Volume 49

Advances in GENETICS

edited by

Jeffrey C. Hall

Jay C. Dunlap

Theodore Friedmann

Volume 49

Advances in Genetics

Advances in Genetics, Volume 49

Volume 49

Advances in Genetics

Edited by

Jeffrey C. Hall
Department of Biology
Brandeis University
Waltham, Massachusetts

Jay C. Dunlap
Department of Biochemistry
Dartmouth Medical School
Hanover, New Hampshire

Theodore Friedmann
Center for Molecular Genetics
University of California at San Diego School of Medicine
La Jolla, California

ACADEMIC PRESS

An imprint of Elsevier Science

Amsterdam Boston London New York Oxford Paris
San Diego San Francisco Singapore Sydney Tokyo

This book is printed on acid-free paper. ♾

Academic Press
An imprint of Elsevier Science.
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.academicpress.com

Academic Press
84 Theobald's Road, London WC1X 8RR, UK
http://www.academicpress.com

International Standard Book Number: 0-12-014769-6

PRINTED IN THE UNITED STATES OF AMERICA
03 04 05 06 07 08 9 8 7 6 5 4 3 2 1

Contents

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Nigel G. Halford (111) Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS41 9AF, United Kingdom

Stephen L. Helfand (67) Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06030

Domenico Lafiandra (111) Department of Agrobiology and Agrochemistry, University of Tuscia, 01100 Viterbo, Italy

Blanka Rogina (67) Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06030

Peter R. Shewry (111) Long Aston Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS41 9AF, United Kingdom

Paul H. Taghert (1) Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110

Jan A. Veenstra (1) Laboratoire Neuroendocrinologie des Insectes, Université Bordeaux I, 33405 Talence Cedex, France

1

Drosophila Neuropeptide Signaling

Paul H. Taghert
Anatomy and Neurobiology
Washington University School of Medicine
St Louis, Missouri 63110

Jan A. Veenstra
Laboratoire Neuroendocrinologie des Insectes
Université Bordeaux I
33405 Talence Cedex, France

Advances in Genetics, Vol. 49

0065-2660/03 $35.00

I. INTRODUCTION

The publication of the *Drosophila* genome has begun a new chapter in the study of the fly. Together with the advanced genetic techniques available for

Drosophila research, genome information provides a basis for a more rational and comprehensive analysis of problems in animal development and physiology. Neuropeptide signaling in insects is an old and august field—in fact, the first observations on neurohormones were made in insects in 1917 (Kopec, 1922). In this review, we have tried to provide something useful for two potential sets of readers. First, we hope this material is useful to *Drosophila* workers who wish for an introduction to current issues and progress in insect neuropeptide signaling. Second, we hope it is also useful to those researchers who study neuropeptides in other model systems or in human physiology, but now intend to examine opportunities afforded by the publication of the fly genome, and wish for more specific information concerning *Drosophila*. With that general intention, we have organized this review into two parts. The first part overviews the major neuropeptide families in *Drosophila* (and related insects). In addition to listing relevant gene identities, we try to provide a biological context for each family, where possible. The second part of the review considers categorical aspects of neuropeptide signaling, such as biosynthesis, release, and degradation. These issues have clear and direct parallels with studies of neuropeptide signaling in diverse animals, including mammals. For each category, we summarize the state of the field, focusing principally on *Drosophila*, and have tried to feature those fly genes that are orthologous to the most relevant genes from mammals. We apologize to colleagues whose work was not reviewed owing to lack of space, or because of our oversight.

II. NEUROPEPTIDE FAMILIES

A. Overview

Most insect neuropeptides are produced by neuroendocrine cells and are released into the blood (hemolymph), and/or are produced by neurons and released into the central nervous system (CNS). However, several biologically active peptides are released, some perhaps exclusively so, by dispersed endocrine cells in the midgut epithelium, or by neuroendocrine cell processes on the surfaces of nerves or trachea. As the distinction between a peripheral neuron and a peripheral endocrine cell is somewhat arbitrary, we have included all such biologically active secretory peptides produced by dispersed neuroendocrine cells in our definition of neuropeptide. The definition of neuropeptide used here also includes the larger neurohormones, which may have molecular mass >10 kDa.

Neuropeptides are variable in structure and are not easily recognized in novel databases such as the recently published *Drosophila* genome sequences (Adams *et al.*, 2000). However, their homologies to related neuropeptides

Table 1.1. *Drosophila* **Neuropeptide Genes**

Gene	CG number	Location	Neuropeptide	Text Section
Dilp-1	CG14173	67C1-2	Insulin-like peptide 1	II.B, II.D
Dilp-2	CG8167	67C1-2	Insulin-like peptide 2	II.B, II.D
Dilp-3	CG14167	67C1-2	Insulin-like peptide 3	II.B, II.D
Dilp-4	CG6736	67C1-2	Insulin-like peptide 4	II.B, II.D
Dilp-5	AE003550.1	67C1-2	Insulin-like peptide 5	II.B, II.D
Dilp-6	CG14049	2F4	Insulin-like peptide 6	II.B, II.D
Dilp-7	CG13317	3F1-2	Insulin-like peptide 7	II.B, II.D
—	CG13687	21C6	PTTH (putative)	II.C
eh	CG6400	90B1	Eclosion hormone	II.C
eth	CG18105	60D15-E1	Ecdysis-triggering hormone	II.C
ccap	CG4910	94C4	Crustacean cardioactive peptide	II.C
akh	CG1171	64A10-12	Adipokinetic hormone	II.F
Dh	CG8348	85E4	CRF-like diuretic hormone	II.F
Dh31	CG13094	29D1	Calcitonin-like diuretic hormone	II.F
leukokinin	CG13480	70E3-4	Leukokinin	II.F
capa	CG15520	99D1	Pyrokinin gene product 1	II.F, II.J
ITP	CG13586	60D5	Intestinal transport peptide-like	II.F
pdf	CG6496	97B2	Pigment-dispersing factor	II.G
fmrfa	CG2346	46C5	dFMRFamide	II.H
dms	CG6440	95F6-9	Dromyosuppressin	II.H
dsk	CG18090	82E1	Drosulfakinin	II.H
npf	CG10342	89D5-6	Neuropeptide F-like	II.H
sNPF	CG13968	38B4	Head peptides	II.H
Ast	CG13633	96A22-23	Allatostatin-A/YXFGLamides	II.I, V.C
MIP	CG6456	74B1	Allatostatin-B/W(X_6) Wamides	II.I
Ast2	CG14919	32D2-3	Allatostatin-C	II.I
hug	CG6371	87C1	Pyrokinin gene product 2	II.J
IFamide	CG4681	60D5-7	IFamide	II.K
tk	CG14734	87A8-9	Tachykinins	II.M, IV.B, V.C
—	CG7105	28D3-4	Proctolin (putative)	II.N
amn	CG11937	19A1	Amnesiac	II.O
crz	CG3302	87B7-8	Corazonin	II.R

(those previously identified by physiological activities in other animals) have permitted the identification of more than 20 neuropeptide-encoding *Drosophila* genes (Table 1.1) (Hewes and Taghert, 2001; Vanden Broeck, 2001). Although a few neuropeptides have been isolated and chemically characterized from *Drosophila* (Table 1.2), all of those have also been found by application of immunological assays for neuropeptides in other insect species. In spite of such similarities, the physiological roles of *Drosophila* neuropeptides remain largely uncertain. With the exception of *amnesiac*, which was defined genetically,

Table 1.2. Neuropeptides Isolated and Chemically Identified in *Drosophila*

Gene	Neuropeptide	Ref.
fmrf	DPKQDFMRFamide	Nambu *et al.*, 1988; Nichols, 1992a
	TPAEDFMRFamide	Nichols, 1992a
	SDNFMRFamide	Nichols, 1992a
dms	TDVDHVFLRFamide	Nichols, 1992a
dsk	FDDYGHMRFamide	Nichols, 1992b
sNPF	SPSLRLRFamide	P. Trifilieff and J. A. Veenstra, unpublished data
npf	KNDVNTMADAYKFLQDLDTYYG DRARVRFamide	P. Brown *et al.*, 1999
leucokinin	NSVVLGKKQRFHSWGamide	Terhzaz *et al.*, 1999
IFa	AYRKPPFNGSIFamide	S. Tehrzaz and J. A. Veenstra, unpublished data

identification of the putative *Drosophila* neuropeptide precursor genes is largely based on homology searches.

In principle, putative neuropeptide-encoding genes, especially those encoding a number of copies of C-terminally amidated neuropeptides, might also be identified by the presence of characteristic precursor sequences. Veenstra (2000) reviewed the sequence information available from insect neuropeptide precursors and showed that the rules established for predicting potential convertase cleavage sites within vertebrate neuropeptide precursors are also valid for insects. We have not attempted to find such genes in the *Drosophila* genome, as their identification is hazardous at best. Rather, as others, we have restricted ourselves in this review to those neuropeptide genes that exhibit strong homology to well-established insect, invertebrate, or vertebrate neuropeptides, or to larger glycoprotein hormones. In this section, we describe ~15 sets of *Drosophila* neuropeptides that are implicated in regulating a variety of developmental and physiological processes.

B. Neuropeptides regulating growth

Until recently the endocrine regulation of insect growth focused almost entirely on molting and metamorphosis. However, molting makes sense only after the size of an insect has exceeded the limits of expansion by its cuticle, or once it has accumulated sufficient reserves to metamorphose. So far, no genes encoding homologs of mammalian leptin, or its receptor, have been identified in the *Drosophila* genome. However, one would expect that a leptin-like hormone could be a useful signal in insects to indicate that sufficient reserves have been

accumulated for molting. It has been recently published that in *Drosophila*, growth itself is regulated by insulin signaling.

A *Drosophila* gene encoding an insulin receptor, *InR* (CG18402 at 93E9), was identified in 1985 (Thompson *et al.*, 1985) and subsequently shown to be essential for normal growth (Chen *et al.*, 1996). Moreover, *chico*, a *Drosophila* homolog of the vertebrate insulin receptor substrates, similarly plays an essential role in the control of *Drosophila* cell growth and size (Bohni *et al.*, 1999). In fact, it appears that the entire signaling pathway comprising phosphatidylinositol 3′-kinase, its downstream target the Akt serine/threonine kinase (protein kinase B), the PTEN tumor supressor, and S6 kinase has been functionally conserved in *Drosophila*. Mutants in any of the *Drosophila* homologs display predictable effects on body and cell size, and on cell numbers, with the exception that the *Drosophila* S6 kinase homolog does not affect cell numbers (Bohni *et al.*, 1999; Leevers *et al.*, 1996; Weinkove *et al.*, 1999; Goberdhan *et al.*, 1999; Huang *et al.*, 1999; Gao *et al.*, 2000; Scanga *et al.*, 2000; Verdu *et al.*, 1999; Montagne *et al.*, 1999). Overexpression of *Dinr* in eye precursor cells leads to an increase in the number of ommatidia and in ommatidial cell size (Brogiolo *et al.*, 2001). These data suggest that the *Drosophila* insulin signaling pathway through the mitogen-activated protein (MAP) kinase induces cell division, and that the one through phosphatidylinositol 3′-kinase stimulates growth (Brogiolo *et al.*, 2001).

It is somewhat surprising that, whereas there appears to be only a single insulin receptor, *Drosophila* has seven different *insulin* genes encoding seven different *Drosophila* insulin-like peptides. The occurrence of multiple insulin-encoding genes in insects has precedent in the silkworm, *Bombyx*, where more than 30 insulin-like (bombyxin) genes have been identified (Kondo *et al.*, 1996). Work on *Caenorhabditis elegans* suggests that under experimental conditions of overexpression, some insulin-like peptides in this species might function as receptor antagonists (Pierce *et al.*, 2001). Such hypotheses of complex gene interactions can now be investigated directly with these model genetic systems.

Five of the *Drosophila insulin* genes—*dilp1-4* (CG14173, CG8167, CG14167, and CG6736, all at 67C1-2) and *dilp-5* (AE003550.1: 259, 432–259,854)—form a gene cluster, whereas *dilp-6* (CG14049 at 2F4) and *dilp-7* (CG13317 at 3F1-2) are found on a different chromosome. As with other invertebrate insulins, sequence similarity to insulin is limited: *dilp-2* is most closely related to human insulin, with 35% identity. Expression of the genes was determined by *in situ* hybridization. *dilp-2 and dilp-3* are coexpressed in seven pairs of brain neurons. *dilp-2* is also expressed in the imaginal discs and salivary glands; *dilp-5* is expressed in the midgut. The neuroendocrine anatomy of these cells has been demonstrated by immunohistology using antiserum to DILP-2 (Cao and Brown, 2001). *dilp-4* is highly expressed in the midgut, and *dilp-6* yields a generally low signal. The expression of *dilp-7* is entirely different, and is

restricted to 10 cells in the ventral nerve cord of the CNS (Brogiolo *et al.*, 2001). Overexpression of *dilp-2* also leads to increases in cell size and numbers, and these effects are dependent on *Dinr* (Brogiolo *et al.*, 2001).

C. Neuropeptides regulating molting

After emerging from the egg, *Drosophila* goes through three larval and one pupal stage before initiating adult development. Each time the insect molts, the old cuticle is detached and a new cuticle is formed beneath the old one. Both the detachment of the old cuticle and the synthesis of a new one are under the control of ecdysone, a steroid hormone produced by the prothoracic gland (PG), also known as the ventral gland. The synthesis of ecdysone by the PG is stimulated by prothoracicotropic hormone (PTTH). This hormone has been isolated and fully identified from the silkworm moth as a 22-kDa dimer (Kawakami *et al.*, 1990). A possible homolog of this hormone may be encoded by GC13687 (at 21C6), but its identity as *Drosophila* PTTH remains to be proved. It should be noted that two other putative *Drosophila* PTTHs have been proposed (Chung *et al.*, 1994; Kim *et al.*, 1997). However, the identity of either of these two as a bone fide PTTH remains unclear: the first cannot be found in the *Drosophila* genome sequence, and the second may represent a breakdown product of a Wnt inhibitory factor 1-like protein.

The type of new cuticle produced (larval, pupal, or adult) depends on the concentrations of juvenile hormone, a sesquiterpenoid produced by the glandular corpora allata (CA). The activity of the CA is regulated by small neuropeptides, which either stimulate (allatotropins) or inhibit (allatostatins) the CA to produce juvenile hormone; see Sections II.I and II.L. From other insect species, there is evidence of both direct innervation of the CA by allatotropinergic neurons and by allatostatinergic neurons, and evidence also of the release of allatotropin and allatostatin into the hemolymph. In *Drosophila*, the CA and the cells producing ecdysone are found in association along with a third glandular tissue, the corpora cardiaca (CC). The three tissues form a unified endocrine structure that lies near the brain and is called the ring gland (RG).

Once the new cuticle has been synthesized, the insect needs to extricate itself from the old one (except for the larval–pupal molt in higher Diptera such as *Drosophila*, when the last larval cuticle is not removed, but becomes a protective layer for the pupa). The behavior associated with extrication from the old cuticle (termed ecdysis, or eclosion; reviewed by Ewer and Reynolds, 2002) in insects is initiated by the neuropeptide eclosion hormone (EH). EH is produced in paired neuroendocrine cells in the brain releasing neuropeptide within the CNS and into the hemolymph (Hewes and Truman, 1991). The structure of EH has been established for two lepidopteran

species (Marti *et al.*, 1987; Kataoka *et al.*, 1987; Kono *et al.*, 1990), and a homologous *Drosophila* gene (CG6400 at 90B1) has been characterized. The cells expressing this gene have been identified by *in situ* hybridization and immunohistology (Horodyski *et al.*, 1993). Although the predicted hormone has not been tested for biological activity, flies expressing Reaper under indirect control of the *eh* gene promoter (by means of the UAS–GAL4 system) have serious problems eclosing, apparently in direct relation to the efficiency of Reaper to induce apoptosis in the eclosion hormone-expressing cells (McNabb *et al.*, 1997).

Eclosion hormone has been particularly well studied in *Manduca sexta*. Among its most important effects is to stimulate Inka cells to release ecdysis-triggering hormones 1 and 2 (ETH1 and ETH2), both of which are produced from a single neuropeptide precursor. Inka cells are dispersed endocrine cells associated with the tracheal system (Zitnan *et al.*, 1996). ETH1 and ETH2, along with EH, induce the complex preprogrammed behaviors that allow the insect to escape from its old cuticle. *eth* (CG18105 at 60D16) has been identified as the homologous *Drosophila* gene, and the predicted *Drosophila* peptides induce premature eclosion behavior (Park *et al.*, 1999). Although the *ETH*-expressing Inka cells have not been definitively identified in *Drosophila*, the Inka cells in *Manduca sexta* are myomodulin immunoreactive, and similar myomodulin-immunoreactive, peritracheal endocrine cells are present in *Drosophila* (O'Brien and Taghert, 1998). Further details concerning EH and ETH interactions are described in Section IV.B.

Once the insect has escaped from its old cuticle, it needs to stretch the new one. This is generally accompanied by increased body hemolymph pressures, as the insect forces both air and hemolymph into body regions and appendages to extend its new cuticle. The associated increase in the rate of heart beat is probably due at least in part to the release of crustacean cardioactive peptide (CCAP) (Ewer and Truman, 1996; Gammie and Truman, 1999), a neuropeptide first identified from crustaceans (Stangier *et al.*, 1988). A gene encoding CCAP has been found in the *Drosophila* genome (CG4910 at 94C4), and neuroendocrine cells and neurons expressing CCAP in *Drosophila* have been described by immunohistology. In *Drosophila* CCAP increases the heart beat rate (Nichols *et al.*, 1999b). After unfolding and stretching the new cuticle the neurohormone bursicon induces its sclerotization (Section II.Q).

D. Neuropeptides regulating reproduction

Work on the endocrine and neuroendocrine regulation of reproduction has focused largely on the regulation of vitellogenesis. Vitellogenesis is stimulated

in most insect species by juvenile hormone, but in Diptera, ecdysone produced by the ovary is much more important (ecdysone is also produced by the testis). In mosquitoes a neurohormone stimulating ecdysone synthesis by the ovary has been identified (Brown *et al.*, 1998) and was found to be a homolog of neuroparsin, a locust neurohormone previously identified for its antijuvenile hormone effects (Girardie *et al.*, 1987, 1989). Neuroparsin retards oocytes growth in adult locusts. In juvenile locusts, injections of antiserum to neuroparsin induce the display of green pigment, characteristic of juvenile hormone action, as well as intermediate pigmentation forms that are characteristic of intermediate, metamorphic stages. Similar injections in adult locusts accelerate sexual maturation in males and oocyte growth in females (Girardie *et al.*, 1987). Because neuroparsin is present not only in adults, but also in larvae, and because neuroparsin antiserum disrupts normal metamorphosis, it may play a role in development.

Locust neuroparsin and its mosquito homolog are both produced by neuroendocrine cells of the pars intermedia region of the brain, and are stained strongly by classic neurosecretory staining methods, such as those involving paraldehyde fuchsin, because of the presence of six intramolecular disulfide bridges in these molecules (Bourâme *et al.*, 1987; Brown and Cao, 2001). Homologous cells are present in the pars intermedia of all insect species studied, including *Drosophila*. Although it seems likely that *Drosophila* has a neuroparsin homolog, genome searches have not yet identified it. Either its homology is too limited, or perhaps it has not been sequenced yet.

Vertebrate insulin in high concentrations is also able to stimulate ecdysone biosynthesis by the mosquito ovary (Graf *et al.*, 1997; Riehle and Brown, 1999) and *dinr* (Section II.C) is abundantly expressed in the *Drosophila* ovary (Garofalo and Rosen, 1988). These facts suggest that the *dilp* genes may also be involved in the regulation of vitellogenesis in *Drosophila*.

In locusts, a third neurohormone, ovary-maturing parsin (OMP), has been identified as a vitellogenesis-stimulating hormone. Specific antisera to OMP recognize neuroendocrine cells in locusts, but not in other insect species (Richard *et al.*, 1994). Therefore, either the structure of this hormone is poorly conserved during evolution, or the hormone is specific to locusts. Indeed, we did not find a homologous *Drosophila* sequence. This hormone is colocalized with the corticotropin-releasing factor (CRF)-like diuretic hormone in locusts (Section II.F). Tamarelle *et al.* (2000) have suggested that it might be a general metabolic hormone, rather than one that specifically regulates reproduction. In fact, we speculate that any neurohormone that stimulates protein synthesis directly or indirectly (e.g., by stimulating feeding activity) will stimulate vitellogenesis in adult females. This reasoning may also be applicable to the reported stimulation of vitellogenesis by the "head peptides" in the locust (Cerstiaens *et al.*, 1999; Section II.J).

E. Homeostatic neurohormones: Regulation of hemolymph carbohydrate

Although glucose is present in insect hemolymph in concentrations similar to those found in vertebrate plasma, carbohydrate metabolism in insects is significantly different from that in vertebrates. The major sugar in insect hemolymph is trehalose (a glucose dimer), which is present in concentrations much higher than glucose (Friedman, 1985). Functionally, trehalose is an intermediate between glycogen and glucose; like glucose, it is soluble in the hemolymph, yet it cannot be taken up directly by insect cells. It must first be cleaved by trehalase into glucose. There is no evidence of the existence of a homeostatic hypoglycemic hormone in insects (Veenstra, 1989a), but all insects produce what is variably called an adipokinetic (AKH), hyperglycemic, or hypertrehalosemic hormone.

In all insect species studied, these hormones stimulate the activation of glycogen phosphorylase in the fat body, which leads to an increase in the concentration of hemolymph trehalose (Friedman, 1985). In several migratory insect species, the same hormones also mobilize lipid; hence the name, adipokinetic hormone, in such species as the migratory locust and the tobacco hornworm moth. AKH peptides are produced by a group of endocrine cells in the CC and tend to be present in rather large quantities. Resting insects do not consume much energy, but active flight muscle demands large amounts of energy substrates. Within this context it is not surprising that the (neuro) endocrine regulation of hemolymph carbohydrate (and/or lipid) emphasizes a hyperglycemic, rather than a hypoglycemic, hormone as in mammals. However, the situation in (adult) flies differs from that generalization; rather than synthesizing trehalose from fat body glycogen, they use carbohydrates stored in the crop. The physiological mechanism that allows for the transfer of the carbohydrates from the crop into the midgut has not been clarified. It likely involves the recurrent nerve of the foregut, as the destruction of the nerve greatly increases hemolymph concentrations of trehalose in flies (cf. Veenstra, 1989b).

The *Drosophila* AKH homolog has been isolated and chemically identified (Schaffer *et al.*, 1990); by *in situ* hybridization, its gene (*Akh*, CG1171 at 64A10-12) is expressed in the RG of third instar larvae (Noyes *et al.*, 1995). The hormone increases the rhythm of heart beat in prepupae (Noyes *et al.*, 1995), an effect also found for several other members of this peptide family. However, it has not yet been shown to be involved in the regulation of carbohydrate and/or lipid metabolism. The endocrine regulation of carbohydrate metabolism may also be important during molting: chitin, the major component of the cuticle, is a polymer of glucose. Indeed, during and 1 day after molting, hemolymph concentrations of trehalose in the cockroach *Periplaneta*

americana are greatly increased (Steele, 1985). Although it has not been shown that AKH homologs are responsible for this increase, it is interesting to note that in locust, the release of AKH is stimulated by CCAP, which is released during molting (Section II.C and III.C).

F. Homeostatic neurohormones: Regulation of water and Ion homeostasis

Water and ion metabolism in insects is regulated in large part by the Malpighian tubules and the hindgut (reviewed by Dow and Davies, 2001). The Malpighian tubules filter the hemolymph and secrete a liquid that is often compared with the primary urine in vertebrates. This liquid passes in the gut and, depending on the species and the "hydration state" of the insect, the primary urine may enter either the midgut (Dow, 1981) or the hindgut, where selective reabsorption of ions and water takes place. Insect neuropeptides, called diuretic or antidiuretic hormones, stimulate fluid secretion by the Malpighian tubules or fluid reabsorption by the tubules or hindgut, respectively. A large number of insect neuropeptides are known to increase the rate of fluid secretion by the Malpighian tubules. They include CRF-related and calcitonin-related diuretic hormones, the leukokinins, and neuropeptides related to the lepidopteran hormone CAP_{2b}. *Drosophila* genes encoding these hormones have been identified for all four putative diuretic hormones (*Dh*, CG8348 at 85E4; *Dh31*, CG13094 at 29D1; *leukokinin*, CG13480 at 70E3-4; *capa*, CG15520 at 99D1).

The main fragment of Malpighian tubules consists of two cell types, the principal cells and the stellate cells (Sözen *et al.*, 1997). Fluid secretion is achieved by a vacuolar H^+-motive ATPase in the apical membrane of the principal cells. In combination with a K^+/H^+ exchanger, the ATPase achieves a net transport of K^+, while chloride flows through maxi-Cl^- channels in the stellate cells (O'Donnell *et al.*, 1998), and water is believed to follow passively through aquaporins (Dow *et al.*, 1995). The vacuolar ATPase can be stimulated by both cAMP and cGMP. The calcitonin-like diuretic hormone II (DH II) stimulates intracellular cAMP and increases transepithelial voltage in the Malpighian tubules, implying that it acts on the principal cells (Coast *et al.*, 2001). None of the CRF-like insect diuretic hormones (DH Is) have been tested so far in *Drosophila*. However, because these hormones all stimulate intracellular cAMP, and because in some cases they increase the transepithelial voltage in the Malpighian tubules (e.g., Reagan, 1994; Furuya *et al.*, 1995; Clark *et al.*, 1998), it is likely that DH I also acts on the principal cells.

The pyrokinin CAP_{2b} (Huesmann *et al.*, 1995) identified from the tobacco hornworm moth *Manduca sexta*, as well as its *Drosophila* homologs

CAPA-1 and CAPA-2, have been shown to stimulate epithelial fluid transport by Malpighian tubules via the cGMP signaling pathway (Davies *et al.*, 1995; Kean *et al.*, 2002). The signaling pathway has been established in detail. CAP_{2b} increases intracellular calcium (Rosay *et al.*, 1997), which activates nitric oxide (NO) synthetase (Davies *et al.*, 1997). That enzyme is present in the principal cells, but not the stellate cells, of the Malpighian tubules (Davies, 2000). Finally, NO stimulates a soluble guanylate cyclase, which leads to an increase in intracellular cGMP in the principal cells (Davies *et al.*, 1997; Kean *et al.*, 2002). Thus, uniquely, the NO produced appears to act within the same cell type in which it is generated. In *Rhodnius*, CAP_{2b} acts through cGMP, but is an antidiuretic hormone (Quinlan *et al.*, 1997), suggesting that CAP_{2b} peptides have wide phylogenetic conservation, but that their roles may be less precisely conserved.

Whereas DH I, DH II, CAPA-1, and CAPA-2 all appear to act on the principal cells, the leukokinins act on the stellate cells to increase Cl^- conductance via an increase in intracellular calcium (O'Donnell *et al.*, 1998). The *Drosophila leukokinin* gene encodes a single peptide (Terhzaz *et al.*, 1999), whereas in other insect species, up to eight leukokinin peptides have been chemically identified. In the case of the mosquito *Aedes aegypti*, all three peptides are encoded by a single gene (Veenstra *et al.*, 1997). *Drosophila* leukokinin has been isolated, and its structure has been chemically determined. It is the largest insect leukokinin known and, as with others, its stimulates fluid secretion by the Malpighian tubules by means of an increase in intracellular calcium in the stellate cells (Terhzaz *et al.*, 1999). The leukokinins also stimulate hindgut contractions in all insect species studied, including the mosquito (Veenstra *et al.*, 1997); this activity has not been determined in *Drosophila*.

Of all the putative diuretic hormones, leukokinin appears to be the most potent, with a significant increase in the rate of fluid secretion at doses of 0.1 nM, whereas CAPA-1, CAPA-2, and Dh31 significantly increase the rate of fluid secretion only at concentrations of 1 nM or higher. Synergistic effects on fluid secretion by the Malpighian tubules have been described for leukokinin and the calcitonin-like diuretic hormone (Coast *et al.*, 2001), as well as for DH I and leukokinin in locusts (Thompson *et al.*, 1995). Cantera and Nässel (1992) used an antiserum to cockroach leukokinin to describe the cellular localization of leukokinin in larval *Drosophila* in seven pairs of neuroendocrine cells in the abdominal neuromeres, as well as six pairs of cerebral neurons and three pairs of suboesophageal neurons. The distribution of DH I has not been reported for *Drosophila*, but in the migratory locust and the tobacco hornworm moth some neuroendocrine cells produce both types of DH (Chen *et al.*, 1994; Thompson *et al.*, 1995).

The CAPA peptides are produced by a limited number of neurons in the *Drosophila* brain, and by two different groups of neuroendocrine cells. The first group consists of a single pair of cells in the labial neuromere and projects to the CC, and the second group consists of three cell pairs in the abdominal neuromeres that project to the abdominal transverse nerves (Kean *et al.*, 2002). All these cells appear different from those producing the leukokinins, or those that are likely to produce DH I.

So far, two putative antidiuretic hormones that can stimulate fluid reabsorption by the hindgut have been identified in locusts: neuroparsin and ion transport peptide (ITP), a homolog of crustacean hyperglycemic hormone (Audsley *et al.*, 1992). The antidiuretic effects of neuroparsin (Section II.D) on the locust hindgut are marginal (Fournier and Girardie, 1988), and it is unlikely to represent a major antidiuretic hormone. In locusts, ITP is found in the CC, and hence may be produced by brain neuroendocrine neurons. The locust *ITP* mRNA is alternatively spliced, with the second mRNA producing a peptide lacking antidiuretic activity (Meredith *et al.*, 1996). In *Bombyx mori*, an *ITP* RNA is expressed in the CC and in five or six putative neuroendocrine neurons of the larval brain (Endo *et al.*, 2000). *Drosophila* prepro-ITP is encoded by CG13586 at 60D5; its functions have not been investigated.

G. Pigment-dispersing factor

The pigment-dispersing factor (PDF) is a *Drosophila* peptide as predicted by a gene encoding its precursor (CG6496 at 97B2). Neuropeptides related to PDF were first studied on the basis of their hormonal activity in a crustacean pigment dispersion assay: in crustacea, α- and β-PDH (pigment-dispersing hormones) contribute to diurnal movements of pigment granules in retinal cells, and cause the dispersion of granules in epithelial chromatophores (Rao and Riehm, 1993). In several insects, β-PDH-like immunoreactivity is found in a limited number of neurons (Nässel *et al.*, 1993; Homberg *et al.*, 1993). Helfrich-Förster (1995) found that a subset of β-PDH-positive neurons in the *Drosophila* brain is also positive for *period* gene expression. That observation suggested PDF-related peptides might be candidate transmitters for such circadian pacemaker neurons. In other insects, β-PDH has pharmacological properties consistent with a circadian role (Pyza and Meinertzhagen, 1997; Petri and Stengl, 1997).

Park and Hall (1998) cloned *Drosophila pdf*, and described a precursor that includes a single predicted PDF peptide and is encoded by a nonrhythmically expressed RNA. Renn *et al.* (1999) described *pdf* mutant animals (apparent nulls) as viable, but with a strong behavioral circadian phenotype. *pdf*-deficient animals can entrain to light:dark cycles, but the

majority of flies are unable to sustain rhythmic daily locomotion under prolonged constant conditions. The PDF neurons survive in the peptide mutant background, and the behavioral deficits can be largely attributed to lack of *pdf* expression. This phenotype is closely mimicked by selective ablation of PDF neurons following their expression of the *hid* and *reaper* cell death genes (Renn *et al.*, 1999). These and other data (e.g., Blanchardon *et al.*, 2001; Helfrich-Förster *et al.*, 2000; Taghert *et al.*, 2001) support the hypothesis that *pdf* encodes the principal circadian transmitter in *Drosophila* (reviewed by Taghert, 2001). Further data describing *pdf* gene regulation are reviewed in Section VII.F. The functions of PDF neuropeptide that is released by noncircadian neurons in *Drosophila* have not been defined (but see Persson *et al.*, 2001).

H. RFamides

The RFamides are biologically active neuropeptides with a C-terminal Arg-Phe-amide sequence. The best known member is the molluscan cardioexcitatory FMRFamide. In *Drosophila* there are at least five different genes that encode neuropeptides with predicted RFamide C termini: *dFMRFamide* (CG2346 at 46C5-6), *dromyosuppressin* (*dms*, CG6440 at 95F6-9), *drosulfakinin* (*dsk*, CG18090 at 82A1), *neuropeptide F* (*NPF/NP-PP*, CG10342 at 89D5-6), and *head peptide* (CG13968 at 38B4). Peptides derived from all five precursors have been isolated and chemically characterized in *Drosophila*.

dFMRFamide is among the best characterized *Drosophila* neuropeptide genes: it encodes a complex precursor with multiple, sequence-related RFamide-like peptides. Three of these have been isolated and chemically identified (Nambu *et al.*, 1988; Nichols, 1992a). It is expressed exclusively in the central nervous system in a limited number of neurons (Schneider *et al.*, 1993a; Nichols *et al.*, 1999a). Although the pattern of gene expression is well established (Schneider *et al.*, 1991, 1993b; O'Brien *et al.*, 1991), the literature disagrees regarding the extent to which the dFMRFamide precursor displays cell type-specific protein processing (reviewed by Taghert, 1999).

There are no mutant alleles of *dFMRFamide* yet reported, but the physiologic effects of dFMRFamide-related peptides at the *Drosophila* larval neuromuscular junction (NMJ) and on the heart have been described. Hewes *et al.* (1998) reported that seven of eight predicted peptides from this precursor all increased nerve-induced, body wall muscle contractions in most segments. These actions likely mimic blood-borne peptide actions, because dFMRFamide peptides are expressed by only two efferent larval neurons (Schneider *et al.*, 1993a). The seven active peptides had similar dose–response curves, and when the peptides were presented in mixtures predicted by their stoichiometric ratios in the precursor, their effects were additive. One of the peptides

(DPKQDFMRFamide) was also tested for effects on synaptic transmission at the NMJ, and it increased the size of the excitatory junctional current. A different result was reached in a study of dFMRFamide peptide action on the *Drosophila* pupal heart (Johnson *et al.*, 2000). Only two of the same eight peptides (DPKQDFMRFamide and PDNFMRFamide) were active in that assay, and both were inhibitory to heart rate. Nichols *et al.*, (1999c) found that only one of three dFMRFamide peptides tested (SDNFMRFamide) was inhibitory to the adult heart rate *in vivo*. It should be noted that the *in vivo* heart assay is not as extensively described for adults as it is for pupae (Johnson *et al.*, 1997). These results together suggest a widespread action of dFMRFamide peptides on both somatic and cardiac muscle, as previously found in other insects (e.g., Walther *et al.*, 1984; Evans and Myers, 1986; and also reviewed by Nässel, 1996). Further, they suggest the likelihood that dFMRFamide peptides interact with more than a single receptor (see Section V.D).

Myosuppressin is encoded by single-copy genes in both the cockroach (Donly *et al.*, 1996) and in *Drosophila* (CG18090 at 82A1). The cockroach gene is expressed by both neurons and midgut endocrine cells (Fuse *et al.*, 1998); in *Drosophila*, specific antibodies indicate its expression in the CNS (Nichols *et al.*, 1997), but its expression in midgut endocrine cells has not been reported. *Drosophila* myosuppressin has been isolated and chemically identified (Nichols, 1992a). In cockroaches, myosuppressin inhibits muscle contractions of the hindgut and the heart. In *Drosophila* the peptide inhibits the heart (Johnson *et al.*, 2000), and it mildly enhances nerve-induced contractures in larval body wall muscle (Hewes *et al.*, 1998).

The sulfakinins have a C-terminal DYsGHMRFamide sequence, in which the *Ys* indicates a sulfated tyrosine. The presence of this sulfated tyrosine, which is essential for biological activity (Nachman *et al.*, 1986a,b), supports the hypothesis that these peptides are evolutionarily related to gastrin and cholecystokinin. This point has also been debated (e.g., Duve *et al.*, 1995). Most if not all insect species appear to have two sulfakinins, which in the case of *Drosophila* have been purified (Nichols, 1992b) and are encoded by the same gene (CG6440 at 95F8, Nichols *et al.*, 1988). Specific antibodies to the *Drosophila* sulfakinins have shown these peptides to be produced by a limited number of peptidergic neurons in the central nervous system, and the gene does not appear to be expressed in the endocrine cells of the midgut (Nichols and Lim, 1996). The sulfakinins stimulate hindgut contractions, but no physiological effects have been reported for *Drosophila*.

In insects, neuropeptide F has been isolated and chemically identified only from *Drosophila*; it is encoded as a single-copy gene (CG10342 at 89D5; Brown *et al.*, 1999). Neuropeptide F has limited sequence similarity to the vertebrate peptides NPY, PP, and peptide YY. Its function is unknown, but it is

expressed by both midgut endocrine cells and by brain neurons. Expression in the brain appears sensitive to gustatory stimulation by sugars (Brown *et al.*, 1999; Shen and Cai, 2001).

Most insect "head peptides" have a C-terminal LRLRFamide sequence. In *Drosophila*, however, CG13968 encodes two predicted head peptides that display the following C terminus: RLRWamide. Sequencing of the *Drosophila* head peptides proved again that peptide isolation is a useful complement to gene sequencing: only SPSLRLRFamide was isolated from fly tissues (P. Trifilieff and J. A. Veenstra, unpublished data), and not the peptides AQRSPSLRLRFamide and WFGDVNQKPIRSPSLRLRFamide that were predicted by consideration of conventional consensus cleavage sites (Vanden Broeck, 2001). Immunohistology for *Drosophila* RLRWamides suggests their common precursor is expressed by a large number of CNS neurons, by peripheral neurons, and by midgut endocrine cells (P. Trifilieff and J. A. Veenstra, unpublished results). Brown *et al.* (1994) suggested that in mosquitoes, a head peptide inhibits host-seeking behavior, whereas Cerstiaens *et al.* (1999) found that locust head peptides accelerate vitellogenesis.

I. Allatostatins

The first allatostatins were identified from the cockroach *Diploptera punctata* (Woodhead *et al.*, 1989; Pratt *et al.*, 1991). These peptides were subsequently shown to be encoded by a single mRNA that yields a putative precursor containing as many as 13 peptides, most of which have the C-terminal sequence YXFGLamide (Donly *et al.*, 1993). The YXFGLamides, also known as the allatostatins-A, are now known to be ubiquitously present in insects. They are produced by CNS neuroendocrine cells and neurons, as well as midgut and peripheral endocrine cells (for review see Bendena *et al.*, 1999). They function as allatostatins in certain hemimetabolous species, but this is unlikely to be the case in *Drosophila*. In most insect species, including flies (Duve and Thorpe, 1994), they inhibit hindgut contractions and this may well be so in *Drosophila*, but it remains to be demonstrated. A *Drosophila* gene (*Ast* at 96A22-23) encoding four putative YXFGLamides is expressed in the CNS, in peripheral neurons, and in midgut endocrine cells (Yoon and Stay, 1995; Lenz *et al.*, 2000). Receptors sensitive to the allatostatin-A peptides are described in Section V.C.

A second family of allatostatins was identified from the cricket *Gryllus bimaculatus* (Lorenz *et al.*, 1995) and found to be homologous to neuropeptides previously identified from a lepidopteran (Blackburn *et al.*, 1995) and a locust (Schoofs *et al.*, 1991a), where they inhibit hindgut contractions. Those peptides are variously called B-allatostatins or galanin-like, for their limited similarity to

galanin [W(X_6)Wamide]. In *Drosophila* a single gene called *MIP* (CG6456 at 74B1) encodes five putative W(X_6)Wamide peptides, and is expressed in both neurons and gut endocrine cells (Williamson *et al.*, 2001). So far, no physiological effects of the predicted peptides have been described for *Drosophila*.

A third allatostatin has been isolated from the tobacco hornworm moth (Kramer *et al.*, 1991). A *Drosophila* homolog of this gene is encoded by gene *Ast2* (CG14919 at 32D2-3), which is expressed in the CNS as well as in gut endocrine cells. No physiological effects are known in *Drosophila*.

