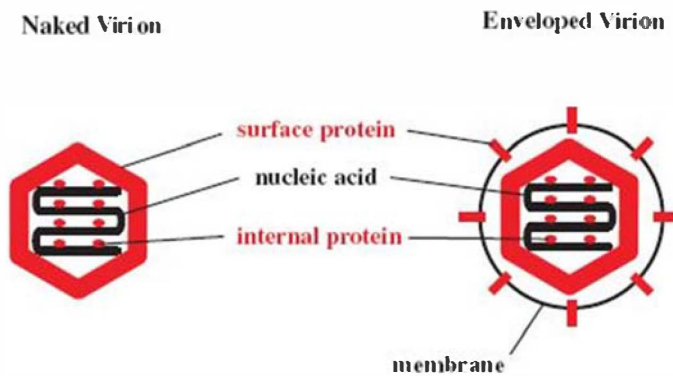
The processes for production of cheese, yogurt, and other dairy products utilize lactic acid bacteria to ferment the milk. There are many phages that infect these bacteria, and a phage infection during a production process can result in delay or failure of the fermentation, or in a poor quality product. To minimize the risk, procedures are required for treatment of the milk or whey to inactivate any phages present, and for the disinfection of the factory environment.

Finally, virologists need supplies of pipettes, cell culture vessels, and other laboratory materials that are guaranteed free from viruses (and other micro-organisms), then after these materials have been used any viruses present must be inactivated before disposal.

24.3 INACTIVATION TARGETS IN VIRIONS

The infectivity of a virion can be destroyed by the alteration or removal of its nucleic acid and/or one or more of its proteins (Figure 24.1). A protein molecule may be altered by the induction of a conformational change or, more drastically, by the breaking of covalent bonds, such as peptide bonds and disulfide bonds.

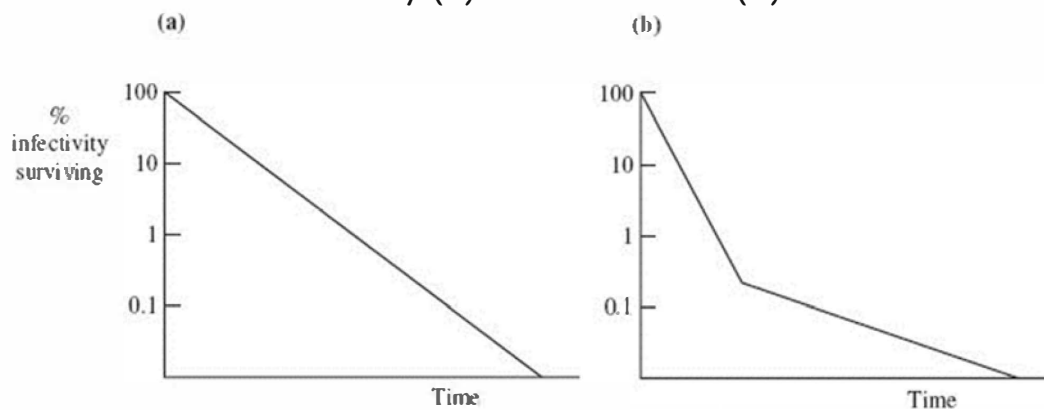
Figure 24.1 Inactivation targets in virions. Infectivity of a virion may be destroyed by damage to a nucleic acid, a protein, and/or a lipid membrane. Alteration of a surface protein might prevent a virion from attaching to its host cell and/or from entering the cell. Stripping the envelope from an enveloped virion removes the surface proteins and achieves the same outcome. Alteration of internal proteins can destroy properties, such as enzyme activities, essential for the replication of the virus.



24.4 INACTIVATION KINETICS

The rate at which a virus loses infectivity under defined conditions (for example, at a particular temperature and pH value) can be determined by incubating aliquots of the virus preparation under those conditions. At various time intervals the surviving infectivity in an aliquot is determined in an infectivity assay appropriate for the virus (Section 2.8). The logarithm of the percentage infectivity surviving is plotted against time (Figure 24.2).

Figure 24.2 Loss of virus infectivity (a) at one rate and (b) at two rates.



In some circumstances infectivity is lost at a constant rate until no further infectivity can be detected (Figure 24.2(a)). The slope of the line is a measure of the rate of infectivity loss. Sometimes there is a change in the rate of infectivity loss after most (usually more than 99%) of the infectivity has been destroyed (Figure 24.2(b)). It is not entirely clear why a small fraction of some virus preparations loses infectivity at a slower rate. It has been suggested that this fraction might contain virions that have a higher degree of resistance to the inactivating factor and/or that the preparation might contain virion clumps, with virions at the centers of the clumps having some protection.

24.5 AGENTS THAT INACTIVATE VIRUS INFECTIVITY

24.5.1 Physical agents

24.5.1.a Temperature

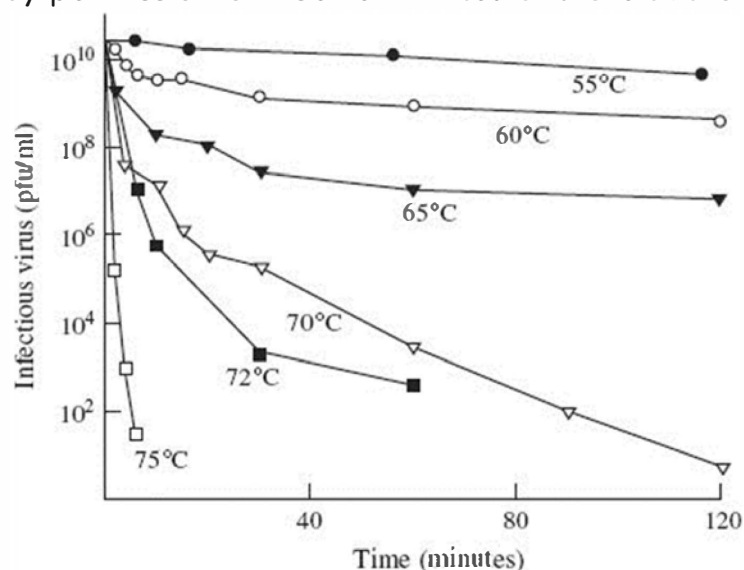
All infective viruses lose infectivity over time, the rate of infectivity loss being highly dependent on temperature. There is a lot of variability between viruses in the rates at which they are inactivated. In general, naked viruses are more stable than enveloped viruses; most naked viruses lose little infectivity after several hours at 37 °C, whereas most enveloped viruses lose significant infectivity under the same conditions. This generalization is broad, as there may be significant differences in stability between closely related viruses. Under dry conditions varicella-zoster virus survives well, while the related herpes simplex viruses survive poorly.

At temperatures below -35 °C most viruses lose infectivity very slowly, so for long-term storage of viruses they are placed in freezers or in liquid nitrogen. Preservation of infectivity may be enhanced by storing virus preparations in a freeze-dried form, rather than as aqueous suspensions. Moving to higher temperatures, most viruses are completely inactivated within 30 minutes at 60 °C or within a few seconds at 100 °C.

Some of the bacteriophages of lactic acid bacteria that cause problems in the dairy industry are quite heat resistant, and some infectivity may survive the heat treatments used in the decontamination procedures (Figure 24.3).

Figure 24.3 Inactivation of *Lactococcus lactis* phage P008 infectivity in a broth at temperatures between 55 and 75 °C.

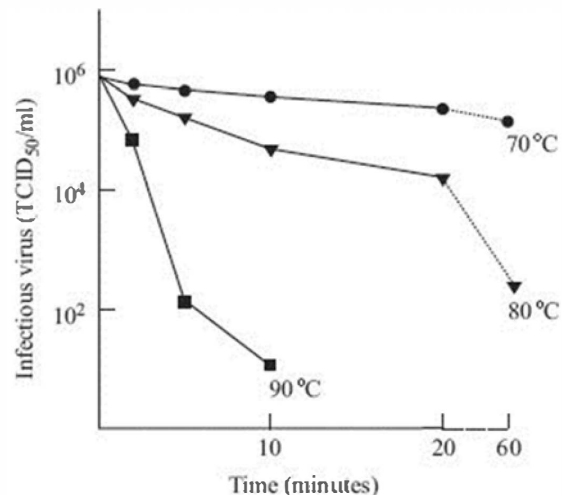
Source: Data from Müller-Merbach et al. (2005) *International Dairy Journal*, 15, 777. Reproduced by permission of Elsevier Limited and the authors.



The virions of some viruses, such as hepatitis A virus (HAV) and parvoviruses, are exceptionally heat stable. Data for the minute virus of mice (a parvovirus) indicate some infectivity surviving after 60 minutes at 80 °C (Figure 24.4).

Figure 24.4 Inactivation of minute virus of mice infectivity in water at 70, 80, and 90 °C.

Source: Data from Boschetti et al. (2003) *Biologicals*, 31, 181. Reproduced by permission of Elsevier Limited and the authors.



Exceptions to the generalization concerning the stability of naked and enveloped virions include those of hepatitis B virus (HBV), which are enveloped but have high stability, and those of poliovirus, which are naked but have low stability. Interestingly, HAV (heat-stable virions) and poliovirus (heat-labile virions) are both members of the family Picornaviridae, so, as with the herpesviruses, there is not always consistency in the degree of stability between viruses with similar structures.

The fact that virus preparations lose infectivity means that special precautions have to be taken to minimize infectivity loss during transport and storage of live vaccines, such as polio, mumps, measles, and rubella. These precautions include a "cold chain" of containers and vehicles that are refrigerated and/or insulated. This adds considerably to the expense of vaccination programs.

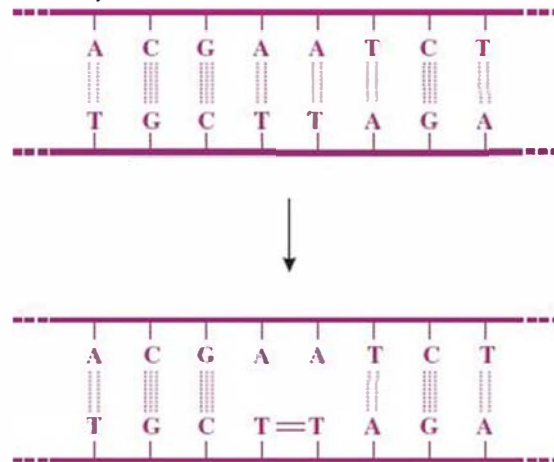
The rate of infectivity loss depends on the nature of the medium containing the virus. The presence of proteins, fats, glycerol, or certain salts may protect infectivity. Magnesium ions reduce the rate of inactivation of many RNA viruses, including poliovirus, so magnesium chloride is added to live polio vaccines to extend their shelf life.

The most heat-resistant viruses known are those that infect the thermophilic prokaryotes that inhabit hot springs. The optimum temperature for the growth of many of the hosts of these viruses is around 85 °C. Prions are also remarkably heat resistant (Section 27.3).

24.5.1.b Radiation

Some forms of radiation inactivate virus infectivity because they damage nucleic acids; for example, X-rays, gamma rays, and ultraviolet rays cause breaks in nucleic acids. Ultraviolet rays also cause other types of damage to nucleic acids, including the formation of thymine dimers in dsDNA (Figure 24.5). The thymine–adenine hydrogen bonds at two adjacent base pairs are broken and covalent bonds are formed between the two thymine residues.

Figure 24.5 Formation of a thymine dimer in dsDNA.



Sunlight inactivates virus infectivity because of its ultraviolet component. A practical aspect of this is the rapid inactivation by sunlight of baculoviruses applied to crops and forests as insecticides.

24.5.2 Chemical agents

Chemicals that inactivate virus infectivity do so with similar kinetics to physical agents, with either one or two inactivation rates (Figure 24.2). When some inactivation data are plotted the initial part of the graph has a small shoulder before the straight line. This could be due to a delay while the chemical reaches its target in the virion. Many chemicals can inactivate virus infectivity, but for many of them the inactivation target(s) are uncertain. We now consider the effects of hydrogen ion concentration (pH) and of a variety of other chemical agents on virus infectivity.

24.5.2.a Acids and alkalis

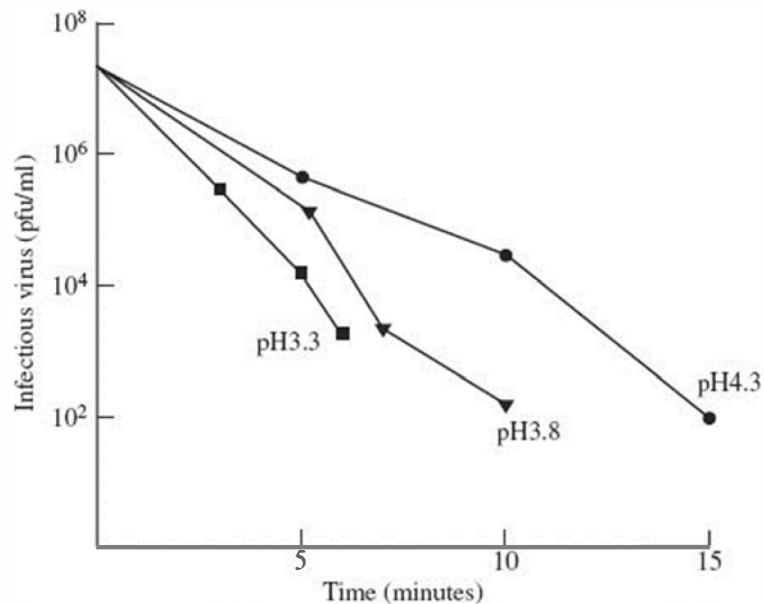
Most viruses are fairly stable at pH values between 5 and 9, but at more extreme pH values virions may undergo conformational changes and rates of infectivity loss usually increase. At high pH values both nucleic acids and proteins can be damaged. Alkali, in the form of dilute sodium hydroxide (sometimes in combination with steam), is used to disinfect manufacturing vessels and tools in the pharmaceutical industry.

The infectivity of some viruses is inactivated at low pH values as a result of alterations to surface proteins. Viruses that infect humans and other mammals via the gastrointestinal route, however, are relatively stable at low pH values. These viruses, which include rotaviruses (Chapter 13), have to survive in the stomach

contents, which are commonly at a pH value between 1 and 2. Furthermore, their virions must be able to withstand the sudden increase in pH when they are moved from the stomach to the intestine (pH about 7.5), and they must be resistant to inactivation by bile salts and proteolytic enzymes.

HAV and many enteroviruses infect their hosts via the gastrointestinal tract. These viruses are picornaviruses (Chapter 14), but, when compared to those picornaviruses (such as rhinoviruses and aphthoviruses) that infect their hosts via other routes, their virions are much more acid-resistant. It can be seen in Figure 24.6 that HAV infectivity can still be detected after 6 minutes at pH 3.3 at a temperature of 85 °C; thus, HAV is an example of a virus with high resistance to inactivation by both heat and acid.

Figure 24.6 Inactivation of hepatitis A virus infectivity at three pH values at 85 °C. Source: Data from Deboosere et al. (2004) *International Journal of Food Microbiology*, 93, 73. Reproduced by permission of Elsevier Limited and the authors.



Phenol, which is a weak acid, destroys the capsids of virions without altering nucleic acids. It is commonly used in the laboratory to extract nucleic acids from virions.

24.5.2.b Hypochlorite

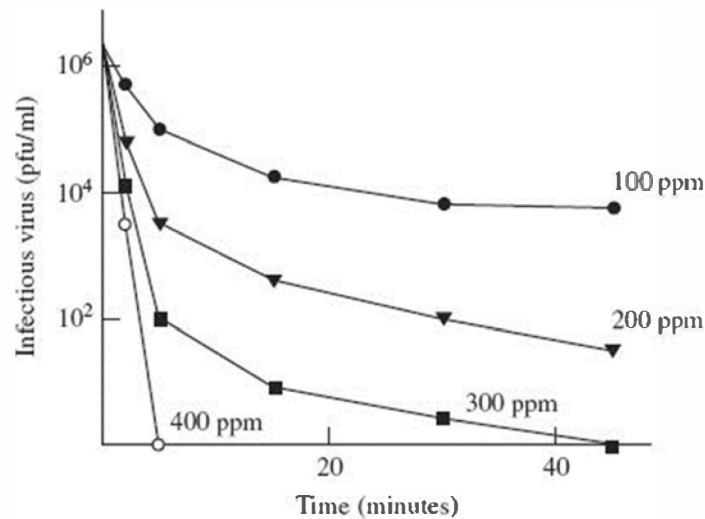
One of the most widely used disinfectants is hypochlorite. Both the hypochlorite ion (OCl^-) and the undissociated hypochlorous acid (HOCl) are excellent disinfectants; they act by oxidizing organic molecules, such as proteins. Hypochlorite ions are either supplied as sodium hypochlorite or they are generated by dissolving chlorine gas in water. The latter process is used to treat drinking water supplies and swimming pool water.

Sodium hypochlorite is one of the disinfectants used in the dairy industry to control phages of lactic acid bacteria such as *Lactobacillus* spp. The inactivation of a *Lactobacillus* phage at four concentrations of hypochlorite is shown in Figure 24.7.

At free chlorine concentrations 100–300 ppm there were slower rates of inactivation for a small percentage of the virions, as discussed in Section 24.4.

Figure 24.7 Inactivation of *Lactobacillus* phage BYM by hypochlorite. The concentrations are shown as parts per million (ppm) of free chlorine. The phage suspensions were at 25 °C and pH 7.

Source: Data from Quiberoni et al. (2003) *International Journal of Food Microbiology*, 84, 51. Reproduced by permission of Elsevier Limited and the authors.



24.5.2.c Aldehydes and alcohols

Formaldehyde and glutaraldehyde damage both enveloped and non-enveloped virions. Formaldehyde cross-links proteins and reacts with amino groups in nucleic acids; it is used to destroy the infectivity of poliovirus and other viruses for use in inactivated vaccines (Section 25.3). Glutaraldehyde is useful for disinfecting metal surfaces because it is non-corrosive.

Ethanol and isopropyl alcohol are used as general purpose disinfectants and in waterless hand washes.

24.5.2.d Lipid solvents

Detergents and many organic solvents, such as chloroform, destroy the infectivity of enveloped viruses by removing their lipid membranes and associated glycoproteins. Detergents and solvents are used in the preparation of clotting factors and other blood products to destroy enveloped viruses (particularly HIV, HBV, and HCV) that may be present in pooled blood donations. Lipid solvents may also damage virions with internal lipid, such as the corticoviruses and the tectiviruses (Section 3.5.2).

Quaternary ammonium compounds are cationic and disrupt virion envelopes. Several of them are used as disinfectants, especially chlorhexidine gluconate, which is widely used in health care institutions. Anionic detergents, such as sodium dodecyl sulfate, disrupt virion envelopes and capsids.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- discuss situations in which preservation of virus infectivity is important;
- discuss situations in which destruction of virus infectivity is important;
- give examples of physical and chemical agents that damage virions and cause them to lose infectivity;
- discuss the kinetics of inactivation of virus infectivity.

SOURCES OF FURTHER INFORMATION

Books

Bogosian, G. (2006) Control of bacteriophage in commercial microbiology and fermentation facilities. Chapter 42 in *The Bacteriophages*, 2nd edition (Calendar R. L., editor), Oxford University Press

Fraise, A. P., Lambert, P. A., and Maillard, J.-Y. (eds) (2004) Russell, Hugo and Aycliffe's *Principles and Practice of Disinfection, Preservation and Sterilization*, 4th edition, Blackwell

Journals

Farshid, M. et al. (2005) The clearance of viruses and transmissible spongiform encephalopathy agents from biologicals. *Current Opinion in Biotechnology*, 16, 561–567

Maillard, J.-Y. (2001) Virus susceptibility to biocides: an understanding. *Reviews in Medical Microbiology*, 12, 63–74

Rzez





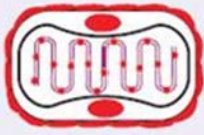


utka, A. and Cook, N. (2004) Survival of human enteric viruses in the environment and food. *FEMS Microbiology Reviews*, 28, 441–453

Takahashi, H. et al. (2011) Effect of food residues on norovirus survival on stainless steel surfaces. PLoS ONE, 6 (8): e21951

CHAPTER 25

Virus Vaccines

CHAPTER 25 AT A GLANCE

 Vaccine Types	Examples
Live attenuated virus 	Polio (Sabin)
Inactivated virus 	Polio (Salk)
Virion subunit 	Influenza (hemagglutinin and neuraminidase)
Live recombinant virus 	Rabies (for wildlife vaccination)
Virus-like particles 	Hepatitis B
DNA 	SARS coronavirus (experimental)

25.1 INTRODUCTION TO VIRUS VACCINES

In this chapter we discuss one of the major applications of virology: the development and use of vaccines against virus diseases in humans and animals.

The term vaccination is derived from the Latin word *vacca*, meaning cow. This is because the original procedure involved the inoculation of material from cowpox lesions into healthy people. Edward Jenner tried the procedure first in 1796 after he noticed that the faces of most milkmaids were unmarked by pocks; this was because milkmaids rarely contracted smallpox. They did, however, commonly contract cowpox, so Jenner inoculated material from a milkmaid's cowpox lesion into the arm of an 8-year-old boy. Six weeks later the boy was inoculated with material from a smallpox scab; he remained healthy. The immune response against cowpox virus had protected against smallpox virus, the protection resulting from related antigens in the two viruses.