J. Pyrokinins

The pyrokinins have the C-terminal structure PRLamide or PRVamide. The first pyrokinins were identified from cockroaches and locusts by their ability to stimulate hindgut contractions in cockroaches (Holman *et al.*, 1986; Schoofs *et al.*, 1991b). Pheromone biosynthesis-activating neuropeptide (PBAN), diapause hormone, and CAP_{2b}, the three pyrokinins isolated from Lepidoptera, have rather different physiological effects. PBAN is a hormone that stimulates the production and release of female sex pheromones in moths, and has been identified from several species (for review see Teal *et al.*, 1996). Diapause hormone is released during the pupal stage by the female silkworm and acts on the ovary to induce embryonic diapause in eggs she will lay as an adult (for review see Yamashita, 1996). Both PBAN and diapause hormone are produced from a single precursor by neuroendocrine cells in the subesophageal ganglion (Sato *et al.*, 1993, 1994). The same precursor also produces three other pyrokinins, at least one of which also stimulates pheromone biosynthesis. CAP_{2b} has been identified from *Manduca sexta* abdominal ganglia as a cardioacceleratory peptide (Huesmann *et al.*, 1995), but in *Drosophila* it stimulates fluid secretion by the Malpighian tubules (Section II.F).

There are two *Drosophila* pyrokinin genes: *hug* (CG6371 at 87C1-C2) and *capa* (CG15520 at 99D1). The *hug* gene product is predicted to produce two pyrokinins and the *capa* gene three pyrokinins. *capa* is expressed in neuroendocrine cells in the abdominal and labial neuromeres, as well as in some neurons in the brain (see Section II.F). Of the three *capa* pyrokinins, only CAPA-1 and CAPA-2 stimulate fluid secretion by the Malpighian tubules (Kean *et al.*, 2002). The structure of the third peptide, CAPA-3, is similar to that of diapause hormone, and related peptides from cockroaches, but no physiological effects of CAPA-3 are known in *Drosophila*. Physiological effects of *hug* gene products have not been reported so far. For fleshflies pyrokinins have been reported to accelerate pupariation (Zdarek *et al.*, 1997); it is not clear

whether these effects are due to activation of receptors of pyrokinins, or to those of ETH peptides (Section II.C), which display similar C-terminal sequences.

K. IFamide

IFamide is the *Drosophila* homolog of a 12-amino acid peptide called "LFamide" that was purified from the blowfly *Neobellieria*, and which stimulates oviduct contractions in *Locusta migratoria* (Janssen *et al.*, 1996). IFamide (predicted by *SIFamide*, CG4681 at 60D5-7) is identical to LFamide, except for its penultimate amino acid residue. The peptide has been isolated from *Drosophila* and its structure confirmed by Edman degradation (S. Terhzaz and J. A. Veenstra, unpublished data). The peptide is expressed exclusively in four peptidergic neurons in the *Drosophila* brain (S. Terhzaz and J. A. Veenstra, unpublished data), as in *Neobellieria* (Janssen *et al.*, 1996). Its function is unknown: as there are no neurohemal release sites, it seems unlikely that the peptide stimulates oviduct contractions directly.

L. Allatotropin

Allatotropin was identified from *Manduca*, where it stimulates the CA to produce juvenile hormone (Kataoka *et al.*, 1987). Homologous peptides have been identified from the migratory locust, the mosquito, and the Colorado beetle (Paemen *et al.*, 1991; Spittaels *et al.*, 1996; Veenstra and Costes, 1999). Allatotropin has been extensively studied in *Manduca*: the gene is alternatively spliced to produce at least three different mRNAs, which could yield three additional allatotropin-related peptides (Horodyski *et al.*, 2001). The gene is expressed both in neurons and neuroendocrine cells. In Lepidoptera, the peptide has been show to have additional effects, including stimulation of the rate of heart beat (Veenstra *et al.*, 1994), inhibition of ion transport in the midgut (Lee *et al.*, 1998), and stimulation of foregut contractions (Duve *et al.*, 1999). In the locust, the allatotropin homolog stimulates oviduct contractions (Paemen *et al.*, 1991), whereas in the cockroach it increases hindgut contractions and increases the frequency of heart beat (Rudwall *et al.*, 2000). The allatotropin antiserum does not detect specific immunoreactive molecules in *Drosophila*, by either competitive enzyme-linked immunosorbent assay (ELISA) or by immunohistology. This is also true in the blowfly *Phormia* (J. A. Veenstra, unpublished data) and in the honeybee (H. Kaatz and J. A. Veenstra, unpublished data). Nevertheless, in the blowfly, oviduct contractions are stimulated at 1 nM concentrations by the mosquito homolog (Veenstra and Costes, 1999). In addition, *Manduca* allatotropin increases the production of juvenile hormone in honeybees,

albeit in relatively high concentrations (Rachinsky and Feldlaufer, 2000). These observations support a hypothesis that, although no allatotropin immunoreactivity can be detected in them, allatotropin-like receptors are functional in these species. If that is true, then perhaps allatotropin-like peptides may be functional as well. Thus, BLAST searches using the various allatotropin homologs (or combinations thereof) have so far proved unsuccessful in finding a *Drosophila allatotropin* gene. However, considering the evolutionary relatedness of blowflies and *Drosophila*, we suggest that such a gene may exist.

M. Locustatachykinins

The locustatachykinins are neuropeptides that stimulate hindgut contractions in cockroaches. Locustatachykinins have been identified from a variety of insect species, and are produced both by neurons and midgut endocrine cells (for review see Nässel, 1999). In *Drosophila*, there is a single tachykinin gene (*tk*, CG14734 at 87 A8-9); it encodes five authentic locustatachykinins, and one peptide with similarity to them (Siviter *et al.*, 2000). *dtk* and its encoded peptides are expressed in both neurons and midgut endocrine cells (Siviter *et al.*, 2000). In insects, these peptides stimulate contractions of the hindgut, and in some species those of the foregut as well. In locusts, locustatachykinins are known to innervate the glandular cells of the CC that produce AKH (Section II.E), and cause AKH release (Section IV.B). The locustatachykinins have also direct effects on identified neurons, and in *Drosophila* the predicted peptides stimulate midgut contractions in the nanomolar range. Two *Drosophila* receptors that are sensitive to tachykinin-related peptides are described in Section V.C.

N. Proctolin

The pentapeptide proctolin (H-Arg-Tyr-Leu-Pro-Thr-OH) has a special place in the pantheon of insect neuropeptides—it was the first to be purified and sequenced (Brown and Starratt, 1975; Starratt and Brown, 1975). Proctolin was purified on the basis of a bioassay that monitored its potent myotropic activity in insect visceral muscle (reviewed by Orchard *et al.*, 1980). For example, when perfused onto an isolated oviduct of the locust, proctolin caused spontaneous contractile activity by increasing muscle tonus, and increasing the amplitude and frequency of phasic contractions (Holman and Cook, 1985). Numerous other neuropeptides also have modulatory effects on visceral muscles and, in many cases, these are independent regulatory systems. For example, both locustatachykinins and proctolin affected spontaneous gut

contractions in locusts, but the two peptide systems did not synergize (Winther *et al.*, 1998). Proctolin is also a transmitter of certain motor neurons that innervate slow somatic muscles in Orthoptera (e.g., Witten and O'Shea, 1985). In such slow muscles, the peptide produced a sustained tonic contracture without affecting membrane potential (Adams and O'Shea, 1983), and in myogenic somatic muscle of the same animals, it caused increases in the contraction rhythm (Evans, 1984). A proctolin analog (α-methyl-L-tyrosine-2-proctolin) has potent and specific antagonist activity (Norohna *et al.*, 1997).

In Diptera, proctolin-like immunoreactivity is found in 500–600 neurons throughout the blowfly CNS (e.g., Nässel and O'Shea, 1987; Cantera and Nässel, 1991). In *Drosophila*, proctolin was identified biochemically (Anderson *et al.*, 1988). It is found by immunostaining in several neurons throughout the larval CNS, including gut efferents, neuroendocrine neurons, and in a class of motor neurons that innervates a specific somatic muscle (no. 12). By the size and appearance of the stained nerve terminals, these proctolin-immunoreactive motor neurons appear identical to those also shown to be insulin-like immunoreactive (Gorczyca *et al.*, 1993). Proctolin decreased the *Drosophila* heart rate at all development stages (Zornik *et al.*, 1999), but increased it in other insect species (Sliwowska *et al.*, 2001). No one has yet cloned a *proctolin* gene in any organism. As discussed below, the small size of bioactive peptides such as proctolin appears to increase the difficulty in identifying their cognate genes by genomic methods alone. Nevertheless, CG7105 (at 28D3-4) encodes a small protein consisting of a signal peptide followed immediately by the proctolin pentapeptide sequence and a single arginine residue, which would be predicted to be cleaved correctly by insect convertases (Veenstra, 2000). Although there is no definite proof that this is an authentic preproproctolin, cleavage by the signal peptidase at the N-terminal site would explain why antiserum directed against the proctolin sequence flanked by arginine residues on both the N and C termini does not recognize proctolinergic neurons in insects (J. A. Veenstra, unpublished data).

O. Amnesiac

The *Drosophila* memory mutant *amnesiac* (*amn*) is one of several single-gene mutants that are defective in memory (Tully and Quinn, 1985). The *amn* gene (CG11937 at 19A1) was cloned (after considerable effort) as a second site suppressor of the female sterility phenotype of the learning/memory mutant *dunce* (Feany and Quinn, 1995). That work described a potential *amn* gene product of 170 amino acids that displays certain features of a neuropeptide

precursor, with regions similar to vertebrate pituitary adenylylcyclase-activating peptide (PACAP). *amn* has since attracted great interest because the hypothesis of neuropeptide involvement provides a specific basis with which to consider cellular interactions that may underlie the consolidation of memories (Kandel and Abel, 1995). The hypothesis derives additional significance from the correspondence between the effects of learning/memory mutations on cAMP metabolism (reviewed by Davis, 1996), and the potential sequence homology of the AMN gene product to PACAP, which is known in mammals to modify that same second-messenger pathway.

In fact, that sequence homology is not strong. In addition, there are several other features of the *amn* gene product, and its expression, that place it in contrast to other, well-established neuropeptide precursors. First, the predicted *Drosophila* propeptide contains several cysteine residues that are likely to change the conformation of the peptide in ways different from that displayed by pro-PACAP. Second, the consensus convertase cleavage site (Lys-Arg) is preceded by two cysteine residues, which are likely to interfere significantly with cleavage of this site (see Veenstra, 2000). Third, the expression of the *amn* RNA is apparently not robust enough to be detected histologically (Waddell *et al.*, 2000): this feature differs markedly from that described for most all other neuropeptide genes. A detailed analysis of AMN propeptide processing will complement and strengthen the ongoing molecular and genetic analyses of *amn*.

The expression of *amn* has been described directly with a specific antibody, and indirectly with several marker P elements inserted at or near the *amn* locus (Waddell *et al.*, 2000). There is an apparent correspondence between several of these complex patterns in two large protocerebral neurons [dorsal paired median (DPM) neurons] that appear to innervate (superficially) the calyxes of the mushroom bodies. Other "AMN-expressing" cells are found near the mushroom bodies, but not in them; their total number is not known. Likewise, potential sites of AMN expression beyond that portion of the anterior brain, and any expression outside the CNS, have not yet been described. Sophisticated genetic manipulations of *amn* cells (that render them conditionally active within minutes) support the hypothesis that some or all *amn*-expressing neurons regulate aspects of olfactory memory consolidation (Waddell *et al.*, 2000). Restoring wild-type *amn* sequences with an *amn*–GAL4 driver rescued the mutant behavior associated with *amn* alleles: that result supported a proposed prominent role of the DPM neurons in mediating *amn* functions (Waddell *et al.*, 2000). However, the same GAL4 lines also produced lethality when used to drive cell death proteins, such as Reaper (Waddell *et al.*, 2000). That result suggests the GAL4 pattern may be more extensive, and also possibly diverse. In a related study, Rosay *et al.* (2001) found that *amn* affected the amplitude of spontaneous rhythmic calcium

waves in mushroom body neurons. Together these observations are consistent with a hypothesis by which putative AMN-derived neuropeptides are delivered in paracrine fashion to mushroom body neurons to regulate their activities.

An alternative hypothesis considers possible developmental roles for the *amn* gene product, as described below. The *amn* gene has also been implicated in *Drosophila* ethanol sensitivity (Moore *et al.*, 1998). A P element allele of *amn* (called *cheapdate*) renders the fly more sensitive to alcohol intoxication. Consistent with the involvement of *amn* in cAMP metabolism, several other single-gene mutations that also perturb cAMP levels can affect the same behavior. Importantly, the *cheapdate* phenotype can be fully reversed by short-term restoration of wild-type *amn* activity (Moore *et al.*, 1998). The latter observation is consistent with the hypothesis that *amn* acts on a short time scale to affect ethanol sensitivity (and does not affect some developmental aspect that controls, for example, ethanol clearance). In contrast, the same rescuing *amn* transgene cannot be used with that induction paradigm to rescue the memory defect associated with *amn* flies. It can rescue such deficits, but only if the *amn* transgene is activated earlier in development (DeZazzo *et al.*, 1999). Thus putative AMN neuropeptides likely have additional and obligate developmental roles in the regulation of memory consolidation.

Hence, these current studies of *amnesiac* have led to the classic consideration of which, if either, contributes more to a particular phenotype—a molecule's physiological or developmental properties? For the case of ethanol sensitivity, *amn* appears to work physiologically without an antecedent developmental function. For the case of memory regulation, it appears less certain. Last, *amn* mutant animals display a slower heart beat in pupal animals (Johnson *et al.*, 1997).

P. Corazonin

Corazonin is a somewhat mysterious peptide that was initially identified from the cockroach *Periplaneta*, where it stimulates the frequency of heart beat at concentrations as low as 0.01 nM (Veenstra, 1989c). However, this pronounced effect appears limited to certain species of cockroaches (Predel *et al.*, 1994). Corazonin at 100 nM increases the frequency of heart beat in *Manduca* by about 10% (Y. Chen and J. A. Veenstra, unpublished data). Its [His^7] homolog has been identified in migratory locusts as responsible in part for the color differences between gregarious and solitary forms (Tawfik *et al.*, 1999). The identification of a lepidopteran cDNA encoding corazonin suggests it is the only biologically active peptide produced from its precursor (Hansen *et al.*, 2001).

Corazonin is produced by both neuroendocrine cells in the pars intercerebralis, and by neurons; its distribution within the CNS has been described for the cockroach *Periplaneta* (Veenstra and Davis, 1993) and the fly *Phormia* (Cantera *et al.*, 1994). Expression of the *Drosophila corazonin* gene (*crz*, CG3302 at 87B7-8) has been established by *in situ* hybridization and immunohistology (J. A. Veenstra and L. Costes, unpublished data) and is essentially identical to that described for *Phormia* (Cantera *et al.*, 1994). An earlier report on the identification of *crz* (Veenstra, 1994) did not correctly identify the intron, which was identified by Adams *et al.* (2000). Elimination of the corazonin-expressing neurons and neuroendocrine cells in *Drosophila*, using *crz*–GAL4, and UAS–*hid* and UAS–*rpr*, does not affect body color in *Drosophila* (G. Isabel and J. A. Veenstra, unpublished data). The function of corazonin in *Drosophila* remains to be determined.

Q. *Drosophila* neuropeptide genes not found

There are three neuropeptide genes that are well established in other insects, but which appear to be absent from the *Drosophila* genome: neuroparsin, allatotropin, and vasopressin. The first two were discussed previously; the vasopressin homolog in locusts (Proux *et al.*, 1987) is produced by two neuroendocrine cells in the subesophageal ganglion and recognized by antibodies to [Arg^8] vasopressin (Rémy and Girardie, 1980). Although [Lys^8] vasopressin-like immunoreactivity is present in the dipteran CNS (Nässel *et al.*, 1989), the staining of homologous [Arg^8] vasopressin-immunoreactive neurons in the subesophageal neuromeres has not been reported in either Lepidoptera or Diptera. The neurons are commonly present in many other insects (e.g., Veenstra *et al.*, 1984; Davis and Hildebrand, 1992). Together these facts suggest that *Drosophila* does not have a vasopressin homolog. There are other *Drosophila* neuropeptides still to be discovered: one known hormone that so far has escaped molecular identification is the large protein hormone bursicon, which is thought to promote wing cuticle sclerotization and wing inflation (Baker *et al.*, 1999). It seems likely that others are still to be identified, because not all neuropeptides are detected by simple and fast bioassays.

R. Vertebrate and invertebrate neuropeptide genes

In the early 1980s, immunoreactivity to a large number of vertebrate neurohormones was reported in numerous insect and other invertebrate species. In a few cases, insect neuropeptides have now been identified that are structurally and/or evolutionarily related to vertebrate neuropeptides (e.g.,

tachykinin-related peptides and calcitonin-related peptides). That no other "vertebrate-like" neuropeptides were recovered by *Drosophila* genomic analysis should not be surprising. The primary reason is the small size of typical bioactive peptides. Conserved neuropeptide precursors typically are divergent in most regions except the bioactive peptide(s), even between closely related species such as *Drosophila melanogaster* and *D. virilis* (Taghert and Schneider, 1990). Hence the likelihood of a homology-based identification is reduced as the lengths of homologous sequences are so restricted. Thus homologs of the smaller neuropeptides, if they are present in *Drosophila*, will need to be characterized chemically before their gene can be identified.

For example, the structural similarity between the locustatachykinins and the vertebrate tachykinins is so limited that, if it were not for their respective receptors (which do show clear homologies), the arguments for their common evolutionary origin would be rather weak. Indeed, it seems likely that the constraints put on the neuropeptide receptors in order to remain functional may be greater than those put on the neuropeptides. We know that there are, for example, two *Drosophila* receptors belonging to the clade that includes the vertebrate gastrin-releasing peptide/bombesin receptors (see Section V.B). Consequently, it seems likely that the *Drosophila* genome encodes one or more ligands related to those vertebrate peptides—however, if they do exist, they have so far escaped detection. Likewise, the relatedness of a particular *Drosophila* neuropeptide to vertebrate neuropeptides may be currently obscure (e.g., PDF). However, such relations may be clarified in future by virtue of the identity of their receptor(s), once defined.

S. Do insect neuropeptides have dedicated functions?

Most of the insect neurohormones discussed so far are produced by either a few cells or cell types, suggestive of a single-gene function. Even in the case of CCAP, which is produced by a larger variety of different cells, there is the "sense" that this is so. For example, perhaps during eclosion, there are a large number of individual muscles that must be activated in order to support the performance of a complex eclosion behavior. In most such cases, an impression is easily gained that a particular neuropeptide is associated with a single function (broad, but well defined) that is conserved in insect phylogeny. This may be a false interpretation in many instances. This issue is apparent in the case of the various diuretic hormones. For example, whereas leukokinins in several insect species act on both the Malpighian tubules and the hindgut, CAP_{2b} (or any of its homologs) is diuretic in *Drosophila* and antidiuretic in the bloodsucking bug *Rhodnius* (Quinlan *et al.*, 1997). For a number of other insect neuropeptides, the situation is even less clear: either they are produced by a

variety of cell types and appear to have a variety of effects, or the only physiological effects known are the (generic) stimulation of visceral muscle contractions.

T. Non-Peptide co-transmitters

Peptidergic neurons often coexpress and corelease non-peptide transmitters (e.g., Church *et al.*, 1993). To what extent do *Drosophila* peptidergic neurons display evidence of expressing non-peptide transmitters? In contrast to the situation in vertebrates, reports of peptide co-transmitter expression in *Drosophila* (and in other insects) are not common. Yasuyama *et al.* (1996) found choline acetyltransferase-like immunosignals in median neurosecretory neurons of *Drosophila*: that observation is consistent with the hypothesis that these peptidergic neurons coexpress a cholinergic phenotype along with a peptidergic phenotype. Stevenson (1999) examined expression of the candidate transmitter taurine in identifiable neurons of the locust and found it colocalized with certain peptidergic (RFamide-positive) neurons in the locust. Likewise, Persson and Nässel (1999) found that a large number of sensory neurons in locusts display peptidergic properties. Because all insect sensory neurons are thought to be cholinergic, these data suggest the hypothesis that, as with vertebrate peripheral neurons, insect peripheral neurons may display different combinations of peptide and non-peptide transmitters. Finally, Rudwall *et al.* (2000) colocalized serotonin immunosignals in some allatotropin neurosecretory neurons that project to and around the heart of the cockroach. This pattern was sexually dimorphic and was displayed in females, but not males.

III. NEUROPEPTIDE BIOSYNTHESIS

A. Overview

Secretory peptides derive from precursor proteins that undergo endoproteolytic cleavage and further processing to generate one or more final bioactive products (reviewed by Zhou *et al.*, 1999). Whereas the basic outline of secretory peptide biosynthesis has advanced from biochemical and molecular studies, genetics has contributed to the understanding of these events *in vivo* (e.g., Naggert *et al.*, 1995; Jackson *et al.*, 1997; Furuta *et al.*, 1997; Wang *et al.*, 1998; Westphal *et al.*, 1999). Such work has provided evidence of functional roles played by specific endoproteases and peptidases, and has also suggested novel, alternative processing mechanisms (e.g., Dong *et al.*, 1999). Further genetic analysis in a model system such as *Drosophila* could help overcome difficulties that arise from

Table 1.3. *Drosophila* **Genes Contributing to, or Potentially Involved in, Neuropeptide Biosynthesis**

Gene	CG number	Location	Alleles
dfurin1	CG10772	96D1	No
dfurin2	CG18734	14C1	No
dPC2 (*amontillado*)	CG6438	97C3	No
dPHM	CG3832	60B1–2	Yes
dPAL1	CG12130	46C6–7	No
dPAL2	CG5472	59F4–6	No
ACER	CG10593	29D1	Yes
ANCE	CG10593	34D7	Yes
ANCE2	CG16869	34D7	No
ANCE3	CG17988	34D8	No
ANCE4	CG8196	45A6	No
ANCE5	CG10142	60E3	No
prolyl endoprotease	CG5355	31E	No
silver	CG4122	1B3	Yes
silver-related	CG4678	15A2	No

studying limiting amounts of critical biosynthetic components. In this section, we highlight studies of *Drosophila* genes encoding putative orthologs of known prohormone-processing enzymes (Table 1.3). We feature those for which there exists evidence to suggest similar functional roles.

B. Endoproteolytic cleavage

Neural and endocrine peptide precursors are cleaved by endoproteolytic enzymes that recognize defined *cis* sequences. In mammals, the best studied are the prohormone convertases (PCs), which are members of the subtilisin-like, serine endopeptidase family of enzymes (reviewed by Seidah *et al.*, 1998). Their substrates include polypeptide hormone precursors, growth factors, and viral envelop glycoproteins. For example, the PC1/3 and PC2 endopeptidases have well-defined roles in processing numerous neuropeptide precursors in both neural and endocrine tissues, and both are packaged within secretory granules. Both enzymes are themselves activated by posttranslational processing, and those processes are regulated by inhibitory accessory proteins [proSAAS for PC1/3 (Fricker *et al.*, 2000) and 7B2 for PC2 (Fortenberry *et al.*, 1999)]. Association with accessory proteins affects the rate of zymogen conversion and enzyme activation.

Veenstra (2000) reviewed the sequence information supporting the hypothesis that insect neuropeptide precursors are often cleaved at dibasic and

monobasic residues. In *Drosophila*, there are three genes highly related to PCs, and at least some of these appear to be involved in neuropeptide biosynthesis. The *dfurin1* (CG10772 at 96D1) and *dfurin2* genes (CG18734 at 14C1) genes were identified by sequence homology to mammalian furin (Roebroek *et al.*, 1991, 1992). When tested with specific substrates *in vitro*, both have enzymatic activities similar to that of mammalian furin, and different from that of mammalian PC2 (De Bie *et al.*, 1995). *dfurin1* is expressed in a small number of identified peptidergic neurons (Roebroek *et al.*, 1995; Jiang *et al.*, 2000). No phenotypic analysis has been reported for either *dfurin* gene.

The third *Drosophila* PC-like gene called *amontillado* (*amon*, CG6438 at 97C3), is highly related to mammalian PC2 (Siekhaus and Fuller, 1999). *amon* RNA is first expressed at embryonic stages in both the CNS and gut, in a pattern likely to include numerous peptidergic neurons. Reported alleles of *amon* include deletions of its genomic region (Siekhaus and Fuller, 1999); a possible allele is also represented by a closely placed P element (Taghert *et al.*, 2001). Deletion analysis indicated that *amon* deficiencies produce lethality due to a distinctive behavioral phenotype: embryos are morphologically normal through late stages, but cannot produce a stereotyped behavioral sequence that helps them to hatch from the eggshell (Siekhaus and Fuller, 1999). It is presumed that *amon* is required for production of one or more secretory peptides needed to display the normal hatching behavior. AMON enzymatic activity has not been examined, but the recombinant protein does display predicted interactions with both *Drosophila* and mammalian 7B2s (Hwang *et al.*, 2000).

Angiotensin-converting enzyme (ACE) is a Zn^{2+} peptidyl-dipeptidase that plays an important role in blood pressure homeostasis in mammals. The BDGP (Berkeley *Drosophila* Genome Project) lists five *Drosophila* gene sequences related to ACE (*ANCE*, CG10593 at 34D7; *ACER*, CG10593 at 29D1; *ANCE2*, CG16869 at 34D7; *ANCE4*, CG8196 at 45A6; and *ANCE5*, CG10142 at 60E3). A sixth (*ANCE3*, CG17988 at 34D8) is present with high BLASTP scores when using either ANCE or ACER as queries. Only ACER and ANCE are predicted to be active enzymes (Coates *et al.*, 2000a). The active sites of ANCE and ACER have structural features that are (respectively) highly similar to those of the two active domains of mammalian somatic ACE. In other insects, ACE protein distribution is consistent with a role in neuropeptide processing; it is especially enriched in neurons producing peptides of the FXPRLamide family (Schoofs *et al.*, 1998). Insect ACE can process neuropeptide precursor intermediates to remove C-terminal dibasic residues (Isaac *et al.*, 1998). In *Drosophila*, there is one recorded allele of *ACER*, a recessive lethal, and two recessive lethals associated with *ANCE*. An ANCE hypomorph displays male sterility and that observation is consistent with the known expression of the enzyme in the insect male gonad (cited in Isaac *et al.*, 2000). The consequences of these mutations on neuropeptide biosynthesis have not yet

been evaluated. On the basis of limited sites of neural and imaginal disc expression, a *Drosophila prolyl endoprotease* (CG5355 at 31E) has been postulated to play a role in either neuropeptide biosynthesis or degradation (Amin *et al.*, 1999).

C. Carboxypeptidase cleavage

After precursor cleavage by the proprotein convertases, C-terminal basic residues are removed by a separate carboxypeptidase (CP) activity. Of the >15 characterized mammalian metallocarboxypeptidase enzymes, only one—CPE/H (Fricker, 1988)—was thought to be dedicated to the processing of neural and endocrine peptides. However, genetic analysis later indicated the possibility that CPs other than CPE/H may also be involved: the mouse mutant *fat* represents a severe-to-null mutation of *CPE*, yet it is viable (Naggert *et al.*, 1997). Among candidate, compensating *CP* genes, *CPD* was judged the most compelling (Dong *et al.*, 1999). CPD contains three potential enzymatic domains sites, only two of which display all the sequence features expected of an active metallocarboxypeptidase catalytic site (Xin *et al.*, 1997). CPD is related to duck gp180, a membrane-bound, multidomain protein that also represents a receptor for duck hepatitis B virus (Eng *et al.*, 1998; Tong *et al.*, 1999). In *Drosophila*, *silver* (CG4122 at 1B3) and a second gene (CG4678 at 15A2) are the two closest orthologs to the mammalian *CP* genes associated with neural and endocrine peptide biosynthesis (Bernasconi, 1994; Settle *et al.*, 1995). *silver* (*svr*) is a complex locus encoding proteins that contain as many as four predicted enzymatic domains (Settle *et al.*, 1995). It is represented by 40 null and hypomorphic alleles; some hypomorphic alleles display alterations in wing development, and in melanization and sclerotization of cuticle (the latter phenotypes generate the gene name). The degree to which neuropeptide biosynthesis is altered in *svr* mutant animals has not been determined. No alleles of CG4678 have yet been recorded.

D. C-terminal amidation

C-terminal peptide α-amidation is a late event in the biosynthesis of secretory peptides and likely to be the rate-limiting step in many instances (Eipper *et al.*, 1993). This modification is widespread: in vertebrates, more than half the known secretory peptides are amidated, whereas in insects greater than 90% are amidated, or predicted by DNA sequence to be amidated (Hewes and Taghert, 2001). Peptide amidation is functionally significant: its absence often disrupts the activity or receptor-binding properties of peptide ligands (reviewed by Kulathila *et al.*, 1999). Finally, secretory peptides are the principal substrates for amidation—the enzymes catalyzing this reaction are exclusively associated

with a lumenal intracellular compartment (Eipper *et al.*, 1993). The modification results from the sequential actions of two enzymes: peptidylglycine α-hydroxylating monooxygenase (PH) and peptidyl α-hydroxyglycine α-amidating lyase (PAL) (Eipper *et al.*, 1993). PHM creates hydroxylated intermediates from prohormone precursor cleavage products that terminate in glycine residues. PAL cleaves the intermediates to produce the final amidated peptides and glyoxylate.

In vertebrates, the two enzymes occupy adjacent domains of a bifunctional protein called peptidylglycine α-amidating monooxygenase (PAM; Eipper *et al.*, 1993), which is incorporated into nascent secretory granules. In *Drosophila*, the PHM and PAL enzyme activities are both present, yet they are physically and genetically distinct (Kolhekar *et al.*, 1997). The *Drosophila* genome predicts one unlinked *PHM* gene (CG3832 at 60B1-2) and two unlinked *PAL* genes (CG12130 at 46C6-7, and CG5472, at 59F4-6). PHM protein is expressed in nearly all tissues, in a heterogeneous, cell type-specific fashion, and PHM enzyme activity mimicks the properties of its mammalian counterpart (Kolhekar *et al.*, 1997). The *PHM* gene mutates to give a lethal phenotype that is at least superficially similar of those of *amon* and *CAPS* (see above, and Section IV.D): a late embryo/early larval lethal phase, with no indication of gross morphological abnormality. *PHM* nulls and severe hypomorphs lack detectable PHM enzymatic activity; they display severely altered neuropeptide biosynthesis at both larval and adult stages (Jiang *et al.*, 2000). *PHM* nulls often die during early larval molts, whereas synthetic hypomorphs die attempting to complete later molts, or pupal head eversion. These features are similar to phenotypes exhibited by several steroid receptor genes (e.g., Lam *et al.*, 1999), and provide novel genetic evidence for already suspected functional overlap between the roles of secretory peptides and of steroid hormones in regulating the molt cycle.

IV. NEUROPEPTIDE RELEASE

A. Overview

In this section we consider four sets of studies that have addressed mechanisms that underlie the release of neuropeptides in insects, especially in *Drosophila* (Table 1.4). We first examine those studies that have identified factors that release or inhibit specific neuropeptides. Second, we describe physiological studies of insect peptidergic neurons that aim to define the specific components of membrane excitability that are associated with peptide release. Next, we feature studies that have used *Drosophila* genetics to define molecular components underlying neuropeptide release. Finally, we consider studies of

Table 1.4. *Drosophila* **Genes Potentially Involved in Neuropeptide Release or Reception**[a]

Gene	CG number	Location	Alleles
Lark	CG8597	65F2–3	Yes
CAPS	CG18026	102F5–8	Yes
IA-2	CG4355	21E1	No
ICA-69	CG10556	78B4	No
dmd NaCl/ppk	CG3478	35A1	No

[a]*Note: A complete list of Drosophila peptide hormone G protein-coupled receptor genes is found in Hewes and Taghert (2001).*

molecules that are specifically enriched within/on peptide-containing dense core vesicles (DCVs) in *Drosophila*.

B. Releasing factors for neuropeptide release

Several peptide factors that influence AKH release from the glandular corpora cardiaca have been identified. In locusts, tachykinin-like peptides (TKs) (Nässel *et al.*, 1995, 1999) and CCAP (Veelaert *et al.*, 1997) both induce AKH release. AKH release in locusts is also inhibited by neuropeptides: SchistoFLRFamide (a myosupressin-related peptide; Section II.H) decreases 3-isoltyl methxant (IBMX)-induced AKH release (Vullings *et al.*, 1998). In moths, two peptides that trigger ecdysial behavior—eclosion hormone (EH) and ecdysis-triggering hormone [ETH; and preecdysis-triggering normone (PETH)]—have mutually excitatory interactions. EH and PETH/ETH stores are rapidly depleted in the minutes to hours before ecdysis. The hypothesis of mutual releasing activities by the two peptides has been tested by applying the peptides to completely isolated EH or ETH cells. PETH and ETH peptides are released in an episodic burst by Inka cells minutes before production of ecdysis-related behaviors in the moth *Manduca* (Zitnan *et al.*, 1999). Those events release >90% of the peptides from cellular stores (~1 pmol) within 15–20 min. EH secreted by the CNS into the blood acts directly on the segmentally repeated, glandular Inka cells of the tracheal system that produce ETH: physiological doses of EH produce a rapid rise in cGMP and a large-scale decrease in stored ETH shortly thereafter (Ewer *et al.*, 1997; Kingan *et al.*, 1997, 2001). Likewise, ETH acts on isolated EH neuronal cell bodies by inducing a large increase in spiking activity (Ewer *et al.*, 1997), consistent with the presumption that EH peptide stores are being released (cf. Truman and Copenhaver, 1989). To produce their releasing activities, EH and ETH act as hormones. In contrast, TK- and

SchistoFLRFamide-secreting neurons project directly to AKH-secreting endocrine cells in the CC, consistent with direct secretomotor AKH-releasing activities (Nässel *et al.*, 1999; Vullings *et al.*, 1998).

A final component of the neuropeptide hierarchy controlling ecdysial behaviors is CCAP, which is released by a set of segmentally repeated, CNS neurons following increases in intracellular cGMP levels (Ewer and Truman, 1996). Gammie and Truman (1999) presented evidence to argue that EH acts directly on CCAP neurons (or on neurons presynaptic to them) to cause CCAP release. In this instance, the route EH peptide takes to reach responsive neurons is thought to be within the CNS as a paracrine modulator (Hewes and Truman, 1991). Interactions between EH, ETH, and CCAP have also been studied in *Drosophila* (Baker *et al.*, 1999; Park *et al.*, 1999), and at least some of these interactions are conserved.

With respect to non-peptide-releasing factors, Aizono *et al.* (1997) demonstrated muscarinic receptor immunosignals on prothoracicotropic hormone (PTTH)-expressing neuronal cell bodies of the silkworm. In *Drosophila*, an antibody to a muscarinic receptor stained neurosecretory neurons of the median protocerebrum (Harrison *et al.*, 1995). These data are consistent with the hypothesis that muscarinic input may regulate PTTH and/or other neuropeptide release by brain peptidergic neurons. Similarly, Richter *et al.* (2000) observed a releasing effect of melatonin on PTTH activity in the cockroach brain.

C. Physiological activity of peptidergic neurons

Physiological recordings of insect peptidergic neurons have been described in different insects including *Drosophila*. Cross-correlation analysis of spiking activity in neurons that express insulin-like peptides (bombyxins) in the silkworm *Bombyx* indicated a weak, mutual coupling (Ichikawa, 2001). Bickmeyer *et al.* (1994) demonstrated multiple inward (calcium) currents in peptidergic cell bodies of the locust. In *Manduca*, the eclosion hormone neurons release their peptide at a stereotyped time in development, just before eclosion or ecdysis. Hewes and Truman (1994) correlated a change in spike activity and in spike threshold with the time of release. Further, Hewes (1999) isolated eclosion hormone cell bodies within hours of normal release and observed calcium-dependent inward currents, and two distinct, outward potassium currents: a $I_{K(Ca)}$ and a transient, voltage-dependent A current (I_A). Gammie and Truman (1997) correlated increased cGMP levels within an identified CCAP neuron in *Manduca* with a lowering of the spike threshold and increased amounts of spiking activity. Likewise, perfusion of a cGMP analog into this

neuron had similar effects; the influence of cGMP on peptidergic neuron excitability appeared to be mediated by effects on calcium levels.

In *Drosophila*, patch-clamp recordings from type III axon terminals on a larval body wall muscle (no. 12) defined macroscopic and microscopic currents present in peptidergic nerve terminals (Martinez-Padron and Ferrus, 1997; Morales *et al.*, 1999). Such terminals are peptidergic by ultrastructure (Martinez-Padron and Ferrus, 1997), and by evidence from light microscopy: they are immunostained by anti-insulin (Gorczyca *et al.*, 1993), anti-proctolin (Anderson *et al.*, 1988), and anti-PHM antibodies (N. Jiang and P. H. Taghert, unpublished data). Their functional role is uncertain because they do not elicit an electrical response from the muscle (Martinez-Padron and Ferrus, 1997). The peptidergic terminals display complex outward currents that can be discriminated by genetics and physiology into at least three types (Martinez-Padron and Ferrus, 1997); these include Shal-like, but not Shaker-type, conductances. The terminals contain cysteine string protein-like immunosignals (as do glutamatergic terminals) that are hypothesized to modulate calcium currents. However, the absence of *csp* in these terminals (in a mutant background) did not prevent the normal appearance of calcium currents (Morales *et al.*, 1999).

D. Genetic studies

Three studies have specifically addressed neuropeptide release by using genetics as the primary method. The release of CCAP neuropeptide implicated in eclosion behaviors has also been studied with respect to circadian control. The *lark* gene (CG8597 at 65F2–3) has an influence on the circadian periodicity of eclosion behavior in *Drosophila*: it delays the phase but does not affect its period (Newby and Jackson, 1996). Furthermore, *Lark* rhythm defects are behavior specific, in that they affect rhythmic eclosion, but not rhythmic daily locomotion. *Lark* is a recessive embryonic lethal, and its circadian behavioral effects are displayed by heterozygous mutant animals. LARK protein is an RNA-binding factor that is ubiquitously expressed in nuclei throughout the CNS and other tissues. However, in many CCAP neurons, LARK protein is distinguished by its cytoplasmic location, and by a circadian periodicity in abundance (McNeil *et al.*, 1998; Zhang *et al.*, 2000). The rhythm in LARK abundance peaks during the day, and is dependent on the circadian clock. In other insects, homologous CCAP peptidergic neurons have been implicated in triggering eclosion behavior motor patterns (Ewer and Truman, 1996; Gammie and Truman, 1999). Together these data are consistent with a hypothesis that invokes daily LARK inhibition of CCAP action by translational regulation of a factor(s) required for CCAP neuropeptide release.

The genetic analysis of CAPS expression also addresses mechanisms of peptide release. CAPS was first discovered on the basis of its activity in reconstituting Ca^{2+}-dependent norepinephrine secretion in permeabilized PC12 cells (Walent *et al.*, 1992). These and later studies showed that CAPS functions at a late stage of regulated exocytosis, and that it is essential for Ca^{2+}-dependent secretion of DCVs, but not of synaptic vesicles (Ann *et al.*, 1997; Elhamdani *et al.*, 1999). In *Drosophila*, genetic analysis indicates that *CAPS* (CG18026 at 102F5-8) mutates to produce a lethal phenotype (late embryo, or early larva), with little morphologic defect (Renden *et al.*, 2001). This phenotype is remarkably similar to those observed in animals deficient in the neuropeptide biosynthetic enzymes encoded by *amon* (dPC2) and *dPHM* (see Section III.D). *CAPS* is normally expressed throughout the CNS, and is present at all synaptic sites. At the neuromuscular junction, *CAPS* mutants present a 50% reduction in glutamatergic neurotransmission. They also display 40% fewer synaptic vesicle profiles, and three times more DCV profiles. Although the lethality is rescued by *CAPS* overexpression, it is not rescued when CAPS expression is limited to motor neurons exclusively. Thus Renden *et al.* (2001) suppose that *CAPS* phenotypic deficits in motor neurons represent a non-cell-autonomous result of disrupting DCV secretion (possibly neuropeptides) by other cells.