A vaccine, therefore, contains material intended to induce an immune response, and this may involve both B cells (which develop into antibody-producing cells) and

T cells (responsible for cell-mediated immunity). Both arms of adaptive immunity can be important for providing protection against a virus infection (Section 9.2.2).

The purpose of most viral vaccines is to induce long-term immunity against the virus by establishing B cell and T cell memory that will be triggered if the virus ever invades the body. In order to establish strong immunological memory there is a requirement for vaccines that induce vigorous immune responses; in other words, highly immunogenic virus materials are required.

Vaccines are in use to protect against diseases caused by many viruses such as polio, rubella, rabies, and foot and mouth disease. Major achievements resulting from vaccine use include the eradication of smallpox in humans and of rinderpest in cattle. This chapter will describe the various categories of virus vaccine that are in medical and veterinary use, and will outline some aspects of their manufacture.

Effective vaccines have yet to be developed against many other viruses, including HIV-1, hepatitis C, Ebola, and the herpes simplex viruses. Those involved in virus vaccine research face many difficulties, such as multiple antigenic variants of target viruses and the requirements for high standards of safety. Some vaccines that have been developed have not been accepted for widespread use because of safety concerns. The great need for new vaccines has spawned attempts to produce novel categories of vaccine, such as replication-defective viruses, peptide vaccines, and DNA vaccines, and some of these will be introduced in this chapter.

25.2 LIVE ATTENUATED VIRUS VACCINES

A live attenuated vaccine contains a mutant strain of a virus that has been derived from a wild-type virulent strain. Vaccines of this type have a number of advantages over most other types of vaccine. One advantage is that there are increasing amounts of virus antigen in the body as the virus replicates. Another is that a wide-ranging immune response is induced that involves B cells, CD4 T cells, and CD8 T cells.

There are two properties that the virus for an attenuated vaccine must possess. First, its antigens must be identical, or very similar, to those of the wild-type virus, so that an immune response against the vaccine virus provides protection from infection with the wild-type virus. Second, the virulence of the wild-type virus must have been attenuated; in other words, the vaccine virus must have little or no virulence.

Most attenuated virus strains have been derived by "hit and miss" procedures such as repeated passage of wild-type virus in cells unrelated to the normal host. The vaccine strains of the three serotypes of poliovirus (Chapter 14) were derived from wild-type strains by passage in monkeys and in monkey kidney cell cultures. Albert Sabin did this pioneering work and the procedure he used to derive one of the attenuated strains is depicted in Figure 25.1. Poliovirus attenuation involves the loss of the ability of the virus to infect neurons (i.e., the loss of neurovirulence). The

mutations responsible for attenuation of the three serotypes are now known and some of them are shown in Figure 25.2.

Figure 25.1 Derivation of attenuated poliovirus strain (Sabin type 1) from wild-type poliovirus strain (Mahoney 1).

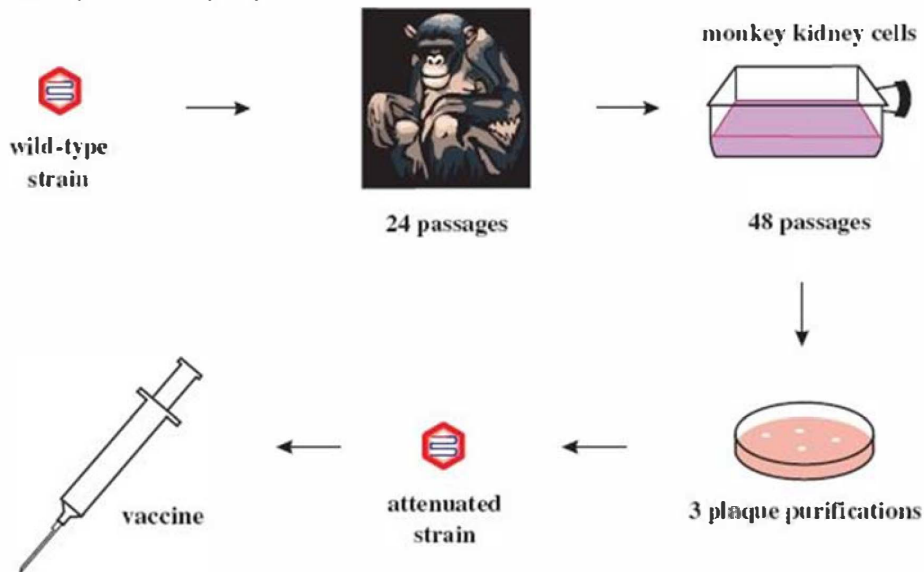
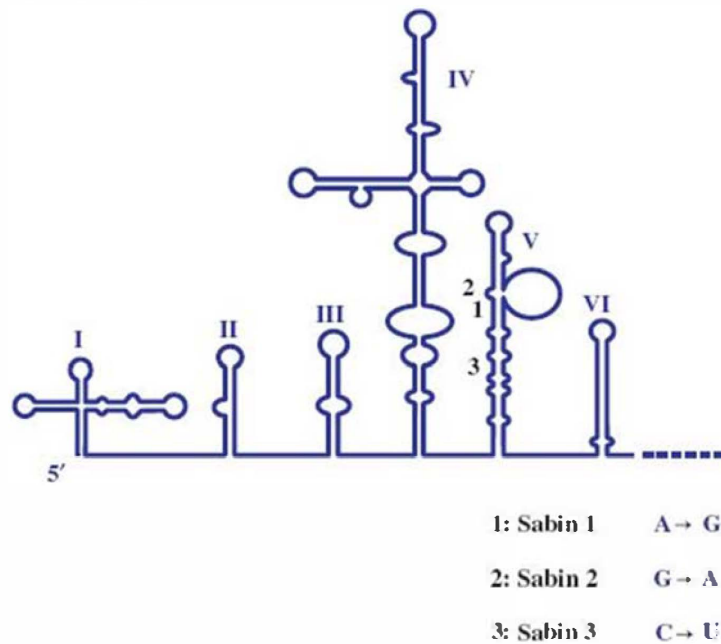


Figure 25.2 5' end of poliovirus RNA with expanded view of domain V. For each of the three Sabin strains a mutation in domain V that contributes to the attenuation of neurovirulence is indicated.



Other attenuated vaccines that have been developed using a similar approach include vaccines for mumps, measles, rubella, yellow fever, canine parvovirus, and canine distemper. The attenuated strains of the first three of these, now combined in the MMR vaccine, were all developed by Maurice Hilleman. The attenuated mumps virus is called Jeryl Lynn after his daughter, from whom the wild-type virus was isolated.

Attenuated vaccines have been developed in response to the huge burden of disease and death caused by rotaviruses (Chapter 13). The vaccines contain

reassortant rotaviruses (Section 21.3.3.c). One vaccine has genes derived from different human viruses, while another has some genes derived from a human virus and some from a bovine virus. These rotavirus vaccines are given by mouth.

Other approaches to the development of attenuated virus vaccines include the following.

- Cold-adapted virus strains. These are virus strains with reduced virulence that have been derived by incubating virus-infected cell cultures at temperatures below the optimum for virus replication. The strains are unable to replicate at normal body temperature, but can replicate at temperatures in the upper respiratory tract. Cold-adapted strains are used in influenza and respiratory syncytial virus vaccines, but their use is restricted because of the risk of side effects. The cold-adapted influenza vaccine is not used in elderly, pregnant, or immunocompromised persons and the respiratory syncytial virus vaccine is not used in children.
- Reverse genetics (Sections 14.6.1 and 15.6). The power of reverse genetics has been used to develop a respiratory syncytial virus vaccine by engineering into the virus genome mutations that attenuate virulence.

One of the risks in using a live vaccine is that, during virus replication, nucleotide substitutions might occur, resulting in reversion to virulence. Poliovirus is an RNA virus with a high mutation rate. When a vaccine strain of poliovirus replicates in the gut there may be nucleotide substitutions when it replicates its genome, or it may undergo recombination with a wild-type strain of poliovirus (Section 21.3.3.b). Either event may result in a virulent strain of poliovirus which can infect susceptible individuals, resulting in poliomyelitis for some. Because of these risks many countries switched from using attenuated polio vaccine to inactivated vaccine.

Each batch of a live virus vaccine is subjected to extensive quality control tests. These include testing for the presence of every known virus that could be a contaminant, and testing for neurovirulence in animals.

25.3 INACTIVATED VIRUS VACCINES

Inactivated, or killed, virus vaccines are made by mass-producing the virulent virus and then inactivating the infectivity, usually by treatment with a chemical such as β -propiolactone, ethylene oxide, or formaldehyde (Section 24.5.2.c). The trick lies in finding the combination of chemical concentration and reaction time that completely inactivates the virus, but leaves its antigens sufficiently unchanged that they can still stimulate a protective immune response.

Jonas Salk developed a treatment for poliovirus that led to the development of the vaccine that bears his name. After treatment with formaldehyde the inactivated virions (a mixture of the three serotypes of poliovirus) are adsorbed onto an adjuvant consisting of aluminum hydroxide or aluminum phosphate. The adjuvant enhances the immunogenicity of the vaccine. Other examples of killed virus

vaccines are those containing inactivated virions of influenza, hepatitis A, and foot and mouth disease viruses.

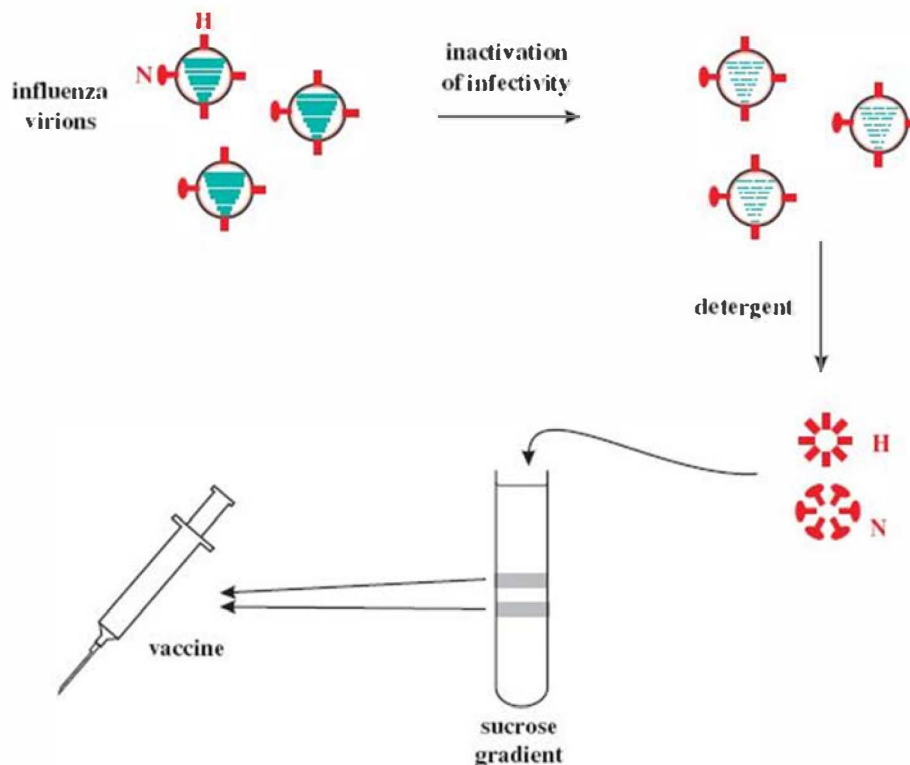
Because the virus used to produce an inactivated vaccine is a virulent strain, it is vital that 100% of the infectivity is destroyed in the production process. As discussed in Section 24.4, there are some situations in which a small proportion of virions is inactivated at a slower rate. It is essential that the kinetics of virus inactivation are understood and that an inactivation procedure is developed that guarantees 100% inactivation.

The need for complete inactivation was underlined in the US in 1955 when four million doses of inadequately inactivated Salk vaccine were inoculated into children. Amongst the vaccinees there were 204 cases of paralytic polio and 11 deaths. A similar problem occurred when there were outbreaks of foot and mouth disease in France and the UK in the period 1979–81 after batches of foot and mouth disease vaccine had been inadequately inactivated.

25.4 VIRION SUBUNIT VACCINES

A subunit vaccine contains purified components of virions. In the case of influenza the vaccines contain the hemagglutinin (H) and neuraminidase (N) surface glycoproteins (Section 16.2). A typical production method is outlined in Figure 25.3. The infectivity of a batch of influenza virions is inactivated with formaldehyde or β -propiolactone, then the virion envelopes are removed with a detergent, such as Triton X-100. This releases the glycoproteins, which form aggregates of H "cartwheels" and N "rosettes." These structures are purified by centrifugation in a sucrose gradient, and then material from three influenza virus strains is combined to form the vaccine.

Figure 25.3 Outline of production method for influenza virus subunit vaccine. Hemagglutinin (H) and neuraminidase (N) are extracted from inactivated influenza virions and purified by sucrose gradient centrifugation. The bands from the gradient are harvested and incorporated into the vaccine.



For protection against influenza the subunit vaccines are preferred over the inactivated vaccines because they cause fewer side effects, especially in children. The subunit vaccines, however, induce poorer immune responses, so two doses are necessary to provide adequate immunity.

25.5 LIVE RECOMBINANT VIRUS VACCINES

A recombinant vaccinia virus engineered to contain the gene for the rabies virus G protein has been used to vaccinate wild mammals against rabies. The use of this vaccine was discussed in Section 15.2.1.

25.6 MASS-PRODUCTION OF VIRUSES FOR VACCINES

The production of the types of vaccine discussed so far (live attenuated, inactivated, subunit, and live recombinant) requires large quantities of virions; for most viruses these are produced in cell cultures. John Enders demonstrated in 1949 that poliovirus can be grown in primary monkey cell culture and this led to procedures for the mass-production of poliovirus for the Salk and Sabin vaccines. It was subsequently found that some monkey cell cultures used for vaccine production had been harboring viruses, including simian virus 40 (Section 23.3). It was therefore recommended that human cell cultures should be used for the production of vaccines for human use, though cell lines of monkey origin, such as Vero cells (African green monkey cells), are used for the production of some.

Another preference is for the use of diploid cell lines, rather than heteroploid cell lines that have been derived from tumors, as there are perceived risks associated

with tumor-derived cells. Examples of human diploid cell lines used for production of vaccine viruses are MRC-5 (developed by the Medical Research Council in the UK) and WI-38 (developed by the Wistar Institute in the US). Both of these cell lines were derived from human lung fibroblasts.

Most of the cells are grown on surfaces such as parallel plates and small plastic particles. The virus inoculum is added to the cell culture at 0.1–10 pfu per cell and the maximum titer is usually reached 12–24 hours after inoculation. The manufacturer might expect a virus yield up to 100 000 times the inoculum.

Some viruses are mass-produced in chick embryos (Figures 2.1 and 25.4), either because no suitable cell culture system is available, or because such a system has been developed only recently. Influenza and yellow fever vaccines are produced in chick embryos that are inoculated by automatic systems. The yield of virus from each embryo provides sufficient virus for a few doses of vaccine.

Figure 25.4 Production of vaccine in chick embryos.

Source: Courtesy of the World Health Organization.



For influenza virus vaccines, decisions have to be made each year as to which H and N subtypes to include, as new influenza strains with altered subtypes are constantly arising through antigenic drift, and sometimes a new strain arises as a result of antigenic shift (Section 16.4). New strains are monitored and characterized by Collaborating Centers for Influenza Reference and Research around the world, and the World Health Organization recommends the strains that should be produced for the next batches of vaccine.

Newly isolated strains of influenza virus rarely grow well in chick embryos, so a reassortant (Section 21.3.3.c) between each new strain and a high-yielding strain is produced. A high-yielding strain that is commonly used is A/Puerto Rico/8/34 (an influenza A virus isolated in Puerto Rico in 1934). This strain evolved during passage in the laboratory; it has roughly spherical virions and it replicates well in chick embryos, in contrast to many wild-type strains, which have filamentous virions and replicate poorly in chick embryos. For vaccine production, reassortants are selected that have the H and N genes of the new strain and the remaining genes of the high-yielding strain.

Although influenza virus has traditionally been produced in chick embryos for incorporation into inactivated and subunit vaccines, cell culture systems for virus production have now been developed. Furthermore, live attenuated influenza virus vaccines have been licensed in some countries, so the types of influenza virus vaccine and the ways in which they are mass-produced may change in the future.

There are major enterprises around the world producing vaccines for human and veterinary use. Some examples of the vaccines produced are given in Table 25.1.

Table 25.1 Examples of viruses that are mass-produced for vaccine manufacture

Vaccines for humans	Vaccines for animals
Polio	Foot and mouth disease
Influenza	Marek's disease
Measles	Newcastle disease
Mumps	Canine distemper
Rubella	Rabies
Rabies	Bluetongue
Yellow fever	

25.7 VIRUS-LIKE PARTICLES

Virus-like particles are structures assembled from virus proteins. The particles resemble virions, but they are devoid of any nucleic acid and can therefore be deemed safer than vaccines containing attenuated or inactivated virions.

Hepatitis B virus (HBV) vaccine is produced in recombinant yeast cells that have the gene for the HBsAg inserted into the genome. The cells are grown in bulk and then broken to release the virus protein. After purification the HBsAg molecules receive a chemical treatment that causes them to aggregate into spherical structures similar to the non-infectious particles produced in the HBV-infected human host (Section 19.4). The structures are adsorbed onto aluminum hydroxide, which acts as an adjuvant.

The major capsid protein of papillomaviruses can self-assemble into virus-like particles that bear the epitopes required for generating neutralizing antibodies. (An epitope is a short sequence of amino acids critical for inducing an immune response.) Human papillomavirus vaccines (Section 23.11) contain virus-like particles that are produced either in insect cells using a baculovirus vector or in yeast cells.

25.8 SYNTHETIC PEPTIDE VACCINES

Each protein antigen has one or more epitopes. These short amino acid sequences can be synthesized in a machine and it was suggested that the resulting peptides might be used as vaccines. Compared with traditional vaccines it would be easier to ensure the absence of contaminants such as viruses and proteins.

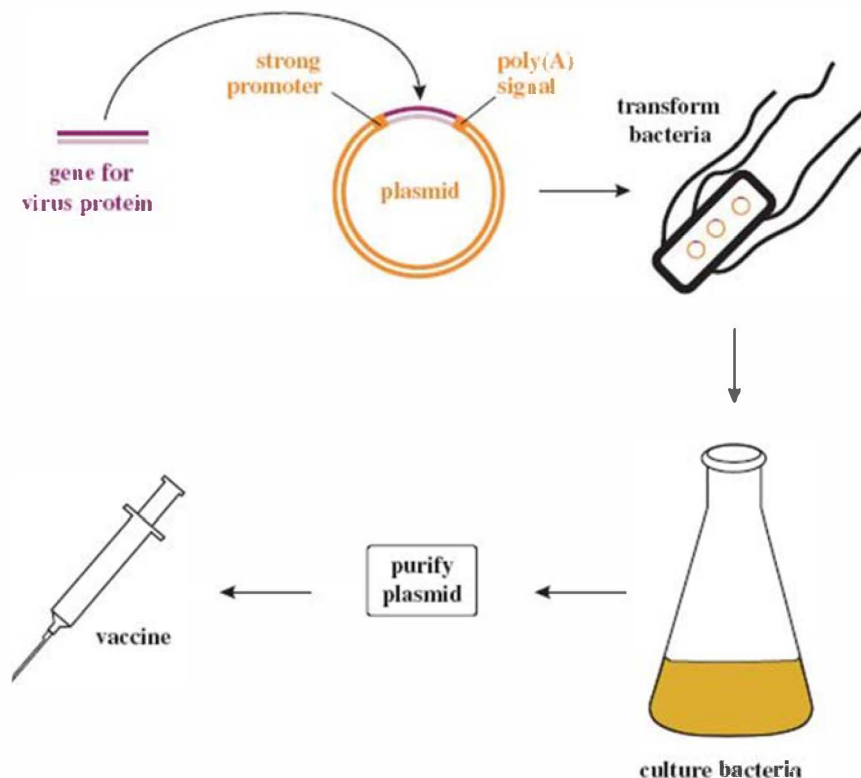
A lot of work has been done to try to develop peptide vaccines against foot and mouth disease virus. In this virus there is an important epitope within the virion protein VP1 (Section 14.3.1). Synthetic peptides of this sequence induced reasonable levels of neutralizing and protective antibodies in laboratory animals, but when vaccine trials were done in farm animals the results were disappointing.

25.9 DNA VACCINES

A revolutionary approach to vaccination is the introduction into the vaccinee of DNA encoding an antigen, with the aim of inducing cells of the vaccinee to synthesize the antigen. One advantage of this approach is that there is a steady supply of new antigen to stimulate the immune system, as with live virus vaccines. Because the antigen (a virus protein in this case) is produced within the cells of the vaccinee, it is likely to stimulate efficient T-cell-mediated responses.