Further insight into the details of neuropeptide release mechanisms has come from parallel studies of synaptic vesicle release. In particular, the knowledge that tetanus toxin can decrease physiological activity in neurons by cleaving synaptobrevin (and so reducing the incidence of synaptic vesicle exocytosis) led to a novel genetic design: the targeted expression of the toxin subunit is designed to specifically silence select neuronal populations (Sweeney *et al.*, 1995). This method has been used successfully to silence diverse neuron cell types in *Drosophila*, including sensory neurons (e.g., Heimbeck *et al.*, 1999), interneurons (e.g., Martin *et al.*, 1998), and motor neurons (e.g., Baines *et al.*, 2001). However, in the one example to date in which the method was applied to defined peptidergic neurons, it failed to produce a phenotype consistent with reduced neuropeptide release.

PDF neuropeptide regulates circadian locomotor rhythms: its release from identified pacemaker neurons has been inferred on the basis of two related behavioral phenotypes (Renn *et al.*, 1999). The first phenotype is displayed by animals genetically deficient in PDF neuropeptide expression, and the second by animals lacking PDF neurons following their targeted ablation. However, targeting tetanus toxin (TeTx) expression to the same PDF pacemaker neurons did not generate a similar phenotype (Kaneko *et al.*, 2000). This result could be explained by at least three alternative hypotheses. First, enough neuropeptide can be released, even under severe toxin loading, such that the behavior can still be properly organized. A second explanation invokes a form of

synaptobrevin expressed by PDF neurons (and perhaps by other peptidergic neurons) that is not efficiently cleaved by TeTx. Finally, it may be that some peptidergic release events utilize a vesicular protein entirely different from synaptobrevin.

E. Specific molecular components of dense core vesicles

A final aspect of neuropeptide release considers molecular components that are specific to, or enriched within, peptidergic neurons. Such molecules could play regulatory roles in peptide biosynthesis, accumulation, or release. Many molecules that are specifically enriched in the secretory granule pool of insulin-producing pancreatic beta cells have been identified as autoantigens from human sera (Tree *et al.*, 2000). Most notably, ICA512/IA-2 and phogrin are a pair of similar membrane proteins, related to tyrosine phosphatases (PTPs), that are specifically enriched in mammalian, peptide-containing DCVs. Both were identified as autoantigens from the sera of patients with insulin-dependent diabetes mellitus (Passini *et al.*, 1995; Hawkes *et al.*, 1996). These molecules normally span the DCV membrane with both intragranular and cytoplasmic domains. They have been hypothesized to act as signaling molecules that link secretion of neuropeptides and peptide hormones with signal transduction pathways, but it is not clear how this relates to tyrosine phosphorylation/dephosphorylation. Both IA-2 and phogrin have an Ala → Asp substitution within the substrate-binding pocket, which may limit their PTPase activities (Cui *et al.*, 1996). IA-2 has an additional Asp → Ala substitution at the opening of the binding cleft, which may limit its activity further. Such inactive enzymes may function to mask or sustain the lifetimes of phosphorylated substrates (Wishart and Dixon, 1998).

IA-2 homologous genes have been identified in both *Caenorhabditis elegans* (Cai *et al.*, 2001; Zahn *et al.*, 2001) and *Drosophila* databases (Cai *et al.*, 2001; CG4355 at 21E1). Remarkably, in the case of the *C. elegans* gene product, at least one of the same critical residues within the catalytic domain is altered. In addition, the fly and worm proteins are both predicted to lack a cysteine residue known to be required for PTP activity (Cai *et al.*, 2001; Wishart and Dixon, 1998). These observations are consistent with the hypothesis that at least some members of the *IA-2* gene family lack PTP activity, and that this may be a phylogenetically conserved feature. In *C. elegans*, reporter constructs of *phogrin*-like gene reveal expression in a subset of peptidergic neurons (Zahn *et al.*, 2001), or panneuronally (Cai *et al.*, 2001). Orthologous gene expression in *Drosophila* embryos is limited to a subset of CNS neurons (Cai *et al.*, 2001); no genetic analysis of *dIA-2* has yet been

reported. ICA69 is another autoantigen from human insulin-secreting cells of unknown function (Pilon *et al.*, 2000): BLASTP reports CG10556 (at 78B4) as a highly related gene.

V. NEUROPEPTIDE RECEPTORS

A. Overview

Neuropeptides and peptide hormones signal primarily via G protein-coupled receptors (GPCRs). These receptors define a large gene family (Bockaert and Pin, 1999) that provides sensitivity to a variety of environmental, developmental, and physiological signals. GPCRs display a uniform topology with seven transmembrane (TM) domains, whose conserved features (e.g., Baldwin *et al.*, 1997; Tams *et al.*, 1998) permit the use of sequence analysis to identify and classify those devoted to (for example) peptide hormones and neuropeptide ligands ("peptide GPCRs"). In a review, Brody and Cravchik (2000) began the process of categorizing *Drosophila* GPCRs by describing some 100 genes, including 21 receptors for classic neurotransmitters and neuromodulators (biogenic amines, related compounds, and purines) and 26–30 peptide GPCRs. More recently, Hewes and Taghert (2001) expanded the set of known and candidate peptide GPCRs to ~44. They reanalyzed the genome annotations according to phylogenetic trees constructed with the aid of the Pfam seven-TM databases (Bateman *et al.*, 2000), and incorporated information deduced by examining gene organizations (http://thalamus.wustl.edu/flyGPCR/peptideGPCR.html). Given such descriptions, we abbreviate remarks here to a brief discussion of the 44 peptide GPCRs, and add references concerning reports that were published more recently.

B. Major peptide G protein-coupled receptor groups

Drosophila peptide GPCRs contain representatives related to at least 15 monophyletic vertebrate GPCR subgroups. Peptide GPCRs are classified within family A (rhodopsin-like) and family B (secretin-like) groups of the GPCRs (http://www.gpcr.org). The set of known and candidate *Drosophila* peptide GPCRs contains 39 representatives of family A and 5 of family B. There are at least 19 *Drosophila* peptide GPCRs belonging to family A/group III-B. Two of these bear strong relation to vertebrate gastrin/cholecystokinin receptors (CG6881 and CG6857, both at 17D3). There are six GPCRs related to vertebrate neurokinin receptors [CG6515 (*NKD*) at 86D, CG7887 (*DTKR*) at 99D, CG5811 (*NepYR*) at 97D14, CG10626 at 67B5, CG10823 at 93D1,

and BACR48G21.1 at 45F4]. Four fly GPCRs bear similarity to the neuropeptide Y receptor subgroup (CG1147 at 83D2, CG7395 at 76F1, CG12610 at 17B1, and CG13995 at 26A8). The bombesin/gastrin-releasing peptide receptor group contains two *Drosophila* representatives (CG14494 and CG14593 at 41F10). At least six *Drosophila* GPCRs are related to neurotensin and neurotensin-related receptors (CG8784 and CG8795 at 87E6, CG14575 at 78C7, CG9918 at 88A7, CG5911 at 93D1, and CG14003 at 25F4). An additional seven genes (CG2114 at 63A1, CG5936 at 17A2, CG6986 at 4C8, CG8985 at 62C4, CG13229 at 47C6, CG13803 at 62C4, and CG16727 at 67B5) may also be members of the neurotensin-related receptor subgroup. In fact, all seven appear to be most closely related to a large set of orphan receptors that were previously identified in *C. elegans* (C. Bargmann, personal communication).

There are 11 representatives of the family A/group V among *Drosophila* peptide GPCRs, and these displayed similarities with 7 vertebrate GPCR subgroups. Of these, the galanin-, somatostatin/opioid-like receptors include four loci [CG2872 (*AlstR*) at 3F1, CG7285 at 75C4, CG10001 (*DAR-2*) at 98E2, and CG13702 at 75C4], whereas the gonadotropin-releasing hormone- and oxytocin/vasopressin-like receptors include three loci [CG6111 at 96F7, CG10698 at 69A4, and CG11325 (*DRHR*) at 26D7]. There are four genes related to the vertebrate glycoprotein hormone-like receptors [CG4187 at 11D1, CG5042 at 96E9, CG7665 (*Fsh*) at 90C1, and CG8930 (*rk*) at 34E1]. Finally, among the five *Drosophila* peptide GPCRs that belong to family B, four are strongly related to one or two vertebrate GPCR subgroups: the calcitonin-like receptor and corticotropin-releasing factor-like receptor [CG4395 at 11D1, CG17415 at 49F2, CG8422 (*DHR1*) at 50F6, and CG12370 (*DHR2*) at 48F5]. The fifth (CG13758 at 3A7) has uncertain origins, but may be distantly related to the corticotropin-releasing factor-like receptor clade. Thus, a large majority of the *Drosophila* and vertebrate neuropeptide GPCR signaling pathways appear to share common evolutionary origins. Functional and genetic analysis of individual receptor genes will be needed to determine whether the functions of these signals have been similarly conserved.

C. Physiologically defined *Drosophila* peptide G protein-coupled receptors

Although more than half of the *Drosophila* peptide GPCRs are orphans, several have been cloned and their ligands identified (discussed by Hewes and Taghert, 2001). Here we briefly review that information. Several groups have now shown that two *Drosophila* receptors that are related to galanin receptors (CG2872 and CG10001) both specifically bind various peptides of the *Drosophila* allatostatin A-type precursors (drostatin A's: Birgul *et al.*, 1999;

Lenz *et al.*, 2000; Larsen *et al.*, 2001). Allatostatin receptors have also been studied in other insect species. A *Bombyx* receptor highly similar to both *DAR*-encoded protein and sensitive to *Bombyx* type A allatostatins was cloned (Secher *et al.*, 2001). Likewise Auerswald *et al.* (2001) reported a cockroach receptor that displays 60% amino acid identity in the TM regions when compared with either *DAR* gene product, and a fragment of a similar receptor from the stick insect. Finally, Bowser and Tobe (2000) used a radioligand-binding assay to study allatostatin-binding sites in the cockroach gut.

Tachykinin signaling in insects has been reviewed (Vanden Broeck *et al.*, 1999). The DTKR protein when functionally expressed in oocytes displays binding to vertebrate tachykinin peptides (Li *et al.*, 1991), whereas the NKD protein similarly binds an insect (locust) tachykinin with specificity (Monnier *et al.*, 1992). A related tachykinin-like receptor from the stable fly binds insectotachykinin peptides with specificity (Torfs *et al.*, 2000). The two corticotrophin-releasing factor-like receptors (CG8422 and CG12370) are orthologous to the diuretic hormone-binding receptors (DIHRs) of the hawkmoth (Reagan, 1994), the cricket (Reagan, 1996), and the silkworm (Ha *et al.*, 2000). Hence those two receptors are good candidates to bind the peptides encoded by the two *Drosophila* diuretic hormone genes (*Dh*, CG8348 at 85E2; and *Dh31*, CG13094 at 29D1). By similar extrapolation, CG10626 is likely to function as a leukokinin receptor because it appears highly similar to the defined molluscan lymnokinin receptor that was shown to bind a leukokinin-like peptide (PSFHSWSamide; Cox *et al.*, 1997). A similar receptor has been cloned from the cattle tick (Holmes *et al.*, 2000). Finally, whereas *NepR* (CG5811) was cloned on the basis of its sequence similarity to a mammalian neuropeptide Y receptor (Li *et al.*, 1992), functional studies demonstrate *NepR* selectivity for peptides related to the RFamide family (e.g., the molluscan PQRFamide-like peptides; St.-Onge *et al.*, 2000).

D. Non-G protein-coupled receptor peptide receptors

Some neuropeptide signaling also occurs via ionotropic signaling (Cottrell, 1997). The molluscan neuropeptide FMRFamide directly gates an amiloride-sensitive sodium channel in *Helix* neurons (Lingueglia *et al.*, 1995). FMRFamide and mammalian RFamide-related neuropeptides also modulate mammalian members of the DEG/ENaC channel family: during acidification, they induce sustained currents in such voltage-insensitive, proton-gated channels (Askwith *et al.*, 2000). Darboux *et al.* (1998) and Adams *et al.* (1998) described a *Drosophila* ortholog of the amiloride-sensitive sodium channel family, *dmdNaCl/ppk* (CG3478 at 35A1), and both groups speculate, on the basis of its cellular

Table 1.5. *Drosophila* **Genes Contributing to, or Potentially Involved in, Neuropeptide Degradation.**[a]

Gene	CG number	Location	Alleles
Nep 1	CG5894	5D1	No
Nep 2	CG9761	82D2	No
Nep 3	CG9565	19A3	No
DPP III	CG7415	84F9	No
IDE	CG5517	75E5	No
IDE-related	CG2025	10F2	No
IDE-related	CG10588	78A2	No
TPP II	CG3991	49F7	No

[a]*Note: Some relevant genes are also listed in Table 1.3.*

distribution, that it could be involved in mechanotransduction. There are no mutants of this gene reported so far.

VI. NEUROPEPTIDE DEGRADATION

A. Overview

Neuropeptide signaling is terminated by a combination of mechanisms that include receptor desensitization or internalization, peptide diffusion, and enzymatic degradation by specific peptidases (reviewed by Grady *et al.*, 1997). The half-lives of a few insect neuropeptides have been studied and found not to be significantly different from those of vertebrates (Reynolds *et al.*, 1979; Quistad *et al.*, 1984; Veenstra, 1989b). In this section we review work on the identification of peptidases that have been implicated in terminating insect peptide signaling (Table 1.5).

B. Neprilysin

Neprilysin (enkephalinase) is a key enzyme in the metabolism of mammalian neuropeptides (Turner *et al.*, 2001). It is enriched on neuronal membranes as a synaptic ectoenzyme, where it can inactivate critical neuropeptides within the brain such as enkephalin and various tachykinins (Matsas *et al.*, 1983), and circulating hormones such as atrial natriuretic peptide. In locusts, NEP-like enzyme activity is enriched in synaptic membranes (Isaac, 1988) and has been implicated in the inactivation of the circulating locust neuropeptide AKH (Rayne and O'Shea, 1992). Isaac *et al.* (2000) report finding 16 *Drosophila*

genes of the NEP family in the annotated genome, and of these 3 are now called *Nep* genes: *Nep1* (CG5894 at 5D2), *Nep2* (CG9761 at 82D2), and *Nep3* (CG9565 at 19A3). By BLASTP, all *Nep* genes are closely related to neprilysin, and *Nep3* is the most closely related to endothelin-converting enzyme I.

C. Angiotensin-converting enzymes

ACE enzymes are zinc metalloproteases that remove dipeptides from the C termini of substrates. Such activity is responsible for activation of angiotensin II from angiotensin I, but it also contributes to the inactivation of particular peptides such as bradykinin. The *ACER/ANCE* family of genes in *Drosophila* was described previously (Section III.B). ACE enzyme activity in insects displays an inhibitor sensitivity similar to that of its mammalian counterpart (Lamango and Isaac, 1993), but the two *Drosophila* enzymes ANCE and ACER differ in their abilities to cleave synthetic substrates (Houard *et al.*, 1998). Several insect peptides, especially members of the insect tachykinin family, are good *in vitro* substrates for ACE from the housefly and for *Drosophila* ANCE (Lamango *et al.*, 1997; Nachman *et al.*, 1998). *ANCE* and *ACER* genes are expressed at nonoverlapping times during embryogenesis and pupal development, suggesting a further distinction between their respective roles (Houard *et al.*, 1998; Taylor *et al.*, 1996; Tatei *et al.*, 1995).

D. Aminopeptidases

In mammalian tissues, prolyl endoprotease (dipeptidyl-peptidase IV; DPP IV) is an inactivator of specific neuropeptides that have a Pro-2 residue (Mentlein, 1999). Numerous insect neuropeptides display a Pro-2 (e.g., *Drosophila* DPKQDFMRFamide) and this occurrence has prompted study of similar activities in cockroaches and in *Drosophila*. In cockroach tissues, a DPP IV-like activity is enriched in brain neuropils such as mushroom bodies, and it displays a similar substrate specificity (Nässel *et al.*, 2000). In *Drosophila*, CG5355 (at 31E; Amin *et al.*, 1999) predicts a prolyl endopeptidase and is expressed in subsets of cells both in the CNS and in imaginal tissues. Mazzocco *et al.* (2001) identified a different membrane-bound aminopeptidase from cockroach tissues by studying the inactivation of a specific neuropeptide, proctolin. The activity cleaves dipeptides from the amino terminus: it is immunologically similar to rat liver dipeptidyl aminopeptidase (DPP III), and was inhibited by a DPP III-specific compound. A DPP III-like sequence is also present in the *Drosophila* genome (Mazzocco *et al.*, 2001; CG7415 at 84F9). An aminopeptidase that removes arginine from proctolin (RYLPT) has been localized to synaptic membranes of the locust CNS (Isaac, 1997).

E. Intracellular degrading enzymes

Insulinase (insulin-degrading enzyme, IDE) is thought to contribute to degradation and clearance of insulin, but its precise actions have not yet been resolved (Authier *et al.*, 1996). IDE is related to the larger family of N-arginine dibasic (NRD) convertases, metalloendopeptidases from rat brain cortex and from testis that cleave peptide substrates on the N terminus of arginine residues in basic doublets (Chesneau *et al.*, 1994). IDE is a cytosolic enzyme, with particular localization to peroxisomes, and so is thought to aid in intracellular insulin clearance, or perhaps even to participate in insulin signaling (Duckworth *et al.*, 1998). *Drosophila* has a closely related *ide* gene (CG5517 at 77B5); by BLASTP search, human IDE recovers two other fly genes (CG2025 at 10F2, and CG10588 at 78A2) that are broadly related to the NRD convertase family. The possible relationship of *Drosophila ide* to insulin signaling (see Section II.B) has not been addressed.

Tripeptidyl peptidase II (TPP II) is an extra lysosomal peptidase with a subtilisin-like catalytic domain that cleaves N-terminal tripeptides (Balow *et al.*, 1986; Tomkinson *et al.*, 1987). It has wide distribution and participates in non-proteosome-mediated protein degradation (Geier *et al.*, 1999). Rose *et al.* (1996) presented evidence that TPP II is a highly specific cholecystokinin (CCK)-degrading peptidase in rat brain and included a demonstration of its specific inhibition by a designed compound, butabindide. A subsequent study (Facchinetti *et al.*, 2000) reexamined TTP II localization and found evidence of widespread TPP II expression in the brain (cf. Tomkinson and Nyberg, 1995). It found some TPP II association with CCK-rich regions, consistent with its alleged function as a CCK-specific peptidase. However, there were many CCK-rich regions lacking TPP II immunosignals, and there was no evidence consistent with enzyme association with the outside plasma membrane. A *Drosophila* TPP II homolog (CG3991 at 49F7) was studied by Renn *et al.* (1998). This serine protease possesses the substrate specificity and catalytic properties of its mammalian counterpart, and a similar inhibitor profile as well. Overexpression of TPP II did not produce an overt phenotype nor affect the viability of the animal (S. C. P. Renn, unpublished data).

VII. DEVELOPMENTAL REGULATION OF NEUROPEPTIDE EXPRESSION

A. Overview

Secretory peptide expression is a highly regulated feature of neural and endocrine tissues. Developmentally, it appears by stereotyped patterns

Table 1.6 *Drosophila* **Genes Contributing to, or Potentially Involved in, Development of Neuropeptide Phenotypes**

Gene	CG number	Location	Alleles
tinman	CG7895	93D10	Yes
buttonless	CG5246	94B9	Yes
apterous	CG8376	41F9–10	Yes
Clock	CG7391	66A12	Yes
cycle	CG8727	76D3	Yes
period	CG2647	3B4	Yes
timeless	CG3234	23F3–5	Yes
Atonal	CG7508	84F6	Yes
ventral veins lacking	CG10037	65C5–D1	Yes

according to rules of cellular differentiation. However, such expression can be plastic and is subject to alteration by later stage changes in hormonal milieu (e.g., Loi and Tublitz, 1993), by influences from postsynaptic targets (e.g., Asmus *et al.*, 2000), and by general physiological parameters such as injury (e.g., Hokfelt *et al.*, 2000). *Drosophila* genetics have been instrumental in addressing a number of central questions in cellular differentiation and tissue development. In this section, we review studies that have defined developmental aspects of neuropeptide expression in *Drosophila*, or that have framed issues for future developmental analysis (Table 1.6).

B. Developmental expression: embryonic

Developmental schedules of neuropeptide RNA and peptide expression in embryos have been described for several different neuropeptides in *Drosophila* (e.g., Schneider *et al.*, 1991, 1993a; Friedman *et al.*, 2001) and also for other insect species (e.g., Broadie *et al.*, 1990; Wall and Taghert, 1991a; Westbrook and Bollenbacher, 1990; Wegerhoff *et al.*, 1996; Ludwig *et al.*, 2001). In *Drosophila* embryos, the first reported appearance of specific neuropeptides is at stage 15 (e.g., Isshiki *et al.*, 2001), and most spatial patterns that are evident in larvae are not fully developed until stage 17, or postembryonically (e.g., Schneider *et al.*, 1993a). The embryonic differentiation of spatial patterns for putative neuropeptide biosynthetic enzymes has been described for *dfur2* (Roebroek *et al.*, 1995) and for *amon* (dPC2; Siekhaus and Fuller, 1999).

C. Developmental expression: metamorphosis

Drosophila metamorphosis presents a remarkable transformation of external body form and internal tissues, including nerve and muscle. It is estimated that more than 90% of neurons of the adult *Drosophila* CNS are generated and differentiated postembryonically (Truman and Bate, 1988). Several neuropeptide systems have been studied with respect to their potential modifications during metamorphosis. including those of *dFMRFamide* (White *et al.*, 1986; Lundquist and Nässel, 1990; O'Brien *et al.*, 1991), and *pdf* (Helfrich-Förster, 1997). Several systems display strong conservation of cellular elements through metamorphosis including the eclosion hormone neurons (Riddiford *et al.*, 1994), and a set of myomodulin-positive tracheal endocrine cells (O'Brien and Taghert, 1998). Taghert *et al.* (2000) analyzed the differentiation of a set of *dFMRFamide*-expressing, tangential neurons of the medulla. That group is composed of both embryonic and imaginal neurons, whose neuropeptide gene expression is differentially regulated at both transcriptional and post transcriptional levels during metamorphosis.

D. Control by steroids

Many physiological and morphological neuronal properties in insects are altered during metamorphosis under the influence of stage-specific hormonal signaling by ecdysteroids and juvenile hormones (Levine *et al.*, 1995). Tublitz and Sylwester (1990) observed such changes in neuropeptide transmitter profiles in identified peptidergic neurons in the moth *Manduca*. Specifically, CAP_{2b}-producing neurons start to produce the protein hormone bursicon in response to changes in ecdysteroid levels at metamorphosis. A comparable situation was suggested for *Drosophila*: the "large" OL2 neurons of the visual system are born in the embryo, but begin to express *dFMRFamide* neuropeptides only at the start of metamorphosis, shortly following known peaks of ecdysteroids (Taghert *et al.*, 2000). The differentiation state of the large OL2 neurons before the onset of *dFMRFamide* expression, and the factors that trigger such delayed expression, have not been defined.

A strong case for the regulation of neuropeptide expression by steroid hormones comes from the study of peptide factors that regulate ecdysis. The relationship between steroid and peptide factors in the regulation of the molt cycle has long been appreciated (Henrich *et al.*, 1999). Ecdysteroids modulate the timing of eclosion and of eclosion hormone release in moths (Truman *et al.*, 1983). Most recently, ecdysteroids were shown to influence EH neuron excitability (Hewes and Truman, 1994) and to induce *ETH* gene expression (Zitnan *et al.*, 1999). A clear suggestion that such regulation of *ETH* may be direct comes from inspection of the *ETH* gene promoter in both *Manduca* and

Drosophila: it contains one or more ecdysone consensus binding sequences near to the transcription start site (Zitnan *et al.*, 1999; Park *et al.*, 1999).

E. Apoptotic death

Certain peptidergic neurons in *Drosophila* exhibit stereotyped cell death following their periods of activation. This has been documented in the cases of specific CCAP neurons (Ewer *et al.*, 1998; Draizen *et al.*, 1999), or suggested in the case of certain PDF neurons (Renn *et al.*, 1999). Such developmental events presumably reflect the fact that, in *Drosophila*, neuropeptides often serve highly prescribed functions, as stage-specific triggers or modulators of behavior.

F. Regulatory mechanisms

Developmental regulation of insect neuropeptide expression has been addressed at both cellular and molecular levels. Cellular experiments have begun to define both intrinsic and extrinsic influences on classic neurotransmitter expression in insect embryos (e.g., Taghert and Goodman, 1984; Huff and Mahowald, 1989; Thor and Thomas, 1997; Lundell and Hirsh, 1998). There are fewer reports that address corresponding control of neuropeptide expression. In *Drosophila*, the Tv neurons of the thoracic neuromeres express the *dFMRFamide* gene and FMRFamide- related peptides (White *et al.*, 1986; Schneider *et al.*, 1993a). In certain genetic backgrounds (e.g., *tinman*), the peripheral nerve in which Tv neurons normally make terminations does not form. In that case, the Tv neurons do not display FMRFamide immunosignals, whereas other FMRFamide-expressing neurons appear normal (Gorczyca *et al.*, 1994). These data suggest that either the survival of the Tv neurons, or their neuropeptide expression, relies on interactions between the peptidergic neurons and some aspect of the periphery. In moth embryos, Wall and Taghert (1991a) tested a related hypothesis—that homologous Tv neurons require interactions with cellular peripheral targets for the normal onset of FMRFamide neuropeptide differentiation. However, such neurons produced a normal schedule of onset, even when the CNS was isolated *in vitro* before the exit of the axons. Those data suggest initial neuropeptide expression is autonomous to events that occur within the CNS.

In contrast to aminergic neurons (Taghert and Goodman, 1984; Huff *et al.*, 1987; Lundell and Hirsh, 1994) and to γ-aminobutyric acid (GABA)-ergic neurons (Witten and Truman, 1991), there are few studies that describe the lineages of specific peptidergic neurons in insect embryos. Isshiki *et al.* (2001) report that a particular corazonin-positive neuron (cf. Cantera *et al.*, 1994) of *Drosophila* derives from the third division of neuroblast 7–3.

Interestingly, other progeny from this stem cell include specific serotonergic neurons (Lundell and Hirsh, 1994).

The transcriptional regulation of neuropeptide genes and associated genes in *Drosophila* has also received experimental attention. Schneider *et al.* (1993b) demonstrated that *dFMRFamide* expression by ~17 different neuronal cell types within the larval *Drosophila* CNS is controlled at a transcriptional level. Further, they and others (Benveniste and Taghert, 1999) defined and placed spatial limits to at least four cell autonomous enhancer regions within the 5′ promoter and intronic regions of the *dFMRFamide* gene. When placed in combination with a heterologous promoter, these putative enhancers were capable of creating authentic, cell-specific patterns of expression *in vivo*. Benveniste *et al.* (1998) used genetics to identify a transcription factor, the LIM homeodomain protein Apterous, as a regulator of *dFMRFamide* expression in two of the normal ~17 *dFMRFamide* cell types. Apterous may regulate *dFMRFamide* in these two cell types directly, but its control is likely to include indirect aspects as well. Together those experiments support a model whereby *dFMRFamide* neuropeptide gene expression is differentially regulated among the 17 diverse neuronal cell types that commonly express that phenotype. McNabb *et al.* (1997) showed that a small 5′ upstream fragment of the *eclosion hormone* gene promoter directed accurate, cell-specific reporter gene expression to the two EH-expressing neurons in transgenic flies. In addition, Moto *et al.* (1999) transformed adult *Bombyx* brains with bombyxin promoter fragments fused to the green fluorescent protein (GFP) gene and observed accurate cell-specific reporter gene expression in eight bone fide bombyxin-expressing neurons.

Similarly, the *Drosophila pdf* neuropeptide gene is differentially regulated by transcription factors depending on cell type. *pdf* is expressed by three cell types: two sets of interneurons in the brain and one set of neuroendocrine neurons in the segmental nerve cord (Helfrich-Förster, 1997). In one set of the brain neurons, *pdf* is positively regulated by the basic helix–loop–helix (bHLH) proteins Clock and Cycle, and by the PAR domain-containing protein Vrille (Blau and Young, 1999; Park *et al.*, 2000). By contrast, *pdf* expression in the other two cell types is not affected by those specific mutant backgrounds. It is not yet known whether any of these factors control *pdf* transcription directly or indirectly. Finally, Rosay *et al.* (1995) studied transcriptional regulation of the *NKD* neuropeptide receptor gene (tachykinin-receptor like) and described its bifunctional character with *in vivo* embryonic assays. A distal region controls late expression within the CNS; a proximal region controls early and transient expression within a set of sensory neuron precursors. In transformed cell lines, the latter regulatory element is directly regulated by the bHLH protein Atonal, which is known to regulate differentiation of the sensory neuron precursors *in vivo*.

The rich *Drosophila* literature concerning embryonic differentiation presents numerous future opportunities to study the specification of neuropeptide phenotypes. One useful approach may be to analyze the functions of genes potentially orthologous to known vertebrate regulators of endocrine and neuroendocrine phenotypes, such as CG10037 (*vvl* at 65C2), similar to mammalian *Pit-1* (Ingraham *et al.*, 1988), or CG1447 (at 100B2), similar to *PTX-1* (Tremblay *et al.*, 1998). Moreover, it is not known whether the different properties that underlie a neuropeptide phenotype are tightly or loosely coordinated within a single cell type. For example the ~17 different *dFMRFamide* cell types display several different transcriptional regulatory mechanisms. Yet they share several generic "peptidergic properties." For example, they share expression of several biosynthetic enzymes such as PHM, among themselves, and with other peptidergic neurons. It is notable that neuropeptide biosynthetic enzymes, dedicated to the production of different neuropeptides, are expressed by many diverse neuronal cell types, but are not ubiquitously expressed in all cells of the CNS (Seikhaus and Fuller, 1999; Jiang *et al.*, 2000). Whether such enzyme-encoding genes are coregulated with neuropeptide genes such as *dFMRFamide* (i.e., differentially and according to cell type), or whether they are regulated by a mechanism common to all peptidergic neurons, remains an intriguing problem in neuronal differentiation.

Acknowledgments

We thank the following colleagues for generously providing personal communications, and/or for reading sections of this review in draft form: Cori Bargmann, Julian Dow, John Ewer, John Hutton, Elwyn Isaac, Rob Jackson, Erik Johnson, Mike Nonet, Phillipe Rosay, Anneliese Schaefer, and Jim Truman. P.H.T. is supported by the NIH and by the Human Frontier Scientific Program Organization.

References

Adams, M. E., and O'Shea, M. (1983). Peptide cotransmitter at a neuromuscular junction. *Science* **221,** 286–289.

Adams, C. M., Anderson, M. G., Motto, D. G., Price, M. P., Johnson, W. A., and Welsh, M. J. (1998). *ripped pocket* and *pickpocket*, novel *Drosophila* DEG/ENaC subunits expressed in early development and in mechanosensory neurons. *J. Cell Biol.* **140,** 143–152.

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X., Brandon, R. C., Rogers, Y. H., Blazej, R. G., Champe, M., Pfeiffer, B. D., Wan, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Gabor, G. L., Abril, J. F., Agbayani, A., An, H. J.,

Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E. M., Beeson, K. Y., Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K. C., Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de Pablos B., Delcher, A., Deng, Z., Mays, A. D., Dew, I., Dietz, S. M., Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C., Dunn, P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A. E., Garg, N. S., Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z., Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z., Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. R., Pacleb, J. M., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M. P., Smith, T., Spier, E., Spradling, A. C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wassarman, D. A., Weinstock, G. M., Weissenbach, J., Williams, S. M., Woodage, T., Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H., Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O., Gibbs, R. A., Myers, E. W., Rubin, G. M., and Venter, J. C. (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287,** 2185–2195.

Aizono, Y., Endo, Y., Sattelle, D. B., and Shirai, Y. (1997). Prothoracicotropic hormone-producing neurosecretory cells in the silkworm, *Bombyx mori*, express a muscarinic acetylcholine receptor. *Brain Res.* **763,** 131–136.

Amin, A., Li, Y., and Finkelstein, R. (1999). Identification of a *Drosophila* prolyl endopeptidase and analysis of its expression. *DNA Cell Biol.* **18,** 605–610.

Anderson, M. S., Halpern, M. E., and Keshishian, H. (1988). Identification of the neuropeptide transmitter proctolin in *Drosophila* larvae: Characterization of muscle fiber-specific neuromuscular endings. *J. Neurosci.* **8,** 242–255.

Ann, K., Kowalchyk, J. A., Loyet, K. M., and Martin, T. F. (1997). Novel Ca^{2+}-binding protein (CAPS) related to UNC-31 required for Ca^{2+}-activated exocytosis. *J. Biol. Chem.* **272,** 19637–19640.

Askwith, C. C., Cheng, C., Ikuma, M., Benson, C., Price, M. P., and Welsh, M. J. (2000). Neuropeptide FF and FMRFamide potentiate acid-evoked currents from sensory neurons and proton-gated DEG/EnaC channels. *Neuron.* **26,** 133–141.

Asmus, S. E., Parsons, S., and Landis, S. C. (2000). Developmental changes in the transmitter properties of sympathetic neurons that innervate the periosteum. *J. Neurosci.* **20,** 1495–1504.

Audsley, N., McIntosh, C., and Phillips, J. E. (1992). Isolation of a neuropeptide from locust corpus cardiacum which influences ileal transport. *J. Exp. Biol.* **173,** 261–274.

Auerswald, L., Birgul, N., Gade, G., Kreienkamp, H. J., and Richter, D. (2001). Structural, functional, and evolutionary characterization of novel members of the allatostatin receptor family from insects. *Biochem. Biophys. Res. Commun.* **282,** 904–909.

Authier, F., Posner, B. I., and Bergeron, J. J. (1996). Insulin-degrading enzyme. *Clin. Invest. Med.* **19,** 149–160.

Baines, R. A., Uhler, J. P., Thompson, A., Sweeney, S. T., and Bate, M. (2001). Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. *J. Neurosci.* **21,** 1523–1531.

Baker, J. D., McNabb, S. L., and Truman, J. W. (1999). The hormonal coordination of behavior and physiology at adult ecdysis in *Drosophila melanogaster*. *J. Exp. Biol.* **202,** 3037–3048.

Baldwin, J. M., Schertler, G. F., and Unger, V. M. (1997). An α-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors. *J. Mol. Biol.* **272,** 144–164.

Balow, R. M., Tomkinson, B., Ragnarsson, U., and Zetterqvist, O. (1986). Purification, substrate specificity, and classification of tripeptidyl peptidase II. *J. Biol. Chem.* **261,** 2409–2417.

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Howe, K. L., and Sonnhammer, E. L. L. (2000). The Pfam protein families database. *Nucleic Acids Res.* **28,** 263–266.

Bendena, W. G., Donly, B. C., and Tobe, S. S. (1999). Allatostatins: A growing family of neuropeptides with structural and functional diversity. *Ann. N. Y. Acad. Sci.* **897,** 311–329.

Benveniste, R. J., and Taghert, P. H. (1999). Cell type-specific regulatory sequences control expression of the *Drosophila FMRF-NH2* neuropeptide gene. *J. Neurobiol.* **38,** 507–520.

Benveniste, R. J., Thor, S., Thomas, J. B., and Taghert, P. H. (1998). Cell type-specific regulation of the *Drosophila FMRF-NH2* neuropeptide gene by Apterous, a LIM homeodomain transcription factor. *Development* **125,** 4757–4765.

Bernasconi, P. (1994). Molecular cloning of a *Drosophila melanogaster* gene coding for an homologue of human carboxypeptidase E. *Arch. Insect Biochem. Physiol.* **27,** 169–178.

Bickmeyer, U., Rossler, W., and Wiegand, H. (1994). Omega AGA toxin IVA blocks high-voltage-activated calcium channel currents in cultured pars *intercerebralis* neurosecretory cells of adult *Locusta migratoria*. *Neurosci. Lett.* **181,** 113–116.

Birgul, N., Weise, C., Kreienkamp, H. J., and Richter, D. (1999). Reverse physiology in *Drosophila*: Identification of a novel allatostatin-like neuropeptide and its cognate receptor structurally related to the mammalian somatostatin/galanin/opioid receptor family. *EMBO J.* **18,** 5892–5900.

Blackburn, M. B., Wagner, R. M., Kochansky, J. P., Harrison, D. J., Thomas-Laemont, P., and Raina, A. K. (1995). The identification of two myoinhibitory peptides, with sequence similarities to the galanins, isolated from the ventral nerve cord of *Manduca sexta*. *Regul. Pept.* **57,** 213–219.

Blanchardon, E., Grima, B., Klarsfeld, A., Chelot, E., Hardin, P. E., Preat, T., and Rouyer, F. (2001). Defining the role of *Drosophila* lateral neurons in the control of circadian rhythms in motor activity and eclosion by targeted genetic ablation and PERIOD protein overexpression. *Eur. J. Neurosci.* **13,** 871–888.

Blau, J., and Young, M. W. (1999). Cycling *vrille* expression is required for a functional *Drosophila* clock. *Cell* **99,** 661–671.

Bockaert, J., and Pin, J. P. (1999). Molecular tinkering of G protein-coupled receptors: An evolutionary success. *EMBO J.* **18,** 1723–1729.

Bohni, R., Riesgo-Escovar, J., Oldham, S., Brogiolo, W., Stocker, H., Andruss, B. F., Beckingham, K., and Hafen, E. (1999). Autonomous control of cell and organ size by CHICO, a *Drosophila* homolog of vertebrate IRS1-4. *Cell* **97,** 865–875.

Bourême, D., Tamarelle, M., and Girardie, J. (1987). Production and characterization of antibodies to neuroparsins A and B isolated from the corpora cardiaca of the locust. *Gen. Comp. Endocrinol.* **67,** 169–177.

Bowser, P. R., and Tobe, S. S. (2000). Partial characterization of a putative allatostatin receptor in the midgut of the cockroach *Diploptera punctata*. *Gen. Comp. Endocrinol.* **119,** 1–10.

Broadie, K. S., Sylwester, A. W., Bate, M., and Tublitz, N. J. (1990). Immunological, biochemical and physiological analyses of cardioacceleratory peptide 2 (CAP2) activity in the embryo of the tobacco hawkmoth *Manduca sexta*. *Development* **108,** 59–71.

Brody, T., and Cravchik, A. (2000). *Drosophila melanogaster* G protein-coupled receptors. *J. Cell Biol.* **150,** F83–F88.

Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., and Hafen, E. (2001). An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* **11,** 213–221.

Brown, B. E., and Starnatt, A. N. (1975). Isolation of proctolin, a myotropic peptide, from *Periplaneta americana*. *J. Insect Physiol.* **21,** 1879–1881.

Brown, M. R., and Cao, C. (2001). Distribution of ovary ecdysteroidogenic hormone I in the nervous system and gut of mosquitoes. *J. Insect Sci.* **1,** 3.

Brown, M. R., Klowden, M. J., Crim, J. W., Young, L., Shrouder, L. A., and Lea, A. O. (1994). Endogenous regulation of mosquito host-seeking behavior by a neuropeptide. *J. Insect Physiol.* **40,** 399–406.

Brown, M. R., Graf, R., Swiderek, K. M., Fendley, D., Stracker, T. H., Champagne, D. E., and Lea, A. O. (1998). Identification of a steroidogenic neurohormone in female mosquitoes. *J. Biol. Chem.* **273,** 3967–3971.

Brown, M. R., Crim, J. W., Arata, R. C., Cai, H. N., Chun, C., and Shen, P. (1999). Identification of a *Drosophila* brain–gut peptide related to the neuropeptide Y family. *Peptides* **20,** 1035–1042.

Cai, T., Krause, M. W., Odenwald, W. F., Toyama, R., and Notkins, A. L. (2001). The *IA-2* gene family: Homologs in *Caenorhabditis elegans, Drosophila* and zebrafish. *Diabetologia* **44,** 81–88.

Cantera, R., and Nässel, D. R. (1991). Dual peptidergic innervation of the blowfly hindgut: A light- and electron microscopic study of FMRFamide and proctolin immunoreactive fibers. *Comp. Biochem. Physiol. C.* **99,** 517–525.

Cantera, R., and Nässel, D. R. (1992). Segmental peptidergic innervation of abdominal targets in larval and adult dipteran insects revealed with an antiserum against leucokinin I. *Cell Tissue Res.* **269,** 459–471.

Cantera, R., Veenstra, J. A., and Nässel, D. R. (1994). Postembryonic development of corazonin-containing neurons and neurosecretory cells in the blowfly, *Phormia terraenovae*. *J. Comp. Neurol.* **350,** 559–572.

Cao, C., and Brown, M. R. (2001). Localization of an insulin-like peptide in brains of two flies. *Cell Tissue Res.* **304,** 317–321.

Cerstiaens, A., Benfekih, L., Zouiten, H., Verhaert, P., De Loof, A., and Schoofs, L. (1999). Led-NPF-1 stimulates ovarian development in locusts. *Peptides* **20,** 39–44.

Chen, C., Jack, J., and Garofalo, R. S. (1996). The *Drosophila* insulin receptor is required for normal growth. *Endocrinology* **137,** 846–856.

Chen, Y., Veenstra, J. A., Hagedorn, H., and Davis, N. T. (1994). Leucokinin and diuretic hormone immunoreactivity of neurons in the tobacco hornworm, *Manduca sexta*, and co-localization of this immunoreactivity in lateral neurosecretory cells of abdominal ganglia. *Cell Tissue Res.* **278,** 493–507.

Chesneau, V., Pierotti, A. R., Prat, A., Gaudoux, F., Foulon, T., and Cohen, P. (1994). *N*-Arginine dibasic convertase (NRD convertase): A newcomer to the family of processing endopeptidases [overview]. *Biochimie.* **76,** 234–240.

Chung, K. W., Oh, H. S., Kim, A. J., Won, Y. J., Kim, N. K., Namkoong, Y., Kim, K., and Lee, C. C. (1994). Isolation of small prothoracicotropic hormone-like gene in *Drosophila melanogaster*. *Korean J. Zool.* **37,** 12–18.

Church, P. J., Whim, M. D., and Lloyd, P. E. (1993). Modulation of neuromuscular transmission by conventional and peptide transmitters released from excitatory and inhibitory motor neurons in *Aplysia*. *J. Neurosci.* **13,** 2790–2800.

Clark, T. M., Hayes, T. K., and Beyenbach, K. W. (1998). Dose-dependent effects of CRF-like diuretic peptide on transcellular and paracellular transport pathways. *Am. J. Physiol.* **274,** F834–F840.

Coast, G. M., Webster, S. G., Schegg, K. M., Tobe, S. S., and Schooley, D. A. (2001). The *Drosophila melanogaster* homologue of an insect calcitonin-like diuretic peptide stimulates V-ATPase activity in fruit fly Malpighian tubules. *J. Exp. Biol.* **204,** 1795–1804.

Coates, D., Siviter, R., and Isaac, R. E. (2000). Exploring the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes to understand neuropeptide and peptidase function. *Biochem. Soc. Trans.* **28,** 464–469.

Cottrell, G. A. (1997). The first peptide-gated ion channel. *J. Exp. Biol.* **200,** 2377–2386.

Cox, K. J., Tensen, C. P., Van der Schors, R. C., Li, K. W., van Heerikhuizen, H., Vreugdenhil, E., Geraerts, W. P., and Burke, J. F. (1997). Cloning, characterization, and expression of a G-protein-coupled receptor from *Lymnaea stagnalis* and identification of a leucokinin-like peptide, PSFHSWSamide, as its endogenous ligand. *J. Neurosci.* **17,** 1197–1205.

Cui, L., Yu, W. P., DeAizpurua, H. J., Schmidli, R. S., and Pallen, C. J. (1996). Cloning and characterization of islet cell antigen-related protein-tyrosine phosphatase (PTP), a novel receptor-like PTP and autoantigen in insulin-dependent diabetes. *J. Biol. Chem.* **271,** 24817–24823.

Darboux, I., Lingueglia, E., Pauron, D., Barbry, P., and Lazdunski, M. (1998). A new member of the amiloride-sensitive sodium channel family in *Drosophila melanogaster* peripheral nervous system. *Biochem. Biophys. Res. Commun.* **246,** 210–216.

Davies, S. A. (2000). Nitric oxide signalling in insects. *Insect Biochem. Mol. Biol.* **30,** 1123–1138.

Davies, S. A., Huesmann, G. R., Maddrell, S. H., O'Donnell, M. J., Skaer, N. J., Dow, J. A., and Tublitz, N. J. (1995). CAP2b, a cardioacceleratory peptide, is present in *Drosophila* and stimulates tubule fluid secretion via cGMP. *Am. J. Physiol.* **269,** R1321–R1326.

Davies, S. A., Stewart, E. J., Huesmann, G. R., Skaer, N. J., Maddrell, S. H., Tublitz, N. J., and Dow, J. A. (1997). Neuropeptide stimulation of the nitric oxide signaling pathway in *Drosophila melanogaster* Malpighian tubules. *Am. J. Physiol.* **273,** R823–R827.

Davis, R. W. (1996). Physiology and biochemistry of *Drosophila* learning mutants [review]. *Physiol Rev.* **76,** 299–317.

Davis, N. T., and Hildebrand, J. G. (1992). Vasopressin-immunoreactive neurons and neurohemal systems in cockroaches and mantids. *J. Comp. Neurol.* **320,** 381–393.

De Bie, I., Savaria, D., Roebroek, A. J., Day, R., Lazure, C., Van de Ven, W. J., and Seidah, N. G. (1995). Processing specificity and biosynthesis of the *Drosophila melanogaster* convertases dfurin 1, dfurin1-CRR, dfurin1-X, and dfurin2. *J. Biol. Chem.* **270,** 1020–1028.

DeZazzo, J., Xia, S., Christensen, J., Velinzon, K., and Tully, T. (1999). Developmental expression of an amn^+ transgene rescues the mutant memory defect in *amnesiac* adults. *J. Neurosci.* **19,** 8740–8746.

Dong, W., Fricker, L. D., and Day, R. (1999). Carboxypeptidase D is a potential candidate to carry out redundant processing functions of carboxypeptidase E based on comparative distribution studies in the rat central nervous system. *Neuroscience* **89,** 1301–1317.

Donly, B. C., Ding, Q., Tobe, S. S., and Bendena, W. G. (1993). Molecular cloning of the gene for the allatostatin family of neuropeptides from the cockroach *Diploptera punctata*. *Proc. Natl. Acad. Sci. USA* **90,** 8807–8811.

Donly, B. C., Fuse, M., Orchard, I., Tobe, S. S., and Bendena, W. G. (1996). Characterization of the gene for leukomyosuppressin and its expression in the brain of the cockroach *Diploptera punctata*. *Insect Biochem. Mol. Biol.* **26,** 627–637.

Dow, J. A. T., and Davies, S. A. (2001). The *Drosophila* melanogaster Malpighian tubule: A genetic model for insect epithelia. *Adv. Insect Physiol.* **28,** 1–83.

Dow, J. A. T., Kelly, D., Davies, S. A., Maddrell, S. H. P., and Brown, D. (1995). A novel member of the major intrinsic protein family in *Drosophila*: Are aquaporins involved in insect Malpighian tubule fluid secretion? *J. Physiol.* **489,** P110–P111.

Draizen, T. A., Ewer, J., and Robinow, S. (1999). Genetic and hormonal regulation of the death of peptidergic neurons in the *Drosophila* central nervous system. *J. Neurobiol.* **38,** 455–465.

Duckworth, W. C., Bennett, R. G., and Hamel, F. G. (1998). Insulin degradation: Progress and potential. *Endocr. Rev.* **19,** 608–624.

Duve, H., and Thorpe, A. (1994). Distribution and functional significance of Leu-callatostatins in the blowfly *Calliphora vomitoria*. *Cell Tissue Res.* **276,** 367–379.

Duve, H., Thorpe, A., Scott, A. G., Johnsen, A. H., Rehfeld, J. F., Hines, E., and East, P. D. (1995). The sulfakinins of the blowfly *Calliphora vomitoria*: Peptide isolation, gene cloning and expression studies. *Eur. J. Biochem.* **232,** 633–640.

Duve, H., East, P. D., and Thorpe, A. (1999). Regulation of lepidopteran foregut movement by allatostatins and allatotropin from the frontal ganglion. *J. Comp. Neurol.* **413,** 405–416.

Eipper, B. A., Milgram, S. L., Husten, E. J., Yun, H. Y., and Mains, R. E. (1993). Peptidylglycine α-amidating monooxygenase: A multifunctional protein with catalytic, processing and routing domains. *Protein Sci.* **2,** 489–497.

Elhamdani, A., Martin, T. F., Kowalchyk, J. A., and Artalejo, C. R. (1999). Ca^{2+}-dependent activator protein for secretion is critical for the fusion of dense-core vesicles with the membrane in calf adrenal chromaffin cells. *J. Neurosci.* **19,** 7375–7383.

Endo, H., Nagasawa, H., and Watanabe, T. (2000). Isolation of a cDNA encoding a CHH-family peptide from the silkworm *Bombyx mori*. *Insect Biochem. Mol. Biol.* **30,** 355–361.

Eng, F. J., Novikova, E. G., Kuroki, K., Ganem, D., and Fricker, L. D. (1998). gp180, a protein that binds duck hepatitis B virus particles, has metallocarboxypeptidase D-like enzymatic activity. *J. Biol. Chem.* **273,** 8382–8388.

Evans, P. D. (1984). Studies on the mode of action of octopamine, 5-hydroxytryptamine and proctolin on a myogenic rhythm in the locust. *J. Exp. Biol.* **110,** 231–251.

Evans, P. D., and Myers, C. M. (1986). The modulatory actions of FMRFamide and related peptides on locust skeletal muscle. *J. Exp. Biol.* **126,** 403–422.

Ewer, J., and Reynolds, S. (2002). Neuropeptide control of molting in insects. *In* "Hormones, Brain, and Behavior" (D. Pfaff, ed.), Academic Press, San Diego, CA.

Ewer, J., and Truman, J. W. (1996). Increases in cyclic 3′, 5′-guanosine monophosphate (cGMP) occur at ecdysis in an evolutionarily conserved crustacean cardioactive peptide-immunoreactive insect neuronal network. *J. Comp. Neurol.* **370,** 330–341.

Ewer, J., Gammie, S. C., and Truman, J. W. (1997). Control of insect ecdysis by a positive-feedback endocrine system: Roles of eclosion hormone and ecdysis triggering hormone. *J. Exp. Biol.* **200,** 869–881.

Ewer, J., Wang, C. M., Klukas, K. A., Mesce, K. A., Truman, J. W., and Fahrbach, S. E. (1998). Programmed cell death of identified peptidergic neurons involved in ecdysis behavior in the moth, *Manduca sexta*. *J. Neurobiol.* **37,** 265–280.

Facchinetti, P., Rose, C., Rostaing, P., Triller, A., and Schwartz, J. C. (2000). Immunolocalization of tripeptidyl peptidase II, a cholecystokinin-inactivating enzyme, in rat brain. *Neuroscience* **88,** 1225–1240.

Feany, M. B., and Quinn, W. G. (1995). A neuropeptide gene defined by the *Drosophila* memory mutant *amnesiac*. *Science* **268,** 869–873.

Fortenberry, Y., Liu, J., and Lindberg, I. (1999). The role of the 7B2 CT peptide in the inhibition of prohormone convertase 2 in endocrine cell lines. *J. Neurochem.* **73,** 994–1003.

Fournier, B., and Girardie, J. (1988). A new function for the locust neuroparsins: Stimulation of water reabsorption. *J. Insect Physiol.* **34,** 309–313.

Fricker, L. D. (1988). Carboxypeptidase E. *Annu. Rev. Physiol.* **50,** 309–321.

Fricker, L. D., McKinzie, A. A., Sun, J., Curran, E., Qian, Y., Yan, L., Patterson, S. D., Courchesne, P. L., Richards, B., Levin, N., Mzhavia, N., Devi, L. A., and Douglass, J. (2000). Identification and characterization of proSAAS, a granin-like neuroendocrine peptide precursor that inhibits prohormone processing. *J. Neurosci.* **20,** 639–648.

Friedman, J., Starkman, J., and Nichols, R. (2001). DPKQDFMRFamide expression is similar in two distantly related *Drosophila* species. *Peptides* **22,** 235–239.

Friedman, S. (1985). Carbohydrate metabolism. *In* "Comprehensive Insect Physiology, Biochemistry and Pharmacology," (G. A. Kerkut, and L. I. Gilbert, eds.), vol. 10, pp. 43–76. Pergamon Press, Oxford.

Furuta, M., Yano, H., Zhou, A., Rouille, Y., Holst, J. J., Carroll, R., Ravazzola, M., Orci, L., Furuta, H., and Steiner, D. F. (1997). Defective prohormone processing and altered pancreatic islet morphology in mice lacking active SPC2. *Proc. Natl. Acad. Sci. USA* **94,** 6646–6651.

Furuya, K., Schegg, K. M., Wang, H., King, D. S., and Schooley, D. A. (1995). Isolation and identification of a diuretic hormone from the mealworm *Tenebrio molitor*. *Proc. Natl. Acad. Sci. USA* **92,** 12323–12327.

Fuse, M., Bendena, W. G., Donly, B. C., Tobe, S. S., and Orchard, I. (1998). *In situ* hybridization analysis of *leukomyosuppressin* mRNA expression in the cockroach, *Diploptera punctata*. *J. Comp. Neurol.* **395,** 328–341.

Gammie, S. C., and Truman, J. W. (1997). An endogenous elevation of cGMP increases the excitability of identified insect neurosecretory cells. *J. Comp. Physiol. A* **180,** 329–337.

Gammie, S. C., and Truman, J. W. (1999). Eclosion hormone provides a link between ecdysis-triggering hormone and crustacean cardioactive peptide in the neuroendocrine cascade that controls ecdysis behavior. *J. Exp. Biol.* **202,** 343–352.

Gao, X., Neufeld, T. P., and Pan, D. (2000). *Drosophila* PTEN regulates cell growth and proliferation through PI3K-dependent and -independent pathways. *Dev. Biol.* **221,** 404–418.

Garofalo, R. S., and Rosen, O. M. (1988). Tissue localization of *Drosophila melanogaster* insulin receptor transcripts during development. *Mol. Cell. Biol.* **8,** 1638–1647.

Geier, E., Pfeifer, G., Wilm, M., Lucchiari-Hartz, M., Baumeister, W., Eichmann, K., and Niedermann, G. (1999). A giant protease with potential to substitute for some functions of the proteasome. *Science* **283,** 978–981.

Girardie, J., Bourâme, D., Couillaud, F., Tamarelle, M., and Girardie, A. (1987). Anti-juvenile effect of neuroparsin A, a neuroprotein isolated from locust corpora cardiaca. *Insect Biochem.* **17,** 977–983.

Girardie, J., Girardie, A., Huet, J. C., and Pernollet, J. C. (1989). Amino acid sequence of locust neuroparsins. *FEBS Lett.* **245,** 4–8.

Goberdhan, D. C., Paricio, N., Goodman, E. C., Mlodzik, M., and Wilson, C. (1999). *Drosophila* tumor suppressor PTEN controls cell size and number by antagonizing the Chico/PI3-kinase signaling pathway. *Genes Dev.* **13,** 3244–3258.

Gorczyca, M., Augart, C., and Budnik, V. (1993). Insulin-like receptor and insulin-like peptide are localized at neuromuscular junctions in *Drosophila*. *J. Neurosci.* **13,** 3692–3704.

Gorczyca, M., Phyllis, R. W., and Budnik, V. (1994). The role of *tinman*, a mesodermal cell fate gene in axon pathfinding during the development of the transverse nerve in *Drosophila*. *Development* **120,** 2143–2152.

Grady, E., Bohm, S., McConalogue, K., Garland, A., Ansel, J., Olerud, J., and Bunnett, N. (1997). Mechanisms attenuating cellular responses to neuropeptides: Extracellular degradation of ligands and desensitization of receptors. *J. Inves. Dermatol. Symp. Proc.* **2,** 69–75.

Graf, R., Neuenschwander, S., Brown, M. R., and Ackermann, U. (1997). Insulin-mediated secretion of ecdysteroids from mosquito ovaries and molecular cloning of the insulin receptor homologue from ovaries of bloodfed *Aedes aegypti*. *Insect Mol. Biol.* **6,** 151–163.

Ha, S. D., Kataoka, H., Suzuki, A., Kim, B. J., Kim, H. J., Hwang, S. H., and Kong, J. Y. (2000). Cloning and sequence analysis of cDNA for diuretic hormone receptor from the *Bombyx mori*. *Mol. Cells* **10,** 13–17.

Hansen, I. A., Sehnal, F., Meyer, S. R., and Scheller, K. (2001). Corazonin gene expression in the waxmoth *Galleria mellonella*. *Insect Mol. Biol.* **10,** 341–346.

Harrison, J. B., Chen, H. H., Blake, A. D., Huskisson, N. S., Barker, P., and Sattelle, D. B. (1995). Localization in the nervous system of *Drosophila melanogaster* of a C-terminus anti-peptide

antibody to a cloned *Drosophila* muscarinic acetylcholine receptor. *J. Neuroendocrinol.* **7,** 347–352.

Hawkes, C. J., Wasmeier, C., Christie, M. R., and Hutton, J. C. (1996). Identification of the 37-kDa antigen in IDDM as a tyrosine phosphatase-like protein (phogrin) related to IA-2. *Diabetes* **45,** 1187–1192.

Heimbeck, G., Bugnon, V., Gendre, N., Haberlin, C., and Stocker, R. F. (1999). Smell and taste perception in *Drosophila melanogaster* larva: Toxin expression studies in chemosensory neurons. *J. Neurosci.* **19,** 6599–6609.

Helfrich-Förster, C. (1995). The *period* clock gene is expressed in central nervous system neurons which also produce a neuropeptide that reveals the projections of circadian pacemaker cells within the brain of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **92,** 612–616.

Helfrich-Förster, C. (1997). Development of pigment-dispersing hormone- immunoreactive neurons in the nervous system of *Drosophila melanogaster*. *J. Comp. Neurol.* **380,** 335–354.

Helfrich-Förster, C., Tauber, M., Park, J. H., Muhlig-Versen, M., Schneuwly, S., and Hofbauer, A. (2000). Ectopic expression of the neuropeptide pigment-dispersing factor alters behavioral rhythms in *Drosophila melanogaster*. *J. Neurosci.* **20,** 3339–3353.

Henrich, V. C., Rybczynski, R., and Gilbert, L. I. (1999). Peptide hormones, steroid hormones, and puffs: Mechanisms and models in insect development. *Vitam. Horm.* **55,** 73–125.

Hewes, R. S. (1999). Voltage-dependent ionic currents in the ventromedial eclosion hormone neurons of *Manduca sexta*. *J. Exp. Biol.* **202,** 2371–2383.

Hewes, R. S., and Taghert, P. H. (2001). Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. *Genome Res.* **11,** 1126–1142.

Hewes, R. S., and Truman, J. W. (1991). The roles of central and peripheral eclosion hormone release in the control of ecdysis behavior in *Manduca sexta*. *J. Comp. Physiol. [A]* **168,** 697–707.

Hewes, R. S., and Truman, J. W. (1994). Steroid regulation of excitability in identified insect neurosecretory cells. *J. Neurosci.* **14,** 1812–1819.

Hewes, R. S., Snowdeal, E. C., III Saitoe, M., and Taghert, P. H. (1998). Functional redundancy of FMRFamide-related peptides at the *Drosophila* larval neuromuscular junction. *J. Neurosci.* **18,** 7138–7151.

Hokfelt, T., Broberger, C., Xu, Z. Q., Sergeyev, V., Ubink, R., and Diez, M. (2000). Neuropeptides—an overview. *Neuropharmacology* **39,** 1337–1356.

Holman, G. M., and Cook, B. J. (1985). Proctolin, its presence in and action on the oviduct of an insect. *Comp. Biochem. Physiol. C* **80,** 61–64.

Holman, G. M., Cook, B. J., and Nachmanm, R. J (1986). Primary structure and synthesis of a blocked myotropic neuropeptide isolated from the cockroach, *Leucophaea maderae*. *Comp. Biochem. Physiol.* **C85,** 219–224.

Holmes, S. P., He, H., Chen, A. C., Ilvie, G. W., and Pietrantonio, P. V. (2000). Cloning and transcriptional expression of a leucokinin-like peptide receptor from the southern cattle tick, *Boophilus microplus* (Acari: Ixodidae). *Insect Mol. Biol.* **9,** 457–465.

Homberg, U., Wurden, S., Dirksen, H., and Rao, K. R. (1993). Comparative anatomy of pigment dispersing hormone-immunoreactive neurons in the brain of orthopteroid insects. *Cell Tissue Res.* **266,** 343–357.

Horodyski, F. M., Ewer, J., Riddiford, L. M., and Truman, J. W. (1993). Isolation, characterization and expression of the *eclosion hormone* gene of *Drosophila melanogaster*. *Eur. J. Biochem.* **215,** 221–228.

Horodyski, F. M., Bhatt, S. R., and Lee, K. Y. (2001). Alternative splicing of transcripts expressed by the *Manduca sexta* allatotropin (Mas-AT) gene is regulated in a tissue-specific manner. *Peptides* **22,** 263–269.

Houard, X., Williams, T. A., Michaud, A., Dani, P., Isaac, R. E., Shirras, A. D., Coates, D., and Corvol, P. (1998). The *Drosophila melanogaster*-related angiotensin-I-converting enzymes Acer

and Ance—distinct enzymic characteristics and alternative expression during pupal development. *Eur. J. Biochem.* **257,** 599–606.

Huang, H., Potter, C. J., Tao, W., Li, D. M., Brogiolo, W., and Hafen, E. (1999). PTEN affects cell size, cell proliferation and apoptosis during *Drosophila* eye development. *Development* **126,** 5365–5372.

Huesmann, G. R., Cheung, C. C., Loi, P. K., Lee, T. D., Swiderek, K. M., and Tublitz, N. J. (1995). Amino acid sequence of CAP2b, an insect cardioacceleratory peptide from the tobacco hawkmoth *Manduca sexta*. *FEBS Lett.* **371,** 311–314.

Huff, R., Furst, A., and Mahowald, A. P. (1989). *Drosophila* embryonic neuroblasts in culture: Autonomous differentiation of specific neurotransmitters. *Dev. Biol.* **134,** 146–157.

Hwang, J. R., Siekhaus, D. E., Fuller, R. S., Taghert, P. H., and Lindberg, I. (2000). Interaction of *Drosophila melanogaster* prohormone convertase 2 and 7B2: Insect cell-specific processing and secretion. *J. Biol. Chem.* **275,** 17886–17893.

Ichikawa, T. (2001). Mutual coupling among insect neurosecretory cells with an ultradian firing rhythm. *Neurosci. Lett.* **299,** 73–76.

Ingraham, H. A., Chen, R. P., Mangalam, H. J., Elsholtz, H. P., Flynn, S. E., Lin, C. R., Simmons, D. M., Swanson, L., and Rosenfeld, M. G. (1988). A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. *Cell* **55,** 519–529.

Isaac, R. E. (1988). Neuropeptide-degrading endopeptidase activity of locust (*Schistocerca gregaria*) synaptic membranes. *Biochem. J.* **255,** 843–847.

Isaac, R. E. (1997). Proctolin degradation by membrane peptidases from nervous tissues of the desert locust (*Schistocerca gregaria*). *Biochem. J.* **245,** 365–370.

Isaac, R., Schoofs, L., Williams, T. A., Veelaert, D., Sajid, M., Corvol, P., and Coates, D. (1998). A novel peptide-processing activity of insect peptidyl-dipeptidase A (angiotensin I-converting enzyme): The hydrolysis of lysyl-arginine and arginyl-arginine from the C-terminus of an insect prohormone peptide. *Biochem. J.* **330,** 61–65.

Isaac, R. E., Siviter, R. J., Stancombe, P., Coates, D., and Shirras, A. D. (2000). Conserved roles for peptidases in the processing of invertebrate neuropeptides. *Biochem. Soc. Trans.* **28,** 460–44.

Isshiki, T., Pearson, B., Holbrook, S., and Doe, C. Q. (2001). *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell* **106,** 511–521.

Jackson, R. S., Creemers, J. W., Ohagi, S., Raffin-Sanson, M. L., Sanders, L., Montague, C. T., Hutton, J. C., and O'Rahilly, S. (1997). Obesity and impaired prohormone processing associated with mutations in the human *prohormone convertase 1* gene. *Nat. Genet.* **16,** 303–306.

Janssen, I., Schoofs, L., Spittaels, K., Neven, H., Vanden Broeck, J., Devreese, B., Van Beeumen, J., Shabanowitz, J., Hunt, D. F., and De Loof, A. (1996). Isolation of NEB-LFamide, a novel myotropic neuropeptide from the grey fleshfly. *Mol. Cell. Endocrinol.* **117,** 157–165.

Jiang, N., Kolhekar, A. S., Jacobs, P. S., Mains, R. E., Eipper, B. A., and Taghert, P. H. (2000). *PHM* is required for normal developmental transitions and for biosynthesis of secretory peptides in *Drosophila*. *Dev. Biol.* **226,** 118–136.

Johnson, E., Ringo, J., and Dowse, H. (1997). Modulation of *Drosophila* heartbeat by neurotransmitters. *J. Comp. Physiol. [B]* **167,** 89–97.

Johnson, E., Ringo, J., and Dowse, H. (2000). Native and heterologous neuropeptides are cardioactive in *Drosophila melanogaster*. *J. Insect Physiol.* **46,** 1229–1236.

Kandel, E. R., and Abel, T. (1995). Neuropeptides, adenylyl cyclase and memory storage. *Science* **268,** 825–826.

Kaneko, M., Park, J. H., Cheng, Y., Hardin, P. E., and Hall, J. C. (2000). Disruption of synaptic transmission or clock-gene-product oscillations in circadian pacemaker cells of *Drosophila* cause abnormal behavioral rhythms. *J. Neurobiol.* **43,** 207–233.

Kataoka, H., Troetschler, R. G., Kramer, S. J., Cesarin, B. J., and Schooley, D. A. (1987). Isolation and primary structure of the eclosion hormone of the tobacco hornworm, *Manduca sexta*. *Biochem. Biophys. Res. Commun.* **146,** 746–750.

Kawakami, A., Kataoka, H., Oka, T., Mizoguchi, A., Kimura-Kawakami, M., Adachi, T., Iwami, M., Nagasawa, H., Suzuki, A., and Ishizaki, H. (1990). Molecular cloning of the *Bombyx mori* prothoracicotropic hormone. *Science* **247,** 1333–1335.

Kean, L., Broderick, K. E., Pollock, V. P., Cazenave, W., Costes, L., Graham, S., Davies, S. A., Veenstra, J. A., and Dow, J. A. T. (2002). Two new members of the CAP2b family of diuretic peptides are encoded by the gene capability in *Drosophila melanogaster*. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **282,** R1297–1307.

Kim, A. J., Cha, G. H., Kim, K., Gilbert, L. I., and Lee, C. C. (1997). Purification and characterization of the prothoracicotropic hormone of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **94,** 1130–1135.

Kingan, T. G., Gray, W., Zitnan, D., and Adams, M. E. (1997). Regulation of ecdysis-triggering hormone release by eclosion hormone. *J. Exp. Biol.* **200,** 3245–3256.

Kingan, T. G., Cardullo, R. A., and Adams, M. E. (2001). Signal transduction in eclosion hormone-induced secretion of ecdysis-triggering hormone. *J. Biol. Chem.* **276,** 25136–25142.

Kolhekar, A. S., Roberts, M. S., Jiang, N., Johnson, R. C., Mains, R. E., Eipper, B. A., and Taghert, P. H. (1997). Neuropeptide amidation in *Drosophila*: Separate genes encode the two enzymes catalyzing amidation. *J. Neurosci.* **17,** 1363–1376.

Kondo, H., Ino, M., Suzuki, H., and Iwami, M. (1996). Multiple gene copies for bombyxin, an insulin-related peptide of the silkmoth *Bombyx mori*: Structural signs for gene rearrangement and duplication responsible for generation of multiple molecular forms of bombyxin. *J. Mol. Biol.* **259,** 926–937.

Kono, T., Nagasawa, H., Kataoka, H., Isogai, A., Fugo, H., and Suzuki, A. (1990). Eclosion hormone of the silkworm *Bombyx mori*: Expression in *Escherichia coli* and location of disulfide bonds. *FEBS Lett.* **263,** 358–360.

Kopec, S. (1922). Studies on the necessity of the brain for the inception of insect metamorphosis. *Biol. Bull.* **42,** 323–342.

Kramer, S. J., Toschi, A., Miller, C. A., Kataoka, H., Quistad, G. B., Li, J. P., Carney, R. L., and Schooley, D. A. (1991). Identification of an allatostatin from the tobacco hornworm *Manduca sexta*. *Proc. Natl. Acad. Sci. USA* **88,** 9458–9462.

Kulathila, R., Merkler, K. A., and Merkler, D. J. (1999). Enzymatic formation of C-terminal amides. *Nat. Prod. Rep.* **16,** 145–154.

Lam, G., Hall, B. L., Bender, M., and Thummel, C. S. (1999). *DHR3* is required for the prepupal–pupal transition and differentiation of adult structures during *Drosophila* metamorphosis. *Dev. Biol.* **212,** 204–216.

Lamango, N. S., and Isaac, R. E. (1993). Metabolism of insect neuropeptides: Properties of a membrane-bound endopeptidase from heads of *Musca domestica*. *Insect Biochem. Mol. Biol.* **23,** 801–880.

Lamango, N. S., Nachman, R. J., Hayes, T. K., Strey, A., and Isaac, R. E. (1997). Hydrolysis of insect neuropeptides by an angiotensin-converting enzyme from the housefly, *Musca domestica*. *Peptides* **18,** 47–52.

Larsen, M. J., Burton, K. J., Zantello, M. R., Smith, V. G., Lowery, D. L., and Kubiak, T. M. (2001). Type A allatostatins from *Drosophila melanogaster* and *Diplotera punctata* activate two *Drosophila* allatostatin receptors, DAR-1 and DAR-2, expressed in CHO cells. *Biochem. Biophys. Res. Commun.* **286,** 895–901.

Lee, K. Y., Horodyski, F. M., and Chamberlin, M. E. (1998). Inhibition of midgut ion transport by allatotropin (Mas-AT) and *Manduca* FLRFamides in the tobacco hornworm *Manduca sexta*. *J. Exp. Biol.* **201,** 3067–3074.

Leevers, S. J., Weinkove, D., MacDougall, L. K., Hafen, E., and Waterfield, M. D. (1996). The *Drosophila* phosphoinositide 3-kinase Dp110 promotes cell growth. *EMBO J.* **15,** 6584–6594.

Lenz, C., Williamson, M., and Grimmelikhuijzen, C. J. (2000). Molecular cloning and genomic organization of an allatostatin preprohormone from *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun.* **273,** 1126–1131.

Lenz, C., Williamson, M., Hansen, G. N., and Grimmelikhuijzen, C. J. (2001). Identification of four *Drosophila* allatostatins as the cognate ligands for the *Drosophila* orphan receptor DAR-2. *Biochem. Biophys. Res. Commun.* **286,** 1117–1122.

Levine, R. B., Morton, D. B., and Restifo, L. L. (1995). Remodeling of the insect nervous system. *Curr. Opin. Neurobiol.* **5,** 28–35.

Li, X. J., Wolfgang, W., Wu, Y. N., North, R. A., and Forte, M. (1991). Cloning, heterologous expression and developmental regulation of a *Drosophila* receptor for tachykinin-like peptides. *EMBO J.* **10,** 3221–3229.

Li, X. J., Wu, Y. N., North, R. A., and Forte, M. (1992). Cloning, functional expression, and developmental regulation of a neuropeptide Y receptor from *Drosophila melanogaster*. *J. Biol. Chem.* **267,** 9–12.

Lingueglia, E., Champigny, G., Lazdunski, M., and Barbry, P. (1995). Cloning of the amiloride-sensitive FMRFamide peptide-gated sodium channel. *Nature* **378,** 730–733.

Loi, P. K., and Tublitz, N. J. (1993). Hormonal control of transmitter plasticity in insect peptidergic neurons. I. Steroid regulation of the decline in cardioacceleratory peptide 2 (CAP2) expression. *J. Exp. Biol.* **181,** 175–194.

Lorenz, M. W., Kellner, R., and Hoffmann, K. H. (1995). A family of neuropeptides that inhibit juvenile hormone biosynthesis in the cricket, *Gryllus bimaculatus*. *J. Biol. Chem.* **270,** 21103–21108.

Ludwig, P., Williams, L., Nässel, D. R., Reichert, H., and Boyan, G. (2001). Primary commissure pioneer neurons in the brain of the grasshopper *Schistocerca gregaria*: Development, ultrastructure, and neuropeptide expression. *J. Comp. Neurol.* **430,** 118–130.

Lundell, M. J., and Hirsh, J. (1994). Temporal and spatial development of serotonin and dopamine neurons in the *Drosophila* CNS. *Dev. Biol.* **165,** 385–396.

Lundell, M. J., and Hirsh, J. (1998). *eagle* is required for the specification of serotonin neurons and other neuroblast 7-3 progeny in the *Drosophila* CNS. *Development* **125,** 463–472.

Lundquist, T., and Nässel, D. R. (1990). Substance P-, FMRFamide-, and gastrin/cholecystokinin-like immunoreactive neurons in the thoraco-abdominal ganglia of the flies *Drosophila* and *Calliphora*. *J. Comp. Neurol.* **294,** 161–178.

Marti, T., Takio, K., Walsh, K. A., Terzi, G., and Truman, J. W. (1987). Microanalysis of the amino acid sequence of the eclosion hormone from the tobacco hornworm *Manduca sexta*. *FEBS Lett.* **219,** 415–418.

Martin, J. R., Ernst, R., and Heisenberg, M. (1998). Mushroom bodies suppress locomotor activity in *Drosophila melanogaster*. *Learn. Mem.* **5,** 179–191.

Martinez-Padron, M., and Ferrus, A. (1997). Presynaptic recordings from *Drosophila*: Correlation of macroscopic and single-channel K^+ currents. *J. Neurosci.* **17,** 3412–3424.

Matsas, R., Fulcher, I. S., Kenny, A. J., and Turner, A. J. (1983). Substance P and [Leu]-enkephalin are hydrolyzed by an enzyme in the pig caudate synaptic membranes that is identical with an endopeptidase of kidney microvilli. *Proc. Natl. Acad. Sci. USA* **80,** 3111–3115.

Mazzocco, C., Fukasawa, K. M., Raymond, A. A., and Puiroux, J. (2001). Purification, partial sequencing and characterization of an insect membrane dipeptidyl aminopeptidase that degrades the insect neuropeptide proctolin. *Eur. J. Biochem.* **268,** 4940–4949.

McNabb, S. L., Baker, J. D., Agapite, J., Steller, H., Riddiford, L. M., and Truman, J. W. (1997). Disruption of a behavioral sequence by targeted death of peptidergic neurons in *Drosophila*. *Neuron* **19,** 813–823.

McNeil, G. P., Zhang, X., Genova, G., and Jackson, F. R. (1998). A molecular rhythm mediating circadian clock output in *Drosophila*. *Neuron* **20,** 297–303.

Mentlein, R. (1999). Dipeptidyl-peptidase IV (CD26): Role in the inactivation of regulatory peptides. *Regul. Pept.* **85,** 9–24.

Meredith, J., Ring, M., Macins, A., Marschall, J., Cheng, N. N., Theilmann, D., Brock, H. W., and Phillips, J. E. (1996). Locust ion transport peptide (ITP): Primary structure, cDNA and expression in a baculovirus system. *J. Exp. Biol.* **199,** 1053–1061.

Monnier, D., Colas, J. F., Rosay, P., Hen, R., Borrelli, E., and Maroteaux, L. (1992). NKD, a developmentally regulated tachykinin receptor in *Drosophila*. *J. Biol. Chem.* **267,** 1298–1302.

Montagne, J., Stewart, M. J., Stocker, H., Hafen, E., Kozma, S. C., and Thomas, G. (1999). *Drosophila* S6 kinase: A regulator of cell size. *Science* **285,** 2126–2129.

Moore, M. S., DeZazzo, J., Luk, A. Y., Tully, T., Singh, C. M., and Heberlein, U. (1998). Ethanol intoxication in *Drosophila*: Genetic and pharmacological evidence for regulation by the cAMP signaling pathway. *Cell* **93,** 997–1007.

Morales, M., Ferrus, A., and Martinez-Padron, M. (1999). Presynaptic calcium-channel currents in normal and *csp* mutant *Drosophila* peptidergic terminals. *Eur. J. Neurosci.* **11,** 1818–1826.

Moto, K., Abdel Salam, S. E., Sakurai, S., and Iwami, M. (1999). Gene transfer into insect brain and cell-specific expression of bombyxin gene. *Dev. Genes Evol.* **209,** 447–450.

Nachman, R. J., Holman, G. M., Cook, B. J., Haddon, W. F., and Ling, N. (1986a). Leukosulfakinin, a sulfated insect neuropeptide with homology to gastrin and cholecystokinin. *Science* **234,** 71–73.

Nachman, R. J., Holman, G. M., Haddon, W. F., and Ling, N. (1986b). Leukosulfakinin-II, a blocked sulfated insect neuropeptide with homology to cholecystokinin and gastrin. *Biochem. Biophys. Res. Commun.* **140,** 357–364.

Nachman, R. J., Isaac, R. E., Coast, G. M., and Holman, G. M. (1998). Aib-containing analogues of the insect kinin neuropeptide family demonstrate resistance to an insect angiotensin-converting enzyme and potent diuretic activity. *Peptides* **18,** 53–57.

Naggert, J. K., Fricker, L. D., Varlamov, O., Nishina, P. M., Rouille, Y., Steiner, D. F., Carroll, R. J., Paigen, B. J., and Leiter, E. H. (1995). Hyperproinsulinemia in obese *fat/fat* mice associated with a carboxypeptidase E mutation which reduces enzyme activity. *Nat. Genet.* **10,** 135–142.

Nambu, J. R., Murphy-Erdosh, C., Andrews, P. C., Feistner, G. J., and Scheller, R. H. (1988). Isolation and characterization of a *Drosophila* neuropeptide gene. *Neuron* **1,** 55–61.

Nässel, D. R. (1996). Neuropeptides, amines and amino acids in an elementary insect ganglion: Functional and chemical anatomy of the unfused abdominal ganglion. *Prog. Neurobiol.* **48,** 325–420.

Nässel, D. R. (1999). Tachykinin-related peptides in invertebrates: A review. *Peptides* **20,** 141–158.

Nässel, D. R., and O'Shea, M. (1987). Proctolin-like immunoreactive neurons in the blowfly central nervous system. *J. Comp. Neurol.* **265,** 437–454.

Nässel, D. R., Holmqvist, B. I., and Moverus, B. J. (1989). Vasopressin- and proctolin-like immunoreactive efferent neurons in blowfly abdominal ganglia: Development and ultrastructure. *J. Comp. Neurol.* **283,** 450–463.

Nässel, D. R., Shiga, S., Mohrherr, C. J., and Rao, K. R. (1993). Pigment-dispersing hormone-like peptide in the nervous system of the flies *Phormia* and *Drosophila:* Immunocytochemistry and partial characterization. *J. Comp. Neurol.* **331,** 183–198.

Nässel, D. R., Passier, P. C., Elekes, K., Dircksen, H., Vullings, H. G., and Cantera, R. (1995). Evidence that locustatachykinin I is involved in release of adipokinetic hormone from locust corpora cardiaca. *Regul. Pept.* **57,** 297–310.

Nässel, D. R., Vullings, H. G., Passier, P. C., Lundquist, C. T., Schoofs, L., Diederen, J. H., and Van der Horst, D. J. (1999). Several isoforms of locustatachykinins may be involved in cyclic AMP-mediated release of adipokinetic hormones from the locust corpora cardiaca. *Gen. Comp. Endocrinol.* **113,** 401–412.