A procedure for the production of a DNA vaccine is outlined in Figure 25.5. The antigen-encoding sequence, obtained directly from a DNA virus or by reverse transcription from an RNA virus, is inserted into a plasmid between a strong promoter and a poly(A) signal. The plasmid is replicated in bacterial cells and is then purified for use as a vaccine. Administration of the vaccine can be by injection into a muscle or by using a gene gun that delivers DNA-coated gold beads directly into skin cells.

Figure 25.5 Production of a DNA vaccine. The virus protein gene is inserted into a plasmid, which is then cloned in bacteria. The plasmid is extracted from the bacterial cells, purified, and incorporated into a vaccine.



Experimental DNA vaccines have been produced for a number of viruses, including HIV-1, SARS coronavirus, West Nile virus, and foot and mouth disease virus; trials have been carried out in animals and humans. Before any DNA vaccine goes into clinical use some important questions about safety must be answered. There must be confidence that the DNA will not trigger an anti-DNA autoimmune disease, and that it will not create cancer-causing mutations by insertion into host genomes.

25.10 STORAGE AND TRANSPORT OF VACCINES

Once a vaccine has been manufactured there is a need to preserve its efficacy until it is used. For live vaccines this means preserving virus infectivity; for vaccines containing inactivated virions, subunits, and virus-like particles it means preserving immunogenicity. The major physical and chemical factors that can reduce the infectivity of a live vaccine were considered in Chapter 24; some of these factors can also reduce the immunogenicity of vaccines.

Most vaccines are stored and transported at low temperatures; this minimizes losses of infectivity and immunogenicity. Substances that reduce the rate of infectivity loss are included in some vaccines, an example being magnesium chloride in live polio vaccines.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- explain how virus diseases are controlled with vaccines;
- evaluate the types of virus vaccine in medical and veterinary use;

- evaluate types of virus vaccine that are experimental;
- describe methods used to manufacture virus vaccines;
- evaluate procedures designed to ensure the safety of virus vaccines.

SOURCES OF FURTHER INFORMATION

Books

Castilho, L. et al. (2008) Chapter 18 in *Animal Cell Technology: From Biopharmaceuticals to Gene Therapy*, Taylor & Francis

Compans, R. W. and Orenstein, W. A. (eds) (2009) *Vaccines for Pandemic Influenza*, Current Topics in Microbiology and Immunology series, Volume 133, Springer

Delves, P. J. et al. (2011) Chapter 13 in *Roitt's Essential Immunology*, 12th edition, Wiley-Blackwell

Rappuoli, R. and Del Giudice, G. (eds) (2011) *Influenza Vaccines for the Future*, 2nd edition, Springer-Verlag

Journals

Amanna, I. J. and Slifka, M. K. (2011) Contributions of humoral and cellular immunity to vaccine-induced protection in humans. *Virology*, 411, 206–215

Berkley, S. et al. (2010) The 2010 scientific strategic plan of the Global HIV Vaccine Enterprise. *Nature Medicine*, 16, 981–989

Campo, M. S. and Roden, R. B. S. (2010) Papillomavirus prophylactic vaccines: established successes, new approaches. *Journal of Virology*, 84, 1214–1220

Frazer, I. H. and Levin, M. J. (2011) Paradigm shifting vaccines: prophylactic vaccines against latent varicella-zoster virus infection and against HPV-associated cancer. *Current Opinion in Virology*, 1, 268–279

Kanga, S. M., Songa, J. M., and Compans, R. W. (2011) Novel vaccines against influenza viruses. *Virus Research*, 162, 31–38

Kwok, R. (2011) Vaccines: the real issues in vaccine safety. *Nature*, 473, 436–438

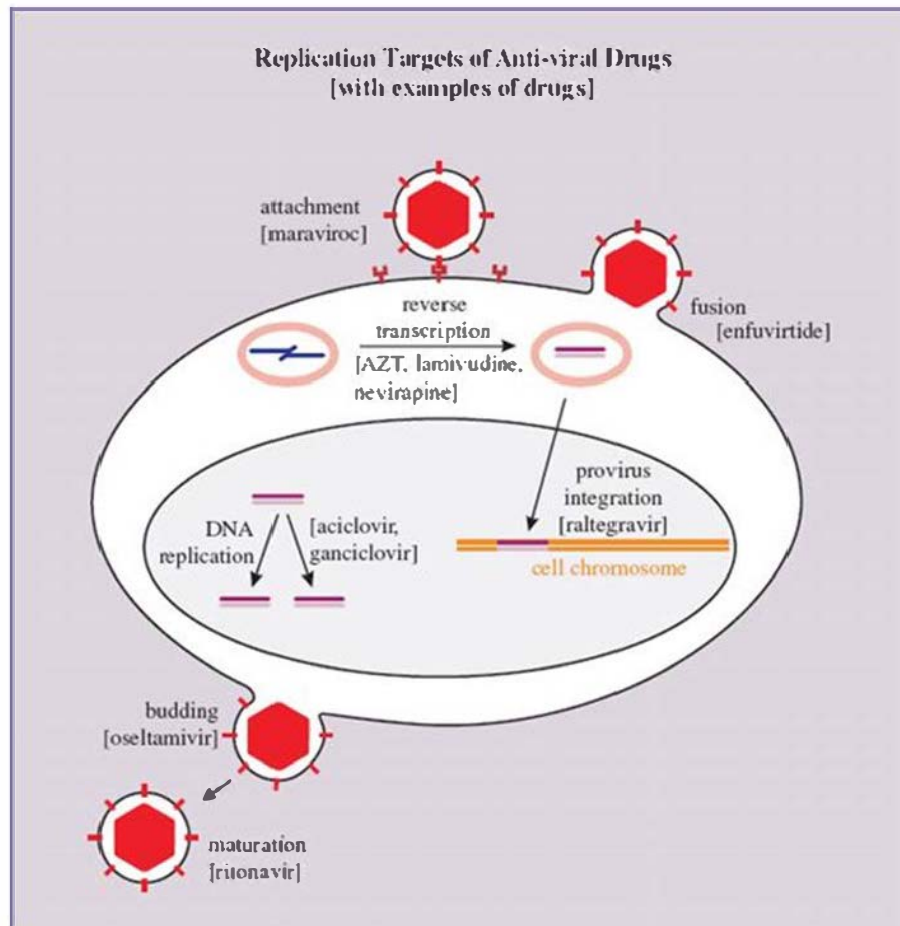
Plummer, E. M. and Manchester, M. (2011) Viral nanoparticles and virus-like particles: platforms for contemporary vaccine design. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 3, 174–196

Victoria, J. G. et al. (2010) Viral nucleic acids in live-attenuated vaccines: detection of minority variants and an adventitious virus. *Journal of Virology*, 84, 6033–6040

CHAPTER 26

Anti-viral Drugs

CHAPTER 26 AT A GLANCE

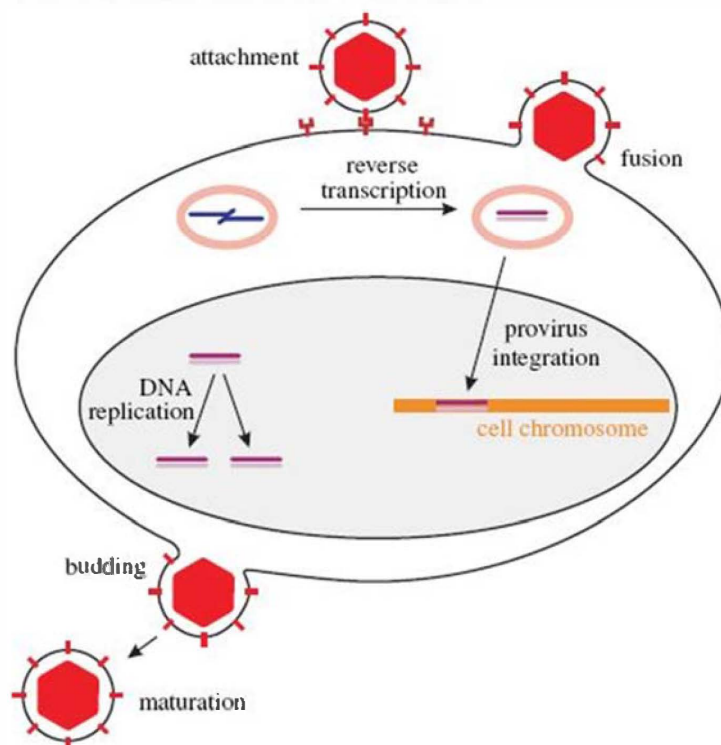


26.1 INTRODUCTION TO ANTI-VIRAL DRUGS

In this chapter we discuss another major application of virology: the development and use of drugs to combat virus diseases in humans. Drugs are available for the treatment of diseases caused by a variety of viruses, and in some circumstances drugs are also used for the prevention of virus infections. We shall discuss the modes of action of several anti-viral drugs, and their clinical uses. We shall also consider some of the problems associated with anti-viral drug use, especially the emergence of drug-resistant virus strains.

For a long time there were very few anti-viral drugs available for clinical use compared with the number of anti-bacterial and anti-fungal drugs. This was because it was difficult to find compounds that interfere specifically with viral activities without causing significant harm to host cell activities. As we have seen in earlier chapters, however, there are many virus-specific activities that are potential drug targets, some of which are summarized in Figure 26.1.

Figure 26.1 Virus activities that are potential drug targets. A generalized view of virus replication is depicted; no single virus performs all of these activities! DNA replication provides a target if the virus produces enzymes (e.g. DNA polymerase, thymidine kinase) distinct from those of the host.



Most of the early anti-viral drugs were found by screening large numbers of compounds, but this approach has now been largely superseded by designing drugs to inhibit specific virus proteins. Both approaches will be outlined in this chapter.

26.2 DEVELOPMENT OF ANTI-VIRAL DRUGS

26.2.1 Screening compounds for anti-viral activity

In this approach to the search for clinically useful anti-viral drugs, large numbers of compounds are tested for possible anti-viral activity. Some of the compounds chosen in the past had already been shown to have anti-tumor cell activity.

The screening procedure involves testing dilutions of the compounds against a range of viruses growing in cell cultures. Evidence that a compound interferes with the replication of a virus might include inhibition of cytopathic effect (CPE) or inhibition of plaque formation. For each potentially useful compound the concentration that inhibits CPE or plaque formation by 50% is determined. This concentration of the drug is known as the 50% inhibitory concentration (IC₅₀).

Screening compounds for anti-viral activity has been largely replaced by the rational design of drugs.

26.2.2 Rational design of anti-viral drugs

The process of designing an anti-viral drug starts with deciding on a virus activity that the drug will target (Figure 26.1). Once this has been selected, it is necessary

to choose a target protein, such as a viral enzyme, that is involved in that activity. A detailed picture of the three-dimensional structure of the protein is derived using techniques such as X-ray crystallography (Section 2.5.3), and a target site in the protein is selected. Computer programs are then used to design compounds that will bind to the target site with the aim of inhibiting the activity of the virus protein. Rational design has been used to develop a large number of anti-viral drugs, though the mode of action of some of them remains unclear.

26.2.3 Safety of anti-viral drugs

If a compound is to have clinical value it must not only inhibit virus replication, but it must cause little or no harm to the host. The research and development procedures therefore include incubating uninfected cell cultures with the compound and examining them for damage such as death or inhibition of cell division.

The potential value of a compound can be assessed by determining the ratio of its cell toxicity to its anti-viral activity. This ratio is known as the selectivity index (SI) and is expressed by the formula

$$SI = \frac{\text{minimum concentration inhibiting cell proliferation (or DNA synthesis)}}{\text{minimum concentration inhibiting virus replication}}$$

A compound with a low IC₅₀ and a high SI is most likely to have value as an anti-viral drug. Some examples of IC₅₀ and SI values are given in Table 26.1. The data indicate that aciclovir would be preferred to ganciclovir for treatment of herpes simplex virus infections.

Table 26.1 Some anti-viral drugs and target viruses. IC₅₀: 50% inhibitory concentration. SI: selectivity index. IC₅₀ and SI vary depending on the virus strain and the cell type. These figures are ranges determined using several cell types. Data from Kinchington et al. (1995)

Drug	Target virus	IC ₅₀ (μM)	SI
Aciclovir	Herpes simplex virus	0.1–1	>1000
	Varicella-zoster virus	2–10	>1000
Ganciclovir	Herpes simplex virus	0.2–2	100–1000
	Cytomegalovirus	2–10	100–1000
Azidothymidine	HIV-1	0.03–0.5	>1000
Dideoxycytidine	HIV-1	0.5–1.5	100–1000
Dideoxyinosine	HIV-1	2–10	≥100

26.3 EXAMPLES OF ANTI-VIRAL DRUGS

26.3.1 Nucleoside analogs

A number of anti-viral drugs are synthetic compounds structurally similar to nucleosides such as guanosine and 2BFI-deoxythymidine (Figure 26.2) and they act

by interfering with the synthesis of virus nucleic acids. After being taken into a cell a nucleoside analog, like a nucleoside, is phosphorylated at the 5' carbon (or its equivalent) to become a nucleotide analog (Figure 26.3). The 5' triphosphate derivative of the nucleoside analog is the active form of the drug and acts as a competitive inhibitor of a viral polymerase such as a reverse transcriptase.

Figure 26.2 Nucleoside analogs used as anti-viral drugs. The analogs are derived from the nucleosides shown on the left.

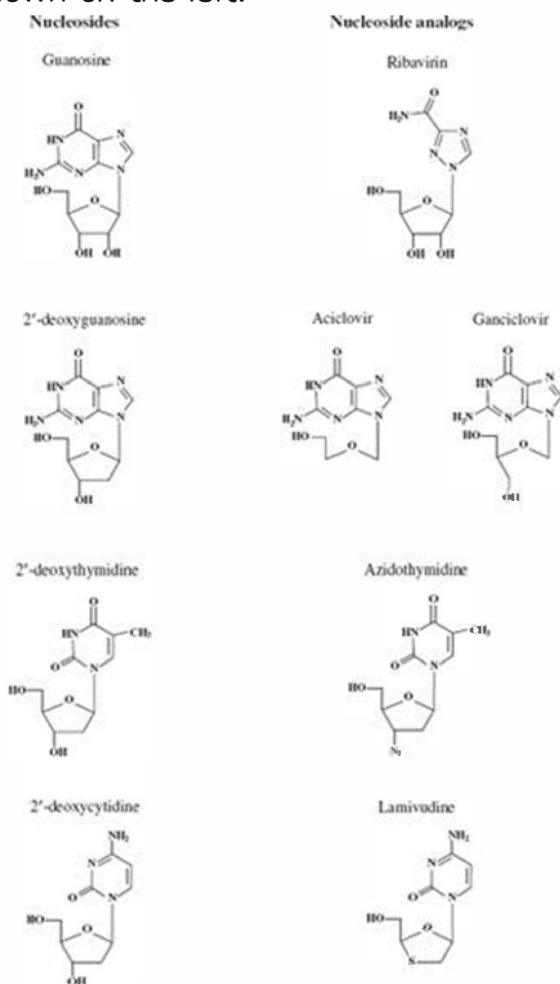
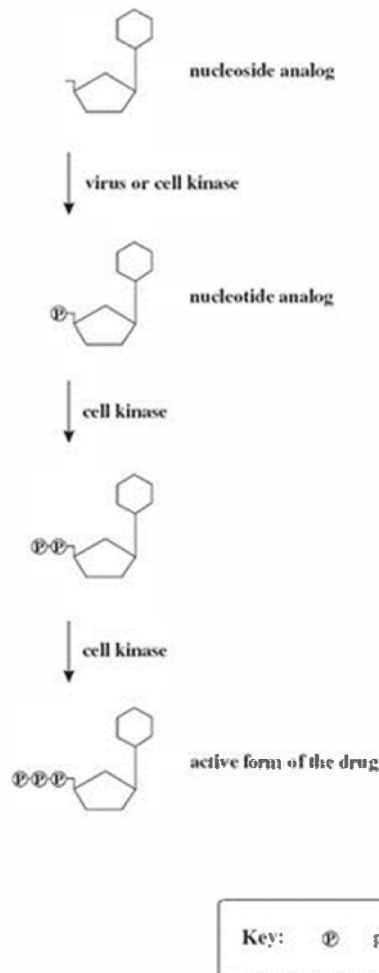


Figure 26.3 Phosphorylation of a nucleoside analog. The active form of the drug is the 5' triphosphate derivative of the nucleoside analog.



During nucleic acid synthesis, if one of these nucleotide analogs is incorporated into a growing strand, then nucleic acid synthesis is terminated, as in the dideoxy method of DNA sequencing. The structure of the nucleotide analog prevents it from accepting the next nucleotide. In some cases the analog lacks the hydroxyl group on the 3' carbon (or lacks the 3' carbon) necessary for the addition of the next nucleotide.

26.3.1.a Ribavirin

Ribavirin is an analog of guanosine (Figure 26.2). It is used for the treatment of infection with several RNA viruses, especially hepatitis C virus (HCV; Section 23.11). Combined treatment with ribavirin and pegylated α -interferon eradicates HCV infection in about 50% of patients, but the treatment often causes significant side effects. Inhibitors of an HCV protease have shown promising results in clinical trials and are to be used, in conjunction with ribavirin and interferon, for treatment of HCV infections.

Ribavirin has been used for treatment of infections with respiratory syncytial virus (Section 15.5), one of the most important causes of severe respiratory tract disease in children. There are, however, doubts about the effectiveness of the drug, and the mode of delivery is expensive and inconvenient (aerosol for 12–18 hours daily), so its use is restricted to some high-risk patients.

The mode of action of ribavirin is not completely understood and several hypotheses have been proposed to explain its anti-viral activity. The drug inhibits the cell enzyme inosine monophosphate dehydrogenase, and has been shown to inhibit several virus activities, including viral RNA synthesis; it is likely that it has more than one mode of action.

26.3.1.b Nucleoside analogs that inhibit herpesvirus DNA synthesis

Aciclovir (also spelt α -acyclovir) is an analog of guanosine (Figure 26.2). It is a very safe drug, with almost no side effects, widely used for the treatment and prevention of diseases caused by three of the herpesviruses: herpes simplex viruses 1 and 2 (HSV-1 and HSV-2) and varicella-zoster virus (VZV) (Chapter 11). Uses of aciclovir include:

- prevention of cold sores;
- prevention and treatment of genital herpes;
- treatment of HSV encephalitis (early treatment reduces the mortality rate);
- protection of immunocompromised patients after exposure to VZV;
- treatment of shingles (pain is reduced significantly).

Aciclovir strongly inhibits virus DNA synthesis but has very little effect on cell DNA synthesis. In a herpesvirus-infected cell the first phosphorylation of aciclovir (Figure 26.3) is carried out by the virus thymidine kinase (TK) (Section 11.5.3). The cell TK is produced only at certain stages of the cell cycle and has a much lower affinity for aciclovir than the viral enzyme, so there is very little phosphorylation of aciclovir in uninfected cells.

During DNA synthesis, aciclovir triphosphate competes with deoxyguanosine triphosphate (dGTP) for incorporation into the new strand. When aciclovir becomes incorporated, DNA synthesis is terminated due to the lack of a 3'-OH group.

Ganciclovir, like aciclovir, is an analog of guanosine, and the first phosphorylation in a cell is carried out by a virus enzyme, in this case a protein kinase. Ganciclovir is more likely than aciclovir to become incorporated into cell DNA. As a result ganciclovir has a lower SI than aciclovir (Table 26.1) and can cause significant side effects, especially reduced numbers of blood cells because of damage to bone marrow cells.

Ganciclovir is used to treat infections with cytomegalovirus, as aciclovir is less effective against this virus than it is against the other herpesviruses. Examples of cytomegalovirus diseases that are treated with ganciclovir include pneumonia and retinitis in immunocompromised patients, such as those who have received organ transplants and those with AIDS.

26.3.1.c Nucleoside analogs that inhibit reverse transcription

Azidothymidine (AZT), also known as zidovudine, is an analog of thymidine (Figure 26.2). Like other nucleoside analogs, it is phosphorylated to the 5'-triphosphate

after uptake into a cell (Figure 26.3). AZT had been investigated as a possible anti-cancer drug and later it was found to inhibit the reverse transcriptase of HIV. AZT triphosphate binds more strongly to the viral reverse transcriptase than to the cell DNA polymerase, and the reverse transcriptase binds AZT triphosphate in preference to deoxythymidine triphosphate. AZT thus inhibits HIV reverse transcription, but it can also interfere with cell DNA synthesis and, like ganciclovir, it can damage the bone marrow. In the early days of AZT treatment HIV-infected patients were administered large doses and many patients suffered severe side effects.

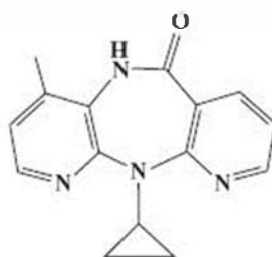
Lamivudine, also known as 3TC, is an analog of cytidine (Figure 26.2). It inhibits reverse transcriptases, including those of HIV and hepatitis B virus (HBV), and is used, usually in combination with other therapies, to treat infections with these viruses. The drug does not eliminate the infections, but it can control them. For treatment of HBV infection lamivudine is often given in combination with α -interferon; after treatment for a year or more inflammation of the liver is reduced in a proportion of patients.