Nässel, D. R., Mentlein, R., Bollner, T., and Karlsson, A. (2000). Proline-specific dipeptidyl peptidase activity in the cockroach brain and intestine: Partial characterization, distribution, and inactivation of tachykinin- related peptides. *J. Comp. Neurol.* **418,** 81–92.

Newby, L. M., and Jackson, F. R. (1996). Regulation of a specific circadian clock output pathway by lark, a putative RNA-binding protein with repressor activity. *J. Neurobiol.* **31,** 117–128.

Nichols, R. (1992a). Isolation and structural characterization of *Drosophila* TDVDHVFLRFamide and FMRFamide-containing neural peptides. *J. Mol. Neurosci.* **3,** 213–218.

Nichols, R. (1992b). Isolation and expression of the *Drosophila* Drosulfakinin neural peptide gene product, DSK-I. *Mol. Cell. Neurosci.* **3,** 342–347.

Nichols, R., and Lim, I. A. (1996). Spatial and temporal immunocytochemical analysis of drosulfakinin (Dsk) gene products in the *Drosophila melanogaster* central nervous system. *Cell Tissue Res.* **283,** 107–116.

Nichols, R., Schneuwly, S. A., and Dixon, J. E. (1988). Identification and characterization of a *Drosophila* homologue to the vertebrate neuropeptide cholecystokinin. *J. Biol. Chem.* **263,** 12167–12170.

Nichols, R., McCormick, J., and Lim, I. (1997). Dromyosuppressin and drosulfakinin, two structurally related *Drosophila* neuropeptides, are uniquely expressed in the adult central nervous system. *Ann. N. Y. Acad. Sci.* **814,** 315–318.

Nichols, R., McCormick, J., and Lim, I. (1999a). Regulation of *Drosophila FMRFamide* neuropeptide gene expression. *J. Neurobiol.* **39,** 347–358.

Nichols, R., Kaminski, S., Walling, E., and Zornik, E. (1999b). Regulating the activity of a cardioacceleratory peptide. *Peptides* **20,** 1153–1158.

Nichols, R., McCormick, J., Cohen, M., Howe, E., Jean, C., Paisley, K., and Rosario, C. (1999c). Differential processing of neuropeptides influences *Drosophila* heart rate. *J. Neurogenet.* **13,** 89–104.

Noronha, K. F., Lange, A. B., and Osborne, R. H. (1997). Proctolin analogues and a proctolin metabolite as antagonist of the peptide proctolin. *Peptides* **18,** 67–72.

Noyes, B. E., Katz, F. N., and Schaffer, M. H. (1995). Identification and expression of the *Drosophila adipokinetic hormone* gene. *Mol. Cell. Endocrinol.* **109,** 133–141.

O'Brien, M. A., Schneider, L. E., and Taghert, P. H. (1991). *In situ* hybridization analysis of the FMRFamide neuropeptide gene in *Drosophila*. II. Constancy in the cellular pattern of expression during metamorphosis. *J. Comp. Neurol.* **304,** 623–638.

O'Brien, M. A., and Taghert, P. H. (1998). A peritracheal neuropeptide system in insects: Release of myomodulin-like peptides at ecdysis. *J. Exp. Biol.* **201,** 193–209.

O'Donnell, M. J., Rheault, M. R., Davies, S. A., Rosay, P., Harvey, B. J., Maddrell, S. H., Kaiser, K., and Dow, J. A. (1998). Hormonally controlled chloride movement across *Drosophila* tubules is via ion channels in stellate cells. *Am. J. Physiol.* **274,** R1039–R1049.

Orchard, I., Belanger, J. H., and Lange, A. B. (1980). Proctolin: A review with emphasis on insects. *J. Neurobiol.* **20,** 470–496.

Paemen, L., Tips, A., Schoofs, L., Proost, P., Van Damme, J., and De Loof, A. (1991). Lom-AG-myotropin: A novel myotropic peptide from the male accessory glands of *Locusta migratoria*. *Peptides* **12,** 7–10.

Park, J. H., and Hall, J. C. (1998). Isolation and chronobiological analysis of a neuropeptide pigment-dispersing factor gene in *Drosophila melanogaster*. *J. Biol. Rhythms.* **13,** 219–228.

Park, J. H., Helfrich-Förster, C., Lee, G., Liu, L., Rosbash, M., and Hall, J. C. (2000). Differential regulation of circadian pacemaker output by separate clock genes in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **97,** 3608–3613.

Park, Y., Zitnan, D., Gill, S. S., and Adams, M. E. (1999). Molecular cloning and biological activity of ecdysis-triggering hormones in *Drosophila melanogaster*. *FEBS Lett.* **463,** 133–138.

Passini, N., Larigan, J. D., Genovese, S., Appella, E., Sinigaglia, F., and Rogge, L. (1995). The 37/40-kilodalton autoantigen in insulin-dependent diabetes mellitus is the putative tyrosine phosphatase IA-2. *Proc. Natl. Acad. Sci. USA* **92,** 9412–9416.

Persson, M. G., and Nässel, D. R. (1999). Neuropeptides in insect sensory neurones: Tachykinin-, FMRFamide- and allatotropin-related peptides in terminals of locust thoracic sensory afferents. *Brain Res.* **816,** 131–141.

Persson, M. G., Eklund, M. B., Dircksen, H., Muren, J. E., and Nässel, D. R. (2001). Pigment-dispersing factor in the locust abdominal ganglia may have roles as circulating neurohormone and central neuromodulator. *J. Neurobiol.* **48,** 19–41.

Petri, B., and Stengl, M. (1997). Pigment-dispersing hormone shifts the phase of the circadian pacemaker of the cockroach *Leucophaea maderae*. *J. Neurosci.* **17,** 4087–4093.

Pierce, S. B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S. A., Buchman, A. R., Ferguson, K. C., Heller, J., Platt, D. M., Pasquinelli, A. A., Liu, L. X., Doberstein, S. K., and Ruvkun, G. (2001). Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* **15,** 672–686.

Pilon, M., Peng, X. R., Spence, A. M., Plasterk, R. H., and Dosch, H. M. (2000). The diabetes autoantigen ICA69 and its *Caenorhabditis elegans* homologue, ric-19, are conserved regulators of neuroendocrine secretion. *Mol. Biol. Cell* **11,** 3277–3288.

Pratt, G. E., Farnsworth, D. E., Fok, K. F., Siegel, N. R., McCormack, A. L., Shabanowitz, J., Hunt, D. F., and Feyereisen, R. (1991). Identity of a second type of allatostatin from cockroach brains: An octadecapeptide amide with a tyrosine-rich address sequence. *Proc. Natl. Acad. Sci. USA* **88,** 2412–2416.

Predel, R., Agricola, H., Linde, D., Veenstra, J. A., and Penzlin, H. (1994). The insect neuropeptide corazonin: Physiological and immunocytochemical studies in Blattariae. *Zoology (ZACS)* **98,** 35–49.

Proux, J. P., Miller, C. A., Li, J. P., Carney, R. L., Girardie, A., Delaage, M., and Schooley, D. A. (1987). Identification of an arginine vasopressin-like diuretic hormone from *Locusta migratoria*. *Biochem. Biophys. Res. Commun.* **149,** 180–186.

Pyza, E., and Meinertzhagen, I. A. (1997). Neurites of period-expressing PDH cells in the fly's optic lobe exhibit circadian oscillations in morphology. *Eur. J. Neurosci.* **9,** 1784–1788.

Quinlan, M. C., Tublitz, N. J., and O'Donnell, M. J. (1997). Anti-diuresis in the blood-feeding insect *Rhodnius prolixus* Stal: The peptide CAP2b and cyclic GMP inhibit Malpighian tubule fluid secretion. *J. Exp. Biol.* **200,** 2363–2367.

Quistad, Q. B., Adams, M. E., Scarborough, R. M., Carney, R. L., and Schooley, D. A. (1984). Metabolism of proctolin, a pentapeptide neurotransmitter in insects. *Life Sci.* **34,** 569–576.

Rachinsky, A., and Feldlaufer, M. F. (2000). Responsiveness of honey bee (*Apis mellifera* L.) corpora allata to allatoregulatory peptides from four insect species. *J. Insect Physiol.* **46,** 41–46.

Rao, K. R., and Riehm, J. P. (1993). Pigment-dispersing hormones. *Ann. N. Y. Acad. Sci.* **680,** 78–88.

Rayne, R. C., and O'Shea, M. (1992). Inactivation of neuropeptide hormones (AKH-I and AKH-II) studied *in vivo* and *in vitro*. *Insect Biochem. Mol. Biol.* **22,** 25–34.

Reagan, J. D. (1994). Expression cloning of an insect diuretic hormone receptor: A member of the calcitonin/secretin receptor family. *J. Biol. Chem.* **269,** 9–12.

Reagan, J. D. (1996). Molecular cloning and function expression of a diuretic hormone receptor from the house cricket, *Acheta domesticus*. *Insect Biochem. Mol. Biol.* **26,** 1–6.

Rémy, C., and Girardie, J. (1980). Anatomical organization of two vasopressin: Neurophysin-like neurosecretory cells throughout the central nervous system of the migratory locust. *Gen. Comp. Endocrinol.* **40,** 27–35.

Renden, R., Berwin, B., Davis, W., Ann, K., Chin, C. T., Kreber, R., Ganetzky, B., Martin, T. F., and Broadie, K. (2001). *Drosophila CAPS* is an essential gene that regulates dense-core vesicle release and synaptic vesicle fusion. *Neuron* **31,** 421–437.

Renn, S. C., Tomkinson, B., and Taghert, P. H. (1998). Characterization and cloning of tripeptidyl peptidase II from the fruit fly, *Drosophila melanogaster*. *J. Biol. Chem.* **273,** 19173–19182.

Renn, S. C., Park, J. H., Rosbash, M., Hall, J. C., and Taghert, P. H. (1999). A *pdf* neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in *Drosophila*. *Cell* **99,** 791–802.

Reynolds, S. E., Taghert, P. H., and Truman, J. W. (1979). Eclosion and bursicon titres and the onset of hormonal responsiveness during the last day of adult development in *Manduca sexta* (*L*). *J. Exp. Biol.* **78,** 77–86.

Richard, O., Tamarelle, M., Geoffre, S., and Girardie, J. (1994). Restricted occurrence of *Locusta migratoria* ovary maturing parsin in the brain–corpora cardiaca complex of various insect species. *Histochemistry* **102,** 233–239.

Richter, K., Peschke, E., and Peschke, D. (2000). A neuroendocrine releasing effect of melatonin in the brain of an insect, *Periplaneta americana* (L.). *J. Pineal Res.* **28,** 129–135.

Riddiford, L. M., Hewes, R. S., and Truman, J. W. (1994). Dynamics and metamorphosis of an identifiable peptidergic neuron in an insect. *J. Neurobiol.* **25,** 819–830.

Riehle, M. A., and Brown, M. R. (1999). Insulin stimulates ecdysteroid production through a conserved signaling cascade in the mosquito *Aedes aegypti*. *Insect Biochem. Mol. Biol.* **29,** 855–860.

Roebroek, A. J., Pauli, I. G., Zhang, Y., and van de Ven, W. J. (1991). cDNA sequence of a *Drosophila melanogaster* gene, *Dfurl*, encoding a protein structurally related to the subtilisin-like proprotein processing enzyme furin. *FEBS Lett.* **289,** 133–137.

Roebroek, A. J., Creemers, J. W., Pauli, I. G., Kurzik-Dumke, U., Rentrop, M., Gateff, E. A., Leunissen, J. A., and Van de Ven, W. J. (1992). Cloning and functional expression of *Dfurin2*, a subtilisin-like proprotein processing enzyme of *Drosophila melanogaster* with multiple repeats of a cysteine motif. *J. Biol. Chem.* **267,** 17208–17215.

Roebroek, A. J., Ayoubi, T. A., Creemers, J. W., Pauli, I. G., and Van de Ven, W. J. (1995). The *Dfur2* gene of *Drosophila melanogaster*: Genetic organization, expression during embryogenesis, and pro-protein processing activity of its translational product Dfurin2. *DNA Cell Biol.* **14,** 223–234.

Rosay, P., Colas, J. F., and Maroteaux, L. (1995). Dual organisation of the *Drosophila* neuropeptide receptor *NKD* gene promoter. *Mech. Dev.* **51,** 329–339.

Rosay, P., Davies, S. A., Yu, Y., Sozen, A., Kaiser, K., and Dow, J. A. (1997). Cell-type specific calcium signalling in a *Drosophila* epithelium. *J. Cell. Sci.* **110,** 1683–1692.

Rosay, P., Armstrong, J. D., Wang, Z., and Kaiser, K. (2001). Synchronized neural activity in the *Drosophila* memory centers and its modulation by *amnesiac*. *Neuron* **30,** 759–770.

Rose, C., Vargas, F., Facchinetti, P., Bourgeat, P., Bambal, R. B., Bishop, P. B., Chan, S. M., Moore, A. N., Ganellin, C. R., and Schwartz, J. C. (1996). Characterization and inhibition of a cholecystokinin-inactivating serine peptidase. *Nature* **380,** 403–409.

Rudwall, A. J., Sliwowska, J., and Nässel, D. R. (2000). Allatotropin-like neuropeptide in the cockroach abdominal nervous system: Myotropic actions, sexually dimorphic distribution and colocalization with serotonin. *J. Comp. Neurol.* **428,** 159–173.

Sato, Y., Oguchi, M., Menjo, N., Imai, K., Saito, H., Ikeda, M., Isobe, M., and Yamashita, O. (1993). Precursor polyprotein for multiple neuropeptides secreted from the suboesophageal ganglion of the silkworm *Bombyx mori*: Characterization of the cDNA encoding the diapause hormone precursor and identification of additional peptides. *Proc. Natl. Acad. Sci. USA* **90,** 3251–3255.

Sato, Y., Ikeda, M., and Yamashita, O. (1994). Neurosecretory cells expressing the gene for common precursor for diapause hormone and pheromone biosynthesis-activating neuropeptide in the suboesophageal ganglion of the silkworm, *Bombyx mori*. *Gen. Comp. Endocrinol.* **96,** 27–36.

Scanga, S. E., Ruel, L., Binari, R. C., Snow, B., Stambolic, V., Bouchard, D., Peters, M., Calvieri, B., Mak, T. W., Woodgett, J. R., and Manoukian, A. S. (2000). The conserved PI3′K/PTEN/Akt signaling pathway regulates both cell size and survival in *Drosophila*. *Oncogene* **19,** 3971–3977.

Schaffer, M. H., Noyes, B. E., Slaughter, C. A., Thorne, G. C., and Gaskell, S. J. (1990). The fruitfly *Drosophila melanogaster* contains a novel charged adipokinetic-hormone-family peptide. *Biochem. J.* **269,** 315–320.

Schneider, L. E., O'Brien, M. A., and Taghert, P. H. (1991). *In situ* hybridization analysis of the FMRFamide neuropeptide gene in *Drosophila*. I. Restricted expression in embryonic and larval stages. *J. Comp. Neurol.* **304,** 608–622.

Schneider, L. E., Sun, E. T., Garland, D. J., and Taghert, P. H. (1993a). An immunocytochemical study of the *FMRFamide* neuropeptide gene products in *Drosophila*. *J. Comp. Neurol.* **337,** 446–460.

Schneider, L. E., Roberts, M. S., and Taghert, P. H. (1993b). Cell type-specific transcriptional regulation of the *Drosophila FMRFamide* neuropeptide gene. *Neuron* **10,** 279–291.

Schoofs, L., Holman, G. M., Hayes, T. K., Nachman, R. J., and De Loof, A. (1991a). Isolation, identification and synthesis of locustamyoinhibiting peptide (LOM-MIP), a novel biologically active neuropeptide from *Locusta migratoria*. *Regul. Pept.* **36,** 111–119.

Schoofs, L., Holman, G. M., Hayes, T. K., Nachman, R. J., and De Loof, A. (1991b). Isolation, primary structure, and synthesis of locustapyrokinin: A myotropic peptide of *Locusta migratoria*. *Gen. Comp. Endocrinol.* **81,** 97–104.

Schoofs, L., Veelaert, D., De Loof, A., Huybrechts, R., and Isaac, E. (1998). Immunocytochemical distribution of angiotensin I-converting enzyme-like immunoreactivity in the brain and testis of insects. *Brain Res.* **785,** 215–227.

Secher, T., Lenz, C., Cazzamali, G., Sorensen, G., Williamson, M., Hansen, G. N., Svane, P., and Grimmelikhuijzen, C. J. (2001). Molecular cloning of a functional allatostatin gut/brain receptor and an allatostatin preprohormone from the silkworm, *Bombyx mori*. *J. Biol. Chem.* **276,** 47052–47060.

Seidah, N. G., Day, R., Marcinkiewicz, M., and Chretien, M. (1998). Precursor convertases: An evolutionarily ancient, cell-specific, combinatorial mechanism yielding diverse bioactive peptides and proteins. *Ann. N. Y. Acad. Sci.* **839,** 9–24.

Settle, S. H., Jr., Green, M. M., and Burtis, K. C. (1995). The *silver* gene of *Drosophila melanogaster* encodes multiple carboxypeptidases similar to mammalian prohormone-processing enzymes. *Proc. Natl. Acad. Sci. USA* **92,** 9470–9474.

Shen, P., and Cai, H. N. (2001). *Drosophila* neuropeptide F mediates integration of chemosensory stimulation and conditioning of the nervous system by food. *J. Neurobiol.* **47,** 16–25.

Siekhaus, D. E., and Fuller, R. S. (1999). A role for *amontillado*, the *Drosophila* homolog of the neuropeptide precursor processing protease PC2, in triggering hatching behavior. *J. Neurosci.* **19,** 6942–6954.

Siviter, R. J., Coast, G. M., Winther, A. M., Nachman, R. J., Taylor, C. A., Shirras, A. D., Coates, D., Isaac, R. E., and Nässel, D. R. (2000). Expression and functional characterization of a *Drosophila* neuropeptide precursor with homology to mammalian preprotachykinin A. *J. Biol. Chem.* **275,** 23273–23280.

Sliwowska, J., Rosinski, G., and Nässel, D. R. (2001). Cardioacceleratory action of tachykinin-related neuropeptides and proctolin in two coleopteran insect species. *Peptides* **22,** 209–217.

Sözen, M. A., Armstrong, J. D., Yang, M., Kaiser, K., and Dow, J. A. (1997). Functional domains are specified to single-cell resolution in a *Drosophila* epithelium. *Proc. Natl. Acad. Sci. USA* **94,** 5207–5212.

Spittaels, K., Verhaert, P., Shaw, C., Johnston, R. N., Devreese, B., Van Beeumen, J., and De Loof, A. (1996). Insect neuropeptide F (NPF)-related peptides: Isolation from Colorado potato beetle (*Leptinotarsa decemlineata*) brain. *Insect Biochem. Mol. Biol.* **26,** 375–382.

Stangier, J., Hilbich, C., Dircksen, H., and Keller, R. (1988). Distribution of a novel cardioactive neuropeptide (CCAP) in the nervous system of the shore crab *Carcinus maenas*. *Peptides* **9,** 795–800.

Starratt, A. N., and Brown, B. E. (1975). Structure of the pentapeptide proctolin, a proposed neurotransmitter in insects. *Life Sci.* **17,** 1253–1256.

Steele, J. E. (1985). Control of metabolic processes. *In* "Comprehensive Insect Physiology, Biochemistry and Pharmacology," Vol. 8 (G. A. Kerkut, and L. I. Gilbert, eds.), pp. 99–145. Pergamon Press, Oxford.

Stevenson, P. A. (1999). Colocalisation of taurine- with transmitter-immunoreactivities in the nervous system of the migratory locust. *J. Comp. Neurol.* **404,** 86–96.

St-Onge, S., Fortin, J.-P., Labarre, M., Steyaert, A., Schmidt, R., Ahmad, S., Walker, P., and Payza, K. (2000). *In vitro* pharmacology of NPFF and FMRFamide-related peptides at the PR4 receptor of *Drosophila melanogaster*. *Soc. Neurosci. Abstr.* **26,** 140.9.

Sweeney, S. T., Broadie, K., Keane, J., Niemann, H., and O'Kane, C. J. (1995). Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron.* **14,** 341–351.

Taghert, P. H. (1999). FMRFamide neuropeptides and neuropeptide-associated enzymes in *Drosophila*. *Microsc. Res. Tech.* **45,** 80–95.

Taghert, P. H. (2001). How does the circadian clock send timing information to the brain?. *Semin. Cell. Dev. Biol.* **12,** 329–341.

Taghert, P. H., and Goodman, C. S. (1984). Cell determination and differentiation of identified serotonin-immunoreactive neurons in the grasshopper embryo. *J. Neurosci.* **4,** 989–1000.

Taghert, P. H., and Schneider, L. E. (1990). Interspecific comparison of a *Drosophila* gene encoding FMRFamide-related neuropeptides. *J. Neurosci.* **10,** 1929–1942.

Taghert, P. H., Roberts, M. E., Renn, S. C., and Jacobs, P. S. (2000). Metamorphosis of tangential visual system neurons in *Drosophila*. *Dev. Biol.* **222,** 471–485.

Taghert, P. H., Hewes, R. S., O'Brien, M. E., Park, J., Han, M., and Peck, M. E. (2001). Multiple amidated peptides are required for normal circadian locomotor activity in *Drosophila*. *J. Neurosci.* **21,** 6673–6686.

Tamarelle, M., Coast, G. M., and Veenstra, J. A. (2000). Ovary maturing parsin and diuretic hormone are produced by the same neuroendocrine cells in the migratory locust, *Locusta migratoria*. *Peptides* **21,** 737–739.

Tams, J. W., Knudsen, S. M., and Fahrenkrug, J. (1998). Proposed arrangement of the seven transmembrane helices in the secretin receptor family. *Receptors Channels* **5,** 79–90.

Tatei, K., Cai, H., Ip, Y. T., and Levine, M. (1995). Race: A *Drosophila* homologue of the angiotensin converting enzyme. *Mech. Dev.* **51,** 157–168.

Taylor, C. A., Coates, D., and Shirras, A. D. (1996). The *Acer* gene of *Drosophila* codes for an angiotensin-converting enzyme homologue. *Gene.* **181,** 191–197.

Tawfik, A. I., Tanaka, S., De Loof, A., Schoofs, L., Baggerman, G., Waelkens, E., Derua, R., Milner, Y., Yerushalmi, Y., and Pener, M. P. (1999). Identification of the gregarization-associated dark-pigmentotropin in locusts through an albino mutant. *Proc. Natl. Acad. Sci. USA* **96,** 7083–7087.

Teal, P. E., Abernathy, R. L., Nachman, R. J., Fang, N., Meredith, J. A., and Tumlinson, J. H. (1996). Pheromone biosynthesis activating neuropeptides: Functions and chemistry. *Peptides* **17,** 337–344.

Terhzaz, S., O'Connell, F. C., Pollock, V. P., Kean, L., Davies, S. A., Veenstra, J. A., and Dow, J. A. (1999). Isolation and characterization of a leucokinin-like peptide of *Drosophila melanogaster*. *J. Exp. Biol.* **202,** 3667–3676.

Thompson, K. L., Decker, S. J., and Rosner, M. R. (1985). Identification of a novel receptor in *Drosophila* for both epidermal growth factor and insulin. *Proc. Natl. Acad. Sci. USA* **82,** 8443–8447.

Thompson, K. S. J., Rayne, R. C., Gibbon, C. R., May, S. T., Patel, M., Coast, G. M., and Bacon, J. P. (1995). Cellular colocalization of diuretic peptides in locusts: A potent control mechanism. *Peptides* **16,** 95–104.

Thor, S., and Thomas, J. B. (1997). The *Drosophila islet* gene govenrs axon pathfinding and neurotransmitter identity. *Neuron* **18,** 397–409.

Tomkinson, B., and Nyberg, F. (1995). Distribution of tripeptidyl-peptidase II in the central nervous system of rat. *Neurochem. Res.* **20,** 1443–1447.

Tomkinson, B., Wernstedt, C., Hellman, U., and Zetterqvist, O. (1987). Active site of tripeptidyl peptidase II from human erythrocytes is of the subtilisin type. *Proc. Natl. Acad. Sci. USA* **84,** 7508–7512.

Tong, S., Li, J., and Wands, J. R. (1999). Carboxypeptidase D is an avian hepatitis B virus receptor. *J. Virol.* **73,** 8696–8702.

Torfs, H., Shariatmadari, R., Guerrero, F., Parmentier, M., Poels, J., Van Poyer, W., Swinnen, E., De Loof, A., Akerman, K., and Vanden Broeck, J. (2000). Characterization of a receptor for insect tachykinin-like peptide agonists by functional expression in a stable *Drosophila* Schneider 2 cell line. *J. Neurochem.* **74,** 2182–2189.

Tree, T. I., O'Byrne, D., Tremble, J. M., MacFarlane, W. M., Haskins, K., James, R. F., Docherty, K., Hutton, J. C., and Banga, J. P. (2000). Evidence for recognition of novel islet T cell antigens by granule-specific T cell lines from new onset type 1 diabetic patients. *Clin. Exp. Immunol.* **121,** 100–105.

Tremblay, J. J., Lanctot, C., and Drouin, J. (1998). The pan-pituitary activator of transcription, Ptx1 (pituitary homeobox 1), acts in synergy with SF-1 and Pit1 and is an upstream regulator of the Lim-homeodomain gene Lim3/Lhx3. *Mol. Endocrinol.* **12,** 428–441.

Truman, J. W., Rountree, D. B., Reiss, S. E., and Schwartz, L. M. (1983). Ecdysteroids regulate the release and action of eclosion hormone in the tobacco hornworm, *Manduca sexta* (L.). *J. Insect Physiol.* **29,** 895–900.

Truman, J. W., and Bate, M. (1988). Spatial and temporal patterns of neurogenesis in the central nervous system of *Drosophila melanogaster*. *Dev. Biol.* **125,** 145–157.

Truman, J. W., and Copenhaver, P. F. (1989). The larval eclosion hormone neurones in *Manduca sexta*: Identification of the brain–proctodeal neurosecretory system. *J. Exp. Biol.* **147,** 457–470.

Tublitz, N. J., and Sylwester, A. W. (1990). Postembryonic alteration of transmitter phenotype in individually identified peptidergic neurons. *J. Neurosci.* **10,** 161–168.

Tully, T., and Quinn, W. G. (1985). Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol. [A]* **157,** 263–277.

Turner, A. J., Isaac, R. E., and Coates, D. (2001). The neprilysin family of zinc metalloendopeptidases: Genomics and function. *Bioessays* **23,** 261–269.

Vanden Broeck, J. (2001). Neuropeptides and their precursors in the fruitfly, *Drosophila melanogaster*. *Peptides* **22,** 241–254.

Vanden Broeck, J., Torfs, H., Poels, J., Van Poyer, W., Swinnen, E., Ferket, K., and De Loof, A. (1999). Tachykinin-like peptides and their receptors. *Ann. N. Y. Acad. Sci.* **897,** 374–387.

Veelaert, D., Passier, P., Devreese, B., Vanden Broeck, J., Van Beeumen, J., Vullings, H. G., Diederen, J. H., Schoofs, L., and De Loof, A. (1997). Isolation and characterization of an adipokinetic hormone release-inducing factor in locusts: The crustacean cardioactive peptide. *Endocrinology* **138,** 138–142.

Veenstra, J. A. (1989). Do insects really have a homeostatic hypotrehalosemic hormone? *Biol. Rev.* **64,** 305–316.

Veenstra, J. A. (1989). Stimulation of the activation of fat body glycogen phosphorylase and of trehalose synthesis by peptide hormones in the American cockroach. *BioSystems* **23,** 31–40.

Veenstra, J. A. (1989). Isolation and structure of corazonin, a cardioactive peptide from the American cockroach. *FEBS Lett.* **250,** 231–234.

Veenstra, J. A. (1994). Isolation and structure of the *Drosophila* corazonin gene. *Biochem. Biophys. Res. Commun.* **204,** 292–296.

Veenstra, J. A. (2000). Mono- and dibasic proteolytic cleavage sites in insect neuroendocrine peptide precursors. *Arch. Insect Biochem. Physiol.* **43,** 49–63.

Veenstra, J. A., and Costes, L. (1999). Isolation and identification of a peptide and its cDNA from the mosquito *Aedes aegypti* related to *Manduca sexta* allatotropin. *Peptides* **20,** 1145–1151.

Veenstra, J. A., and Davis, N. T. (1993). Localization of corazonin in the nervous system of the cockroach *Periplaneta americana*. *Cell Tissue Res.* **274,** 57–64.

Veenstra, J. A., Romberg-Privee, H. M., and Schooneveld, H. (1984). Immunocytochemical localization of peptidergic cells in the neuro-endocrine system of the Colorado potato beetle, *Leptinotarsa decemlineata*, with antisera against vasopressin, vasotocin and oxytocin. *Histochemistry* **81,** 29–34.

Veenstra, J. A., Lehman, H. K., and Davis, N. T. (1994). Allatotropin is a cardioacceleratory peptide in *Manduca sexta*. *J. Exp. Biol.* **188,** 347–354.

Veenstra, J. A., Pattillo, J. M., and Petzel, D. H. (1997). A single cDNA encodes all three *Aedes* leucokinins, which stimulate both fluid secretion by the malpighian tubules and hindgut contractions. *J. Biol. Chem.* **272,** 10402–10407.

Verdu, J., Buratovich, M. A., Wilder, E. L., and Birnbaum, M. J. (1999). Cell-autonomous regulation of cell and organ growth in *Drosophila* by Akt/PKB. *Nat. Cell Biol.* **1,** 500–506.

Vullings, H. G., Ten Voorde, S. E., Passier, P. C., Diederen, J. H., Van Der Horst, D. J., and Nässel, D. R. (1998). A possible role of SchistoFLRFamide in inhibition of adipokinetic hormone release from locust corpora cardiaca. *J. Neurocytol.* **27,** 901–913.

Waddell, S., Armstrong, J. D., Kitamoto, T., Kaiser, K., and Quinn, W. G. (2000). The *amnesiac* gene product is expressed in two neurons in the *Drosophila* brain that are critically for memory. *Cell* **103,** 805–813.

Walent, J. H., Porter, B. W., and Martin, T. F. (1992). A novel 145 kd brain cytosolic protein reconstitutes Ca^{2+}-regulated secretion in permeable neuroendocrine cells. *Cell* **70,** 765–775.

Wall, J. B., and Taghert, P. H. (1991a). The timing of initial neuropeptide expression by an identified insect neuron does not depend on interactions with its normal peripheral target. *J. Neurobiol.* **22,** 935–956.

Wall, J. B., and Taghert, P. H. (1991b). Segment-specific modifications of a neuropeptide phenotype in embryonic neurons of the moth, *Manduca sexta*. *J. Comp. Neurol.* **309,** 375–390.

Walther, C., Schiebe, M., and Voigt, K. H. (1984). Synaptic and non-synaptic effects of molluscan cardioexcitatory neuropeptides on locust skeletal muscle. *Neurosci. Lett.* **45,** 99–104.

Wang, W., Cain, B. M., and Beinfeld, M. C. (1998). Adult carboxypeptidase E-deficient *fat/fat* mice have a near-total depletion of brain CCK 8 accompanied by a massive accumulation of glycine and arginine extended CCK: Identification of CCK 8 Gly as the immediate precursor of CCK 8 in rodent brain. *Endocrinology* **9,** 329–332.

Wegerhoff, R., Breidbach, O., and Lobemeier, M. (1996). Development of locustatachykinin immunopositive neurons in the central complex of the beetle *Tenebrio molitor*. *J. Comp. Neurol.* **375,** 157–166.

Weinkove, D., Neufeld, T. P., Twardzik, T., Waterfield, M. D., and Leevers, S. J. (1999). Regulation of imaginal disc cell size, cell number and organ size by *Drosophila* class I(A) phosphoinositide 3-kinase and its adaptor. *Curr. Biol.* **9,** 1019–1029.

Westbrook, A. L., and Bollenbacher, W. E. (1990). The development of identified neurosecretory cells in the tobacco hornworm, *Manduca sexta*. *Dev. Biol.* **140,** 291–299.

Westphal, C. H., Muller, L., Zhou, A., Zhu, X., Bonner-Weir, S., Schambelan, M., Steiner, D. F., Lindberg, I., and Leder, P. (1999). The neuroendocrine protein 7B2 is required for peptide hormone processing in vivo and provides a novel mechanism for pituitary Cushing's disease. *Cell* **96,** 689–700.

White, K., Hurteau, T., and Punsal, P. (1986). Neuropeptide-FMRFamide-like immunoreactivity in *Drosophila*: Development and distribution. *J. Comp. Neurol.* **247,** 430–438.

Williamson, M., Lenz, C., Winther, A. M., Nässel, D. R., Grimmelikhuijzen, C. J., and Winther, M. E. (2001). Molecular cloning, genomic organization, and expression of a B-type (cricket-type) allatostatin preprohormone from *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun.* **281,** 544–550.

Winther, A. M., Muren, J. E., Lundquist, C. T., Osborne, R. H., and Nässel, D. R. (1998). Characterization of actions of *Leucophaea* tachykinin-related peptides (LemTRPs) and proctolin on cockroach hindgut contractions. *Peptides* **19,** 445–458.

Wishart, M. J., and Dixon, J. E. (1998). Gathering STYX: Phosphatase-like form predicts functions for unique protein-interaction domains. *Trends Biochem. Sci.* **23,** 301–306.

Witten, J. L., and O'Shea, M. (1985). Peptidergic innervation of insect skeletal muscle: Immunochemical observations. *J. Comp. Neurol.* **242,** 93–101.

Witten, J. L., and Truman, J. W. (1991). The regulation of transmitter expression in postembryonic lineages in the moth *Manduca sexta*. II. Role of cell lineage and birth order. *J. Neurosci.* **11,** 1990–1997.

Woodhead, A. P., Stay, B., Seidel, S. L., Khan, M. A., and Tobe, S. S. (1989). Primary structure of four allatostatins: Neuropeptide inhibitors of juvenile hormone synthesis. *Proc. Natl. Acad. Sci. USA* **86,** 5997–5601.

Xin, X., Varlamov, O., Day, R., Dong, W., Bridgett, M. M., Leiter, E. H., and Fricker, L. D. (1997). Cloning and sequence analysis of cDNA encoding rat carboxypeptidase D. *DNA Cell Biol.* **16,** 897–909.

Yamashita, O. (1996). Diapause hormone of the silkworm, *Bombyx mori*: Structure, gene expression and function. *J. Insect Physiol.* **42,** 669–679.

Yasuyama, K., Kitamoto, T., and Salvaterra, P. M. (1996). Differential regulation of choline acetyltransferase expression in adult *Drosophila melanogaster* brain. *J. Neurobiol.* **30,** 205–218.

Yoon, J. G., and Stay, B. (1995). Immunocytochemical localization of *Diploptera punctata* allatostatin-like peptide in *Drosophila melanogaster*. *J. Comp. Neurol.* **363,** 475–488.

Zahn, T. R., Macmorris, M. A., Dong, W., Day, R., and Hutton, J. C. (2001). IDA-1, a *Caenorhabditis elegans* homolog of the diabetic autoantigens IA-2 and phogrin, is expressed in peptidergic neurons in the worm. *J. Comp. Neurol.* **429,** 127–143.

Zdarek, J., Nachman, R. J., and Hayes, T. K. (1997). Insect neuropeptides of the pyrokinin/PBAN family accelerate pupariation in the fleshfly (*Sarcophaga bullata*) larvae. *Ann. N. Y. Acad. Sci.* **814,** 67–72.

Zhang, X., McNeil, G. P., Hilderbrand-Chae, M. J., Franklin, T. M., Schroeder, A. J., and Jackson, F. R. (2000). Circadian regulation of the lark RNA-binding protein within identifiable neurosecretory cells. *J. Neurobiol.* **45,** 14–29.

Zhou, A., Webb, G., Zhu, X., and Steiner, D. F. (1999). Proteolytic processing in the secretory pathway. *J. Biol. Chem.* **274,** 20745–20748.

Zitnan, D., Kingan, T. G., Hermesman, J. L., and Adams, M. E. (1996). Identification of ecdysis-triggering hormone from an epitracheal endocrine system. *Science* **271,** 88–91.

Zitnan, D., Ross, L. S., Zitnanova, I., Hermesman, J. L., Gill, S. S., and Adams, M. E. (1999). Steroid induction of a peptide hormone gene leads to orchestration of a defined behavioral sequence. *Neuron.* **23,** 523–535.

Zornik, E., Paisley, K., and Nichols, R. (1999). Neural transmitters and a peptide modulate *Drosophila* heart rate. *Peptides* **20,** 45–51.

This Page Intentionally Left Blank

2

From Genes to Aging in *Drosophila*

Stephen L. Helfand* and Blanka Rogina
Department of Genetics and Developmental Biology, School of Medicine, University of Connecticut Health Center, Farmington, Connecticut 06030, USA

If we kept throughout life the same resistance to stress, injury and disease which we had at the age of ten, about one-half of us here today

*To whom correspondence should be addressed. E-mail: shelfand@neuron.uchc.edu

0065-2660/03 $35.00

> might expect to survive in 700 years' time. The reason that we cannot is that in man, and in many, but probably not all, other animals, the power of self-adjustment and self-maintenance declines with the passage of time, and the probability of disease and death increases. The increase in man becomes eventually so steep that while exceptional individuals may outlast a century, there is an effective limit, depending upon our present age, upon the number of years for which any of us can reasonably expect to go on living. The uniformity of this process is one of the earliest unpleasant discoveries which every individual has to make, and although we have many psychological expedients to blunt its impact, the fact of this effective fixity of life-span, and of the decline in activity and health which often determine it, is always in the background of the human mind. (Comfort, 1979)

ABSTRACT

Despite the intimate nature of the aging process we actually know little about it. In more recent years, work on a variety of organisms, utilizing approaches including demography, molecular genetics, and epidemiology, have challenged some of the more commonly held assumptions about the aging process. These studies have served to reinvigorate the field of aging research and are beginning to lead the way in a renaissance in aging research (Helfand and Inouye, 2002). Invertebrate model systems such as *Drosophila* and *Caenorhabditis elegans* that permit extensive genetic analysis are at the forefront of this renaissance.

I. DEFINING AGING

How to define the aging process? There are a number of reasons why aging is a difficult subject. Since "aging" is a personal experience and something we all share, we naturally develop our own notions as to why or how it takes place. These unrecognized prejudices manifest themselves in subtle and not so subtle ways. They often come into play when discussing the mechanisms of aging.

What do we mean by aging? A half a century ago Sir Peter Medawar noted that

> It is a curious thing that there is no word in the English language that stands for the mere increase of years; that is, for ageing silenced of its overtones of increasing deterioration and decay. At present we are obliged to say that Dorian Gray did not exactly "age", though to admit that he certainly grew older. (Medawar, 1957)

In his book *Longevity, Senescence, and the Genome*, Caleb Finch wrestles with this problem and attempts to define aging in a context that is more neutral, or at least "silenced of its overtones" of deterioration (Finch, 1990). He notes that aging is a word used to describe a variety of time-dependent changes to which many different biological phenomena are subject, including changes from the level of the molecule to the ecosystem. The causes of these changes may be entirely different, implying no common mechanism, yet often implying an element of deterioration. Finch proposes that we use the term "age-related changes." This term is more neutral and does not suggest that a particular change is necessarily deleterious.

The term "senescence" is reserved for age-related changes that take place in an organism, leading to a reduction in its vitality and function and specifically increase its mortality rate or the chances of its dying in the next interval of time. Senescent changes are nonreversible and deteriorative in nature. A problem arises, however, in defining which changes will lead to an increase in the risk of mortality. There are many common age-related changes such as graying of the hair that do not lead to any obvious increase in mortality, barring a societal decision to kill anyone who begins to show the signs of gray hair. The problem of predicting what changes are of a senescent nature and what changes are neutral is similar to the problem of predicting fitness for an organism or species. Changes that increase mortality in one context may not do so in another. For example, tens of thousands of years ago when humans lived on the African savannah, an age-related decrease in the ability to run away from predators would likely lead to a significant increase in mortality. The same slowing down with age, except in its most terminal phase of immobility, however, would not necessarily be much of a disadvantage in parts of the modern industrialized world. Despite this, a population that shows senescence or "aging" is presently defined as one having an increase in age-dependent mortality, older individuals having a higher rate of mortality than younger ones (after excluding infant mortality, for example).