Several other nucleoside analogs that inhibit reverse transcription are used for treatment of HIV infection (examples are dideoxycytidine and dideoxyinosine) or HBV infection (an example is adefovir).

26.3.2 Non-nucleoside inhibitors of reverse transcription

The approach of rational design was used to develop several HIV reverse transcriptase inhibitors that are not based on nucleosides. These non-nucleoside inhibitors target different sites in the reverse transcriptase to those targeted by the nucleoside analogs. An example of a non-nucleoside inhibitor that is used against HIV is nevirapine (Figure 26.4).

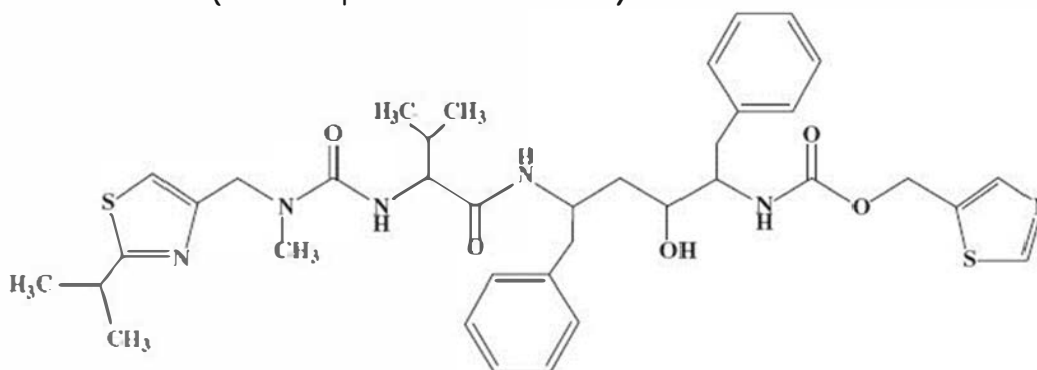
Figure 26.4 Nevirapine (a non-nucleoside inhibitor of reverse transcriptase).



26.3.3 HIV protease inhibitors

The maturation of a retrovirus virion involves the cleavage by the virus protease of the Gag and Gag-Pol proteins to form the virion proteins (Section 17.3.6). If this processing does not take place the virion does not acquire infectivity. Peptide mimics of the cleavage site in the protein have been developed and these compounds can fit into the active site of the HIV protease. HIV-infected cells treated with one of these protease inhibitors release fewer virions, and those virions that do bud are unable to initiate a successful replication cycle in another cell. An example of an HIV protease inhibitor is ritonavir (Figure 26.5).

Figure 26.5 Ritonavir (an HIV protease inhibitor).



26.3.4 HIV attachment and fusion inhibitors

Binding of an HIV-1 virion to a cell involves attachment to a co-receptor, usually CCR5 or CXCR4 (Section 18.4.1). A drug (maraviroc) has been developed that blocks attachment to CCR5, so it is active against R5 strains of HIV-1. A major limitation of maraviroc is its inability to block attachment to CXCR4, so it is not active against X4 strains.

After an HIV-1 virion has bound to a cell the trans-membrane glycoprotein gp41 must fuse the envelope with the membrane of the cell if the virion contents are to be delivered into the cytoplasm. In order for this to take place gp41 must undergo a conformational change initiated by interaction between different regions of the molecule.

A number of drugs have been developed that can inhibit fusion between the membranes of an HIV-1 virion and a potential host cell. They do this by binding to gp41 and inhibiting the conformational change. An example is enfuvirtide, a 36-amino acid peptide with the sequence of a gp41 region that interacts with another gp41 sequence to cause the conformational change. Enfuvirtide has much lower activity against HIV-2, which has a trans-membrane glycoprotein only distantly related to that of HIV-1. A disadvantage of enfuvirtide is that it has to be given by subcutaneous injection, whereas the other anti-HIV drugs can be given orally.

Unlike most anti-HIV drugs, maraviroc and enfuvirtide prevent infection of a cell with HIV-1.

26.3.5 HIV integrase inhibitors

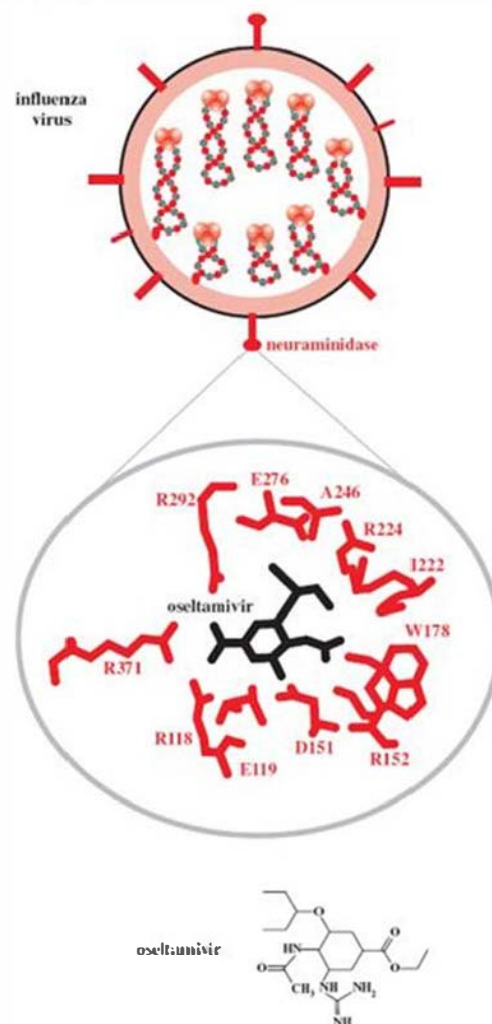
As a retrovirus, HIV takes into a host cell an integrase, the role of which is to integrate a DNA copy of the virus genome (the provirus) into a cell chromosome (Section 17.3.3). Inhibitors of this enzyme have been developed, one of which is raltegravir.

26.3.6 Influenza virus neuraminidase inhibitors

One of the surface proteins of influenza A and B viruses is a neuraminidase. This enzyme plays a vital role during the final stages of virion budding from infected cells (Section 16.3.4); if the enzyme is inhibited virions are not released. Each

neuraminidase spike on the virion surface is made up of four monomers, each consisting of a balloon on a stick.

In 1983 an influenza virus neuraminidase was crystallized and its structure was determined. It was found that the monomer has a deep cleft and that this forms the active site of the enzyme. Compounds were designed to bind in the cleft and several were found to inhibit neuraminidase activity. One such compound is oseltamivir (Figure 26.6), which is approved for use as an anti-influenza virus drug. Figure 26.6 Binding of oseltamivir to influenza virus neuraminidase. The inset shows the amino acid residues (one-letter abbreviations together with their codon numbers) that bind oseltamivir.



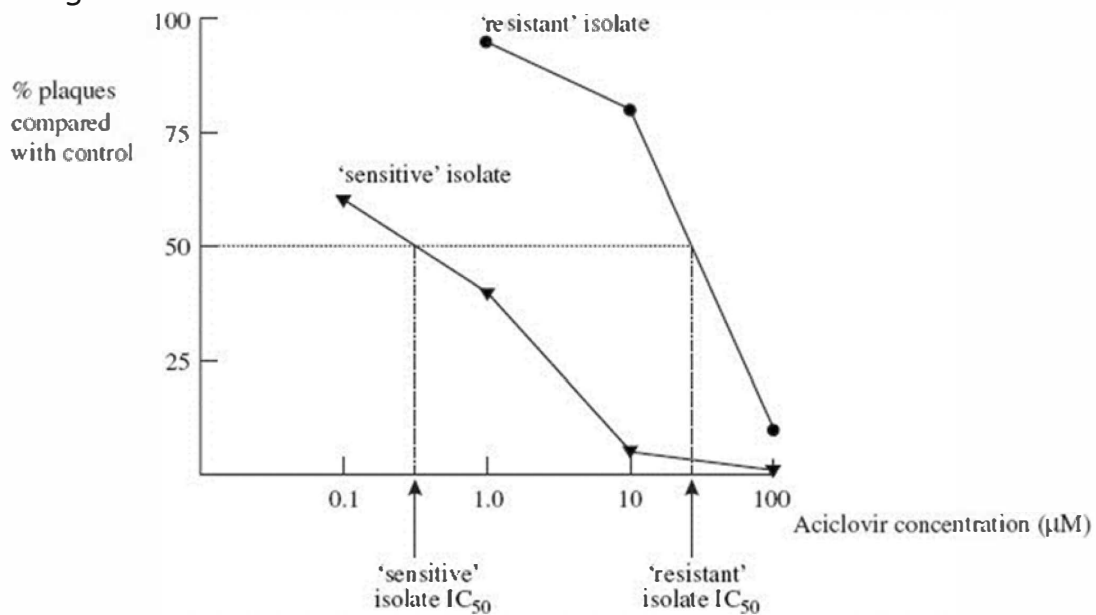
26.4 DRUG RESISTANCE

Soon after the introduction of anti-viral drugs into clinical practice drug-resistant strains of the target viruses began to emerge. Many strains of influenza A virus are now resistant to amantadine and rimantadine, two early anti-influenza drugs, and some strains are now resistant to neuraminidase inhibitors. The emergence of drug-resistant viruses should not have been surprising, as antibiotic-resistant bacteria, insecticide-resistant insects, and rodenticide-resistant rats had emerged as a result of natural selection. Viruses, especially RNA viruses, can mutate at high frequencies and evolve rapidly (Section 21.3.3.a) so genotypes encoding drug

resistance can arise rapidly. Drug-resistant genotypes may be at an advantage in hosts where the drug is present and may become the dominant genotypes in those hosts.

Drug resistance of viruses is relative, rather than absolute. A measure of the degree of resistance can be obtained by determining the IC₅₀ (Figure 26.7). A virus strain is considered to be resistant to a drug if it is able to replicate in the body in the presence of a concentration of the drug that inhibits replication of sensitive strains.

Figure 26.7 Susceptibility of sensitive and resistant isolates of HSV-1 to aciclovir. The resistant isolate has an aciclovir IC₅₀ approximately 100 times greater than the sensitive isolate.



Drug-resistant viruses are found to have one or more mutations in the genes encoding the proteins that are the drug targets. The vast majority of aciclovir-resistant HSV-1 isolates have a mutated tk gene, while a small number of isolates have mutations in the DNA pol gene. HIV-1 strains resistant to protease inhibitors such as ritonavir are mutated in the gene encoding either the enzyme (PR) or the substrate (gag). Further examples of mutations in drug-resistant viruses are given in Table 26.2 and Figure 26.8.

Figure 26.8 Examples of mutations to drug resistance in the genome of HIV-1.

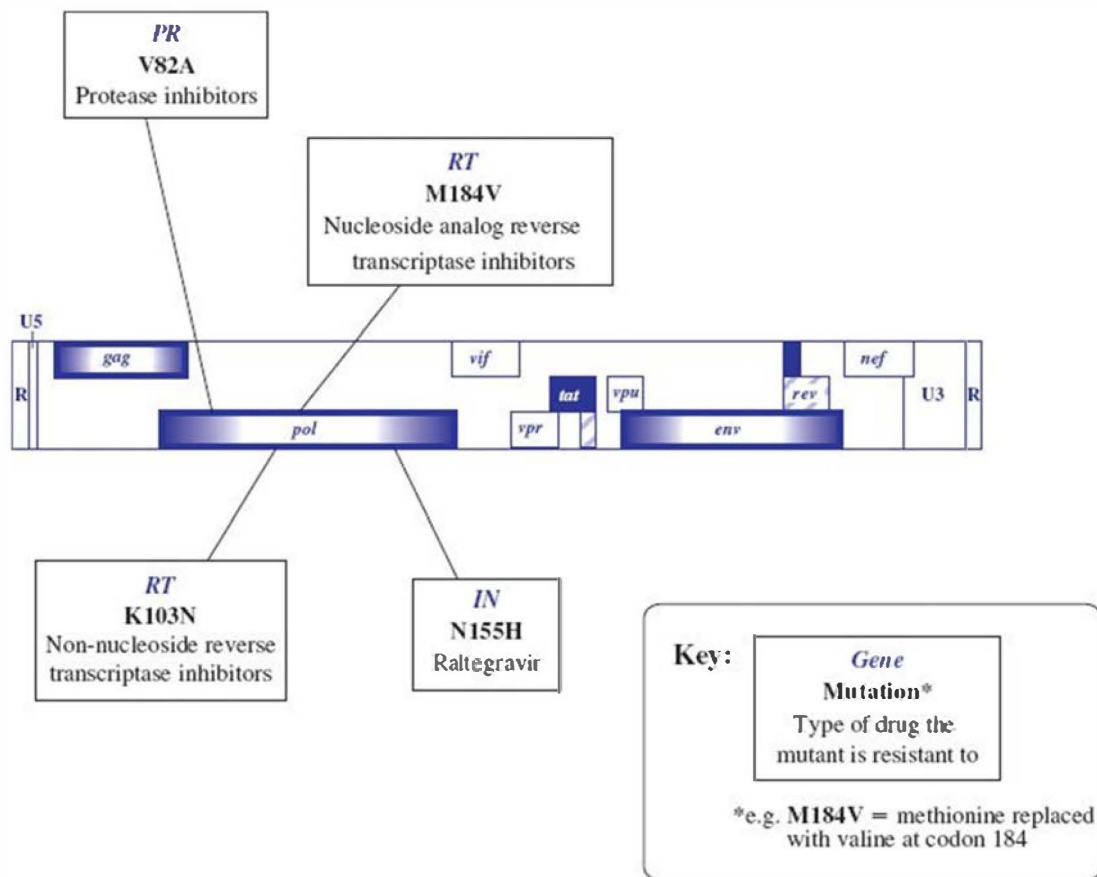


Table 26.2 Examples of genes mutated in drug-resistant virus strains.

Drug	Virus	Gene mutated
Aciclovir	HSV-1	tk (thymidine kinase) or DNA pol (DNA polymerase)
Amantadine	Influenza A	M2 (membrane 2)
Oseltamivir	Influenza A	NA (neuraminidase)
Ritonavir	HIV-1	PR (protease) or gag
AZT	HIV-1	RT (reverse transcriptase)
Nevirapine	HIV-1	RT (reverse transcriptase)
Enfuvirtide	HIV-1	env (envelope; gp41)

Most HIV-1 mutations to drug resistance are amino acid substitutions, like the examples in Figure 26.8, but some mutations are deletions or insertions. Many mutations confer cross-resistance to other drugs of the same type; for example, the V82A mutation in the PR gene confers resistance to a number of protease inhibitors. Mutations in the HIV-1 RT gene that confer resistance to a nucleoside analog (e.g. AZT) are in different codons to those that confer resistance to a non-nucleoside reverse transcriptase inhibitor (e.g. nevirapine). This reflects the different target sites in the reverse transcriptase for the two classes of drug.

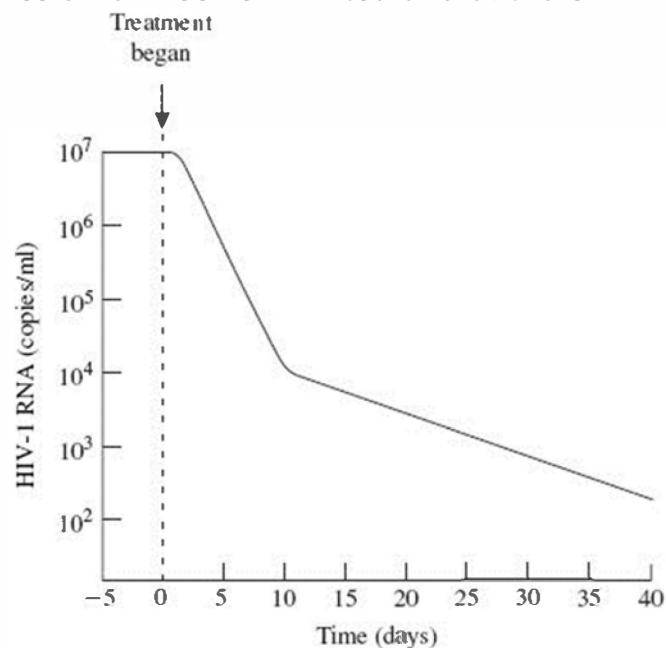
Clinical problems arise when drug-resistant virus strains emerge in patients undergoing treatment, and when resistant strains are transmitted to other individuals. When such problems arise patients may be treated with alternative drugs. Infections with aciclovir-resistant HSV strains can be treated with foscarnet or cidofovir, neither of which depends on the HSV TK for their action.

When AZT was first used for treatment of HIV infection, AZT-resistant strains of HIV rapidly emerged and a similar problem was encountered when other anti-HIV drugs were used alone. Now the standard treatment for HIV infection involves a combination of drugs. The rationale is the same as that used for chemotherapy of tuberculosis and cancer; there is a very small probability of a mutant arising that is resistant to all drugs in the combination. Various combinations of anti-HIV drugs are used, such as two reverse transcriptase inhibitors and a protease inhibitor. This form of treatment is known as highly active anti-retroviral therapy (HAART).

The effect of drug treatment on an HIV infection can be monitored by measuring the concentration of HIV RNA (and hence of virions) in the patient's blood. HAART usually results in a rapid decline in the HIV RNA concentration for about the first 10 days, and then a slower decline for a period of weeks (Figure 26.9). Over a longer period the HIV RNA concentration reaches a steady state in the range 5×10^2 – 5×10^3 copies/ml in some patients, while in others it falls to less than 5 copies/ml. HAART also reduces the levels of HIV in the seminal fluid of infected men and in the genital secretions of infected women.

Figure 26.9 Effect of HAART on the concentration of HIV-1 in the blood.

Source: Redrawn from Di Mascio et al. (2004) *Mathematical Biosciences*, 188, 47. Reproduced by permission of Elsevier Limited and authors.



HAART does not eliminate HIV infection from the body; the infection persists in latently infected macrophages and memory CD4 T cells. It may also persist in sanctuary sites, such as the brain and the testes, where infected cells may evade

the drugs and/or immune clearance. HAART has, however, markedly reduced death rates from AIDS, and treatment of HIV-infected women has significantly reduced the risk of mother-to-child transmission (Section 18.7). Those receiving HAART may experience difficulties, especially when the medication causes unpleasant side effects, but if patients lapse from their daily routine of taking the drugs they increase the risk of drug-resistant strains of HIV emerging.

26.5 ANTI-VIRAL DRUG RESEARCH

In addition to the emergence of drug-resistant strains of viruses there are other problems related to the use of anti-viral drugs. Most of the drugs in use do not eliminate virus infections from the body, and many can cause significant side effects. Both of these problems can occur, for example, during treatment of HBV infection with a nucleoside analog and/or α -interferon.

There is therefore a great need for continued research programs aimed at extending the range of drugs available. One objective of these programs is to reduce the severity of the problems discussed above; another objective is to develop drugs for virus diseases that are currently untreatable. In addition to the approaches outlined in this chapter alternative avenues are being explored, including exploitation of some of the body's anti-viral defenses, such as RNA silencing (Section 9.2.3).

LEARNING OUTCOMES

After review of this chapter you should be able to:

- evaluate procedures used to develop new anti-viral drugs;
- describe the modes of action of selected anti-viral drugs;
- evaluate anti-viral drugs in clinical use;
- discuss virus resistance to drugs.

SOURCES OF FURTHER INFORMATION

Books

Flint, S. J. et al. (2009) Chapter 9 in Principles of Virology, Volume 2, Pathogenesis and Control, 3rd edition, ASM Press

Kinchington, D. et al. (1995) Design and testing of antiviral compounds. Chapter 6 in Medical Virology: A Practical Approach (Desselberger, U., editor), IRL Press

LaFemina, R. L. (ed) (2009) Antiviral Research: Strategies in Antiviral Drug Discovery, ASM Press

Tan, S.-L. and He, Y. (eds) (2011) Hepatitis C: Antiviral Drug Discovery and Development, Caister Academic Press

Torrence, P. F. (ed) (2007) Combating the Threat of Pandemic Influenza: Drug Discovery Approaches, Wiley

Journals

Ciesek, S. and Manns, M. P. (2011) Hepatitis in 2010: the dawn of a new era in HCV therapy. *Nature Reviews Gastroenterology and Hepatology*, 8, 69-71

Colman, P. M. (2011) New antivirals and drug resistance. *Annual Review of Biochemistry*, 78, 95-118

De Clercq, E. (2009) The history of antiretrovirals: key discoveries over the past 25 years. *Reviews in Medical Virology*, 19, 287-299

Grimm, D. et al. (2011) HBV life cycle and novel drug targets. *Hepatology International*, 5, 644-653

Ison, M. G. (2011) Antivirals and resistance: influenza virus. *Current Opinion in Virology*, 1, 563-573

Kuritzkes, D. R. (2011) Drug resistance in HIV-1. *Current Opinion in Virology*, 1, 582-589

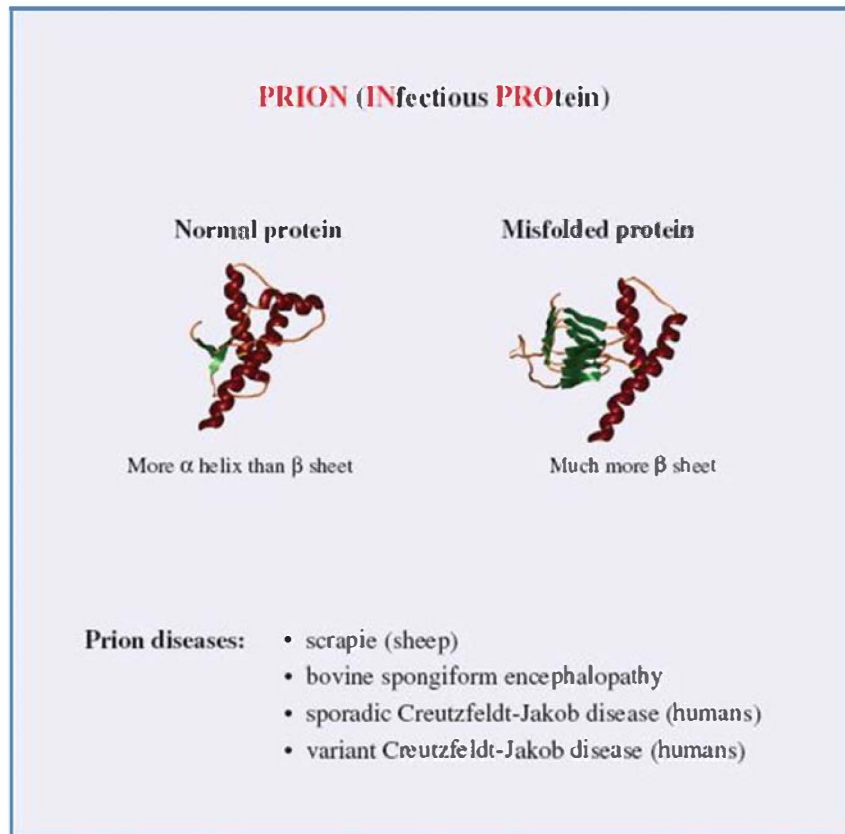
Tsibris, A. M. N. and Hirsch, M. S. (2010) Antiretroviral therapy in the clinic. Journal of Virology, 84, 5458-5464

CHAPTER 27

Prions

CHAPTER 27 AT A GLANCE

Illustrations from Cervia et al. (2006) Transfusion Medicine Reviews, 20, 190. Reproduced by permission of Elsevier.



27.1 INTRODUCTION TO PRIONS

This chapter is not about viruses, but about some infectious diseases of man and animals that are known as prion diseases. Much of the research on prions is published in virology journals; hence their appearance in this book. The evidence suggests that the causative agents are protein molecules from within the cells of the host; no nucleic acid has been found associated with them.

The prion diseases are characterized by very long incubation periods, measured in years. (The incubation period of an infectious disease is the time that elapses between the infectious agent entering the host and the first appearance of signs and symptoms of the disease.) Signs of prion disease include dementia and loss of coordination; the patient gradually deteriorates, and death is inevitable.

This subject assumed particular importance toward the end of the twentieth century with the onset in the UK of bovine spongiform encephalopathy (BSE), or

mad cow disease as it is popularly known, and its apparent transmission to humans as variant Creutzfeldt-Jakob disease (vCJD).

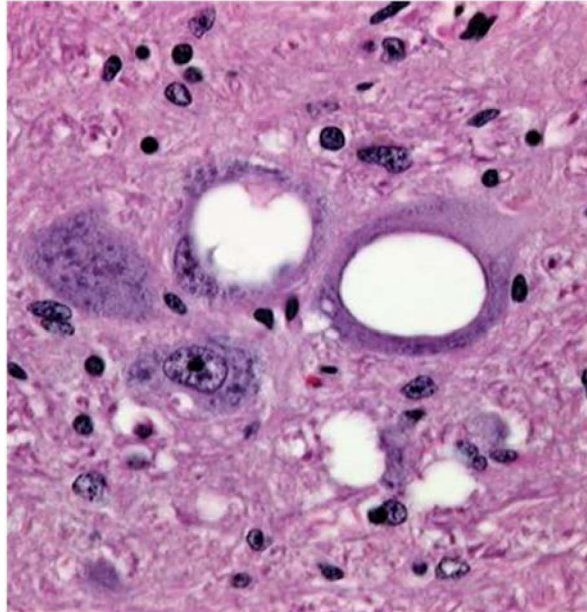
27.2 TRANSMISSIBLE SPONGIFORM ENCEPHALOPATHIES

The prion diseases are known as transmissible spongiform encephalopathies (TSEs). Taking these words in reverse order:

- encephalopathy means disease of the brain;
- spongiform refers to the development of holes in the brain, making it appear like a sponge (Figure 27.1);

Figure 27.1 Brain section from a sheep with scrapie. The spongiform appearance (holes in the tissue) is evident. Magnification Γ — 500.

Source: Courtesy of Dr R. Higgins, University of California, Davis.



- transmissible refers to the fact that the causative agent is infectious. It can be transmitted to members of the same species, and sometimes to other species.

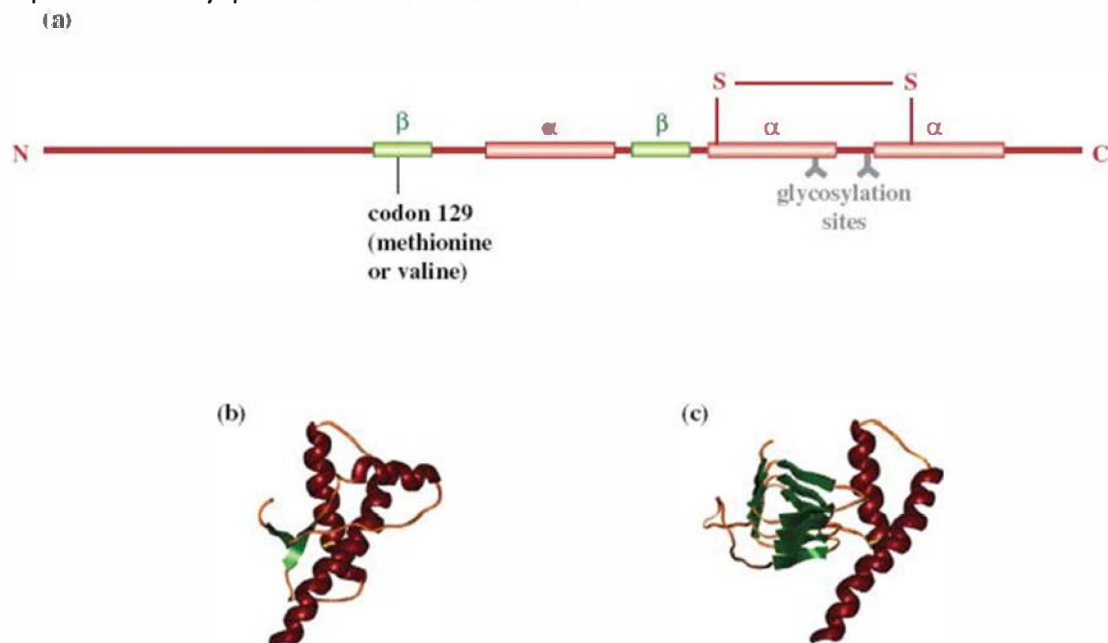
27.3 THE NATURE OF PRIONS

The agents that cause TSEs appear to be misfolded forms of normal cell proteins; there is no evidence that they contain nucleic acid. This protein-only hypothesis was proposed by Stanley Prusiner, who also suggested the term prion (pronounced pree-on), derived from infectious protein. Versions of the normal protein have been found in mammals, birds, reptiles, and fish; in humans it is encoded by the PRNP gene. The role of the protein is not yet clear. It cycles between endosomes and the cell surface, where it is held in the plasma membrane by a glycosyl-phosphatidyl-inositol anchor at its C terminus. It is found on many cell types, but especially on cells of the central nervous system.

The prion protein molecule has a loop as a result of a disulfide bond (Figure 27.2). The loop contains two asparagine residues, either or both of which can be N-glycosylated, so the protein exists as three glycoforms (unglycosylated, monoglycosylated, and diglycosylated); the sugars attached to the protein may be of various types.

Figure 27.2 Normal and misfolded forms of prion. (a) and (b) representations of the normal form; (c) the misfolded form has more β sheet (shown in green) than the normal form.

Source: (b) and (c) from Cervia et al. (2006) *Transfusion Medicine Reviews*, 20, 190. Reproduced by permission of Elsevier.



The conformation of much of the normal protein is α helix (Figure 27.2). When the protein misfolds there is a decrease in α helix structure and an increase in β sheet. This change in conformation is accompanied by changes in properties of the protein, including susceptibility to digestion by proteinase K. The normal protein is completely digested, but the misfolded protein is largely resistant to the enzyme, which is able to digest only 90%–100 amino acid residues from the N terminus, the remaining 27%–30 kD portion of the protein remaining intact. Misfolding also renders the protein insoluble in nonionic detergents. In prion-diseased tissues molecules of the misfolded protein aggregate as fibrils, rods, or other forms, depending on the host and the prion strain.

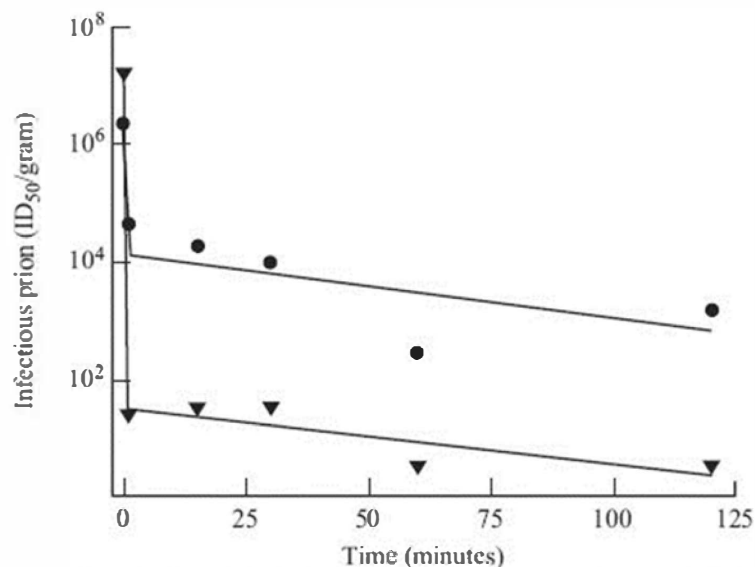
Various terminologies are used for the normal and the misfolded forms of the protein. The normal protein is commonly designated as PrP_c (PrP=prion protein; c=cell), while the misfolded form is designated as PrP_{Sc} (Sc=scrapie) or PrP_{res} (res=resistant to proteinase).

The presence of certain amino acids at particular sites in the prion protein can affect a number of aspects of prion disease, including susceptibility to prion diseases and incubation period. For example, the PRNP gene encodes either

methionine or valine at codon 129 (Figure 27.2). There can be very different outcomes depending whether the individual is homozygous (methionine/methionine or valine/valine) or heterozygous (methionine/valine). So far, all cases of vCJD have been homozygous for methionine at codon 129. Small variations in the prion gene sequence can also strongly influence the clinical course of a prion disease.

Prion infectivity is remarkably heat resistant and some infectivity can survive autoclaving for prolonged periods (Figure 27.3). It is interesting to note that a small subpopulation of the infective material has a higher level of resistance to inactivation. This subpopulation is inactivated at a slower rate, mirroring the situation with many viruses (Section 24.4). Prion infectivity is also very resistant to inactivation by irradiation and by some chemicals that inactivate virus infectivity. Treatments that are used to inactivate prion infectivity include exposure to 5% sodium hypochlorite solution or 1 M NaOH for 16–24 hours.

Figure 27.3 Inactivation of two strains of scrapie prion in an autoclave at 126 B°C. Source: Data from Somerville (2002) Trends In Biochemical Sciences, 27, 606. Reprinted by permission of Elsevier and the author.



27.3.1 Prion replication

The mechanism by which prion replication takes place is not fully understood. It has been suggested that a form of evangelism is involved, whereby molecules of the misfolded protein convert normal protein molecules, causing them to misfold. In other words, the misfolded protein causes a change in the conformation of the normal protein.

A further suggestion is that initiation of the replication process requires a seed consisting of an aggregate of misfolded protein molecules. There are three ways in which this seed could arise: it could be introduced into the body, it could be formed after a rare conformational change in a normal protein molecule, or it could be the inevitable outcome of the expression of a mutant prion gene.

The formation of an infectious pathogenic agent as a result of host protein molecules undergoing a conformational change is a very different process to virus replication, where new virions are constructed from amino acids and nucleotides. It has been difficult for many to accept the idea that an infectious agent that is able to replicate is devoid of any nucleic acid, but all attempts to demonstrate the presence of a nucleic acid have failed.

The following work, published in 2004, provides evidence that prions are infectious proteins. Mouse prion protein was produced in recombinant bacterial cells and polymerized into fibrils, which were injected into the brains of mice. The mice subsequently developed signs of a TSE and extracts of their brains transmitted the disease to other mice.

The misfolded protein accumulates in endosomes and lysosomes, and quantities build up especially in neurons. In some prion diseases the misfolded protein can also be found in a number of other organs and tissues, including the spleen and lymph nodes.

27.4 PRION DISEASES

Some cases of prion disease arise spontaneously, some are inherited and some are acquired as a result of a TSE agent entering the body. In the latter case entry may be through the digestive tract, followed by transport to the lymphoreticular system (e.g. spleen, lymph nodes, Peyer's patches), where the molecules are amplified. From these sites they can be transferred into the central nervous system. Replication takes place slowly but, because the misfolded protein is not degraded, concentrations gradually build up. The molecules form insoluble aggregates, often visible as plaques in sections of central nervous tissue. The accumulation of these aggregates is thought to lead to dysfunction and death of neurons, leading to the development of the holes in the brain (Figure 27.1). The inevitable outcome is the death of the host.

27.4.1 Prion diseases in animals

Scrapie is a disease of sheep and goats that has been known in Britain and other parts of Europe for hundreds of years. Many affected animals scrape against hard objects such as fence posts, hence the name of the disease, and many grind their teeth, stumble, and fall; all eventually die. In the 1930s it was demonstrated that scrapie can be transmitted from sheep to sheep by injection of brain tissue. It was mentioned earlier that small variations in the prion gene sequence can affect susceptibility to infection. It has been found that sheep are more susceptible to scrapie if the prion gene has valine rather than alanine at codon 136.

More recently a distinct TSE has been reported in sheep and is referred to as atypical scrapie.

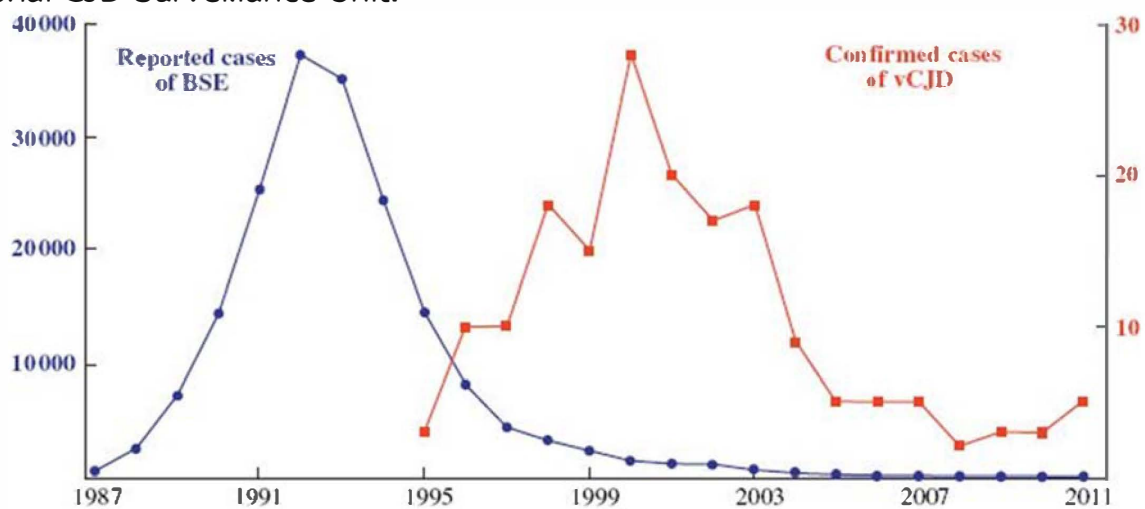
In the US a TSE in farmed mink was first recognized in 1947, then in the early 1980s chronic wasting disease was described in mule deer and elk in captivity. The

latter disease has since been found in wild mule deer, white-tailed deer, elk, and moose; it is the only TSE known to occur in free-ranging animals.

BSE was first reported in the UK in 1986 and a massive outbreak rapidly developed (Figure 27.4). The disease was spread by feeding meat and bone meal to cattle as a protein supplement. Body parts from animals incubating the disease were incorporated into meat and bone meal and this served to spread the disease on a large scale. Cattle that appeared to be healthy could in fact be incubating BSE and there could be large quantities of misfolded protein in their brains and spinal cords. It is uncertain whether the original infectious material was from a sheep with scrapie or from a cow in which BSE had arisen spontaneously.

Figure 27.4 Numbers of cases of BSE and vCJD in the UK.

Source: BSE data from World Organization for Animal Health. vCJD data from UK National CJD Surveillance Unit.



In the early years of the BSE outbreak many other species of domestic and captive mammal developed spongiform encephalopathies. These animals included domestic cats, big cats (such as puma and tiger), and herbivores (such as bison and eland). It is likely that these cases resulted from feeding animals meat from BSE-infected cattle.

Live cattle and large quantities of meat and bone meal were exported from the UK in the early days of the BSE outbreak. BSE has since been reported in many countries around the world.

27.4.2 Prion diseases in humans

Human prion diseases fall into three categories; some examples are given in Table 27.1. Sporadic CJD is the most common, occurring throughout the world at an incidence of about 1.7 cases per million people per year. Inherited prion diseases have high incidences in families whose genomes encode certain amino acids at particular codons in the PRNP gene.

Table 27.1 Examples of human prion diseases

Category of disease	Examples
---------------------	----------

Spontaneous	Sporadic CJD
Inherited	Familial CJD
	Fatal familial insomnia
Acquired	Kuru
	vCJD

Kuru and vCJD are normally acquired by ingesting prions. In the case of kuru the source of the prions was human brain tissue. This disease used to occur in areas of New Guinea where it was the custom to eat tissue from relatives who had died.

vCJD is a relatively new disease. The first case was reported in the UK in 1995 with the death of a young man called Stephen Churchill, and since then there have been some cases each year (Figure 27.4). There is no evidence that these cases represent an inherited form of prion disease, and they have characteristics that distinguish them from sporadic CJD. Amongst these characteristics are:

- the relatively young age of the victims;
- the relatively short duration of the disease;
- distinct pathological changes in the brain;
- the presence of infective prion in the spleen, lymph nodes, and tonsils.

These features indicated that vCJD was a new disease, and because its appearance coincided with the BSE outbreak in the UK it was suggested that it had resulted from humans ingesting the BSE prion. Further evidence that the BSE prion is the causative agent of vCJD came from laboratory studies, which demonstrated that the BSE prion and the vCJD prion have identical molecular characteristics, and that they induce identical responses in inoculated mice. A few cases of vCJD appear to have resulted from transmission of the agent via blood transfusion, rather than through ingestion.

There have been cases of vCJD in countries other than the UK. Some of these victims had lived in the UK, so may have acquired the disease before moving abroad. Others had never visited the UK, but they could have become infected from meat exported from the UK or from BSE in local cattle.

27.5 PRION STRAINS

For some TSE agents a number of strains can be differentiated, each having phenotypic characters that are consistently transmitted from generation to generation: for example, different strains of the scrapie agent can induce different clinical syndromes in goats. Differences may also be observed in laboratory studies. Wild-type mice can easily be infected by inoculating them with vCJD prions, but these animals are highly resistant to infection with sporadic CJD prions. Prion strains may be distinguished by reproducible properties such as incubation period in

mice after injection into the brain, and degree of heat resistance. Figure 27.3 illustrates two strains of scrapie agent that differ in heat resistance.

Virus strains are defined by differences in the sequences of their nucleic acids, but TSE agents do not contain nucleic acids. The different strains of TSE agent from one host are proteins with an identical sequence of amino acids, so how can prion strains with different phenotypic characters be explained? One proposal is that the prion protein can misfold into different conformations, one for each strain. Alternatively, prion strain differences may result from the extent to which the molecules are glycosylated at the two glycosylation sites; in other words, a particular strain may have a particular ratio of the three glycoforms (Section 27.3).

27.6 PRION TRANSMISSION

Infective prions have been found in saliva, nasal secretions, milk, urine, feces, and the placenta of infected hosts. These are potential sources of infection for other hosts, which might become infected through direct contact with infected hosts, or indirectly through acquiring prions shed into the environment.