II. MEASURING AGING

A. Populations, survivorship curves, and mortality rates

Defining aging and senescence is problematic. How we measure a phenomenon often may help to better define it. There are two different methods that can be used to measure aging and senescence. The most common measurement used in aging research is derived from data obtained by a life table for a cohort of individuals of the same age. The life table is the collected information concerning the time or age of death of each of the individuals in a given

population. From these data a survivorship curve can be graphed, showing the percentage of individuals surviving versus age for a cohort of same-aged individuals. A derivative of this is the rate of death or mortality rate for the population, which is graphed as the mortality rate—using the natural logarithm. From the survivorship and mortality curves a number of features concerning the "aging" of a population can be illustrated. In the wild, the shape of the survivorship curve usually shows an approximately continuous decay or constant loss of life over time due to predation, injury, and disease (Finch, 1990) (Fig. 2.1)

When plotted as a rate of death or mortality curve it shows a constant rate of mortality or a straight horizontal line (Fig. 2.2).

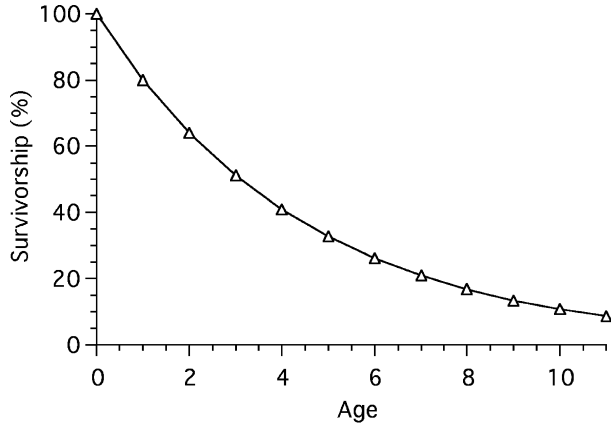


Figure 2.1. A stylized example of a survivorship curve for a population in the wild, based on Finch (1990).

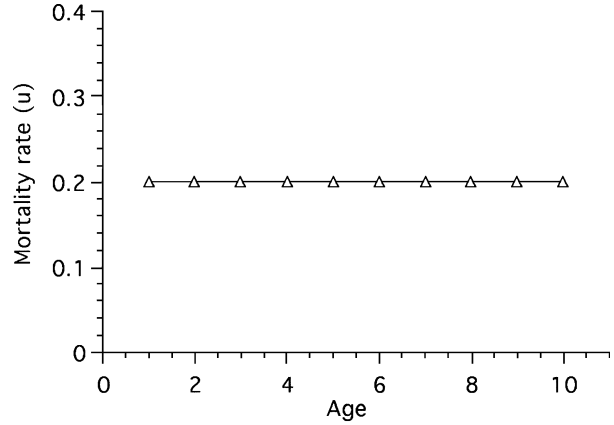


Figure 2.2. The mortality rate of the population shown in the survivorship curve of Fig. 2.1.

This type of curve implies that death is not related to the age of the mature adult but to some cause that does not distinguish between ages such as predators, accidents, or starvation (although such things may in fact be age related). Thus, according to the definition of senescence, this population would not be considered to represent a population that is showing senescence. The individuals in the population are not dying in an age-dependent or senescent manner. That is, the chances of dying as they get older are not increasing. Another type of survivorship curve can be seen when a population shows age-dependent death but still does not show senescent changes—an ever-increasing chance of dying in the next age interval. Figure 2.3 shows an example of a population in which the individuals all live to some particular age and then rapidly die off; Fig. 2.4 shows the accompanying mortality rate.

In the above-described cases (Figs. 2.3 and 2.4), a dramatic example of morbidity compression or rectangularization of the survivorship curve, there is an age-related change in survivorship. However, the mortality rate in Fig. 2.4 does not show an increase over a large portion of the life span but rather a short period of massive death. The Pacific salmon is an example of a species that shows such a survivorship curve after spawning (Finch, 1990).

In Figs. 2.5 and 2.6 are survivorship and mortality curves that show the classic elements of aging in a population.

B. Late life plateau

After an initial early period of low mortality (not including infantile mortality), as individuals get older there is an increasing rate of mortality, or decreasing rate of surviving to the next period of time. This idea of an increase in mortality rate, more specifically, an exponential increase in the rate of mortality with age was first developed in 1825 by the British actuary Benjamin Gompertz. He referred to it as the "law of mortality" (Gompertz, 1825). An exponential increase in mortality with age demonstrated not only the increasing burden of age on survival, but also provided the underpinning for another commonly known phenomenon, species-specific life span. The idea of a species-specific life span is that each species has a maximal life expectancy that it does not surpass. For example, mice live 2 to 3 years and dogs live 15 to 20 years. The idea was that if mortality rates increase exponentially there will have to be an age beyond which no individuals in the population will be alive. In more recent years the work of Vaupel and colleagues has cast serious doubt on both the idea of an exponential increase in mortality and species-specific life spans (Carey *et al.*, 1992; Curtsinger *et al.*, 1992; Vaupel *et al.*, 1998; Vaupel, 1997).

Drosophila melanogaster has played a critical role in these studies, demonstrating that for many invertebrate species, including yeasts, nematodes, medflies, fruit flies, and parasitic wasps, there is in fact a late-life mortality

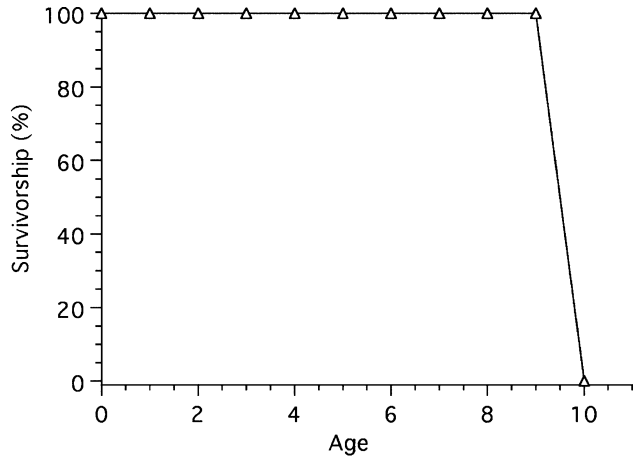


Figure 2.3. A stylized example of a population that has a sudden dramatic increase in death at a particular age.

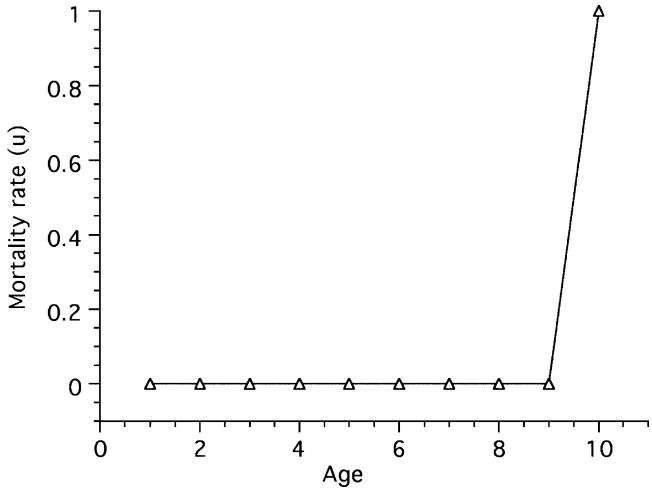


Figure 2.4. The mortality rate of the population shown in the survivorship curve of Fig. 2.3.

plateau and this plateau is not due to genetic heterogeneity (Curtsinger *et al.*, 1992) (Fig. 2.7).

C. Mortality curves and aging

Although there is a late life plateau in mortality rates, for most of the life span the mortality curve does show an exponential increase. Examination of the

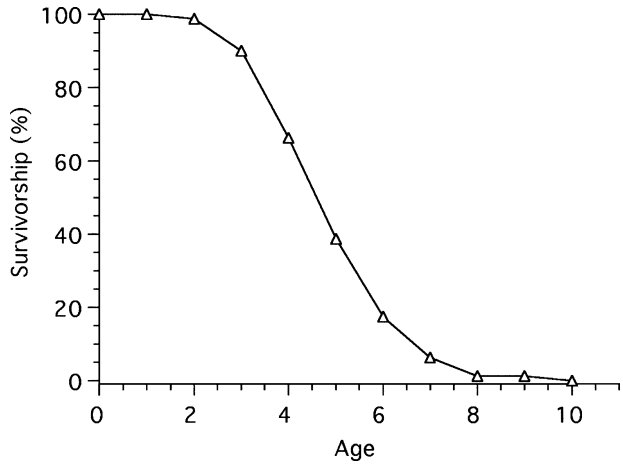


Figure 2.5. A survivorship curve from flies carrying the *drop dead* mutation and having a short adult life span (Rogina *et al.*, 1997).

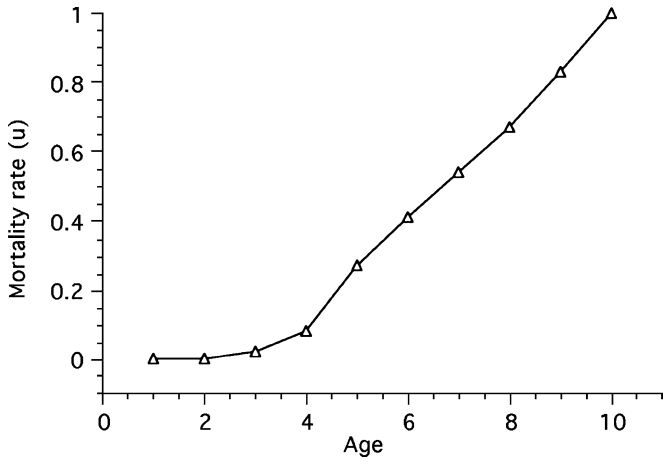


Figure 2.6. The mortality rate of the population shown in the survivorship curve of Fig. 2.5.

slope and y intercept of the exponential portion of mortality curve can provide important information about how different interventions may relate to the process of aging. Interventions that increase or decrease life span may cause a change in the slope and/or a shift in the time of initiation of increasing mortality (reflected in the y intercept). Some interventions alter the slope of the curve while others selectively shift the entire curve to the right (in cases of life span extension) or left (in cases of life span shortening) without changing the slope (Figs. 2.8 and 2.9).

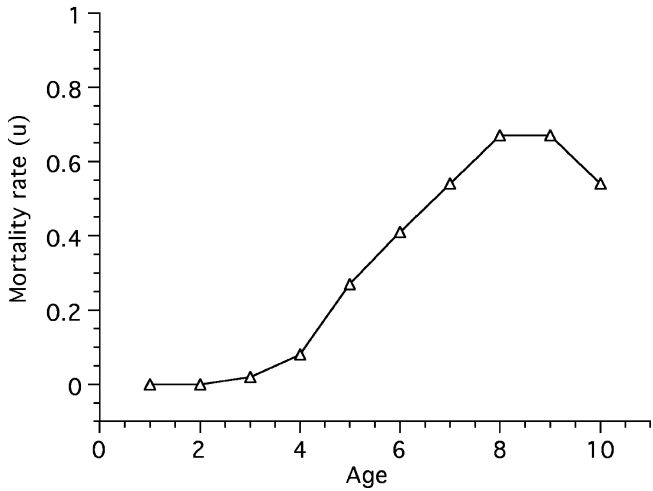


Figure 2.7. The late life deceleration in mortality rates and, as occurs in some cases, exemplified here, a decline in mortality with age.

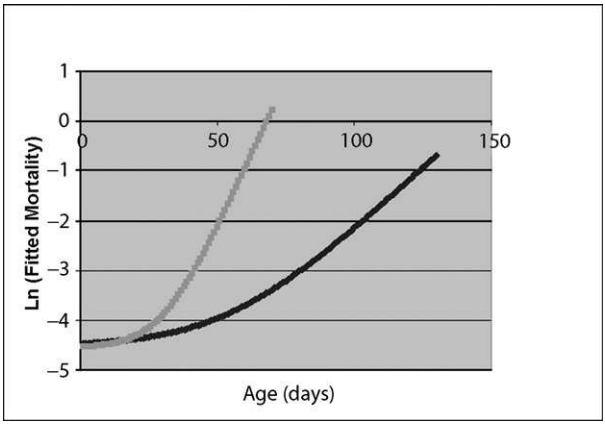


Figure 2.8. A difference in slope of the mortality curves. These data are from survivorship data from adult male *Drosophila melanogaster* grown at 25°C versus 18°C.

It is thought that a change in the slope of the mortality curve represents a change in the rate of aging, while a shift in the curve reflects a change in the hazard function without actually changing the rate of aging (Wilmoth, 1997). Interventions that shift the mortality curve while leaving the slope unchanged may delay or shorten the time of initiation of the aging process but do not affect the rate at which aging proceeds once it is initiated. One of the major goals of aging research has been to look for interventions that alter the slope of the mortality curve; such interventions are thought to represent a

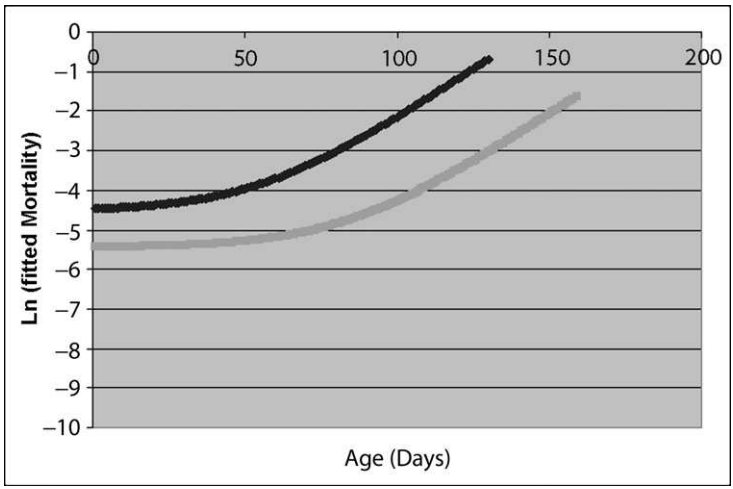


Figure 2.9. The difference in y intercept but not in the slope of the mortality curves. It can be seen that although the increase in mortality rate begins at a later time, once initiated its "rate" as reflected by the slope is nearly the same in both cases (S. Helfand and B. Rogina, unpublished data).

change in the rate of aging and thus provide insights into the normal process of aging. Interventions that extend life span and decrease the slope of the mortality curve include caloric restriction in mammals and decreases in ambient temperature in poikilotherms (Finch, 1990; Bartke *et al.*, 2001). Three mutations in *Drosophila*, *Hyperkinetic*, *Shaker*, and *drop dead*, all change the slope of the mortality curve but do so by increasing the slope and shortening life span (Trout and Kaplan, 1970). So far only one known mutation in *Drosophila*, *Indy*, increases life span and decreases the slope of the mortality curve (Rogina *et al.*, 2000; S. L. Helfand, and B. Rogina, unpublished data). All the other known life-extending mutations in *Drosophila* including the *methuselah*, *chico*, and *insulin receptor* mutations appear to have either a significant or primary effect by shifting the mortality curve to the right leaving the slope nearly the same as in controls (Lin *et al.*, 1998; Tatar *et al.*, 2001; Clancy *et al.*, 2001). Other interventions that alter life span in flies, such as reproduction and the laboratory-selected long-lived lines, also appear to shift the mortality curve without altering its slope (S. L. Helfand, and B. Rogina, unpublished data). Finally, although caloric restriction in mammals, the only known means of extending life span in mammals, has been reported to change the slope of the mortality curve (Bartke *et al.*, 2001) caloric restriction in female flies, which also increases life span, appears to shift the mortality curve and not change the slope (S. L. Helfand and B. Rogina, unpublished data).

III. AGING OR SENESCENCE IN THE INDIVIDUAL

The use of survivorship and mortality curves is of great value in estimating the rate of death of a population. While this is certainly something that the life insurance companies want to know, it is not what most people want to know. What most people want to know is "How long am I going to live?" Survivorship and mortality data cannot reveal this type of information; they are only concerned with populations and cannot tell us anything about the elements of aging at the level of the individual. Did those individuals that died early, helping to shape the survivorship or mortality curves, die from some acceleration of the aging process or did they die from accidents, predators, or disease?

Measuring age at death in a population does not tell us much about the process of aging in the individual. Life tables and their subsequent analysis, through survivorship and mortality rate curves, are not measuring any physiological element of the aging process but only the age of death of individuals in a cohort. To really begin to understand the process of aging it is important to develop methods of assessing physiological changes associated with aging: what is the physiological age of individuals or populations under different environmental or genetic conditions?

To determine aging in the individual we need some means of measuring and then assessing age-related changes in individuals—the physiological age of the individual, not just the chronological age. The notion of measuring aging in the individual led to the field of biomarkers of aging. Biomarkers of aging are those measurable changes that can provide us with a better idea of the age of the individual than calendar time itself.

IV. BIOMARKERS OF AGING

Biomarkers of aging should measure physiological age and not simply chronological age. What is the distinction between physiological age and chronological age? As an example, all of us have probably experienced individuals that looked much older or younger than their chronological age. The distinction between chronological age and physiological age can be demonstrated by examining poikilothermic animals such as *Drosophila melanogaster*. Life span and changes in gross measures of physiological age such as female fertility–reproduction and motility scale inversely with temperature (Miquel *et al.*, 1976) (Fig. 2.10). Adult flies living at 18°C live threefold longer than those living at 29°C and two- or more fold longer than those living at 25°C (Miquel *et al.*, 1976). Associated with the differences in life span are differences in the timing of reproductive and motor decline. Peak reproduction is much earlier for female flies living at

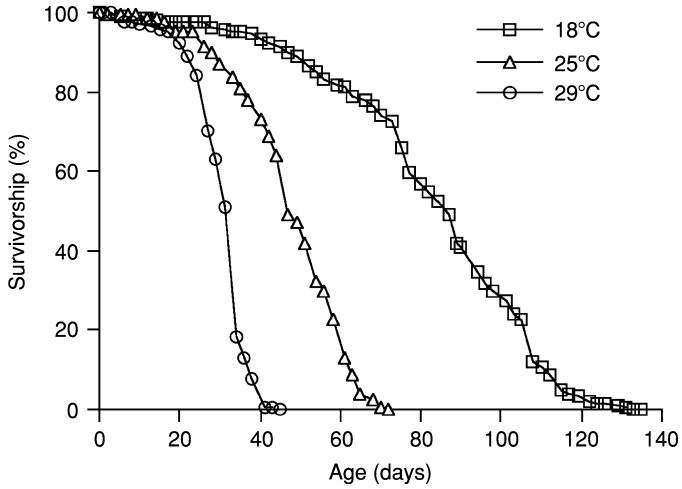


Figure 2.10. Life span of fruit flies at three different temperatures. Adapted from Helfand *et al.* (1995).

29°C than for female flies living at 18°C. Likewise motor function, as measured by the ability to climb in negative geotaxis assays or phototaxis assays, shows the same differences in loss of ability dependent on the temperature at which the animals are living (Lamb, 1978). Hence, physiological age and chronological age can be readily dissociated in poikilothermic animals, using temperature.

Although it may be more difficult to measure in other animals, particularly mammals, the notion that some individuals age at a different rate than others of the same cohort leads to the need to make objective measures based on other physiological information than simply calendar age or chronological age. The ability to measure some biological marker that reflects physiological age, and does so over a short period of time, is essential for studying the biology of aging as well as assessing interventions that may alter aging and life span.

A. Potential method for tracking physiological age: Temporal patterns of gene expression

The use of enhancer trap and reporter-marked genes in *Drosophila* has demonstrated the possibility of utilizing gene regulation as a means of providing biomarkers of aging. Examining the temporal patterns of expression of many different genes has shown that for some genes their temporal pattern of expression shows a complex dynamic pattern that is somehow linked to or associated with life span, so-called age-dependent gene expression, while a

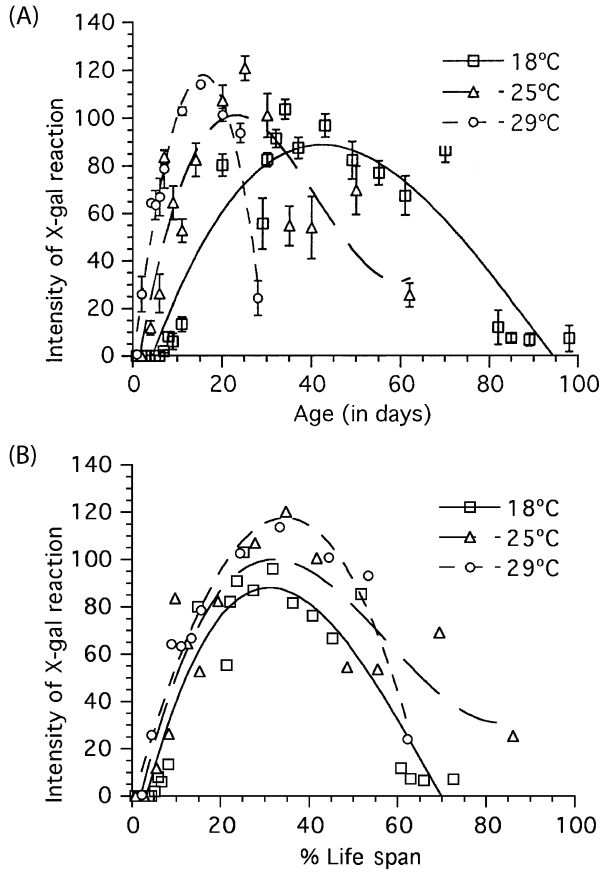


Figure 2.11. (A) and (B) illustrate an example of a gene (marked as enhancer trap line 1085) whose temporal pattern of expression in the antennae of the adult fly shows age-dependent expression. (A) When examined as a function of chronological age the temporal pattern of expression of this gene in the antenna follows a similar but temporally different pattern of expression depending on whether the adult animals are living at 18, 25, or 29°C. (B) If the temporal pattern of expression of this gene is now plotted as a function of percent life span it is seen that the three previously disparate curves now appear to largely overlap, showing that the expression of this gene is somehow linked to life span, and perhaps to the rate of aging of the adult fly. Figures taken from Helfand *et al.* (1995).

smaller set of genes shows an age-independent temporal pattern of gene expression (Helfand and Rogina, 2000). (Fig. 2.11).

In principle, age-dependent and age-independent gene expression should be present in all organisms, including humans. Examination of the trajectory in the temporal pattern of gene expression of age-dependent and

age-independent genes should provide the much-needed biomarkers of aging allowing for the measurement of aging on an individual basis. This will be more useful for both understanding aging, and for the more practically based testing of interventions that extend life span in a shorter period of time, than having to wait for age at death. The ability to measure gene expression for many different genes simultaneously, as the microarray system affords, should permit this approach to be readily used in flies and mammals.

V. GENETICS AND AGING

Genetics is known to play a substantial role in longevity and the process of aging (Finch, 1990). Knowledge that a major element of the aging process is hereditary or under genetic control, however, does not necessarily mean that aging is determined by a fixed genetic pathway. Two distinctly different means by which genetic elements could determine longevity can be imagined. One is that genetic elements could be providing the instructive and driving forces that result in the stereotypic changes we see as aging. During development, a similarly complex biological phenomenon, genetic elements do provide the instructive and driving forces for the characteristic changes that take place in developing from a fertilized egg to a mature individual. Given this precedence it is possible that genetic elements may also play a substantial role in driving some aspects of the stereotypic changes seen during the aging process.

In contradistinction to genetic elements providing directed control of the aging process, the genetic contribution to aging could be limited to providing a predisposition for the organism to respond to changes during life. In this case genetic elements are not controlling the process but providing, in a passive manner, the limitations an organism may possess for responding to change or other challenges life requires. For example, if it were to be imagined that aging was due to the accumulation of damage from toxic elements, such as oxidative damage to macromolecules (see below), the ability to defend against these degenerative changes could determine the length and pace of aging. By helping to set the amount of protective anti-toxic systems through a genetic endowment, genetic elements could determine life span and the pace of aging without directly setting the course of the aging process itself.

Regardless of whether genetic elements are participating through an active (instructive pathway) or passive (endowment) manner, genetic approaches can be used as a tool to identify those physiological systems and agents responsible for the process of aging and perhaps to provide direct insight into the means for its amelioration. The successes that genetic approaches have had in helping to understand similarly complex biological phenomena, such as

development, strongly suggest that these same approaches will be of enormous use in penetrating the mysteries of aging and longevity.

VI. USE OF FLIES FOR AGING RESEARCH

One of the most powerful model organisms used to study genetic elements and in the application of genetic approaches to understand complex biological phenomena is the fruit fly *Drosophila melanogaster*. There are a number of obvious advantages to the use of *Drosophila melanogaster* for studying aging. These include its relatively short life span, about 3 months at 25°C (Miquel *et al.*, 1976); ease of maintenance; environmental and genetic interventions that alter life span; a large body of information on aging in *Drosophila*; well-defined anatomy; availability of stocks with altered genes; powerful molecular genetic techniques; proven success in dissecting complex biological phenomena such as development; and knowledge of the full sequence of the *Drosophila* genome.

Two additional aspects of the biology of the fruit fly stand out as advantageous for studying aging. The life history of the fruit fly is such that the major elements of life, development and reproduction, have been separated into two morphologically distinct stages. After embryogenesis, during the three larval phases the organism eats and grows, eats and grows. At the end of the third larval stage metamorphosis takes place, during which the wormlike larval organism is transformed into the morphologically distinct adult fly. Within 24 h, the adult fly is sexually mature and can reproduce. It is during the "adult" phase of life that aging is usually thought to take place. While it may often be difficult to tell when maturity has been reached in some organisms, this is not a problem when using fruit flies because the sexually mature state of the adult starts within hours of its emergence from the pupal case as a fly.

A second notable feature of the adult fly is that it consists almost entirely of postmitotic cells (Bozuck, 1972; Ito and Hotta, 1992). Only cells in the gonads and a few cells in the gut continue to divide. The reason for pointing out this distinction is that many other organisms, including humans, have organs that are made up of tissues in which cellular replacement is a prominent feature. For example, the gastrointestinal and erythropoietic systems are two examples of organs in which cell division and replacement is an essential feature. Neurons in the brain and cardiac and skeletal muscle cells, however, rarely divide throughout the entire life of the individual. Organs in which cell replacement is a prominent aspect of normal physiology may be "aging" in a different manner than those composed of postmitotic cells. It may be asked for example, whether a "new" cell born in an older individual is older than the same cell born in a younger individual? And is this what we mean when we talk

of aging, particularly at the cellular level? *Drosophila melanogaster* is almost exclusively a postmitotic organism. Except for the two exceptions, gonads and gut, all the cells it has at the time of emergence as an adult are all it will have for its entire life. For this reason it has been suggested by R. Arking that for purposes of aging, the adult fly might be considered to consist of a set of synchronously aging cells (Arking, 1991). If one of the primary purposes of aging research is to understand how cells that are born with the organism change with age, then the adult fly, being almost entirely postmitotic, is an excellent model system.

A. History of aging and genetics of aging in the fly

1. Aging obeys normal chemicophysical laws

The benefits of using the fruit fly as a model system for studying aging is not a new concept. At least as early as 1915 scientists employed fruit flies to demonstrate fundamental features of the aging process. Loeb and Northrup performed a series of experiments using *Drosophila pseudoobscura* to show that aging, like life, obeys the physicochemical laws of nature. Loeb and Northrup showed that the length of life, and presumably the rate of aging in poikilothermic animals, such as *Drosophila*, is inversely related to ambient temperature (Loeb and Northrop, 1916, 1917). They concluded that the same Q10 relationship between kinetic reactions (i.e., for every 10°C increase in temperature, the rate of reaction doubles) in the organism was true for aging as well as other physiological features of life known at that time.

2. Life span is an inherited trait

In the 1920s Pearl and colleagues used *Drosophila* to formally demonstrate that longevity was an inherited trait. In a series of papers they showed that different stocks of *Drosophila* and lines derived from individual flies from wild-type *Drosophila* populations have different but reproducible life spans (Pearl and Parker, 1921, 1922; Pearl *et al.*, 1923). In addition to showing that life span is a heritable trait Pearl and colleagues used *Drosophila* to develop the "rate of aging" hypothesis (Pearl, 1928; Alpatov and Pearl, 1929). Studies by Pearl also demonstrated the effects of hybridization on life span (Pearl, 1922). He showed that crossing two different strains of *Drosophila* often led to an increase in life span for the F_1 hybrids. Hybrid vigor, as this is known, is an important phenomenon that needs to be taken into consideration when investigating the life span of genetically altered animals that involve the crossing of two different stocks. As is noted below, the current use of overexpression or ectopic expression systems (e.g., Gal4–UAS system) requires the crossing of two

different stocks and the subsequent examination of life span in the F_1 hybrid. Selecting the correct controls for genetic background and for the effects of hybrid vigor needs to be considered in interpreting life span results from such studies.

3. Fitness characteristics are important in determining life span and the rate of living hypothesis is challenged

J. Maynard Smith in the 1950–1960s used *Drosophila* to try to understand the nature of the aging process by examining the life span of *Drosophila pseudoobscura* at different temperatures and under different reproductive schedules (Smith, 1958, 1962). This work, directed at determining the effect of reproduction on life span in *Drosophila*, has provided much of the initial empirical evidence for evolutionary biological theories on aging and senescence. In work utilizing temperature shift experiments the "rate of living" hypothesis was tested. This hypothesis proposes that a fixed amount of some "vital" substance is used up over the life span. The temperature shift experiments did not appear to support this hypothesis, suggesting that vitality was not used up in a continuous manner, and Maynard Smith proposed an alternative "threshold" theory (Smith, 1958, 1962; Clarke and Smith, 1961a,b; Lamb, 1978). The work on temperature shifts has been controversial, and a number of investigators have attempted to repeat it with limited success (Miquel *et al.*, 1976; Lamb, 1978). For the most part, however, it does appear that in poikilothermic animals such as *Drosophila melanogaster* there is an inverse relationship between ambient temperature and life span.

4. Life span plasticity is demonstrated by selection experiments

In the 1980s two groups of population genetics/evolutionary biologists examined the genetic plasticity of longevity, using *Drosophila melanogaster* (Rose and Charlesworth, 1980, 1981; Luckinbill *et al.*, 1984; Rose, 1984; Luckinbill and Clare, 1985, 1987). Repeated selection of outbred stocks of flies for late-life female fertility over many generations (20–50) led to a dramatic increase in life span for both males and females. The selected long-lived lines lived up to twice as long as the original unselected parental lines. Among other things, these studies demonstrated that there are genes in outbred stocks of flies that through selection can be combined to lead to a dramatic increase in life span. Although selection methods clearly demonstrate the plasticity of heritable life span, like many other selection experiments on complex biological phenomena in flies, it has been exceedingly difficult to identify how many and which genes are involved in extending life span. These studies and others

performed with outbred stocks reveal the interplay between environment and genome but are less well suited to determine the actual mechanisms by which these interactions take place.

B. Manipulations that extend life span in flies

1. Nongenetic interventions

There are a number of different nongenetic interventions that extend life span in *Drosophila*. These include (1) decreasing ambient temperature or physical activity, (2) decreasing reproductive activity, (3) exposure to low levels of a stressor such as brief (less than 60 min) heat shock (37°C) or mild X-irradiation, and (4) dietary restriction. Many of these interventions may affect the same or similar physiological functions. For example, the mechanism by which mild irradiation increases life span, especially in females, is thought to be due in large part to its effect on reducing fertility. The dose of radiation that extends life span also causes a temporary sterility (Lamb, 1964, 1978). Furthermore, females that are sterile, due to a genetic lesion affecting the reproductive system, do not show a further increase in life span when irradiated. Some of the effects of heat shock may also be associated with a decrease in female reproduction.

a. Heat shock, irradiation, and other stressors: Resistance to stress

A mild–nonlethal stress (heat shock, irradiation, cold stress) is thought to result in an increase in life span (Johnson *et al.*, 1996). The mechanism by which this occurs may be through the induction of protective and/or repair systems (Masoro, 2000). For example, in *Drosophila* life span extension following heat shock stresses has been postulated to be due in part to the induction of the heat shock chaperone system, in particular *hsp70*. It was reported that a heat shock too brief to extend life span in normal flies could decrease mortality rates in *Drosophila* that had many additional copies of the *hsp70* gene (Tatar *et al.*, 1997). Follow-up studies to examine the effect of overexpression of *hsp70* on life span have not been able to demonstrate any increase in mean or maximal life span (Minois *et al.*, 2001). Many of the single-gene mutations that extend life span in *Caenorhabditis elegans* have also been found to be stress resistant (Johnson *et al.*, 1996). In *Drosophila* the long-lived *methuselah* (*mth*) mutant and the laboratory-selected long-lived lines of Luckinbill and of Rose also appear to be more stress resistant (Rose *et al.*, 1992; Djawdan *et al.*, 1998; Lin *et al.*, 1998; Luckinbill, 1998; Nghiem *et al.*, 2000). However, it appears that this correlation does not extend to some of the other long-lived mutants such as *InR*, *chico*, or *Indy* (Clancy *et al.*, 2001; Tatar *et al.*, 2001; S. L. Helfand and B. Rogina, unpublished data). The use of resistance to

stress as a phenotype in screening for life span-extending genes has been proposed (Lin *et al.*, 1998). Protective or repair systems are likely to be involved in preserving function in adult life, and further studies will be required to determine how significant an effect increasing expression or function of these protective or repair systems may have on life span and aging.

b. Ambient temperature and physical activity

It has been known since at least 1915 that there is an inverse relationship between ambient temperature and longevity in poikilothermic organisms such as *Drosophila* (see above). Pearl's rate of living hypothesis was partially based on experiments with different temperatures in *Drosophila* (Pearl, 1928; Sohal, 1986). In the late 1950s and early 1960s J. Maynard Smith used the relationship between temperature and life span to examine elements of the process. Maynard Smith, along with Clarke, proposed an alternative hypothesis to the rate of living hypothesis, in which no role is played by a temperature- or activity-dependent loss of a "longevity substance." Their hypothesis was called the "threshold" theory (Clarke and Smith, 1961a,b). This conclusion was based on studies from a temperature shift experiment. The life span of flies that spent a portion of their life at a high temperature, such as 30°C, and their remaining life at a lower temperature, such as 18°C, lived almost as long as those spending their entire life at 18°C, but more importantly much longer than would be predicted if the expected life span at 30°C was simply added to the expected life span at 18°C (Clarke and Smith, 1961a,b). If life span is determined by the loss of a certain amount of "vitality" then it would be expected that under this regimen, in which an animal lived what would be 50% of its life at 30°C and then switched to 18°C, the flies should live significantly less than the expected life span for animals constantly kept at 18°C. These data were interpreted to represent two phases of aging: an "irreversible aging" phase in which the rate is largely independent of temperature (from 15 to 30°C) and "dying" process late in life that is in part temperature dependent (Clarke and Smith, 1961a,b). Further investigation by others has led to conflicting results, but for the most part it appears that temperature does have a quantitative additive or subtractive effect on aging in *Drosophila* (reviewed by Lamb, 1978).

One reason lower temperatures increase life span in flies and other poikilotherms may be that chemical reactions and physical activity are decreased at these lower temperatures. Therefore, a decrease in physical activity would be expected to increase life span while an increase in physical activity should decrease life span. In house flies, Sohal and colleagues have demonstrated that by housing them under conditions in which they are unable to fly or compete with each other, their life spans will increase to the same amount as with

decreases in ambient temperature (Sohal and Buchan, 1981). Conversely, mutations in potassium channel-encoding genes—*Shaker* and *Hyperkinetic*—that increase physical activity and metabolic rate result in decreased life span and an increase in the slope of the mortality curve, suggesting that the process of aging has been accelerated (Trout and Kaplan, 1970). The conclusions from these studies are that similar to ambient temperature, physical activity, or more likely metabolic activity, shows an inverse relationship to length of life.

c. Cost of reproduction

The cost of reproduction on life span has been the subject of detailed investigation. Both the cost of egg laying for females and the cost of mating for males and females shorten life span. Virgin females can live twice as long as fully mated females (Smith, 1958, 1962). The effect on males is usually less, but this depends on the strain of flies examined and may involve the stresses males encounter in competing for and courting females (Partridge and Farquhar, 1981). In addition to the obvious cost of egg production female longevity studies have also shown that a significant part of the loss in life span of fully mated females is due to the transfer of seminal fluid from the males during mating. The removal of the accessory gland cells, the cells responsible for seminal fluid, using molecular genetic methods or cauterization, showed that transmission of seminal fluid from the male to the female during intercourse makes a major contribution to the decrease in life span of mated females (Chapman *et al.*, 1993, 1995). Females mated to males that still had sperm but not seminal fluid, and therefore still produced offspring, lived almost as long as virgin females (Chapman *et al.*, 1995). It is thought that some of the components of the male seminal fluid are involved in female life span reduction (Lung *et al.*, 2002). Whether reproduction imposes a direct cost on females, resulting in a shorter life span, or whether mating induces an acceleration of the normal aging process is not known. Preliminary data on mortality curves suggest that the major difference between virgins and fully mated females is not related to the slope of the mortality curve but to the y intercept, or as noted above to the hazard function (S. L. Helfand and B. Rogina, unpublished data). This suggests that the effects of mating on shortening female longevity are not due to an acceleration of the normal process of aging but to some other, perhaps toxic effect.

d. Diapause or a hibernation-type state

A. Comfort noted the idea that reproductive diapause or overwintering of short-lived species might be a metabolic state with importance for understanding aging.

> The selectionist argument which regards senescence as the decline of evolved survival-power through successive age groups is most convincing when we apply it to mammals and birds: among invertebrates, reservations require to be made. In those which are predominantly seasonal, with a total life-span less than one year, and which winter as fertilized adults, it is by no means true that at all times of the year young individuals must outnumber old in a free-running population. The autumn contingent of overwintering animals will consist of 'old' individuals. In such forms, the selective advantage of different genotypes will vary from season to season, and there will be an ultimate requirement that the adult be capable of living long enough to overwinter. Forms producing two broods annually will tend to select fertility in the spring brood and longevity in the autumn, but with a time lag of one generation between selection and potential expression. The mechanism of selection in such a system must be very complicated. (Comfort, 1979)

The relevance of diapause-like states to aging has been highlighted by the great successes in aging research seen with mutations in *Caenorhabditis elegans* associated with dauer formation, a state with similarities to diapause in insects. In a review Tatar and Yin (2001) describe the similarities between insect diapause and *Caenorhabditis elegans* dauer state and point out that understanding the physiology of diapause in insects may potentially clarify the process of aging. The relationships between insect diapause, longevity, and endocrinology suggest that this is an important physiological system that should yield a great deal of information concerning aging in *Drosophila*, as it has already done in *Caenorhabditis elegans*.

e. Caloric restriction/dietary restriction

The only known way of extending life span in mammals is through a restriction of calories while maintaining normal nutrition (Masoro, 2000). Life span increases in mice and rats of as much as 50–60% have been demonstrated. There are several studies using *Drosophila melanogaster* that report increases in life span with diets that decrease the total amount of calories in the food that the flies are living in or alter the levels of yeast available to the flies (Chippindale, 1993; Chapman and Partridge, 1996; Good and Tatar, 2001). There is one dissenting report that was unable to show an increase in life span (Bourg and Minois, 1996).

Interestingly, it appears that life extension in female *Drosophila*, induced by a reduction in the amount of calories in the fly food, does not show a large change in the slope of the mortality curve. Instead, a shift of the curve so that mortality starts at a later time, indicated by a change in *y* intercept, is primarily

seen (S. L. Helfand and B. Rogina, unpublished data). This appears to be unlike the case for mammals, for which reports suggest that life extension is associated with a shift in the slope of the mortality curve (Bartke *et al.*, 2001). However, in the article by Bartke *et al.* (2001) the data on mortality rates are not presented, only the survivorship curves. Extensive data on mammalian mortality rates under caloric restriction regimens have not been published and are sorely needed.

2. Genetic interventions that alter life span: Which will be helpful for understanding the biology of aging?

The use of genetic methods for dissecting the complexities of development suggests that these methods could also be applied to studying the process of aging. As with any genetic screen, one issue that needs to be considered is, while it may be possible to isolate mutations that increase or decrease life span, which of these mutations will provide useful information about aging? For example, decreasing metabolic activity, resulting in a fly that barely moves, may dramatically increase life span but at a cost to the animal that might be considered unacceptable. Mutations such as these, which decrease metabolism and increase life span, do not add much to our understanding of the process of aging. In addition, animals that forfeit a "normal life style," falling below the level of acceptable function, are unlikely to be the kinds of alterations that we would be interested in making in humans. Typically it is thought that two criteria need to be addressed when trying to determine whether an intervention that extends life span, such as a single-gene alteration, is of value for helping to understand the process of aging in *Drosophila*. Does the intervention lead to an extension in life span primarily through its effect on reproduction or metabolism? Decreases in reproduction, as noted above, for any number of reasons, can increase life span, and by themselves are less likely to reveal new insights into the biology of aging. If a decline in metabolic rate is associated with both an extension in life span and a significant slowing of physical activity then here, too, it is unlikely to be of great interest. Changes in metabolic rate that increase life span yet show little to no change in physical activity, however, may be of considerable interest. For a single-gene alteration that increases life span to be of great interest it should fulfill the criteria that it does not significantly decrease reproduction or metabolism/physical activity.

a. Studying genetic alterations that shorten or lengthen life span as a means of dissecting the process of aging

The great power of the *Drosophila* model lies in its unprecedented molecular genetic approaches and techniques. These approaches can be used to test present aging theories or to discover new ones. In the candidate gene

approach, the activity of a specific gene or genes associated with a particular physiological system suspected of being involved in aging is either augmented or attenuated and then its effect on life span is examined. For a field such as aging research, where so little is really understood, an even more valuable tool in the *Drosophila* arsenal is the random alteration of individual genes. Here the activity of individual genes is altered randomly, one at a time, and the resulting effect on the phenotype of interest—life span being one—is examined. Theoretical and practical considerations have delayed the use of single-gene mutagenesis for isolating life-extending genes.

b. Theoretical problems with single-gene mutagenesis for life span extension

Population and evolutionary biologists have argued for some time that a single-gene mutation extending life span without a loss in fertility or activity will not exist (Partridge and Harvey, 1993). The isolation of the *daf-2* mutation in *Caenorhabditis elegans* (Kenyon *et al.*, 1993) demonstrated that single-gene mutations that extend life span without a loss in major fitness characteristics can be isolated, and there are now more than 100 different genes in *Caenorhabditis elegans* that can be altered to extend life span in the nematode (Hekimi, 2000). The other concern of population and evolutionary biologists has been that the need for genetically homogeneous populations of flies for single-gene mutagenesis complicates the identification of "true" long-lived mutations. It is thought that inbred genetically homogeneous stocks will have accumulated random mutations, many of which may be causing an artificial decrease in life span in the stock. This may be further confounded by the usual practice of culturing flies that likely selects for early high reproduction, a fitness characteristic that has been predicted to be associated with a shortened life span. Hence, either through accumulation of random mutations or selection for early reproduction all inbred stocks could be short lived, compared with what should be its theoretical normal—a true wild-type stock. Using this argument, it is thought that any single-gene alteration that is shown to increase the life span of an inbred stock could just be rescuing deleterious mutations or reverting the effects of selected short-lived genes, thereby only returning the inbred line to its "normal" life span. Such extensions in life span may lead to the erroneous conclusion that these genes are involved in life span extension and aging. One method of checking for such a possibility is by crossing the putative long-lived mutant gene into different genetic backgrounds, both inbred and outbred stocks. This has been successfully done for *Indy*, *chico*, and *InR* mutations (Rogina *et al.*, 2000; Clancy *et al.*, 2001; Tatar *et al.*, 2001), and while caution is always warranted there is no strong reason to think that the use of inbred stocks for identifying genes important in aging will be overly problematic.

c. Procedural problems with single-gene mutagenesis for life span extension

The second major problem in attempting to isolate single-gene alterations by standard mutagenesis approaches is entirely procedural. There is no acceptable surrogate marker for assessing life span other than survivorship, measuring age at death of individuals in a population. Because of the stochastic nature of the life span curve, demographic studies and statistical analyses have suggested that at least 200 and preferably more than 300 flies are required for a reasonably accurate life span determination. Because of gender differences in life span at least 200 males and 200 females are required. Since the phenotype being examined requires a population, and hence all animals carrying the same mutation, then an F_1 screen cannot be done. Therefore, an F_2 screen is required, with all the accompanying problems of maintaining hundreds and thousands of stocks. In addition, if the goal is to look for a greater than 20% increase in mean or maximum life span, then for most wild-type strains growing at 25°C, such as Canton-S or Oregon-R, more than 2 to 3 months will be needed for mean life span and greater than 3 to 5 months or more for maximum life span assessment.

A combination of the number of flies and the extended period of time needed per life span assessment presents any individual laboratory with a major procedural or technical challenge. Depending on the amount of effort placed into the screen only a relatively small number (200–300) of lines could be expected to be tested per year. The usual expectation of screening through tens of thousands of mutagenized chromosomes cannot be achieved without a significant commitment or major procedural compromise or change. Approaches to streamline or accelerate this process introduce their own inherent problems. Using a higher temperature for life span tests, such as 29°C, can reduce the overall time to test each line, but usually only on the order of a few weeks, not months. In our experience, the mean and maximal life spans of many different inbred stocks of flies, when cultured at 29°C, differ by only a few days to 1 week; therefore, mean life spans are not dramatically shorter at 29 than 25°C, although maximal life spans may often be almost half. This appears to be due to a "rectangularization" of the survivorship curve at 29°C under optimal culture conditions, so that there is a smaller difference between mean and maximal life span at 29°C. In addition, there is anecdotal evidence that culturing animals at 29°C is not normal and introduces other problems when assessing a complex biological phenomenon such as aging. Another means of accelerating the screen would be to use fewer flies per life span test. Decreasing the number of individuals in each life span test has two potential benefits: it decreases the number of flies that need to be manipulated and counted in each life span study and it also decreases the number of flies that need to be grown up to start the life span study. Usually, obtaining enough age-matched cohorts to

do a life span study requires a significant set of cultures per line, and this too takes time and resources. The disadvantages of using fewer flies is that, as noted above, taking into account the stochastic nature of life span, a smaller population will likely result in a greatly increased rate of false positives and negatives. With life span as the phenotype, retesting all the false positives will be an additional major commitment.

The complexity of the aging process, however, may be a reason for optimism in genetic screens. If there are a number of different genes or physiological systems that, when altered, can significantly extend life span, this would increase the number of potential targets. With more targets the chances of finding a mutation that extended life span may not require having to examine tens of thousands of mutant chromosomes as is expected when there is only a single target gene of interest. It may be then that screening as few as 1000–2000 different lines may yield a gene of interest for aging. Still, even this can be a considerable undertaking.

d. Type of mutagenizing agent

Traditionally the generation of random gene alterations was through the use of various mutagenizing agents: chemical, radiation, or transposable elements. These agents would usually result in a decrease in function of that gene (of course, many examples of a gain in function are also known). In addition to creating loss-of-function and the occasional gain-of-function mutations, new schemes have been developed to overexpress random genes. The development and use of the yeast Gal4–UAS system in *Drosophila* (Brand and Perrimon, 1993) has led to screens in which overexpression and/or temporal/spatial misexpression of otherwise normal genes can also be utilized as a means of identifying genes or physiological systems important in life span determination. Taking advantage of the tendency of P-elements to insert in the 5′ regulatory region of genes, Rorth *et al.* developed a P-element construct with Gal4-binding sites (UAS) and (unidirectionality) that when inserted upstream of a native gene can drive the expression of that native gene in the presence of Gal4 (Rorth, 1996; Rorth *et al.*, 1998). By using Gal4 drivers—lines in which a specific promoter region has been set up to drive Gal4—the experimenter can express or overexpress the native gene in the spatial and temporal pattern determined by the promoter associated with Gal4 and hence determined by the experimenter. Many different Gal4 drivers are available, allowing the experimenter enormous power and flexibility. Using this system and other available genetics in *Drosophila* it is possible to increase or decrease the expression of any gene or group of genes in any subset of cells at any time in the entire life span of the fly—development through adult life. This is a particularly useful tool for examining adult life span since alterations, either increases or decreases, in the activities of many genes during

development could lead to lethality or dramatically alter development in such a way as to negate, obscure, or mask the effects of the gene during adult life (Johnston, 2002).

There are inherent benefits and problems with each of these approaches. Chemical mutagenesis (using by the application of ethyl methane sulfonate in *Drosophila*) leads to a high rate of mutation and, more important, maintenance of stocks in the same background as the original parental stock, an important advantage, but suffers from the difficulty of identifying the gene of interest. If life span extension is the phenotype being sought, then it will be difficult to perform classic meiotic recombination or deficiency mapping approaches to narrow down the chromosomal location of the gene of interest. The problem with using P-elements as a mutagen is that the experimenter is forced to cross a number of different stocks; hence there arises the possibility that any effect on life span seen could be due to a change in genetic background. With life span as a phenotype it may be laborious to prove that any life span extension seen is due to the insertion of the P-element and not due to a change in genetic background. For example, one of the means of determining causal relationship with P-element mutations is through the excision of the P-element and reversion or rescue of the phenotype. While this is relatively straightforward when the phenotype is a deleterious one, with life span extension, where "reversion" would mean a shortening of the life span to the original parental length, it is more complicated. Many different genetic events can lead to a shortening in life span during the crosses necessary to perform the excision of the P-element, make it homozygous, and then test it for reversion to normal life span. A reversion could indicate that the P-element was causally involved in the life span extension, but it could also be a result of any changes in genetic background that have occurred during the process of P-element excision. The use of the Gal4-UAS system is powerful but has problems in that up until now most Gal4 lines were made in different laboratories and almost certainly have been derived from different stocks. The site of insertion or any other chromosomal alterations involved in making the Gal4 stock might contribute to making each Gal4 driver line unique. Therefore, it is not entirely clear what the correct controls would be to determine whether a particular combination of Gal4 driver and UAS expression system was the cause of an increase in life span. Using inducible promoters, assuming that the inducible agent will not itself affect life span, may be the best choice for controlling for genetic background effects when analyzing life span.

The use of inducible promoters, or similar systems, that allow for the activation or inactivation of a particular gene or set of genes during adult life, and at particular times during aging, may turn out to be one of the best approaches. In addition to the ability to control the time of implementation of the genetic change, by definition the inducible promoter systems allow for

studies in which the control animals are genetically identical to the experimental animals, with the only difference being the agent of induction: drug, metals, or heat shock, for example. Tower and colleagues have pioneered this important approach for studying aging. As noted in the section on testing the oxidative stress hypothesis (Section VI.D.1) by using a heat shock-inducible FLP recombinase-based system the investigators set up the wherewithal to turn on, or off, genes of interest, such as those encoding superoxide dismutase (SOD) and catalase (Sun and Tower, 1999). They also initiated studies using the tetracycline-inducible system to turn on or off genes during adult life, which also shows great promise for ameliorating some of the problems noted here (Bieschke *et al.*, 1998). Additional newer approaches, such as the RU486-inducible promoter, have been utilized in *Drosophila* to study neuronal development and function (Osterwalder *et al.*, 2001; Roman *et al.*, 2001). In addition to the flexibility of using tissue-specific promoters the control of gene expression is reported to be tighter than for the tetracycline system.

C. Mutations that shorten or lengthen life span

1. Short-lived mutants

In principle, obtaining mutations or gene alterations that either decrease or increase life span could provide information about the normal process of aging. However, although obtaining and examining short-lived mutations could allow for more rapid progress, confounding problems have severely restricted the use of screens for short-lived mutants in *Drosophila* and consigned examination of already available short-lived mutations to the background. There are several reasons for this. A number of pathological conditions, most of which have little to reveal about the normal process of aging, can result in a shortened life span (e.g., neuronal degeneration, developmental defects in limbs and cuticle causing accidental death—becoming stuck in the food—or death by infection). Few mutations that result in a "sick" fly will provide information about the rate or character of the aging process. In the absence of criteria that can identify those mutations or alterations that "accelerate" aging from among the larger number of short-lived sick flies, screens for short-lived flies are usually thought to be of little help. This situation could dramatically change, however, if methods of assessing physiological age could be developed instead of having to rely on age at death as the only phenotype.

There are a few short-lived *Drosophila* mutations that have been used to test theories of aging. The neurobehavioral mutations *Shaker* and *Hyperkinetic*, both of which alter potassium channels, have been used to examine the relationship between metabolism and life span (Trout and Kaplan, 1970, 1981). Sh^5 and Hk^1 showed a 20–40% reduction in life span and are

associated with an increase in physical activity and metabolic rate (increased oxygen consumption). Examination of the mortality curve revealed a change in the slope of the curve, suggesting acceleration in the normal process of aging. Trout and Kaplan proposed that the association of increased metabolic rate, shortened life span, and change in slope of the mortality curve supports the rate of living hypothesis. The modern manifestation of this hypothesis is the oxidative stress hypothesis; and loss-of-function mutations in three components of the antioxidant system, Cu, Zn-SOD, catalase, and thioredoxin reductase (TxR), all cause a shortened life span, further supporting the oxidative stress hypothesis (Phillips *et al.*, 1989; Phillips and Hilliker, 1990; Orr *et al.*, 1992; Missirlis *et al.*, 2001). A confounding concern of these studies, however, is that in all cases a decrease in or lack of the antioxidant enzyme also occurred during development. How much of the shortening of adult life span is due to an acceleration of aging, and how much is due to the residual effects of damage during development? A study using a single putative biomarker of aging, the temporal pattern of expression of the *wingless* gene in the antennae, showed decline in this putative biomarker in SOD-deficient animals. This result is consistent with the interpretation that the short adult life span of SOD-deficient animals is due to an increase in the rate of aging of the adult fly. This suggests that studying the SOD-deficient animal may provide useful information about the normal process of aging.

The *drop dead* mutant is caused by an X-linked mutation that leads to a short adult life span, on the order of 2 weeks. Three different putative biomarkers of aging were shown to accelerate their progression in the *drop dead* background, suggesting that the early demise of the *drop dead* animal may be associated with an acceleration in the rate of aging (Rogina *et al.*, 1997). It is not thought that the *drop dead* mutation itself is an antiaging gene, but rather its mutation may somehow trigger an increase in the rate of aging (the slope of the mortality curve is consistent with interpretation of an increase in the rate of aging; the slope of the mortality curve is dramatically altered) (B. Rogina and S. L. Helfand, unpublished results). Since animals mutant for *drop dead* have a short life span, and there is at least some evidence from biomarkers that this may be due to an acceleration in the rate of aging, these animals may allow for a rapid means of looking for genes that affect aging by looking for genetic alterations that suppress or enhance the *drop dead*-shortened life span (Rogina *et al.*, 1997).

2. Long-lived mutants

Some of the first concerted attempts to identify genes important in extending life span in *Drosophila* were through examination of the laboratory-selected long-lived lines of Rose and Luckinbill (Arking *et al.*, 1988, 1993, 1996; Arking

and Dudas, 1989; Arking and Wells, 1990; Dudas and Arking, 1995; Mockett *et al.*, 2001). Despite several efforts, including studies using quantitative trait locus (QTL) analyses, it has been exceedingly difficult to disentangle the effects of individual genes (Fleming *et al.*, 1993; Kurapati *et al.*, 2000; Curtsinger and Khazaeli, 2002). Making use of the power of *Drosophila* molecular genetics, the search for genes that extend life span in *Drosophila* has included three complementary approaches: candidate gene approach, single-gene mutagenesis, and serendipity.

D. Candidate gene approach

1. Oxidative stress hypothesis: Tests using *Drosophila*

One of the most prominent theories of aging is the oxidative stress hypothesis (Sohal, 1986; Sohal and Weindruch, 1996; Sohal *et al.*, 2000). Many correlative data have demonstrated a relationship between the rate of accumulation of oxidative damage and aging, including work in *Drosophila* (e.g., short-lived loss-of-function mutations in SOD and catalase; increased accumulation of carbonyls in flies, etc.) (Sohal *et al.*, 2000). Loss-of-function mutants in each of the three known antioxidant enzyme systems—catalase, SODs, and TxR—have each exhibited shortenings of adult life span (Phillips *et al.*, 1989; Phillips and Hilliker, 1990; Orr *et al.*, 1992; Missirlis *et al.*, 2001). As noted above, some of this may be due to lack of the antioxidant enzymes during development so that a shortening of life span could reflect either the necessity of the enzymes during adult life or a "sick" adult due to accumulated developmental defects. Hence a more direct test of the oxidative stress hypothesis would be to use the transgenic systems in *Drosophila* to increase antioxidant systems and investigate whether there is a causal link between oxidative damage and aging. Several studies have addressed this and shown at best a moderate effect on life span extension with increasing the known antioxidant enzymes throughout life. Overexpression of either Cu, Zn-SOD, Mn-SOD, glutathione reductase, or catalase alone throughout life shows either a decrease in life span or no change (Mockett *et al.*, 1999a,b; Sohal *et al.*, 1995; Orr and Sohal, 1992, 1993; Orr *et al.*, 1992; Parkes *et al.*, 1998; Phillips *et al.*, 2000). Combined increases in both Cu, Zn-SOD and catalase expression were reported to cause a 30% increase in life span (Orr and Sohal, 1994). However, the complicated crosses necessary to engineer two different transgenes into the same animal raised concerns over what the correct genetically matched controls should be. Subsequent studies have shown a modest or minimal effect on life span in flies with SOD–catalase dual overexpression, despite a decrease in the accumulation of oxidative damage (Sohal *et al.*, 1995). Although there is only a minimal increase in life

span with a measured decrease in oxidative damage a more prolonged period of activity later in life, compared with controls, was reported.

Because of the possibility that overexpression of these antioxidant enzymes during development may be detrimental and thus somehow abrogate any positive effect on adult life span, molecular genetic approaches for limiting overexpression to only the adult portion of life have been developed. The FLP recombinase system to activate the overexpression of SOD and/or catalase during adult life was used to increase expression in the adult (Sun and Tower, 1999). A brief heat shock, delivered to activate FLP recombinase during adult life, led to a measurable increase in antioxidant activity (two- to three-fold for catalase and 1.5-fold for SOD). For catalase overexpression, despite a two- to three-fold increase in activity in the adult after induction there were "neutral or slightly negative effects on mean life span" (Sun and Tower, 1999). SOD overexpression may show a small positive effect on life span, approximately 15–20%. Of the several transgenic lines examined two showed consistent increases in SOD activity. One of these two lines showing a 1.5-fold increase in SOD overexpression and a 10–15% increase in mean life span; while the other line, with the same 1.5-fold overexpression of SOD, showed no increase in one genetic background and a 15–20% increase in another one. However, along with these increases in life span was a large reduction in physical activity in these same lines, as determined by assays of negative geotaxis and a significant reduction in fertility as measured by number of offspring produced. The possibility that the 10–20% increase in life span is related to a reduction in physical activity or fertility needs to be further examined. Interestingly, the addition of a second SOD transgene to the line showing an increase in life span led to the same induction in SOD activity (1.5-fold); but these flies no longer exhibited an increase in life span. All of these studies were performed on females. One conclusion may be that either a 1.5-fold increase is an insufficient increase in SOD to affect life span, or perhaps antioxidant activity via these enzyme systems is not a primary factor in aging. Knowledge of the rate of accumulation of oxidative damage in these lines will be important in making these distinctions. In the final section of this report (Sun and Tower, 1999) a different FLP recombinase construct was used; males from two lines were examined, both containing the same SOD insert, one having an additional SOD insert. In the line with the single insert, but not in the one with an additional SOD insert, a 48% increase in life span was noted (the other stock containing the same SOD insert and an additional one showed a 14% increase in life span). There are several concerns about this result. For example, the uninduced control life span for this line, used to determine the 48% score, had the lowest life span reported among all the lines reported. The mean life span of the induced life extended line reported was itself still lower than the uninduced control life span of this same transgenic line in all the other genetic backgrounds.

Although the inducible expression of antioxidants allows for restricting expression during the adult phase, the constructs so far used have not allowed for tissue-specific expression. It may be that another confounding issue is that overexpression of antioxidants in all tissues causes counteracting effects. The use of the yeast Gal4–UAS system for specific tissue expression has been utilized in the fly. In a fascinating study it was reported that overexpression of human SOD1 in a subset of tissues, the motor neurons of the adult fly, could increase life span by as much as 40–50% (Parkes *et al.*, 1998). This unexpected finding is of great interest for several reasons. From a theoretical perspective evolutionary theories of aging had predicted that all physiological systems should fail at or around the same time and that improvement in one organ system should not have a significant effect on overall life span. The fact that overexpression in only a subset of the nervous system led to an increase in adult life span is of great interest. Further studies have confirmed that overexpression of human SOD2 in these same tissues can also increase life span (Phillips *et al.*, 2000). Disconcertingly, however, when catalase was simultaneously overexpressed with SOD1, life span extension was lost (Phillips *et al.*, 2000). Additional studies will need to be done to sort out these fascinating but seemingly contradictory findings.

To date, the hope that the powerful molecular genetic tools available for the *Drosophila* model would be able to demonstrate the causal relationship between oxidative damage and aging has not been realized. Whether there are specific technical issues that have prevented this, or whether a more balanced or directed alteration in antioxidant levels is required, is not clear.

2. Insulin-like signaling pathway

The discovery that mutations in genes involved in the insulin pathway can dramatically increase life span in *Caenorhabditis elegans* (*daf-2*, *daf-23/age-1*) led to the examination of similar genes in the fly. Mutations in the genes encoding insulin-like receptor (*InR*), the insulin receptor substrate (*chico*), phosphatidylinositol 3-kinase (PI3K) Dp110/p60, and PI3K target protein kinase B (PKB or Dakt1) were each examined for possible effects on life span (Clancy *et al.*, 2001; Tatar *et al.*, 2001). Confounding these experiments was the fact that in most cases mutations in these genes are lethal or cause severe growth defects. For example, viable *InR* alleles or transheterozygote *InR* animals as well as homozygous null *chico* mutations all give rise to small (dwarf) and fragile-appearing flies. To control for genetic background and inbred line effects, flies from each of these single-gene mutations were outcrossed to outbred stocks before life span examination. One particular heteroallelic *InR* hypomorphic mutant combination (InR^{P5545}/InR^{E19}) showed an increase in median female life span of 85%

with maintenance of normal metabolic rates (Tatar *et al.*, 2001). The mortality curve showed a shift to the right with little change in slope. Male life span was not increased although in later ages, after an initial period of significantly higher than normal mortality, a decline in later life age-specific mortality was seen. Interestingly, the long-lived females were sterile, with egg development arrested before vitellogenesis (similar to juvenile hormone-deficient animals). Treatment with a juvenile hormone analog initiated vitellogenesis although it did not restore fertility; but more importantly it did revert the long-lived female to the shorter (normal) life span. The rescue of the life extension without resumption of normal reproduction led the authors to suggest that the increase in adult life span due to a decrement in juvenile hormone may not simply be due to the decrease in reproduction.

In another report, dwarf *chico* homozygous null mutants when crossed into the background of an outbred stock showed an increase in female mean life span of up to 48% with little to no increase in male life span (Clancy *et al.*, 2001). Females heterozygous for the *chico* null mutation, which are not dwarf, had an increase in mean life span of up to 36%. Similar to the *InR* females, there is greatly reduced fertility in *chico* mutants of that sex. *chico* homozygotes are essentially sterile, and heterozygotes have much reduced fertility. The authors reasoned that, if the cause of the increase in female life span in *chico* mutant animals is due to egg production, then a comparison with other female–sterile mutants with extended life spans may be informative. Therefore the life span of dominant female–sterile mutant ovo^{D1} flies (ovo^{D1} females are sterile with nonvitellogenic egg chambers—similar to *chico*) was compared with the life span of animals heterozygous for the *chico* mutation and ovo^{D1} (ovo^{D1}/+;*chico*/+), *chico* heterozygote (*chico*/+), and *chico* homozygote (*chico*/*chico*) animals. The ovo^{D1}/+;*chico*/+ and *chico*/+ females had the same mean life span, approximately 20% greater than the ovo^{D1} animals alone. The authors concluded that if the life span-extending effect of *chico* mutations on females is related to an effect on female reproduction it is not via the same mechanism by which ovo^{D1} increases life span. However, the fact that *chico*/+ and ovo^{D1}/+; *chico*/+ animals both have the same life span suggests that the cause of the life span extension in each case may be due to alterations in the same physiological system—in this test they appear epistatic to each other. *chico* and ovo^{D1} life span extension is not additive, as might be expected if they affected two independent physiological pathways.

The demonstration of life span extension in one specific heteroallelic combination involving an *InR* mutation and a *chico* mutation suggests a role of the insulin-like signaling pathway in life span determination in *Drosophila*. Before concluding that the insulin-like signaling pathway is as important in *Drosophila* life span determination as it is for *Caenorhabditis elegans*, certain issues need to be resolved. First is the fact that in all cases life span extension is

restricted to females, and these females are either sterile (*InR* heteroalleles or *chico* homozygotes) or subfertile (*chico* heterozygotes). Whether this is related to the well-known effect of reduced fertility on extending life span or to a more interesting alteration involving signaling between the germ line and the soma for determining life span, similar to what is seen in *Caenorhabditis elegans* (Hsin and Kenyon, 1999; Lin *et al.*, 2001), will need further investigation. In addition, except in the case of the *chico* heterozygote animals, there are severe pleiotropic effects of these mutations that confound the benefits of life span extension. The long-lived mutants (except for the *chico* heterozygotes) have retarded developmental periods, over twice as long as normal, with high mortality during development, and give rise to dwarfed fragile adults that would be unlikely to survive without the great care they received in the laboratory.

3. Protein repair mechanisms

In the sections on stressors and oxidative damage it was noted that the accumulation of abnormal or defective macromolecules, particularly DNA and protein, might play an important role in the rate of aging and the determination of life span. Interventions to limit the accumulation of damaged macromolecules include a reduction in initial damage, an increase in repair processes, and an increase in removal and replacement of the abnormal or defective macromolecules. Studies on oxidative damage have focused on attempts to decrease oxidative damage and limit the damage at its initial point of occurrence. Another approach is to boost the repair systems, in particular protein repair systems.

a. Chaperones and heat shock proteins

One general approach to increasing repair or removal of damaged proteins is through the chaperone system, of which the heat shock proteins hsp70 and hsp40 are a major component. These proteins are known to be important in protecting other such macromolecules from certain denaturing conditions, such as heat, and help in the refolding of misfolded proteins, preventing aggregation of defective molecules or dissolving aggregates when formed. Chaperones have been utilized to "treat" neurodegenerative disorders in flies that were induced by expression of abnormal human proteins in the fly—the fly equivalent of Huntington's disease, spinocerebellar degeneration 1 and 3, and Parkinson's disease (Warrick *et al.*, 1999; Chan *et al.*, 2000; Kazemi-Esfarjani and Benzer, 2000; Auluck *et al.*, 2002). In the case of fly Parkinson's disease the level of chaperones was directly related to the severity of the neurodegeneration. Increasing chaperone levels prevented the specific neurodegeneration, while a reduction in chaperone levels by the use of specific mutations led to more severe neurodegenerative effects.

The role of chaperone proteins in aging is less clear. The levels of several different chaperone proteins, such as hsp70 and hsp22, have been shown to increase with age (Wheeler *et al.*, 1995; King and Tower, 1999). Analysis of one of the laboratory-selected long-lived lines showed an increase in heat shock proteins in the long-lived line compared with lines with normal life spans (Kurapati *et al.*, 2000). The long-life mutation, *methuselah* (*mth*), confers increased resistance to stresses, including heat shock (Lin *et al.*, 1998). Taken together these data suggest a possible role for chaperones or heat shock proteins as a protective agent during aging. Despite this, there have been few reports showing any positive effect on life span with an increase in chaperone levels. Tatar *et al.*, showed that a mild heat shock in a strain carrying multiple additional copies of the *hsp70* gene led to a significant decline in age-specific mortality in later ages compared with controls (Tatar *et al.*, 1997). Follow-up studies by Curtsinger and colleagues have failed to demonstrate any significant positive effect on life span or physical activity when *hsp70* expression was increased (Khazaeli *et al.*, 1997; Minois *et al.*, 2001).

b. Protein carboxymethyltransferase

Another approach that has been explored is to increase the protein carboxymethyltransferase (PCMT) activity of flies (Chavous *et al.*, 2001). PCMT is thought to be an important initial step in the repair of atypical protein isoaspartyl residues that arise during aging spontaneously (Chavous *et al.*, 2001). There are at present no data detailing whether the accumulation of such residues is a feature of aging in the fly or other organisms. An increase in mean life span only when flies were cultured at 29°C, but not at 25°C, was reported when PCMT was overexpressed using the Gal4–UAS system and driven by an *hsp70* driver or a spatially ubiquitous actin driver. The authors speculate that the increase in mean life span at 29°C suggests that PCMT overexpression is of value when flies live under certain environmental conditions, such as high temperatures, or under other mild stress conditions. The fact that the effect is not seen at 25°C is interesting but of concern in trying to generalize this as an important element in normal aging.

4. Chromatin structure

The demonstration that histone deacetylatases and other elements important in modifying or maintaining chromatin structure are important in the determination of life span in yeast (Kim *et al.*, 1999; Guarente and Kenyon, 2000; Jazwinski, 2000) has led to the examination of such effects in *Drosophila*. It was reported that the feeding of a drug known to inhibit histone deacetylation, 4-phenylbutyrate (PBA), increases life span and alters the expression of a number of different genes without a significant loss of fertility or physical activity (Kang

et al., 2002). In addition to providing information about a possible effect of PBA on life span, this study also suggests that a high-throughput drug screen for extending life span in *Drosophila* is a plausible goal. As with genetic interventions, any drug that is found to have a significant effect on life span will need to be examined to determine whether this is due to a secondary effect on fertility, physical activity, or metabolism. Initial studies on flies fed PBA suggest that it does not cause a decrease in fertility or physical activity (Kang *et al.*, 2002).

5. Candidate gene approaches using newer inducible promoter/drivers

Two of the possible concerns in interpreting the studies in which a particular gene has been overexpressed is the tissue specificity of the expression—is it being expressed in the correct tissues or subcellular portion of the cell, and is it expressed at the right time for the overexpression to have its maximal positive effect? Thus a number of investigators have begun to develop other means of controlling gene overexpression, both temporally and spatially. Several techniques have been devised and are only now beginning to be used. These include the use of FLP recombinase to activate expression of a gene in the adult phase (Sun and Tower, 1999) or of inducers that can activate or inactivate promoters by the feeding of the aforementioned chemicals tetracycline (Bieschke *et al.*, 1998) or RU486 (Osterwalder *et al.*, 2001; Roman *et al.*, 2001). The use of inducible promoters should help alleviate some of the problems with the genetic background of controls. In this case the experimental and control animals could be genetically identical. The presence or absence of the inducing agent will constitute the main difference.

E. Single-gene alterations that extend life span

Perhaps the most valuable approach to exploring complex biological phenomena in *Drosophila* is to alter single genes and examine them for changes in the phenotype of interest. The selection of an appropriate phenotype for studying the process of aging is somewhat problematic. As noted at the beginning of this review, the methods of measuring aging fall into two different categories. The most common approach is to perform demographic studies on life span. However, this approach does not examine aging, or the rate of aging, but instead provides a population-based analysis of age at death of individuals in a cohort, which indirectly reflects elements of the aging process. Although the relationship between these measurements and the aging process is not perfect, it is still the best phenotype available for understanding aging. Other surrogate

phenotypes for rate of aging or for longevity have been proposed but are not yet validated. Therefore, the most common means of looking for genes that affect aging is to look for genes that alter life span, usually by extending it, since as noted in an earlier section, alterations that shorten life span may often do so by pathological changes that do not reveal useful information about the normal process of aging. Therefore, most studies look for genes that increase life span.

The benefits of single-gene alterations for exploring elements of complex biological phenomena are well known. Of principal importance for studying aging is that approaches utilizing single-gene alterations make use of random alterations and do not require any *a priori* knowledge of the physiology of the system as candidate approaches do. Only the ability to measure changes in the phenotype being examined, such as life span, are important.

Unlike in laboratory selection experiments, the ability to alter one or a few genes at a time allows the experimenter to determine rapidly the relative contribution each genetic alteration may make to the process being studied. The difficulty of disentangling the various genetic elements associated with laboratory selection experiments is a well-known problem. In addition, laboratory selection approaches can select only from among the allele pool originally available. Single-gene alterations from mutageneses have the additional potential, arguably superior to selection experiments, of creating new or unique proteins as well as altering the expression profile, temporally or spatially, of gene products in a manner that may never normally be found in the wild.

There are two notable examples of genes that extend life span that were discovered through a screen looking for long-lived mutations or by serendipity. In both cases the mutagen was a transposable element, an enhancer trap P-element (O'Kane and Gehring, 1987).

1. *methuselah*

A reduction in the level of the *methuselah* (*mth*) gene, which encodes a member of a unique subfamily of the G-coupled transmembrane receptor-like proteins, results in a 35% increase in mean life span (Lin *et al.*, 1998). (Fig. 2.12). The life-extending effect of *mth* is seen when it has been partially reduced; null mutations are embryonic lethal and heterozygote *mth*/+ animals also have extended life spans. Life extension is seen in both males and females and females are reported to have normal reproduction (Lin *et al.*, 1998). *mth* mutant long-lived animals are also resistant to various stresses, including heat, starvation, desiccation, and paraquat (an oxygen radical generator). The mechanism by which mutations in *mth* lead to life span extension is not

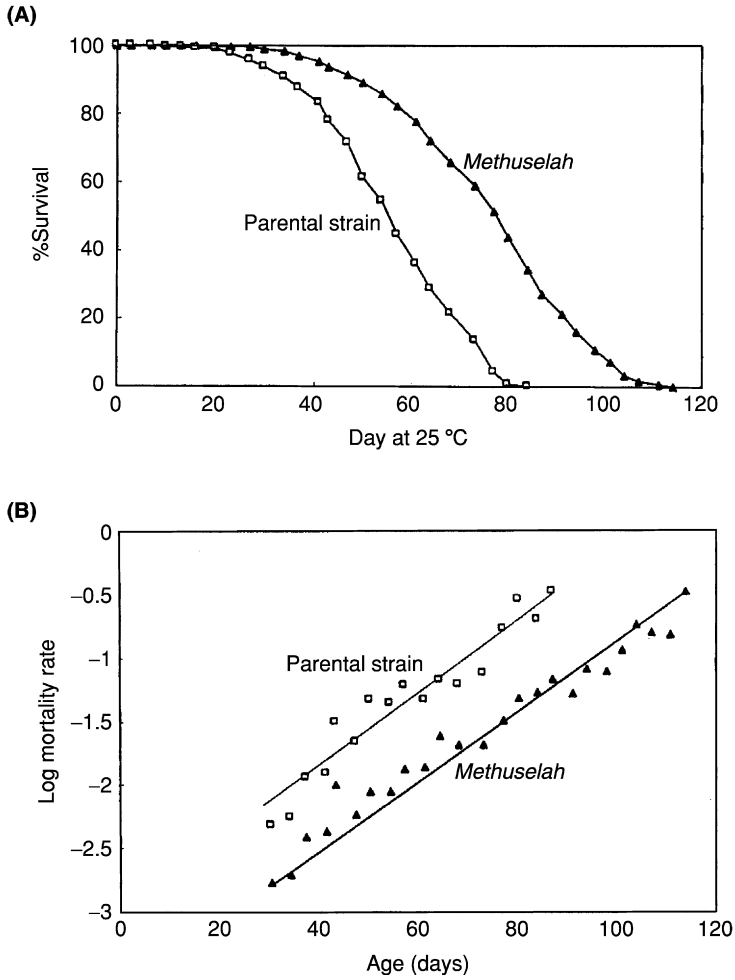


Figure 2.12. Mutations in the *mth* gene result in a 35% increase in average life span (A). Life span extension is associated with a shift in the hazard function in the mortality curve (B). From Lin *et al.* (1998).

understood. The ligand, which binds to the MTH protein, and the pathway it stimulates are not yet known.

2. *Indy*

A reduction in the level of the *Indy* gene, which encodes a putative transporter of Krebs cycle intermediates, leads to a near doubling of the average life span

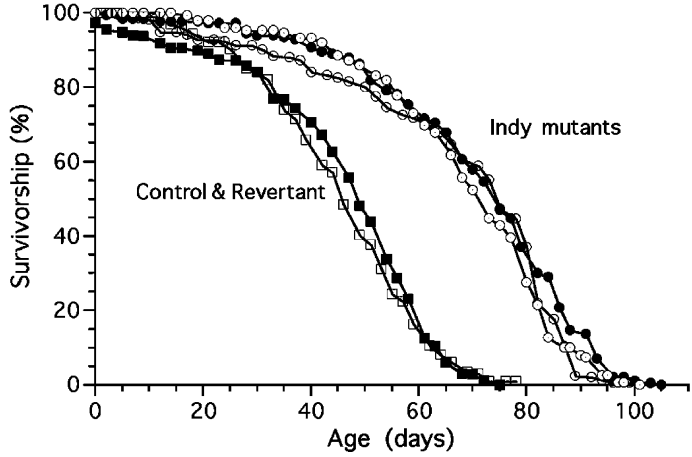


Figure 2.13. Male or female animals heterozygous for *Indy* (mutated at a locus encoding a Krebs cycle transporter protein) have a greater than 90% increase in average life span. From Rogina *et al.* (2000).

and a more than 40% increase in maximal life span of both male and female flies without a loss of reproduction or physical activity (Rogina *et al.*, 2000). (Fig. 2.13). Expression of *Indy* is most prominently found, by *lacZ* reporter expression, in the fat body, oenocytes, and a small subsegment of the midgut in adult flies. Gene dosage experiments showed that a partial reduction in *Indy* activity—in animals heterozygous for *Indy* (*Indy* mutant/wild-type *Indy*)—resulted in a near doubling of the life span. Further reduction of the gene product—in *Indy* homozygous mutant flies—caused a much smaller increase in life span compared with controls. More severe reductions in *Indy* activity, in flies hemizygous for an *Indy* mutation (*Indy* mutant/deletion of the locus), resulted in a shortening of life span compared with controls. The predicted homology of *Indy* to mammalian dicarboxylate cotransporters, localization in tissues associated with intermediary metabolism, and gene dose effects on life span suggest that a reduction in nutrient intake in *Indy* mutants may be the cause of the life span extension. It was postulated that *Indy* mutations might be inducing a state similar to caloric restriction.

VII. CONCLUSIONS

Aging research has begun to experience a true renaissance. A combination of new and traditional molecular genetic approaches has provided new insights and stimulated great optimism in the field. Research derived from *Drosophila*

melanogaster and other model organisms, such as *Caenorhabditis elegans*, has been among the major forces driving this renaissance. The ability to rapidly test the impact of alterations in different physiological systems on the process of aging is one reason *Drosophila* and *Caenorhabditis elegans* are useful for aging research. In the longer term, understanding the process of aging will require a broad unbiased method of identifying genes and physiological systems important in aging. It is likely that this will first be provided by well-developed molecular genetic systems such as *Drosophila*. For example, the powerful new molecular genetic techniques in *Drosophila*, allowing for control of the level of expression of individual genes in any tissue at any time during life, permits unprecedented control over the molecular genetic life of the fly. This will surely be of great utility in examining complex biological phenomena such as aging. Another challenge to aging researchers is to identify additional methods of assessing aging that augment measures of age at death. Measures of physiological change that are able to distinguish between young and old individuals and populations are of critical importance. It is likely that here too *Drosophila* will provide a useful lead.