As we have already discussed, prions can be transmitted to other members of the same species and often to other species. The efficiency of transmission depends on the route by which the prion molecules enter the body; injection into the brain is much more efficient than ingestion. Dose is also important; the larger the dose, the more likely it is that transmission will occur.

Many attempts to transmit TSE agents to distantly related species have been reported as difficult. When transmission does occur the minimum infective dose is commonly larger than for the same or closely related species, and incubation periods are commonly longer. This effect has been called **the species barrier**. Most tests for cross-species transmission have used the development of prion disease in recipient animals as the criterion for successful transmission. There is, however, evidence for subclinical infections, where prions replicate without development of disease during the lifespan of the infected animal. It is therefore possible that the strength of **the species barrier** has been overestimated.

If a TSE agent is transmitted to a different species, and then passaged in that species, progressive changes take place; the incubation period becomes shorter over a number of passages and then stabilizes.

The transmission of BSE into humans as vCJD involved a **species barrier** being crossed. As the BSE outbreak progressed, the UK government introduced increasingly strict measures to reduce the risk of the BSE prion being present in food. Animal materials likely to constitute a risk were banned; these include brain, spinal cord, and mechanically recovered meat from the spine. Governments introduced additional measures, including the testing of cattle and meat for the BSE prion using serological methods.

The possibility of human-to-human transmission is also a matter for serious concern, especially since there have been several cases of vCJD transmission in blood transfusions. Measures are taken to minimize the risk of prion transmission via donated blood and blood products, and via surgical instruments.

The cases of TSE in cats and other species, coincident with the BSE outbreak, are thought to have been transmitted in meat containing the BSE agent (Section 27.4.1). A number of additional species have been infected with the BSE agent under experimental conditions, thus the BSE agent appears to have a wide host range.

27.7 THE PROTEIN-ONLY HYPOTHESIS

If the protein-only hypothesis is correct, and the infectious agent is misfolded prion protein, then transmission of a prion disease implies that the introduction of the misfolded protein into the body of a new host initiates the misfolding of protein molecules in that host. If the agent is transmitted to other species then, because each species makes a specific prion protein, the molecules that become misfolded have the amino acid sequence of the recipient prion, not that of the donor.

The protein-only hypothesis suggests that the agent is derived from one of the body's own proteins, and this can explain the inherited forms of TSEs (Table 27.1). According to the hypothesis, these diseases develop as a consequence of inheriting a prion gene encoding a protein with an amino acid sequence that has a high probability of misfolding.

There remain many controversies and there is still much to be learned about the TSEs and their causative agents. Progress in research has been slow as a result of the long incubation periods of the diseases. There are now available cell lines that can support prion replication and these are being used in studies at the cellular and molecular levels. Furthermore, prion-like molecules have been found in yeasts and in filamentous fungi, and these are proving to be useful models for mammalian prions.

LEARNING OUTCOMES

After review of this chapter you should be able to:

- define the terms prion and transmissible spongiform encephalopathy;
- evaluate the protein-only hypothesis;
- describe the characteristics of prions;
- discuss a theory of prion replication;
- describe prion diseases in animals and man;
- discuss the transmission of prions.

SOURCES OF FURTHER INFORMATION

Books

Brown, P. and Detwiler, L. A. (2007) Bovine spongiform encephalopathy: consequences for human health. Chapter 28 in *Food Microbiology: Fundamentals and Frontiers*, 3rd edition (Doyle, M. P. and Beuchat, L. R., editors), ASM Press

Geschwind, M. D. and Legname, G. (2008) Transmissible spongiform encephalopathies. In *Protein Misfolding in Neurodegenerative Diseases* (Smith, H. J., Simons, C., and Sewell, R. D. E., editors), CRC Press

Hilf, M., Ruesch, D., and Kretzschmar, H. (2007) *Prions in Humans and Animals*, De Gruyter

Ingrusso, L., Pocchiari, M., and Cardone, F. (2007) Bioanalytical diagnostic test for measuring prions. In *Immunoassay and Other Bioanalytical Techniques* (van Emon, J. M., editor), CRC Press

Tatzelt, J. (ed.) (2011) *Prion Proteins*, Springer

Journals

Aguzzi, A., Sigurdson, C., and Heikenwaelder, M. (2008) Molecular mechanisms of prion pathogenesis. *Annual Review of Pathology: Mechanisms of Disease*, 3, 11-40

Andreatti, O. et al. (2011) Atypical/Nor98 scrapie infectivity in sheep peripheral tissues. *PLoS Pathogens*, 7 (2): e1001285

Colby, D. W. and Prusiner, S. B. (2011) De novo generation of prion strains. *Nature Reviews Microbiology*, 9, 771-777

Collinge, J. and Clarke, A. R. (2007) A general model of prion strains and their pathogenicity. *Science*, 318, 930-936

Imran, M. and Mahmood, S. (2011) An overview of human prion diseases. *Virology Journal*, 8: 559

Maddison, B. C. et al. (2010) Environmental sources of scrapie prions. *Journal of Virology*, 84, 11560-11562

Sandberg, M. K. et al. (2011) Prion propagation and toxicity in vivo occur in two distinct mechanistic phases. *Nature*, 470, 540-542

Savistchenko, J. et al. (2011) Mammalian prions: tracking the infectious entities. *Prion*, 5, 84-87

Virologists' Vocabulary

Brief definitions of terms used in virology, plus a selection of relevant terms from cell biology, molecular biology, immunology and medicine.

abortive infection	Infection of a cell where the virus replication cycle is not completed and no progeny virions are formed.
ambisense genome	Virus genome composed of ssRNA or ssDNA that is partly plus sense and partly minus sense.
antibody	Glycoprotein synthesized in a plasma cell, which is derived from a B cell that has interacted with a specific antigen. The antibody molecule can bind specifically to this antigen.
antigen	Molecule that: (a) triggers synthesis of antibody and/or a T cell response; (b) binds specifically to an antibody or a lymphocyte receptor.
antiserum (plural antisera)	Blood serum from an animal that has been injected with an antigen. The serum contains antibodies specific to that antigen.
anti-termination	Mechanism involved in the control of transcription in which termination is overcome at specific terminator sites. Anti-terminator proteins allow RNA polymerase to read through these sites into genes downstream.
apoptosis	Programmed cell death. Cell suicide. A process controlled by a cell that results in the death of the cell.
archaea	One of two major groups of prokaryotes, the other being the bacteria.

assembly	Stage in the virus replication cycle when components come together to form virions.
asymptomatic infection	Infection without symptoms or signs of disease.
attachment	Binding of a virion to specific receptors on the host cell.
attenuated strain of a virus	Strain of a pathogenic virus that has much reduced virulence, but can still replicate in its host. The virulence is said to have been attenuated.
avirulent strain of a virus	Strain of a virus that lacks virulence, but can still replicate in its host.
B lymphocyte (B cell)	Cell with surface receptors that can recognize a specific antigen. Antigen binding can trigger a B cell to develop into an antibody-secreting plasma cell.
bacteriophage (phage)	Type of virus that infects bacteria.
Baltimore classification	Scheme that classifies viruses into seven groups on the basis of the nature of the genome and the way in which it is transcribed.
bicistronic mRNA	mRNA with two open reading frames.
bronchiolitis	Inflammation of the bronchioles.
burst size	Average number of new virions released from a single infected cell.
cap	Methylated guanosine triphosphate joined to the 5' end of an mRNA by a 5'-5' bond.
capsid	The protein coat that encloses the nucleic acid of a virus.
capsomere	Discrete component of a capsid, constructed from several identical protein molecules.

CD (cluster of differentiation)	CD antigens are cell surface molecules recognized by monoclonal antibodies. CD4 is the receptor for HIV.
chemokine	Cytokine that stimulates the migration and activation of cells in the animal body, especially cells involved in inflammation.
cirrhosis	Liver disease characterized by scarring and loss of function.
cohesive end (cos)	Single-stranded end of a dsDNA molecule with a base sequence complementary to another single-stranded end of the same or a different DNA molecule, allowing hydrogen bonding between the two ends.
coliphage	Bacteriophage that infects E. coli.
complement	Series of proteins that is activated when the body is infected. Activation has a number of anti-viral effects, including lysis of infected cells and enhancement of phagocytosis.
concatemer	Very long DNA molecule composed of multiple repeats of a nucleotide sequence.
conditional lethal mutant	Mutant virus that expresses wild-type characteristics under certain (permissive) conditions, but mutant (lethal) characteristics under other (non-permissive/restrictive) conditions, e.g. a temperature-sensitive mutant.
confocal microscopy	Use of a microscope that excludes light from out-of-focus regions of the specimen, producing clear images of thick and fluorescently labeled specimens.
conservative replication	The replication mode of some dsRNAs. The parental double strand is conserved and both strands of the progeny molecule are newly synthesized.

continuous cell line	Clone of cells that is immortal and can be subcultured indefinitely.
co-receptor	Second cell receptor to which some viruses must bind in order to infect a cell.
cryo-electron microscopy	Examination of frozen specimens in an electron microscope with a cold stage.
cytokine	Protein that is secreted from a cell and has a specific effect on other cells, including cells of the immune system.
cytopathic effect (CPE)	Change in the appearance of a cell culture induced by virus infection.
defective interfering particle	Virus that lacks part of its genome and interferes with the replication of standard virus.
defective virus	Virus that lacks part of its genome and is unable to complete the replication cycle without the aid of a helper virus.
diploid	Having two copies of the genome.
DNA-dependent RNA polymerase	Enzyme that synthesizes RNA from a DNA template.
domain	Discrete portion of a protein or a nucleic acid with its own structure and function.
elongation factor (EF)	Specific protein required for polypeptide elongation during translation.
encephalitis	Inflammation of the brain.
endemic	Describes a disease that is constantly present or commonly present in a geographical area.
endocytosis	Process whereby eukaryotic cells take in extracellular materials by engulfing them in endosomes.
endogenous retroviruses	Retrovirus sequences present in the genomes of vertebrate animals.

endosome	Vesicle formed during endocytosis by membrane pinching off from the plasma membrane.
enhancer	Short DNA sequence to which a transcription factor can bind, resulting in enhanced transcription of a gene.
envelope	Lipid bilayer and associated protein forming the outer component of an enveloped virion.
enzyme-linked immunosorbent assay (ELISA)	Serological method used to assay antigens and antibodies. Positive results detect an enzyme label.
epidemic	Rapid increase in the number of cases of a disease that spreads over a larger geographical area.
epitope	The part of an antigen that binds to an antibody or a lymphocyte receptor.
Escherichia coli (E. coli)	Gram-negative bacterium that inhabits the human colon.
exogenous retrovirus	Normal infectious retrovirus, cf. endogenous retroviruses.
exon	Part of a eukaryotic gene that is transcribed and usually translated into protein.
exonuclease	Enzyme that digests a DNA or RNA strand one nucleotide at a time from either the 5' end or the 3' end.
F pilus	Specialized pilus, encoded by the F plasmid, required for conjugation in E. coli.
fluorescence microscopy	Use of a microscope fitted with an appropriate source of light, e.g. an arc lamp or a laser, to examine fluorescent specimens.

fusion protein	(a) Virus protein that fuses the membrane of an enveloped virus with a cell membrane. (b) Protein generated by joining two genes together e.g. a foreign protein fused to a coat protein in phage display.
genome	The DNA or RNA that encodes the genes of an organism or a virus.
genotype	Complete set of genes of an organism or a virus.
glycoprotein	Protein with one or more oligosaccharide groups covalently attached.
glycosylation	Process of adding oligosaccharide groups to proteins.
Golgi complex	Membranous organelle found in most eukaryotic cells. Its primary function is to process and sort proteins.
green fluorescent protein (GFP)	Jellyfish protein that fluoresces green. The GFP gene can be fused to a gene of interest, producing a "tagged" fusion protein, and the localization or movement of the protein in living cells can be visualized. The GFP gene can also be used to monitor the expression of a gene of interest in living cells; the control sequences for the gene of interest are linked to the GFP gene.
guanylyl transferase	Enzyme that adds guanosine 5'-monophosphate to the 5' end of mRNA when it is capped.
hairpin	Structure formed by hydrogen bonding within a single-stranded nucleic acid molecule, producing a double-stranded stem and a loop of unpaired nucleotides.
helical symmetry	Type of capsid symmetry present in many ssRNA viruses where the RNA forms a helix that is coated with protein.

helicase	Enzyme that unwinds a DNA duplex at replication forks.
helper virus	Virus that can provide function(s) missing from a defective virus, thereby enabling the latter to complete its replication cycle.
hepatocyte	Main cell type in the liver.
heteroploid cell	Cell with an abnormal number of chromosomes.
hexon	Type of capsomere surrounded by six other capsomeres.
histones	Proteins comprised mainly of basic amino acids, associated with DNA in the cells of eukaryotes and some archaea. Histones play roles in gene regulation.
horizontal gene transfer	Transfer of genetic material laterally from one organism to another, mediated in bacteria by conjugation, transduction, and transformation.
host	Cell or organism in which a virus or a plasmid can replicate.
host-controlled restriction	Mechanism by which bacteria can cleave unmodified foreign DNA by a restriction endonuclease that recognizes a specific nucleotide sequence.
icosahedral symmetry	Type of symmetry present in viruses where the capsid is constructed from protein molecules arranged to form 20 triangular faces.
immunofluorescence	Detection of an antigen using an antibody labeled with a fluorescent dye.
immunogenic	Ability of a substance to induce an immune response.
immunoglobulin	Glycoprotein that functions as an antibody.

incubation period	Time between infection of a host and the appearance of the first signs and/or symptoms of disease.
induction	Activation of a latent infection or an inactive gene.
infectivity	Ability of virions to initiate an infection.
integral membrane protein	Protein that is closely associated with a membrane. Most integral membrane proteins have one or more sequences that span the lipid bilayer.
integrase	Enzyme that integrates virus DNA into host DNA.
interferon	Protein produced by animal cells in response to virus infection. There are several types of interferon and they interfere with virus replication in a variety of ways.
intergenic (IG) region	Sequence of nucleotides between two genes.
internal ribosome entry site (IRES)	Binding site for eukaryotic ribosomes, present in the RNAs of some plus-strand viruses. The site is internal (near the 5' end of the RNA), in contrast to most eukaryotic mRNAs, which bind ribosomes at the 5' end.
intron	Non-coding sequence interrupting coding sequences (exons) in a gene. The intron is transcribed into RNA, but is subsequently excised by a splicing reaction during processing of the primary transcript into mRNA. Introns are common in eukaryotes and their viruses, but rare in prokaryotes and their viruses.
isometric virion	Virion that has the same size from all perspectives. Virions that have icosahedral symmetry are isometric or almost isometric.

kilobase (kb)	Measure of length for DNA and RNA equal to 1000 bases.
kinase	Enzyme that catalyses phosphorylation.
laboratory strain	Virus strain derived from a wild strain after propagation in the laboratory.
latent infection	Infection of a cell where the replication cycle is not completed, but the virus genome is maintained in the cell.
ligase	Enzyme that can catalyze bond formation between two similar types of molecule. A DNA ligase can join two DNA molecules by catalyzing the formation of a phosphodiester bond between the 5' end of a polynucleotide chain and the 3' end of another (or the same) polynucleotide chain.
ligation	Joining of two molecules, e.g. DNA.
lipopolysaccharide (LPS)	Compound comprising lipid and polysaccharide that is a key component of the outer membrane of Gram-negative bacteria.
lysis	Destruction of a cell caused by the rupture of its membrane and release of contents, e.g. the bursting of a cell and release of progeny virions at the end of the replication cycle.
lysogen	Bacterium containing the genome of a phage that is in the prophage state and thus repressed for lytic functions. Such a bacterium is said to be lysogenic for [name of the phage].
lysogenic conversion (phage conversion)	When a cell becomes lysogenized, occasionally extra genes (unrelated to the lytic cycle or lysogeny) carried by the prophage are expressed in the host cell. These genes can change the phenotype of the host.

lysogeny	Latent infection of a bacterial cell with a phage.
lysosome	Membrane-bounded organelle in the cytoplasm of eukaryotic cells. Lysosomes contain enzymes, e.g. proteases and nucleases, that can digest virions.
major histocompatibility complex (MHC)	Region of the vertebrate genome that encodes major histocompatibility proteins. MHC class I and class II molecules are cell surface proteins that play important roles in immune responses.
meningitis	Inflammation of the meninges (membranes covering the brain and the spinal cord).
meningoencephalitis	Inflammation of the meninges and the brain.
methyl transferase	Enzyme that catalyses the transfer of methyl groups from one molecule to another. Methyl transferases play a role in capping eukaryotic mRNA.
microarray	Substrate with an orderly array of small spots of material (e.g. DNA or protein) attached. Each spot contains a specific probe that can detect specific target molecules in a sample.
microtubule	Protein structure in eukaryotic cells, forming a component of the cytoskeleton, with roles in intracellular transport and mitosis.
minus-strand (negative strand, (-) strand)	Nucleic acid strand that has the nucleotide sequence complementary to that of the mRNA.
monocistronic mRNA	mRNA with one open reading frame.
monoclonal antibody	Single type of antibody produced by a clone of identical cells.

monolayer	Layer of cells growing on the surface of a plastic or glass vessel.
multiplicity of infection (m.o.i.)	Ratio of virions to cells in a virion–cell mixture.
mutation	Alteration in the sequence of DNA or RNA. Mutations may occur spontaneously or they may be induced.
myristylation (myristoylation)	Addition of a myristyl (myristoyl) group to a molecule.
naked virus	Virus that does not have an envelope on the virion.
negative staining	Use of a compound containing a heavy metal to reveal the structure of specimens in a transmission electron microscope.
neuron	Main type of cell in the nervous system.
neurovirulence	Measure of the severity of nervous system disease that a virus (or any micro-organism) is capable of causing.
neutralization	Inactivation of virus infectivity by reaction with specific antibody.
non-structural protein	Virus protein that is not a component of the virion, but has one or more roles in the replication cycle.
northern blotting	Transfer of RNA molecules to a membrane after gel electrophoresis. Specific RNAs can be detected on the membrane using probes.
nuclear envelope	Structure, composed of two membranes, that separates the nucleus from the cytoplasm in a eukaryotic cell.
nuclear localization signal	Positively charged sequence in a protein that directs the protein through nuclear pores into the nucleus of a cell.

nuclear pore	Protein complex in the nuclear envelope of a eukaryotic cell through which materials are transported in and out of the nucleus.
nucleic acid hybridization	Formation of double-stranded nucleic acids by base-pairing between complementary single strands.
nucleocapsid	Virus genome enclosed in a protein capsid.
nucleoside	Nitrogen base (purine or pyrimidine) joined to a sugar (ribose or deoxyribose).
nucleotide	Nitrogen base (purine or pyrimidine) joined to a sugar, which is joined to one or more phosphate groups.
Okazaki fragment	Short fragment of DNA with the RNA primer attached, produced during discontinuous DNA synthesis. Okazaki fragments are joined by a DNA ligase to form the lagging strand.
oncogene	Cell or virus gene, the expression of which can cause a cell to become transformed. This may lead to the development of a cancer.
open reading frame (ORF)	Sequence of nucleotides, starting with an initiation codon and finishing with a termination codon, that encodes the amino acids of a protein.
origin (ori)	Specific sequence of DNA at which replication is initiated.
packaging signal (PS)	Nucleotide sequence in a virus genome that is recognized by a virus protein during virion assembly.
pandemic	Disease outbreak throughout the world.
pararetrovirus	Virus with a DNA genome that is replicated via RNA.

passage	"Subculture" of a virus or a prion in cell culture or in an organism.
penton	Type of capsomere surrounded by five other capsomeres.
perforin	Protein present in CD8 T cells and NK cells. These cells can kill a virus-infected cell by releasing perforins, which form pores in the plasma membrane of the target cell.
persistent infection	Long-term, possibly lifelong, infection of a host. A persistent virus infection may be productive or latent.
phage display	Display of recombinant proteins on the surface of a phage, e.g. M13.
phagemid	Cloning vector that combines features of a phage and plasmid replicon and can replicate in either mode. The vector carries a plasmid ori for replication in plasmid mode, and a filamentous phage ori, leading to production of single-stranded copies of the phagemid when the cell is infected with the relevant helper phage.
phenotype	Observable characteristics of an organism or virus, determined by its genotype and environmental factors.
phosphoprotein	Protein with one or more phosphate groups covalently attached.
phosphorylation	Addition of one or more phosphate groups to a molecule.
phylogenetic tree	Diagram showing evolutionary relationships between viruses or other organisms.
pilin	Protein subunit that is polymerized into the pilus in bacteria.

plaque	(a) Area of lysis in a layer of cells, usually initiated by a single virion infecting a cell, followed by the spread of infection to surrounding cells. (b) Aggregate of prion fibrils in the central nervous system of a human or animal with a transmissible spongiform encephalopathy.
plaque-forming unit	Quantity of virus (often a single virion) that can initiate formation of one plaque.
plaque-purified virus	Clone of virus derived from an individual plaque.
plasma membrane	Membrane at the surface of a eukaryotic cell.
plasmid	Self-replicating, extrachromosomal dsDNA molecule that is generally circular, though can be linear. Plasmids are common in prokaryotes and rare in eukaryotes. Artificial plasmids can be created to clone DNA sequences.
plus-strand (positive strand, (+) strand)	Nucleic acid strand that has the same sequence as the mRNA.
poly(A) polymerase	Enzyme that adds adenylate residues to the 3' end of a eukaryotic mRNA.
polyadenylation	Process of adding adenylate residues to the 3' end of a eukaryotic mRNA.
polycistronic mRNA	mRNA with more than one open reading frame.
polymerase chain reaction (PCR)	In vitro technique for amplifying specific DNA sequences.
polyprotein	Large protein that is cleaved to form smaller functional proteins.
post-translational modification	Modification to a protein after it has been translated. Examples: glycosylation, cleavage.

primary transcript	RNA molecule synthesized during transcription. Introns may be removed from the primary transcript to form functional RNAs.
primase	Enzyme that synthesizes RNA primers for DNA synthesis.
primer	Molecule (often RNA, sometimes DNA or protein) that provides a free —OH group at which a polymerase can initiate synthesis of DNA or RNA.
primosome	Complex of proteins involved in the priming action that initiates DNA replication on ϕ X-type origins.
probe	Specific sequence of DNA or RNA used to detect nucleic acids by hybridization. A probe may be labeled, e.g. with a fluorescent molecule, to enable detection and quantification of the hybrid.
procapsid	Precursor of a capsid.
processivity factor	Protein that increases the efficiency of DNA synthesis by enhancing the ability of a DNA polymerase to remain associated with, and to process along, its template.
productive infection	Infection of a cell that results in the production of progeny virions.
promoter	Specific DNA sequence, usually upstream of the transcription start point of a gene, to which DNA-dependent RNA polymerase binds to initiate transcription.
proofreading	Mechanism for correcting errors in nucleic acid or protein synthesis.
prophage	Phage genome that resides in a bacterial host in a latent state. The prophages of most phages are integrated into the bacterial genome, but some are not.

protease	Enzyme that cleaves protein molecules by breaking peptide bonds.
proto-oncogene	Cell gene that can become an oncogene (tumor gene) if activated or over-expressed.
provirus	dsDNA copy of a retrovirus genome RNA.
pseudoknot	ssRNA secondary structure with two loops, formed when a sequence in a loop base-pairs with a complementary sequence outside the loop.
reading frame	Phase in which the nucleotides of a nucleic acid are read in triplets.
reassortment	Category of recombination that may occur with those segmented genome viruses that have all the segments packaged in one virion. Reassortment occurs in a cell co-infected with two virus strains, and is the formation of progeny virions containing mixtures of genome segments from the two parental strains.
receptor	Molecule on a cell surface to which a virus specifically attaches.
recombinant	Organism or virus with a new genome, produced as a result of recombination. May also be used to describe the genome of the organism or virus.
recombination	Process of combining genetic sequences that results in the production of a new genome, derived from two parental genomes. For a cellular organism, recombination can take place between DNA molecules within a cell. For viruses, recombination may occur when a cell is co-infected with two related viruses.
replicase	RNA polymerase that replicates the genome of an RNA virus by synthesizing both (+) RNA and (-) RNA.

replication cycle	Process whereby a virus is replicated. For most viruses the process starts with attachment to a host cell and this is followed by entry into the cell. The process ends with exit of progeny virions from the cell. Some authors use the terms "infection cycle" and "life cycle."
replicative intermediate	Structure formed during ssRNA replication, consisting of a template RNA associated with nascent RNAs of varying length. The nascent RNAs have the opposite polarity to the template RNA.
restriction endonuclease	Enzyme, usually from a bacterium, that cuts dsDNA at a specific site.
retrovirus	Member of the family Retroviridae, so named because these viruses carry out reverse transcription.
reverse genetics	Generation of an RNA virus genome from a cloned copy DNA.
reverse transcriptase	Enzyme that can synthesize DNA using (a) an RNA template and (b) a DNA template.
reverse transcription	Synthesis of DNA from an RNA template.
reverse transcription–polymerase chain reaction (RT–PCR)	In vitro technique for amplifying the data in RNA sequences by first copying the RNA to DNA using a reverse transcriptase. The DNA is then amplified by a PCR.
ribonuclease (RNase)	Enzyme that hydrolyses RNA.
ribonuclease H (RNase H)	Ribonuclease that specifically digests the RNA in an RNA–DNA duplex.
ribonucleoprotein (RNP)	RNA with bound protein molecules.

ribosomal frameshifting	Mechanism that allows a ribosome to read two overlapping open reading frames (ORF1 and ORF2) in an mRNA. A ribosome reading ORF 1 shifts into a different reading frame towards the end of the ORF. The ORF 1 stop codon is therefore not recognized and the ribosome now reads ORF 2 to produce an elongated version of the ORF 1 protein.
RNA-dependent RNA polymerase	Enzyme that synthesizes RNA from an RNA template.
RNA interference (RNAi, RNA silencing)	Process that interferes with the expression of a specific gene. The process is induced by a dsRNA and results in the destruction of mRNA that has the same sequence as the dsRNA.
RNA polymerase	Enzyme that synthesizes RNA.
RNA polymerase II	Eukaryotic cell enzyme that synthesizes mRNA.
rough endoplasmic reticulum	System of membranes and ribosomes in the cytoplasm of eukaryotic cells.
satellite virus	Defective virus that depends on a helper virus to provide one or more functions.
scaffolding protein	Protein that facilitates the assembly of a procapsid. The scaffolding protein is removed in the final stages of assembly and is therefore not present in the mature virion.
segmented genome	Virus genome that is composed of two or more nucleic acid molecules.
self-assembly	Ability of a biological structure, e.g. a virus particle, to form spontaneously from its component parts.

semi-conservative replication	Replication mode of dsDNA and some dsRNAs in which the parental double strand is not conserved. Each progeny molecule consists of one parental strand and one newly synthesized strand.
serology	Study of antigen–antibody reactions and their use in tests to detect specific antigens and antibodies.
seropositive	A person or an animal is said to be seropositive for an antigen, e.g. a virus, if their blood contains antibodies specific for that antigen.
serotype	Strain of a virus or micro-organism distinguished by its antigens.
sex pilus	Thin protein appendage for bacterial mating.
Shine–Dalgarno (SD) sequence	Purine-rich sequence in prokaryotic mRNA just upstream of the translation start. The sequence can base-pair with a sequence near the 3' end of 16S ribosomal RNA and facilitates the initiation of protein synthesis.
sigma (σ) factor	Subunit of bacterial RNA polymerase, responsible for the recognition of a specific promoter. Different sigma factors allow recognition of different promoter sequences.
single-stranded binding (ssb) protein	Basic protein with a high affinity for ssDNA. Ssb protein protects ssDNA from nuclease attack and prevents it from re-annealing into dsDNA.
Southern blotting	Transfer of DNA molecules to a membrane after gel electrophoresis. Specific DNA can be detected on the membrane using probes.

splicing	Process of removing introns from primary transcripts and joining the exons to form mRNA.
start (initiation) codon	Codon on mRNA for initiation of protein synthesis: commonly AUG, less commonly GUG, and rarely UUG.
stop (termination) codon	Codon on mRNA for termination of protein synthesis: UAG, UAA or UGA. When a ribosome encounters a stop codon a termination factor interacts with the ribosome, causing polypeptide synthesis to stop and the ribosome to dissociate from the mRNA.
structural protein	Protein that is a virion component.
superinfection	Infection by a virus of a cell that is already infected (often as a latent infection) with that virus or a related virus.
superinfection immunity (homoimmunity)	Immunity of a lysogen to superinfection by a phage with a similar regulatory mechanism.
synchronous infection	Near simultaneous infection of all cells in a culture with a virus. This can be achieved by, for example, using a high multiplicity of infection, or by limiting the time of attachment then diluting the culture so that virions and cells are unlikely to make contact.
syncytium (plural syncytia)	Multinucleated giant cell formed by the fusion of membranes of a number of individual cells. A number of viruses can cause the formation of syncytia when they infect cells.

T lymphocyte (T cell)	Cell with surface receptors that can recognize a specific antigen. Antigen binding can trigger a T cell to perform one of several roles, including helper T cell or cytotoxic T cell.
tailed phage	Phage that has a tail attached to the head, which contains the genome.
TATA box	DNA sequence in a eukaryotic promoter, typically 15–25 nucleotides upstream from the transcription start.
TCID50	Dose of virus that infects 50% of tissue cultures inoculated with aliquots of a virus preparation.
tegument	Layer of protein and RNA between the capsid and the envelope of a herpesvirus particle.
temperate phage	Phage capable of either establishing a lysogenic state in a susceptible bacterial host or of entering the lytic cycle.
temperature-sensitive mutant	Virus mutant that is unable to replicate at some temperatures (non-permissive temperatures) at which the wild-type virus can replicate, but is able to replicate at other temperatures (permissive temperatures).
titer	Concentration of virus, antibody, or other material in a preparation, determined by titration.
transcriptase	Enzyme that carries out transcription.
transcription	Synthesis of mRNA from a DNA or an RNA template.
transcription factor	Protein that binds specifically to a promoter or enhancer to control gene expression.
transfection	Process for introducing nucleic acids (e.g. a virus genome) into cells.

transformation	(a) Changes in an animal cell that result in it developing into a tumor cell. (b) Bacterial gene transfer process involving uptake of naked DNA and acquisition of an altered genotype.
translation	Synthesis of protein from the genetic information in mRNA.
transmission electron microscope	Microscope in which the image is formed by electrons transmitted through the specimen.
transovarial transmission	Transmission of a virus from a female, e.g. insect, to the next generation within eggs.
uncoating	Release of the virus genome (and associated protein in some cases) from a virion when a cell is infected.
upstream	Region of a nucleic acid extending in the 5' direction from a gene; in the opposite direction to that of transcription.
vector	(a) Organism that transmits a virus from an infected host to an uninfected host. (b) Virus or plasmid DNA, into which foreign DNA can be inserted for the purpose of transferring it to a host cell for amplification/cloning or gene therapy.
vertical transmission	Transmission of a genetic element, virus or micro-organism to the next generation of the host.
viremia	Presence of infectious virus in the bloodstream.
virion	Virus particle.
viroplasm	Morphologically distinct region in which virus replication occurs within an infected cell.

virulence	Measure of the severity of disease that a virus (or micro-organism) is capable of causing.
virulent strain of a virus	Virus strain that can cause disease when it infects a host.
virulent phage	Phage that can produce only a lytic infection; it is unable to induce lysogeny.
virus attachment site	Surface regions of one or more virion proteins, which form a site that attaches to the receptor on a host cell.
western blotting	Transfer of protein molecules to a membrane after gel electrophoresis. Specific proteins can be detected on the membrane.
wild strain	Virus strain isolated from a naturally infected host, cf. laboratory strain.
wild-type virus	Parental strain of a mutant or laboratory-adapted virus.
zinc finger	Region of a nucleic acid-binding protein that resembles a finger and binds a zinc ion.

Index

3TC

50% inhibitory concentration

abortive infection

Acholeplasma laidlawii virus L2, budding

aciclovir

acid-triggered membrane fusion

Acidianus convivator phage

acquired immune deficiency syndrome, see AIDS

acylation

adaptive immunity

adefovir

adeno-associated virus

gene vector

adenovirus

DNA replication

E1A protein

endocytosis

genome

genome packaging

helper virus

oncogenic

primer

structure

adjuvant

adult T cell leukemia

AIDS

cancer

cytomegalovirus

alfalfa mosaic virus, virions

alpha-interferon

Alphaflexiviridae

Alpharetrovirus

alphavirus

budding

hosts

amantadine

ambisense genome

Andes virus

animal cell culture

anti-restriction
anti-termination
antibody
antigen
antigenic drift
antigenic shift
aphid vectors
Aphthovirus
apical surface of epithelial cell
aplastic anemia
APOBEC3 proteins
apoptosis
arbovirus
Arenaviridae
arenavirus
genome
assay
infectivity
quantal
quantitative
assembly
nucleocapsid
virion
asymptomatic infection
attachment (virion to receptor)
attenuated strain
autonomous parvovirus
autoradiography
avian influenza
avian leukosis virus
avian myeloblastosis virus
avian myelocytomatosis virus
azidothymidine (AZT)
B lymphocyte (B cell)
Epstein-Barr virus
B19 parvovirus
DNA
host cell
OI-propiolactone
bacteriophage, see phage

Baculoviridae
baculovirus
budding
capsid
envelope
gene expression
genome
genome packaging
inactivation
occlusion body
pesticide
recombinants
transmission
vector
Baltimore classification
Baltimore, David
barley yellow dwarf virus, genome
basolateral surface of epithelial cell
bats, viruses in
coronaviruses
filoviruses
Hendra virus
Nipah virus
rabies virus
Bcl-2 protein
bean yellow mosaic virus, transmission
beet yellows virus
structure
transmission
beetle vectors
Beijerinck, Martinus
beta-interferon
Betacoronavirus
Betaretrovirus
bicistronic messenger RNA
biological control
bioterrorism
Birnaviridae
black beetle virus, genome
blood group P antigen

blood products
blood transfusion, prion transmission
blood, virus in
bluetongue virus
cytopathic effect
serotypes
Bocavirus
Bombyx mori densovirus
Bornaviridae
bovine spongiform encephalopathy
brain cancer
brome mosaic virus
Bromoviridae
BSE
budding of virions
hepatitis B virus
HIV
influenza virus
retrovirus
rhabdovirus
Bunyaviridae
bunyavirus
genome
buoyant density
Burkitt's lymphoma
Burkitt, Denis
burst size
bushmeat
c-myc gene
cancer
canine parvovirus
canyon, virion surface
cap on RNA
messenger RNA
virus genome
cap snatching
capsid
hepatitis B virus
herpesvirus
HIV

parvovirus
picornavirus
retrovirus
rotavirus
capsomere
carcinoma
Carroll, James
Caspar, Donald
cations, virion components
Caudovirales
Caulimoviridae
cauliflower mosaic virus
transmission
caveolin
caveolin-mediated endocytosis
CCR5
CCR5 gene
CD4
CD4 T cell
CD155
cell culture
virus production
cell cycle
cancer
parvoviruses
cell line
diploid
heteroploid
MRC-5
prion replication in
Vero
WI-38
cell-free synthesis of virus
central DNA flap
central dogma
centrifugation
density gradient
differential
equilibrium
isopycnic

rate zonal
cervical carcinoma
Chase, Martha
chemokine receptor
chick embryo
chicken viruses
chickenpox
chikungunya virus
cholera toxin gene
chronic wasting disease
Chrysoviridae
cidofovir
Circoviridae
circular nucleic acid
circulative transmission
cirrhosis
classification of viruses
clathrin
clathrin-mediated endocytosis
parvovirus
rhabdovirus
cloning vector
M13 mp series
phagemid
Closteroviridae
closterovirus
co-evolution of virus and host
co-receptor
B19 virus
cold sore
cold-adapted strain
Collaborating Centers for Influenza Reference and Research
common cold
Comovirinae
complement
complex retroviruses
concatemer
connector
conservative replication of double-stranded RNA
continuous cell line

Coronaviridae
Coronavirinae
coronavirus
capping enzymes
cowpea chlorotic mottle virus
cowpea mosaic virus
capsid
genome
transport through plasmodesmata
cowpox
coxsackieviruses
medical conditions
Creutzfeldt-Jakob disease (CJD)
familial
sporadic
variant
Crick, Francis
cryo-electron microscopy
cucumber mosaic virus
genome
transport through plasmodesmata
Culicoides
cultivation of viruses
CXCR4
cyclophilin A
cypovirus
Cypovirus
cytokine
storm
cytomegalovirus
cytopathic effect
cytotoxic T cell
Cystoviridae
Darwin, Charles
dB $\bar{\tau}$ TMHerelle, Felix
deer mouse
defective interfering particle
defective virus
Deltaretrovirus
Densovirinae

dependovirus
DNA integration
gene vector
genome
latent infection
replication
detergent, effect on infectivity
diarrhea, rotaviral
Dicer
dideoxycytidine
dideoxyinosine
dimerization initiation site
Dinovernavirus
DIP
direct terminal repeat in nucleic acid
hepatitis B virus
herpes simplex virus
directed assembly of virion
disinfection
hypochlorite
DNA phage
DNA replication
filamentous single-stranded DNA phage
herpesvirus
parvovirus
phage T7
phage Φ X174
DNA-dependent DNA polymerase
DNA-dependent RNA polymerase
eukaryotic
bacterial
double-layered particle, rotavirus
double-stranded RNA phage
double-stranded RNA replication
drugs, anti-viral
design
development
hepatitis B virus
HIV
resistance

safety
targets
DTR
duck hepatitis B virus
E6 and E7 proteins, papillomavirus
E. coli
O157 toxin genes
Ebola virus
eclipse period
electron microscope
electron microscopy
electrophoresis
emerging viruses
encephalitis
herpes simplex virus
Nipah virus
West Nile virus
Enders, John
endocytosis
endogenous retrovirus
endosomal sorting complex required for transport protein
enfuvirtide
enhancer
hepatitis B virus
enterocyte
Enterovirus
entry of animal viruses into cells
env gene
HIV
retrotransposon
retrovirus
envelope
hepatitis B virus
herpesvirus
HIV
retrovirus
rhabdovirus
enzyme-linked immunosorbent assay
epidemic influenza
epidermodysplasia verruciformis

epitope
Epsilonretrovirus
Epstein, Anthony
Epstein-Barr virus
cancer
latent infection
erythema infectiosum
Erythrovirus
ESCRT protein, see endosomal sorting complex required for transport protein
eukaryotic initiation factor
evolution of viruses
exanthem subitum
exit of virion from cell
exocytosis
exogenous retrovirus
exon
F pilus, see sex pilus
F protein
F-specific filamentous phage, see filamentous single-stranded DNA phage
family (taxonomic group)
Fc (Fragment crystallizable)
feline leukemia virus
feline panleukopenia virus
Ff phage, see filamentous single-stranded DNA phage
fifth disease
filamentous single-stranded DNA phage
assembly
DNA replication
genome
replication cycle
transcription
virion
Filoviridae
filoviruses
Flaviviridae
foot and mouth disease
economic consequences
vaccine
foot and mouth disease virus
evolution

phylogenetic tree
serotypes
transmission
formaldehyde
foscarnet
fox, rabies
frameshift signal
Frosch, Paul
fungal viruses
fungus-transmitted plant viruses
furin
Furovirus
fusion inhibitor
fusion protein
gag gene
HIV
retrotransposon
retrovirus
Gag, Gag-Pol proteins
HIV
retrotransposon
retrovirus
gamma-interferon
Gammaretrovirus
ganciclovir
Geminiviridae
geminivirus
capsid
genome
gene expression, control
HIV
herpesvirus
rhabdovirus
gene therapy
gene vector
genetics
reverse
genital herpes
genome, virus
manipulation

packaging
sequencing
size
genome, replication
hepatitis B virus
herpesvirus
influenza virus
initiation
location in cell
parvovirus
picornavirus
retrovirus
rhabdovirus
rotavirus
genotype
genus (taxonomic group)
glandular fever
glycoform
glycoprotein
hepatitis B virus
HIV
herpesvirus
retrovirus
rotavirus
glycosylation of proteins
hepatitis B virus
HIV
retrovirus
rhabdovirus
rotavirus
Golgi complex
granulovirus
grapevine fanleaf virus, transmission
green fluorescent protein
guanylyl transferase
rotavirus
H5N1 influenza virus
HAART
HAV, see hepatitis A virus
Haemophilus influenzae

Hantaan virus
hantavirus
hantavirus pulmonary syndrome
Harvey murine sarcoma virus
HBV, see hepatitis B virus
HCV, see hepatitis C virus
heat shock protein
HeLa cell line
helical symmetry
helicase
herpesvirus
helicase-primase
helper component
helper-component proteinase
helper T cell
helper virus
hemagglutination
hemagglutinin
hemorrhagic fever
with renal syndrome
Hendra virus
Henipavirus
Hepadnaviridae
hepadnaviruses
see also hepatitis B virus
heparan sulfate
hepatitis
hepatitis A virus
stability
hepatitis B e antigen (HBeAg)
hepatitis B surface antigen (HBsAg)
hepatitis B virus
assembly
budding
C (core) protein
capsid
enhancers
envelope proteins
evolution
genetic groups

genome
genome synthesis
heat stability
importance
L (large) protein
liver cancer
M (medium) protein
non-infectious particles
open reading frames
P (polymerase) protein
polyadenylation signal
post-translational modifications
pregenome
promoters
replication
reverse transcription
S (small) protein
structure
transcription
translation
transmission
treatment of infection
vaccine
world distribution
X protein
hepatitis C virus
evolution
genome diversity
liver cancer
NS3-4A protein
treatment of infection
hepatitis delta virus
genome
hepatocellular carcinoma
hepatocyte
Hepatovirus
heron hepatitis B virus
herpes simplex virus
capsid
cold sore

DNA polymerase
early (E) proteins
encephalitis
genome
genome replication
glycoproteins
helicase
immediate early (IE) proteins
introns
latent infection
LATs
ori
packaging signal
primase
procapsid
pUL6
receptor
recombinants
replication
scaffolding proteins
tegument proteins
terminase
thymidine kinase
transcription
translation
virion assembly
virion host shutoff (vhs) protein
VP5
VP16
Herpesvirales
Herpesviridae
herpesvirus(es)
concatemer
genome
human
latent infection
proteins
structure
turkey
Hershey, Alfred

HERV

heteroploid cell line

hexon

herpesvirus

highly active anti-retroviral therapy

Hilleman, Maurice

histone

HIV

antigens

assembly

auxiliary genes

capsid

co-receptor

drug resistance

early gene expression

Env protein

evolution

genome

gp41

gp120

HIV-2

host cells

late gene expression

Nef protein

nucleocapsid (NC) protein

origins

p6 protein

phylogenetic tree

R5 strain

R5X4 strain

receptor

recombinants

replication

Rev protein

RRE

structure

SU protein

TAR element

Tat protein

TM protein

transcription
transmission
treatment of infection
Vif protein
Vpr protein
Vpu protein
X4 strain
Hoc
Hodgkin's lymphoma
Hoffmann-Berling, Hartmut
horizontal transmission
host-controlled modification and restriction system, see restriction system
HPV, see human papillomavirus
Hsc70
HSV, see herpes simplex virus
HTLV-1
human cytomegalovirus
human endogenous retrovirus
human herpesvirus
human herpesvirus
human immunodeficiency virus, see HIV
human metapneumovirus
human papillomavirus
vaccine
human T-lymphotropic viruses
hybridization, nucleic acid
hydrops fetalis
hydroxy-methyl-cytosine
IC50, see 50% inhibitory concentration
icosahedral symmetry
icosahedron
ICTV
IG region, see intergenic region
IgA
IgG
IgM
immune response
immunodeficiency
immunofluorescence
immunoglobulin

immunological memory
inactivation of infectivity
incubation period
induction of latent infection
infectious bursal disease virus
infectious hematopoietic necrosis virus
infectious pancreatic necrosis virus, genome
infectivity
assay
destruction
effect of lipid solvents
effect of pH
effect of radiation
effect of temperature
inactivation kinetics
preservation
prion
survival
influenza virus
assembly
budding
cap snatching
cold-adapted strain
genome
genome replication
hemagglutinin
hosts
importance
M1 protein
M2 protein
mass production
neuraminidase
neuraminidase inhibitor
NS1 protein
pandemics
reassortment
receptor
replication
RNA polymerase
strains

structure
subtypes
surveillance
transcription
translation
transmission
vaccine
influenza A virus
influenza B virus
influenza C virus
Influenzavirus
inhibition of host gene expression
initiation codon
innate immunity
Inoviridae
insect vectors
reoviruses
rhabdoviruses
integrase
inhibitor
integration of virus genomes into host genomes
integrin
interferon
intergenic region (sequence)
internal ribosome entry site
picornavirus
International Committee on Taxonomy of Viruses
intracellular transport
intron
invertebrate iridescent virus
virion
inverted terminal repeat in nucleic acid
parvovirus DNA
ion channel
IRES, see internal ribosome entry site
iridovirus
envelope formation
structure
iron lung
isolation of viruses

Iteravirus
ITR, see inverted terminal repeat in nucleic acid
Ivanovski, Dimitri
Jaagsiekte sheep retrovirus
Jenner, Edward
Jeryl Lynn strain
Kaposi's sarcoma
Kaposi's sarcoma-associated herpesvirus
latency-associated nuclear antigen
split gene
kinase, protein
cell
herpesvirus
rhabdovirus
kinase, thymidine
cell
herpesvirus
kissing-loop complex
Klug, Aaron
kuru
lactic acid bacteria, phage
lamivudine
late (L) domains of proteins
HIV Gag
retrovirus Gag
rhabdovirus M
latency-associated transcript (LAT)
latent infection
herpesvirus
HIV
Kaposi's sarcoma-associated herpesvirus
retrovirus
LDI
leader RNA
leafhopper vectors
leaky scanning
Lentivirus
leukemia
Leviviridae
ligase

linear nucleic acid
lipid
lipid raft
liver cancer
Loeb, Timothy
Löffler, Friedrich
long distance interaction
long terminal repeat (LTR)
HIV
retrotransposon
retrovirus
Longidoridae
louping ill virus, transmission
Luteoviridae
luteovirus, RNA replication
lymphocyte
lymphoma
lysis
lysis inhibition
lysogenic conversion
lysogeny
lysozyme, phage
Lyssavirus
lytic cycle
M protein
rhabdovirus
MHC, see major histocompatibility complex
M13 mp series vector
macrophage
maize chlorotic mottle virus, transmission
maize streak virus
major histocompatibility complex
malaria
malignant mesothelioma
maraviroc
Marburg virus
Marek's disease
Marek's disease virus
Marvin, Don
matrix protein

rhabdovirus
measles
measles virus
classification
M protein
receptor
structure
transmission
vaccine
meat and bone meal
megavirus
genome
membrane, virion
formation
membrane anchor sequence
membrane lipid
membrane protein
membrane synthesis
memory cell
memory T cell
meningoencephalitis
Merkel cell carcinoma
mesothelioma
messenger RNA, see mRNA
Messing, Joachim
methyl transferase
rhabdovirus
rotavirus
microarray
microscopy
confocal
cryo-electron
electron
fluorescence
microtubule
parvovirus
HIV
Microviridae
midge vectors
mimivirus

capsid
genome
possible origin
proteins
translation machinery
minimum infective dose
minus-strand RNA viruses
reverse genetics
minute virus of mice
genome
heat stability
mite vectors
MMR vaccine
Moloney murine leukemia virus
monkeypox virus
monkeys
monocistronic messenger RNA
Mononegavirales
mononucleosis
Morbillivirus
mosquito vectors
motor protein
mouse mammary tumor virus
movement protein (MP)
mRNA
bicistronic
capped
eukaryotic
monocistronic
polycistronic
prokaryotic
transport
uncapped
multicentric Castleman's disease
multipartite genome
mumps
mumps virus
structure
vaccine
murine leukemia virus

gene vector
receptor
murine sarcoma viruses
mushroom virus X
mutant
conditional lethal
temperature-sensitive
mutation
drug resistance
Myoviridae
myristylation of proteins
hepatitis B virus
picornavirus
retrovirus
rotavirus
myxoma virus
myxomatosis
N-glycosylation
naïve lymphocyte
nasopharyngeal carcinoma
natural killer cell
nectin
negative effector of transcription
negative staining
negative-strand RNA viruses, see minus-strand RNA viruses
nematode vectors
nepovirus
neuraminidase
inhibitor
neuron
herpes simplex virus
prion
rabies virus
neutralization of infectivity
neutrophil
nevirapine
Nidovirales
Nipah virus
NK cell, see natural killer cell
nomenclature of viruses

non-Hodgkin's lymphoma
non-permissive cell
non-productive infection
Norovirus
norovirus
northern blotting
Nucleorhabdovirus
nuclear envelope
nuclear export signal
nuclear localization signal
HIV proteins
influenza virus
parvovirus proteins
nuclear pore
hepatitis B virus
parvovirus
nucleocapsid
assembly
helical symmetry
icosahedral symmetry
nucleopolyhedrovirus
occlusion body
virion
nucleorhabdovirus
envelope formation
nucleoside analog
O-glycosylation
occlusion body
Oct-1
Okazaki fragment
oncogene
retrovirus
oncogenesis
oncogenic virus
one-step growth curve
open reading frame (ORF)
Orbivirus
order (taxonomic group)
ori, see replication origin
origins of viruses

Orthomyxoviridae
Orthoreovirus
oseltamivir
overlapping genes
p53
packaging of virus genomes
packaging signal
Ff phage
hepatitis B virus
HIV
retrovirus
rhabdovirus
pandemic influenza
panhandle in RNA
papilloma
Papillomaviridae
papillomavirus
cancer
E6 and E7 proteins
genome packaging
structure
Paramecium bursaria Chlorella virus, structure
Paramyxoviridae
Paramyxovirinae
paramyxoviruses
M protein
pararetrovirus
location of genome replication
recombination
Partitiviridae
Parvoviridae
Parvovirinae
parvovirus
assembly
B19
capsid
DNA replication
genome
heat stability
non-structural proteins

procapsid
proteins
receptor
replication
structure
transcription
translation
pattern recognition receptor
PBS, see primer binding site
PCR
Penicillium chrysogenum virus
penton
herpesvirus
perforin
permissive cell
persistent infection
hepatitis B virus
herpesviruses
HIV
papillomavirus
pet passport scheme
pH-independent fusion
phage
phage BYM, inactivation
phage CTX Π ⁺
phage display
phage f1
phage fd
phage lambda (O \gg)
cos site
gene products
genome
lytic cycle/lysogeny
promoters
phage M13
phage MS2
genome
lysis protein
virion
phage P008, inactivation

phage P4
phage PM2, membrane
phage QOI
genome
read-through protein
phage T1
phage T2
recombination
phage T3
phage T4
burst size
classification
genome
lysis inhibition
promoters
scaffolding protein
transcription
virion
phage T5
phage T6
phage T7
DNA replication
genome
structure
transcription
phage typing
phage Π +6
envelope
RNA L
RNA M
RNA S
phage Π X174
DNA replication
genome
rolling circle replication
scaffolding protein
virion
phagemid
phagocyte
phagocytosis

phospholipase
phosphorylation of proteins
HIV
parvovirus
rhabdovirus
phylogenetic tree
foot and mouth disease viruses
human papillomaviruses
lentiviruses
Reoviridae
Phytoreovirus
Picornavirales
Picornaviridae
picornavirus
assembly
burst size
canyon
capsid
genome
genome replication
internal ribosome entry site (IRES)
myristylation of proteins
polyprotein
primer for RNA synthesis
procapsid
proteases
proteins
recombination
replication
replication complex
replicative intermediate
reverse genetics
RNA polymerase
structure
transcription
translation
VPg (virus protein, genome linked)
pilus, see sex pilus
plaque, prion
plaque, virus

plasma cell
plasmid
plasmodesmata
plus-strand RNA viruses
reverse genetics
pock
Podoviridae
pol gene
retrotransposon
retrovirus
poliomyelitis
poliovirus
cell culture
cell-free synthesis
cytopathic effect
evolution
heat stability
receptor
recombinant
reverse genetics
serotypes
vaccines
poly(A) polymerase
rhabdovirus
polyadenylate tail (poly(A) tail)
polyadenylation of messenger RNA
rhabdovirus
polyadenylation signal
hepatitis B virus
HIV
influenza virus
retrovirus
polyamines, virion components
polycistronic messenger RNA
polymerase chain reaction
polymerase processivity factor
Polyomaviridae
polyomavirus
cancer
genome packaging

human
Merkel cell
monkey
structure
polyphage
polyprotein
HIV
picornavirus
retrovirus
polypurine tract, retrovirus
HIV
polyribonucleotidyl transferase
porcine circovirus
portal
herpesvirus
positive effector of transcription
positive-strand RNA viruses, see plus-strand RNA viruses
post-transcriptional gene silencing, see RNA silencing
post-translational modification of proteins
hepatitis B virus
influenza virus
picornavirus
retrovirus
rhabdovirus
rotavirus
post-transplant lymphoproliferative disorder
potato virus X
potato virus Y
transmission
potato yellow dwarf virus
Potyviridae
potyvirus, helper-component proteinase
Poxviridae
poxvirus
capping enzymes
envelope formation
PPT, see polypurine tract
prairie dog
pRb
pre-integration complex, retrovirus

HIV

pregenome, hepadnavirus

Pribnow box

primary effusion lymphoma

primase

primer

hepatitis B virus

influenza virus

picornavirus

retrovirus

primer binding site

prions

protein-only hypothesis

replication

species barrier

strain

transmission

PRNP gene

probe, DNA

procapsid

herpesvirus

parvovirus

picornavirus

processivity factor

productive infection

programmed cell death

prohead

promoter

bacterial

eukaryote

hepatitis B virus

phage T4

phage O»

retrovirus

proofreading

prophage

protease

cell

HIV

picornavirus

retrovirus
protease inhibitor
protein-only hypothesis
proteins, virus
nomenclature
proto-oncogene
provirus, retrovirus
HIV
PrPc
PrPres
PrPSc
Prusiner, Stanley
pseudoknot
Pteropus
quasi-equivalence theory
quasispecies
quaternary ammonium compound
rabies
vaccine
rabies virus
M protein
transmission
raltegravir
re-emerging viruses
read-through protein
real-time PCR
reassortant
influenza virus
rotavirus
reassortment
receptor
recombinant
picornavirus
vaccinia virus
recombination
phage T2
picornavirus
template switching mechanism
Reed, Walter
Reed-Muench method

Reoviridae

genera

phylogenetic tree

reoviruses

repeat sequences

herpes simplex virus

replicase

replication compartments

replication complex, picornavirus

replication origin

replicative form

replicative intermediate

picornavirus

rhabdovirus

single-stranded RNA phage

reservoir

Marburg and Ebola viruses

rabies virus

resistance to anti-viral drugs

respiratory syncytial virus

treatment of infection

vaccines

restriction system

retinoblastoma protein, see pRB

retrotransposition

retrotransposon

Retroviridae

retrovirus

assembly

budding

capsid (CA) protein

endogenous

env gene

gag gene

gene vectors

genome

genome replication

matrix (MA) protein

nucleocapsid (NC) protein

oncogene

packaging signal
pol gene
recombination
replication
structure
surface (SU) protein
transcription
translation
transmembrane (TM) protein
Rev response element
reverse genetics
influenza virus
minus-strand RNA viruses
plus-strand RNA viruses
vaccine
reverse transcriptase
hepatitis B virus
inhibitors
molecular biology applications
retrotransposon
retrovirus
reverse transcription
error rate
hepatitis B virus
HIV
retrotransposon
retrovirus
reverse transcription complex, retrovirus
HIV
reverse transcription-polymerase chain reaction, see RT-PCR
RF, see replicative form
Rhabdoviridae
rhabdovirus
assembly
budding
fish
genome
genome replication
glycoprotein (G)
host gene inhibition

intergenic sequence
large protein (L)
matrix protein (M)
nucleoprotein (N)
phosphoprotein (P)
plant
replication
replicative intermediate
RNA polymerase
structure
transcription
translation
rhinovirus
capsid
receptors
serotypes
rho factor
RI, see replicative intermediate
ribavirin
ribonuclease H
hepatitis B virus
retrovirus
ribosomal frameshift
HIV
retroviruses
ribosomal subunits
eukaryotic
prokaryotic
ribozyme
rice dwarf virus, transmission
rice tungro bacilliform virus
rinderpest
RISC
ritonavir
RNA in DNA viruses
RNA interference (RNAi)
RNA phage
RNA polymerase
see also DNA-dependent RNA polymerase
see also RNA-dependent RNA polymerase

RNA polymerase II
RNA replication
picornavirus
rhabdovirus
rotavirus
single-stranded RNA phages
RNA silencing
RNA triphosphatase
RNA-dependent RNA polymerase
influenza virus
picornavirus
rhabdovirus
rotavirus
RNA-induced silencing complex
RNase H, see ribonuclease H
rolling circle DNA replication
herpesvirus
DNA phages
rolling hairpin DNA replication
rooted phylogenetic tree
rotavirus
disease
double-layered particle
replication
RNA polymerase
stability
structure
transcription
translation
triple-layered particle
vaccines
Rous sarcoma virus
oncogene
Rous, Peyton
RRE
RT-PCR
SARS coronavirus
West Nile virus
rubella virus
transmission

ryegrass mosaic virus, transmission
Sabin, Albert
Salk, Jonas
Sanger, Fred
sarcoma
SARS
SARS coronavirus
genome
satellite tobacco mosaic virus, capsid
satellite tobacco necrosis virus
satellite virus
dependovirus
scaffolding protein
herpesvirus
phage T4
phage Φ X174
scrapie
atypical
SDS-PAGE
rotavirus RNA
secondary structure of nucleic acid
Secoviridae
seed-transmitted viruses
segmented genome
ambisense RNA viruses
bunyaviruses
influenza virus
minus-strand RNA viruses
plus-strand RNA viruses
reassortment
reoviruses
tobacco rattle virus
selectivity index
self-assembly of virion
self-priming of DNA replication
semi-conservative replication
double-stranded DNA
double-stranded RNA
Semliki Forest virus
sequencing genomes

SARS coronavirus
serotype
severe acute respiratory syndrome
sex pilus
sexual transmission
Shiga toxin
Shine-Dalgarno sequence
shingles
sialic acid
sigma factor
sigma replication of DNA
sign of disease
signal sequence
simian foamy virus
simian immunodeficiency virus
phylogenetic tree
simian virus
endocytosis
T antigen
simple retroviruses
Sin Nombre virus
single-stranded DNA binding protein
single-stranded DNA phage
single-stranded RNA phage
Min Jou (MJ) interaction
replicase
replication cycle
van Duin (VD) interaction
single-stranded RNA replication
Siphoviridae
SIV, see simian immunodeficiency virus
slippery sequence
smallpox
vaccination
virus
Soc
Southern blotting
species (taxonomic group)
spike, envelope
herpesvirus

influenza virus
HIV
rhabdovirus
spliced transcript
HIV
influenza A virus
parvovirus
retrovirus
split gene
Spongospora subterranea
spraying
Spumavirus
squamous cell carcinoma
steps in virus replication
sterilization
stomatitis
storage of viruses
structure of viruses
subclinical infection
subfamily (taxonomic group)
subgenomic messenger RNAs
superinfection immunity
suppressor transfer RNA
surveillance of virus disease
swine flu
SV40, see simian virus
symmetry
helical
icosahedral
symptom, definition
syncytium
T even phage
T lymphocyte (T cell)
HIV infection
T7 DNA polymerase
T7 RNA polymerase
tailed phage
TAR element
TATA box
Tax protein

taxonomic groups
TCID50
tequivirus
DNA replication
genome
tegument
temperate phage
Tenuivirus
terminally redundant
terminase
tetherin
theta replication of DNA
thymidine kinase
cell
herpesvirus
thymine dimer
tick-borne encephalitis virus
tick vectors
tobacco necrosis satellite virus,
genome
tobacco mosaic virus
assembly
genome
structure
transport through plasmodesmata
tobacco rattle virus
Tobamovirus
Tobravirus
tobravirus transmission
Togaviridae
Toll-like receptor
tomato mosaic virus
tomato yellow leaf curl virus, transmission
Tombusviridae
tombusvirus P19 protein
torque teno virus
transactivation response element
transcriptase
see also reverse transcriptase
transcription

bacterial
eukaryotic
hepatitis B virus
herpesvirus
HIV
influenza virus
initiation
parvovirus
phage T4
phage T7
picornavirus
retrovirus
rhabdovirus
rotavirus
termination
transcription factor
transcription factor IID (TFIID)
transduction
transfection
transfer RNA in retrovirus virion
transformation
transgenic mice
translation
bacterial
eukaryotic
hepatitis B virus
herpesvirus
HIV
influenza virus
initiation
parvovirus
picornavirus
retrovirus
rhabdovirus
rotavirus
translational coupling
transmissible spongiform encephalopathy
transmission electron microscope
transmission of prions
transmission of viruses

bluetongue virus
chikungunya virus
filoviruses
Hendra virus
hepatitis B virus
HIV
invertebrate viruses
Nipah virus
plant viruses
vertebrate viruses
West Nile virus
transovarial transmission
transport of proteins in cells
influenza virus
transposition
transposon
Trichodoridae
triple-layered particle, rotavirus
turnip crinkle virus, capsid
turnip yellow mosaic virus
Twort, Frederick
U3 sequence
U5 sequence
ultraviolet radiation, effect on infectivity
uncoating
unrooted phylogenetic tree
uridylylation of poliovirus VPg
vaccination
wild mammals
vaccines
contamination
DNA
hepatitis B virus
HIV
human papillomavirus
inactivated
influenza
live, attenuated
Marek's disease virus
MMR

peptide
poliovirus
preservation
quality control
recombinant
subunit
virus-like particle
virus production
vaccinia virus
recombinant
variant Creutzfeldt-Jakob disease (vCJD)
varicella
varicella-zoster virus
variola virus
vector
transmission of viruses
see also cloning vector
see also gene vector
Vero cells
vertical transmission
plant viruses
vertebrate viruses
invertebrate viruses
vesicular stomatitis virus
budding
phosphoprotein
Vesiculovirus
viremia
HIV
see also blood, virus in
virion
assembly
definition
detection
size
structure
virological synapse
viroplasm
virulence
virulence factor, bacterial

virulent phage
virus definition
virus replication, overview
virus-like particle
VPg (virus protein, genome linked)
walleye dermal sarcoma virus
wart
water treatment
Watson, James
West Nile virus
western blotting
white clover cryptic virus
whitefly vectors
winter vomiting disease
woodchuck hepatitis virus
World Health Organization
X-linked severe combined immunodeficiency (X-SCID)
X-ray crystallography
yeast two-hybrid system
yellow fever virus
budding
structure
transmission
vaccine
zidovudine
zinc finger
Zinder, Norton
zoster