Acknowledgments

The authors thank Marvin L. Tanzer, Joseph Jack, Robert A. Reenan, and Marc Lalande for helpful comments and discussion and for support from NIH Grants AG-14532 to S.L.H., ES-11463 and AG-20816 to B.R., a Donaghue Investigator Award from the Patrick and Catherine Weldon Donaghue Medical Foundation (S.L.H.), and an Ellison Medical Foundation Senior Investigator Award (S.L.H.).

References

Alpatov, W. W., and Pearl, R. (1929). Experimental studies on the duration of life. XII. Influence of temperature during the larval period and adult life on the duration of the life of the imago of *Drosophila melanogaster*. *Am. Nat.* **63,** 37–67.

Arking, R. (1991). "Biology of Aging: Observations and Principles." Prentice Hall, Englewood Cliffs, NJ.

Arking, R., and Dudas, S. P. (1989). Review of genetic investigations into the aging processes of *Drosophila*. *J. Am. Geriatr. Soc.* **37,** 757–773.

Arking, R., and Wells, R. A. (1990). Genetic alteration of normal aging processes is responsible for extended longevity in *Drosophila*. *Dev. Genet.* **11,** 141–148.

Arking, R., Buck, S., Wells, R. A., and Pretzlaff, R. (1988). Metabolic rates in genetically based long lived strains of *Drosophila*. *Exp. Gerontol.* **23,** 59–76.

Arking, R., Dudas, S. P., and Baker, G. T., III (1993). Genetic and environmental factors regulating the expression of an extended longevity phenotype in a long lived strain of *Drosophila*. *Genetica.* **91,** 127–142.

Arking, R., Force, A. G., Dudas, S. P., Buck, S., and Baker, G. T., III (1996). Factors contributing to the plasticity of the extended longevity phenotypes of *Drosophila*. *Exp. Gerontol.* **31,** 623–643.

Auluck, P. K., Chan, H. Y. E., Trojanovski, J. Q., Lee, V. M.-L., and Bonini, N. (2002). Chaperone suppression of α-synuclein toxicity in a *Drosophila* model for Parkinson's disease. *Science*. **295,** 865–868.

Bartke, A., Wright, J. C., Mattison, J. A., Ingram, D. K., Miller, R. A., and Roth, G. (2001). Extending the lifespan of long-lived mice. *Nature*. **414,** 412.

Bieschke, E. T., Wheeler, J. C., and Tower, J. (1998). Doxycycline-induced transgene expression during *Drosophila* development and aging. *Mol. Gen. Genet*. **258,** 571–579.

Bourg, E. L., and Minois, N. (1996). Failure to confirm increased longevity in *Drosophila melanogaster* submitted to a food restriction procedure. *J. Gerontol. A Biol. Sci. Med. Sci*. **51,** B280–B283.

Bozuck, A. N. (1972). DNA synthesis in the absence of somatic cell division associated with ageing in *Drosophila subobscura*. *Exp. Gerontol*. **7,** 147–156.

Brand, A. H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*. **118,** 401–415.

Carey, J. R., Liedo, P., Orzoco, D., and Vaupel, J. W. (1992). Slowing of mortality rates at older ages in large medfly cohorts. *Science*. **258,** 457–461.

Chan, H. Y., Warrick, J. M., Gray-Board, G. L., Paulson, H. L., and Bonini, N. M. (2000). Mechanisms of chaperone suppression of polyglutamine disease: Selectivity, synergy and modulation of protein solubility in *Drosophila*. *Hum. Mol. Genet*. **9,** 2811–2820.

Chapman, T., and Partridge, L. (1996). Female fitness in *Drosophila melanogaster*: An interaction between the effect of nutrition and of encounter rate with males. *Proc. R. Soc. Lond. B Biol. Sci*. **263,** 755–759.

Chapman, T., Hutchings, J., and Partridge, L. (1993). No reduction in the cost of mating for *Drosophila melanogaster* females mating with spermless males. *Proc. R. Soc. Lond. BBiol. Sci*. **253,** 211–217.

Chapman, T., Liddle, L. F., Kalb, J. M., Wolfner, M. F., and Partridge, L. (1995). Cost of mating in *Drosophila melanogaster* females is mediated by male accessory gland products. *Nature* **373,** 241–244.

Chavous, D. A., Jackson, F. R., and O'Connor, C. M. (2001). Extension of the *Drosophila* lifespan by overexpression of a protein repair methyltransferase. *Proc. Natl. Acad. Sci. USA* **98,** 14814–14818.

Chippindale, A. K. (1993). Phenotypic plasticity and selection in *Drosophila* life-history evolution. I. Nutrition and the cost of reproduction. *J. Evol. Biol*. **6,** 171–193.

Clancy, D. J., Crems, D., Harshman, L. G., Oldham, S., Stocker, H., Hafen, E., Leevers, S. J., and Partridge, D. (2001). Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. *Science* **292,** 104–106.

Clarke, J. M., and Smith, J. M. (1961a). Independence of temperature of the rate of ageing in *Drosophila subobscura*. *Nature* **190,** 1027–1028.

Clarke, J. M., and Smith, J. M. (1961b). Two phases of ageing in *Drosophila subobscura*. *J. Exp. Biol*. **38,** 679–684.

Comfort, A. (1979). "The Biology of Senescence." Churchill Livingstone, Edinburgh, UK.

Curtsinger, J. W., Fukui, H. H., Townsend, D. R., and Vaupel, J. W. (1992). Demography of genotypes: Failure of the limited life-span paradigm in *Drosophila melanogaster*. *Science* **258,** 461–463.

Curtsinger, J. W., and Khazaeli, A. A. (2002). Lifespan, QTLs, age-specificity, and pleiotropy in *Drosophila*. *Mech. Ageing Dev*. **123,** 81–93.

Djawdan, M., Chippindale, A. K., Rose, M. R., and Bradley, T. J. (1998). Metabolic reserves and evolved stress resistance in *Drosophila melanogaster*. *Physiol. Zool*. **71,** 584–594.

Dudas, S. P., and Arking, R. (1995). A coordinate upregulation of antioxidant gene activities is associated with the delayed onset of senescence in a long-lived strain of *Drosophila*. *J. Gerontol. A Biol. Sci. Med. Sci*. **50,** B117–B127.

Finch, C. E. (1990). "Longevity, Senescence, and the Genome." University of Chicago Press, Chicago.
Fleming, J. E., Spicer, G. S., Garrison, R. C., and Rose, M. R. (1993). Two-dimensional protein electrophoretic analysis of postponed aging in *Drosophila*. *Genetica*. **91,** 183–198.
Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and a new mode of determining the value of life contingencies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **115,** 513–585.
Good, T. P., and Tatar, M. (2001). Age-specific mortality and reproduction respond to adult dietary restriction in *Drosophila melanogaster*. *J. Insect Physiol.* **47,** 1467–1473.
Guarente, L., and Kenyon, C. (2000). Genetic pathways that regulate ageing in model organisms. *Nature* **408,** 255–262.
Hekimi, S. (2000). Crossroads of aging in the nematode *Caenorhabditis elegans*. *Results Probl. Cell Differ.* **29,** 81–112.
Helfand, S. L., and Inouye, S. K. (2002). Rejuvenating views of the ageing process. *Nat. Rev. Genet.* **3,** 149–153.
Helfand, S. L., and Rogina, B. (2000). Regulation of gene expression during aging. *Results Probl. Cell Differ.* **29,** 67–80.
Helfand, S. L., Blake, K. J., Rogina, B., Stracks, M. D., Centurion, A., and Naprta, B. (1995). Temporal patterns of gene expression in the antenna of the adult *Drosophila melanogaster*. *Genetics* **140,** 549–555.
Hsin, H., and Kenyon, C. (1999). Signals from the reproductive system regulate the lifespan of *C. elegans*. *Nature* **399,** 362–366.
Ito, K., and Hotta, Y. (1992). Proliferation pattern of postembryonic neuroblasts in the brain of *Drosophila melanogaster*. *Dev. Biol.* **149,** 134–148.
Jazwinski, S. M. (2000). Metabolic control and ageing. *Trends Genet.* **16,** 506–511.
Johnson, T. E., Lithgow, G. J., and Murakami, S. (1996). Hypothesis: Interventions that increase the response to stress offer the potential for effective life prolongation and increased health. *J. Gerontol. A Biol. Sci. Med. Sci.* **6,** B392–B395.
Johnston, D. S. (2002). The art and design of genetic screens: *Drosophila melanogaster*.. *Nat. Rev. Genet.* **3,** 176–188.
Kang, H. L., Benzer, S., and Min, K. T. (2002). Life extension in *Drosophila* by feeding a drug. *Proc. Natl. Acad. Sci. USA* **99,** 838–843.
Kazemi-Esfarjani, P., and Benzer, S. (2000). Genetic suppression of polyglutamine toxicity in *Drosophila*. *Science* **287,** 1837–1840.
Kenyon, C. J., Chang, J., Gensch, E., Rudner, A., and Tabtiang, R. (1993). A *C. elegans* mutant that lives twice as long as wild type. *Nature* **366,** 461–464.
Khazaeli, A. A., Tatar, M., Pletcher, S. D., and Curtsinger, J. W. (1997). Heat-induced longevity extension in *Drosophila*. I. Heat treatment, mortality, and thermotolerance. *J. Gerontol. A Biol. Sci. Med. Sci.* **52,** B48–B52.
Kim, S., Benguria, A., Lai, C. Y., and Jazwinski, S. M. (1999). Modulation of life-span by histone deacetylase genes in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **10,** 3125–3136.
King, V., and Tower, J. (1999). Aging-specific expression of *Drosophila* hsp22. *Dev. Biol.* **207,** 107–118.
Kurapati, R., Passananti, H. B., Rose, M. R., and Tower, J. (2000). Increased hsp22 RNA levels in *Drosophila* lines genetically selected for increased longevity. *J. Gerontol. A Biol. Sci. Med. Sci.* **55,** B552–B559.
Lamb, M. J. (1964). The effects of radiation on the longevity of female *Drosophila subobscura*. *J. Insect Physiol.* **10,** 487–497.
Lamb, M. J. (1978). Ageing. *In* "The Genetics and Biology of *Drosophila*" (M. Ashburner, and T. R. F. Wright, eds.), Vol. 2c, pp. 43–104. Academic Press, London.
Lin, K., Hsin, H., Libina, N., and Kenyon, C. (2001). Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nat. Genet.* **28,** 139–145.

Lin, Y. J., Seroude, L., and Benzer, S. (1998). Extended life-span and stress resistance in the *Drosophila* mutant *methuselah*. *Science* **282,** 943–946.

Loeb, J., and Northrop, J. H. (1916). Is there a temperature coefficient for the duration of life? *Proc. Natl. Acad. Sci. USA* **2,** 456–457.

Loeb, J., and Northrop, J. H. (1917). On the influence of food and temperature upon the duration of life. *J. Biol. Chem.* **32,** 103–121.

Luckinbill, L. S. (1998). Selection for longevity confers resistance to low-temperature stress in *Drosophila melanogaster*. *J. Gerontol. A Biol. Sci. Med. Sci.* **53,** 147–153.

Luckinbill, L., and Clare, M. J. (1985). Selection for life span in *Drosophila melanogaster*. *Heredity* **55,** 9–18.

Luckinbill, L. S., and Clare, M. J. (1987). Successful selection for increased longevity in *Drosophila*: Analysis of the survival data and presentation of a hypothesis on the genetic regulation of longevity. [letter to the editor]. *Exp. Gerontol.* **22,** 221–226.

Luckinbill, L., Arking, R., Clare, M. J., Cirocco, W. C., and Buck, S. (1984). Selection for delayed senescence in *Drosophila melanogaster*. *Evolution* **38,** 996–1003.

Lung, O., Tram, U., Finnerty, C. M., Eipper-Mains, M. A., Kalb, J. M., and Wolfner, M. F. (2002). The *Drosophila melanogaster* seminal fluid protein Acp62F is a protease inhibitor that is toxic upon ectopic expression. *Genetics* **160,** 211–224.

Masoro, E. J. (2000). Caloric restriction and aging: An update. *Exp. Gerontol.* **35,** 299–305.

Medawar, P. B. (1957). An unsolved problem in biology. *In* "The Uniqueness of the Individual," Vol. 45, pp. 44–70. Basic Books, New York.

Minois, N., Khazaeli, A. A., and Curtsinger, J. W. (2001). Locomotor activity as a function of age and life span in *Drosophila melanogaster* overexpressing hsp70. *Exp. Gerontol.* **36,** 1137–1153.

Miquel, J., Lundgren, P. R., Bensch, K. G., and Atlan, H. (1976). Effects of temperature on the life span, vitality and fine structure of *Drosophila melanogaster*. *Mech. Ageing Dev.* **5,** 347–370.

Missirlis, F., Phillips, J. P., and Jackle, H. (2001). Cooperative action of antioxidant defense systems in *Drosophila*. *Curr. Biol.* **11,** 1272–1277.

Mockett, R. J., Orr, W. C., Rahmandar, J. J., Benes, J. J., Radyuk, S. N., Klichko, V. I., and Sohal, R. S. (1999a). Overexpression of Mn-containing superoxide dismutase in transgenic *Drosophila melanogaster*. *Arch. Biochem. Biophys.* **371,** 260–269.

Mockett, R. J., Sohal, R. S., and Orr, W. C. (1999b). Overexpression of glutathione reductase extends survival in transgenic *Drosophila melanogaster* under hyperoxia but not normoxia. *FASEB J.* **13,** 1733–1742.

Mockett, R. J., Orr, W. C., Rahmandar, J. J., Sohal, B. H., and Sohal, R. S. (2001). Antioxidant status and stress resistance in long- and short-lived lines of *Drosophila melanogaster*. *Exp. Gerontol.* **36,** 441–463.

Nghiem, D., Gibbs, A. G., Rose, M. R., and Bradley, T. J. (2000). Postponed aging and desiccation resistance in *Drosophila melanogaster*. *Exp. Gerontol.* **35,** 957–969.

O'Kane, K., and Gehring, W. (1987). Detection in situ of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **84,** 9123–9127.

Orr, W. C., and Sohal, R. S. (1992). The effects of catalase gene overexpression on life span and resistance to oxidative stress in transgenic *Drosophila melanogaster*. *Arch. Biochem. Biophys.* **297,** 35–41.

Orr, W. C., and Sohal, R. S. (1993). Effects of Cu-Zn superoxide dismutase overexpression of life span and resistance to oxidative stress in transgenic *Drosophila melanogaster*. *Arch. Biochem. Biophys.* **301,** 34–40.

Orr, W. C., and Sohal, R. S. (1994). Extension of life-span by overexpression of superoxide dismutase and catalase in *Drosophila melanogaster*. *Science* **263,** 1128–1130.

Orr, W. C., Arnold, L. A., and Sohal, R. S. (1992). Relationship between catalase activity, life span and some parameters associated with antioxidant defenses in *Drosophila melanogaster*. *Mech. Ageing Dev.* **63,** 287–296.

Osterwalder, T., Yoon, K. S., White, B. H., and Keshishian, H. (2001). A conditional tissue-specific transgene expression system using inducible GAL4. *Proc. Natl. Acad. Sci. USA* **98,** 12596–12601.

Parkes, T. L., Elia, A. J., Dickinson, D., Williker, A. J., Phillips, J. P., and Boulianne, J. L. (1998). Extension of *Drosophila* lifespan by overexpression of human SOD1 in motor neurons. [see comments]. *Nat. Genet.* **19,** 171–174.

Partridge, L., and Farquhar, M. (1981). Sexual activity reduces lifespan of male fruit flies. *Nature* **294,** 580–582.

Partridge, L., and Harvey, P. H. (1993). Gerontology: Methuselah among nematodes. *Nature* **366,** 404–405.

Pearl, R. (1922). "The Biology of Death." J. B. Lippincott, Philadelphia.

Pearl, R. (1928). "The Rate of Living: Being an Account of Some Experimental Studies on the Biology of Life Duration." University of London Press, London.

Pearl, R., and Parker, S. L. (1921). Experimental studies on the duration of life. I. Introductory discussion of the duration of life in *Drosophila*. *Am. Nat.* **60,** 481–509.

Pearl, R., and Parker, S. L. (1922). Experimental studies on the duration of life. II. Hereditary differences in duration of life in line-breed strains of *Drosophila*. *Am. Nat.* **56,** 174–187.

Pearl, R., Parker, S. L., and Gonzales, B. M. (1923). Experimental studies on the duration of life. VII. The Mendelian inheritance of duration of life in crosses of wild type and quintuple stocks of *Drosophila melanogaster*. *Am. Nat.* **57,** 153–192.

Phillips, J. P., and Hilliker, A. J. (1990). Genetic analysis of oxygen defense mechanisms in *Drosophila melanogaster*. *Adv. Genet.* **28,** 43–71.

Phillips, J. P., Campbell, S. D., Michaud, D., Charbonneau, M., and Hilliker, A. J. (1989). Null mutation of copper/zinc superoxide dismutase in *Drosophila* confers hypersensitivity to paraquat and reduced longevity. *Proc. Natl. Acad. Sci. USA* **86,** 2761–2765.

Phillips, J. P., Parkes, T. L., and Hilliker, A. J. (2000). Targeted neuronal gene expression and longevity in *Drosophila*. *Exp. Gerontol.* **35,** 1157–1164.

Rogina, B., Benzer, S., and Helfand, S. L. (1997). *Drosophila drop-dead* mutations accelerate the time course of age-related markers. *Proc. Natl. Acad. Sci. USA* **94,** 6303–6306.

Rogina, B., Reenan, R. A., Nilsen, S. P., and Helfand, S. L. (2000). Extended life-span conferred by contransporter gene mutations in *Drosophila*. *Science* **290,** 2137–2140.

Roman, G., Endo, K., Zong, L., and Davis, R. L. (2001). P[Switch], a system for spatial and temporal control of gene expression in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **98,** 12602–12607.

Rorth, P. (1996). A modular misexpression screen in *Drosophila* detecting tissue-specific phenotypes. *Proc. Natl. Acad. Sci. USA* **93,** 12418–12422.

Rorth, P., Szabo, K., Bailey, A., Laverty, T., Rehm, J., Rubin, G. M., Weigmann, K., Milan, M., Benes, V., Ansorge, W., and Cohen, S. M. (1998). Systematic gain-of-function genetics in *Drosophila*. *Development* **125,** 1049–1057.

Rose, M. (1984). Laboratory evolution of postponed senescence in *Drosophila melanogaster*. *Evolution* **38,** 1004–1009.

Rose, M., and Charlesworth, B. (1980). A test of evolutionary theories of senescence. *Nature* **287,** 141–142.

Rose, M. R., and Charlesworth, B. (1981). Genetics of life history in *Drosophila melanogaster*. II. Exploratory selection experiments. *Genetics* **97,** 187–196.

Rose, M. R., Vu, L. N., Park, S. U., and Graves, J. L., Jr. (1992). Selection on stress resistance increases longevity in *Drosophila melanogaster*. *Exp. Gerontol.* **27,** 241–250.

Smith, J. M. (1958). The effects of temperature and of egg laying on the longevity of *Drosophila subobscura*. *J. Exp. Biol.* **35,** 832–842.

Smith, J. M. (1962). The causes of ageing. *Proc. R. Soc. Lond. B Biol. Sci.* **157,** 115–127.

Sohal, R. S. (1986). The rate of living theory: A contemporary interpretation. *In* "Insect Aging: Strategies and Mechanisms" (K.-G. Collatz, and R. S. Sohal, eds.), Springer-Verlag, Berlin.

Sohal, R. S., and Buchan, P. B. (1981). Relationship between physical activity and life span in the adult housefly, *Musca domestica*. *Exp. Gerontol.* **16,** 157–162.

Sohal, R. S., and Weindruch, R. (1996). Oxidative stress, caloric restriction, and aging. *Science.* **273,** 59–63.

Sohal, R. S., Agarwal, A., Agarwal, S., and Orr, W. C. (1995). Simultaneous overexpression of copper- and zinc-containing superoxide dismutase and catalase retards age-related oxidative damage and increases metabolic potential in *Drosophila melanogaster*. *J. Biol. Chem.* **270,** 15671–15674.

Sohal, R. S., Mockett, R. J., and Orr, W. C. (2000). Current issues concerning the role of oxidative stress in aging: A perspective. *Results Probl. Cell Differ.* **29,** 45–66.

Sun, J., and Tower, J. (1999). FLP recombinase-mediated induction of Cu/Zn-superoxide dismutase transgene expression can extend the life span of adult *Drosophila melanogaster* flies. *Mol. Cell Biol.* **19,** 216–228.

Tatar, M., and Yin, C. (2001). Slow aging during insect reproductive diapause: Why butterflies, grasshoppers and flies are like worms. *Exp. Gerontol.* **36,** 723–738.

Tatar, M., Khazaeli, A. A., and Curtsinger, J. W. (1997). Chaperoning extended life. *Nature* **390,** 30.

Tatar, M., Kopelman, A., Epstein, D., Tu, M. P., Yin, C. M., and Garofalo, R. S. (2001). A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science* **292,** 107–110.

Trout, W. E., and Kaplan, W. D. (1970). A relation between longevity, metabolic rate, and activity in *Shaker* mutants of *Drosophila melanogaster*. *Exp. Gerontol.* **5,** 83–92.

Trout, W. E., and Kaplan, W. D. (1981). Mosaic mapping of foci associated with longevity in the neurological mutants *Hk1* and *Sh5* of *Drosophila melanogaster*. *Exp. Gerontol.* **16,** 461–474.

Vaupel, J. W. (1997). Trajectories of mortality at advanced ages. *In* "Between Zeus and the Salmon" (K. Wachter, and C. E. Finch, eds.), pp. 17–37. National Academy Press, Washington, D. C.

Vaupel, J. W., Carey, J. R., Christensen, K., Johnson, T. E., Yashin, A. I., Holm, N. V., Iachine, I. A., Kannisto, V., Khazaeli, A. A., Liedo, P., Longo, V. D., Zeng, Y., Manton, K. G., and Curtsinger, J. W. (1998). Biodemographic trajectories of longevity. *Science* **280,** 855–860.

Warrick, J. M., Chan, H. Y., Gray-Board, G. L., Chai, Y., Paulson, H. L., and Bonini, N. M. (1999). Suppression of polyglutamine-mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70. *Nat. Genet.* **23,** 425–428.

Wheeler, J. C., Bieschke, E. T., and Tower, J. (1995). Muscle-specific expression of *Drosophila* hsp70 in response to aging and oxidative stress. *Proc. Natl. Acad. Sci. USA* **92,** 10408–10412.

Wilmoth, J. R. (1997). In search of limits. *In* "Between Zeus and the Salmon" (K. Wachter, and Finch, eds.), pp. 38–64. National Academy Press, Washington, D.C.

This Page Intentionally Left Blank

3

Genetics of Wheat Gluten Proteins

Peter R. Shewry and Nigel G. Halford
Long Ashton Research Station
Department of Agricultural Sciences
University of Bristol
Long Ashton, Bristol BS41 9AF, United Kingdom

Domenico Lafiandra
Department of Agrobiology and Agrochemistry
University of Tuscia
01100 Viterbo, Italy

Advances in Genetics, Vol. 49

0065-2660/03 $35.00

I. INTRODUCTION

These are good grounds for considering wheat to be the most important crop in the world. First, the total annual production over the period 1996–1998 averaged almost 595 million tonnes, narrowly exceeding that of the two other major cereals, rice (571 million tonnes) and maize (592 million tonnes) (Food and Agriculture Organization, 1999). Second, it has the widest geographical and climatic range of all major crops, being grown between Scandinavia and Argentina, including higher elevations in the tropics and subtropics. This wide cultivation is made possible by an immense amount of genetic variation, with some 2500 cultivars having been produced so far (Feldman *et al*., 1995). Third, unlike maize, a high proportion of the total wheat produced worldwide, about 65%, is consumed by humans, with 21% used for livestock, 8% as seed, and 6% for other uses including industrial raw material (Orth and Shellenberger, 1988).

Whereas maize and rice are frequently consumed by humans without any processing except cooking, wheat is almost solely consumed after processing, to form bread, other baked goods (cakes, cookies, crackers etc.), pasta and noodles, and a wide range of other products. Within breads alone the range of diversity is enormous, including yeast-leavened pan and hearth-baked

breads in Europe and North America, steamed breads in China, and flat breads (which may be single or double layered and either unleavened, yeasted, or sour dough) in the Middle East and the Indian subcontinent. Similarly, a range of types of Asian noodles and pasta is produced from bread and pasta wheats, respectively.

The ability to process wheat into such a wide range of products is determined largely by one group of grain components, the gluten proteins, which confer unique viscoelastic properties to wheat doughs. Consequently, the structures, properties, and genetics of these proteins have been widely studied in order to determine the biochemical and molecular basis for their functional (i.e., processing) properties and to make improvements by plant breeding, optimizing agronomy and processing conditions and, more recently, genetic engineering. This chapter focuses on genetic studies, which have not been comprehensively reviewed since Konzak (1977). However, before discussing these it is necessary to first briefly describe the genetics of the wheat plant and the properties and nomenclature of the wheat gluten proteins.

A. Genetics of the wheat plant

Cultivated wheat (*Triticum*) is not a single species but consists of diploid, tetraploid, and hexaploid forms, the polyploids having arisen by amphiploidy between the diploids and related wild species of *Triticum* or *Aegilops*. The basic chromosome number (n) for both genera is 7, with $2n = 14$ for diploids, 28 for tetraploids, and 42 for hexaploids. The precise nomenclature and relationships of the various species of these genera are still not agreed by all authorities and the following description is based largely on Feldman *et al.* (1995), which is also the source of Table 3.1.

Diploid wheat (einkorn) comprises three species: *T. boeoticum*, *T. urartu*, and *T. monococcum*, the first two of which are wild and the last the cultivated form. All share the A genome with all polyploid wheats. Tetraploid *T. turgidum* contains the A and B genomes, the origin of the latter being uncertain. The cultivated form var. *durum* (often called *T. durum*) is widely grown in regions with a Mediterranean climate as durum or pasta wheat. A second tetraploid species, *T. timopheevii*, contains the G genome, which is also of uncertain origin. It is still grown to a limited extent in Armenia and Transcaucasia. A single hexaploid species, *T. aestivum*, contains the AB genomes of *T. turgidum* combined with the D genome from the related wild species *Aegilops squarrosa* (also called *Triticum tauschii*). This "bread wheat" is the dominant species in world agriculture. Most bread wheat is the free threshing var. *aestivum*, but other free threshing types (vars. *compactum* and *sphaerococcum*) are grown in restricted areas. Similarly, the hulled var. *spelta*

Table 3.1. Classification of Cultivated Wheats and Closely Related Wild Species[a]

Species	Genome	Wild (hulled)	Cultivated	
			Hulled	Free-threshing
Diploid ($2n = 14$)				
Aegilops speltoides	S (G)	All		
Aegilops bicornis	S^b	All		
Aegilops longissima	S^l	All		
Aegilops searsii	S^s	All		
Aegilops squarrosa	D	All		
Triticum urartu	A	All		
Triticum monococcum	A	var. *boeoticum* (wild einkorn)	var. *monococcum* (cultivated einkorn)	var. *sinskajae* (cultivated einkorn)
Tetraploid ($2n = 28$)				
Triticum timopheevii	AG	var. *araraticum*	var. *timopheevi*	var. *militinae*
Triticum turgidum	AB	var. *dicoccoides* (wild emmer)	var. *dicoccum* (cultivated emmer)	var. *durum* var. *turgidum* var. *polonicum* var. *carthlicum* var. *turanicum*
Hexaploid ($2n = 42$)				
Triticum aestivum	ABD		var. *spelta* var. *macha* var. *vavilovii*	var. *aestivum* var. *compactum* var. *sphaerococcum*

[a]Taken from Feldman *et al.* (1995), with permission.

(spelt) has received some attention as a "healthy" alternative to other wheats, although there is no clear scientific basis for this claim.

Hexaploid bread wheat has a comparatively recent origin, having originated and entered cultivation only after the simultaneous domestication of diploid and tetraploid forms, about 10,000 years ago. Although the A, B, and D genomes are related, the tetraploid and hexaploid wheats behave as diploids during meiosis because of the action of a specific gene (*Ph1*) on chromosome 5B (Riley and Chapman, 1958). The related chromosomes of the three genomes are usually referred to as "homoeologous."

The polyploid nature of bread and durum wheats has two important consequences for the composition and analysis of the gluten proteins. First, all three genomes contain genes that encode related proteins and are often located in similar positions on the chromosomes, resulting in a complex protein composition and complex patterns of protein segregation during genetic analysis. However, the close relationship between the genomes is also advantageous in that it provides "buffering" to the genome, allowing the

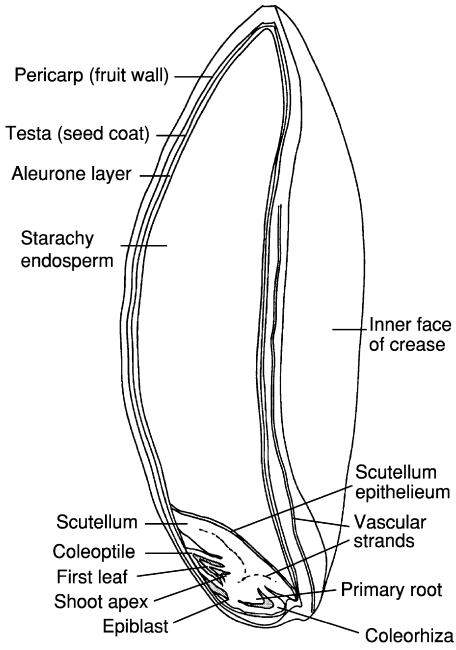


Figure 3.1. Median longitudinal section of a wheat grain. Taken from Gooding and Davies (1997) with permission (redrawn from Barnes, 1989).

development of genetic stocks in which whole chromosomes, chromosome arms, or chromosome segments can be lost or replaced by duplicated material from the homoeologous chromosomes of other genomes or by related genetic material from alien species (e.g., rye or *Aegilops* spp.). The development of such stocks, which was pioneered by Sears (see Sears, 1972), has greatly facilitated the mapping of a range of genes, including those encoding gluten proteins, in wheat.

B. The wheat grain

The mature wheat grain consists of two genetically different organs, the endosperm and embryo, that arise from separate fertilization events. The pollen tube delivers two sperm cells into the embryo sac, one of which fuses with the egg cell to give the diploid zygote. This gives rise in the mature grain to the embryonic axis and a single cotyledon, called the scutellum (Fig. 3.1). Although the scutellum contains some storage compounds (notably oil and globulin storage proteins) it does not contain gluten protein. The second sperm cell fuses with two polar nuclei to form a triploid endosperm cell that expands, cellularizes, divides, and differentiates to give two tissues: the starchy

endosperm, which comprises about 80% of the dry matter and 72% of the protein in the mature grain (Gooding and Davies, 1997), and the aleurone, which consists of a single outer layer of thick-walled cells. Although they differ in origin and are triploid rather than diploid, the aleurone cells appear to store compounds similar to those in the scutellum. In contrast, the starchy endosperm cells contain high levels of starch (60–70% dry weight) and prolamin storage proteins.

II. WHEAT GLUTEN PROTEINS

A. Classification and nomenclature

The analysis of wheat gluten proteins has a long history, extending back to 1745 when Giacomo Beccari, Professor of Chemistry at the University of Bologna, reported the preparation of a water-insoluble fraction which he called "glutinis" from wheat flour (Beccari, 1745). However, the study of plant proteins was put on a sound scientific basis by T. B. Osborne working between 1886 and 1928 (see Osborne, 1924). He classified proteins into groups on the basis of their solubility: albumins (water), globulins (dilute saline), prolamins (alcohol–water mixtures), and glutelins (dilute acid or alkali). Furthermore, he showed that prolamins were restricted to the grain of cereals and other grasses.

Gluten is still routinely produced by washing dough made with white flour (i.e., derived from starchy endosperm cells) with water, the resulting fraction comprising about 70% protein with the remainder being starch, lipids, and other components. The gluten produced in this fashion forms a cohesive viscoelastic mass. However, it must be borne in mind that gluten forms a continuous dispersed network in the dough.

Gluten is classically divided into alcohol-soluble (gliadin) and insoluble (glutenin) fractions, which are further separated by electrophoresis. The gliadins consist of monomeric proteins, which are separated into α, β, γ, and ω groups by polyacrylamide electrophoresis at low pH (Fig. 3.2). The glutenins consist of polymeric proteins stabilized by interchain disulfide bonds. These bonds need to be reduced before the component subunits can be separated into two groups, high molecular weight (HMW) and low molecular weight (LMW) subunits, with the latter being further divided into B-, C-, and D-type subunits according to size, isoelectric points, and composition (Jackson *et al*., 1983) (Fig. 3.2).

It was long considered that the gliadin and glutenin fractions comprised different types of proteins that corresponded to the prolamins and glutelins, respectively, as defined by Osborne. However, a range of biochemical and molecular studies carried out over the past two decades have demonstrated

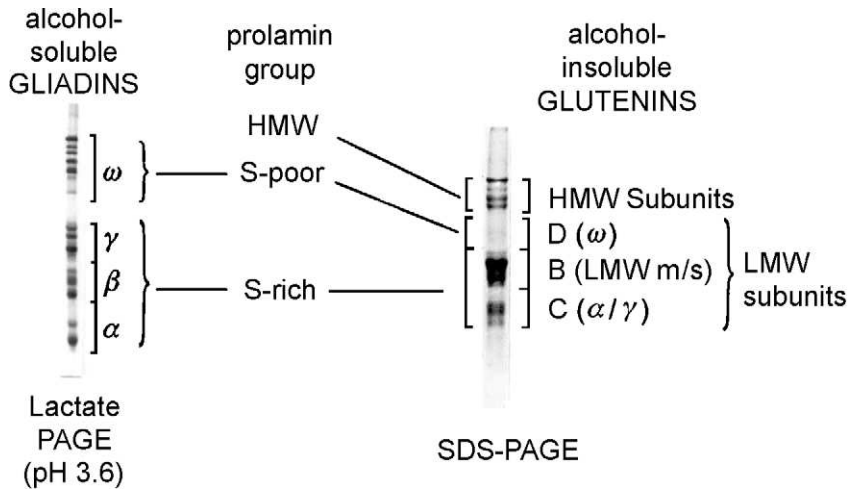


Figure 3.2. The classification and nomenclature of wheat gluten proteins separated by SDS–PAGE and electrophoresis at low pH. The D-type LMW subunits are only minor components and are not clearly resolved in the separation shown. Taken from Shewry *et al.* (1999) with permission.

that this is not the case and this conclusion is supported by the genetic studies discussed below. We now know that all gluten proteins are structurally and evolutionarily related and can be defined as prolamins in that they are soluble in alcohol–water mixtures either as protein monomers (gliadins) or as reduced subunits (glutenins). Furthermore, it is possible to define three groups of prolamins, which contain gliadin and/or glutenin proteins. These are the high molecular weight prolamins, which comprise only the HMW subunits of glutenin, the sulfur-poor (S-poor) prolamins, which comprise ω-gliadins and D-type LMW subunits of glutenin, and the S-rich prolamins, which comprise α-type (i.e., α and β) gliadins, γ-type gliadins, and the B- and C-type LMW subunits of glutenin. Furthermore, the C and D groups of LMW subunits are highly similar in sequence to individual α-type/γ-type and ω-gliadins, respectively, and are considered to be derived from these components by mutations resulting in the presence of additional cysteine residues, which are able to form interchain disulfide bonds (see below). In contrast, the B-type LMW subunits form a discrete group and no closely related gliadin components have been identified. The relationships between groups of gluten proteins defined by classic and molecular approaches are summarized in Fig. 3.2.

It is clear from the discussion above that the main distinction between the gliadin and glutenin proteins is that the former are monomeric and the latter polymeric. Nevertheless, this classification has been retained by cereal

chemists for two reasons. First, the two groups of proteins are relatively easy to prepare. Second, they have functional significance, with the glutenins being primarily responsible for the elasticity (strength) of the gluten and the gliadins for viscosity.

B. Wheat gluten: a highly complex mixture

Storage protein fractions are generally characterized by a high level of polymorphism (Shewry, 1995). This is amplified in polyploid wheats by the presence of storage protein genes on the different genomes. Consequently, it is necessary to use two-dimensional electrophoresis to achieve good resolution. The most widely used system combines sodium dodecyl sulfate–polyacrylamide gel electrophoresis, (SDS–PAGE) as the second dimension with isoelectric focusing (IEF) as the first, with a nonequilibrium IEF system (nonequilibrium pH gradient electrophoresis, NEPHGE) sometimes being used to give improved resolution of high p*I* components (as shown in Fig. 3.3). In this system the gliadins and LMW glutenin subunits are not clearly separated. A second system combining electrophoresis at low and high pH (pH 3.2 followed by pH 9.2) provides excellent resolution of gliadins as shown in Fig. 3.4 (Lafiandra and Kasarda, 1985).

The more recent application of other analytical tools, such as high-performance liquid chromatography (HPLC), capillary electrophoresis (CE), and mass spectroscopy, has contributed further to our knowledge of the composition of wheat gluten and of the genetics and biochemical properties of the individual components (Kruger and Bietz, 1994; Dworschak *et al*., 1998; Bean *et al*., 1998; Bean and Lookhart, 2000).

C. Prolamin structure: A brief overview

We now have a detailed understanding of the structures, sequences, and properties of many gluten proteins, based on analyses of purified proteins and of cloned cDNAs and genes. These studies show that the individual proteins vary widely in their molecular masses (from about 30 to 90,000 Da) and structures. Nevertheless, all share two properties. First, their sequences can be divided into at least two parts (usually called domains), which have different amino acid compositions, adopt different structures, and probably have different origins. Second, in all cases one of these domains consists of repeated sequences based on the reiteration of one or more short (three- to nine-residue) peptide motifs. In addition, there is clear similarity between the repetitive sequences in the S-rich and S-poor prolamins and between the nonrepetitive domains of the S-rich and HMW groups (Kreis *et al*., 1985; Shewry and Tatham, 1999), implying that they have common evolutionary origins.

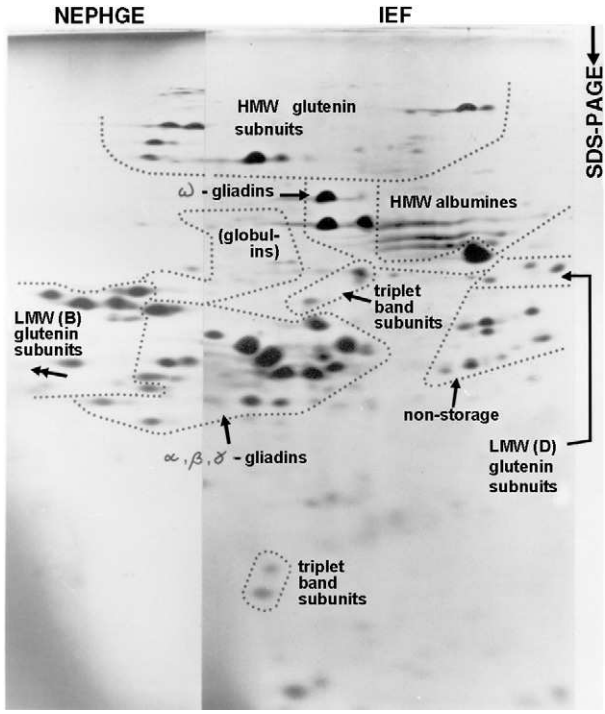


Figure 3.3. Fractionation of endosperm proteins by nonequilibrium pH gradient electrophoresis × sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) and isoelectric focusing × SDS–PAGE. The two first dimensions have a partially overlapping pH gradient so that a few proteins occur on both systems. The overlapping areas of the gel photographs have been removed to give a continuous pH gradient. Taken from Payne *et al.* (1985) with permission.

The sequences of the HMW subunits can be divided into three domains, with short N-terminal and C-terminal domains (of 81–104 and 42 residues, respectively) flanking a repetitive domain which varies in length from about 480 to 700 residues in the subunits commonly found in commercially grown wheats (see Shewry *et al.*, 1992) (Fig. 3.5). Furthermore, two types of subunit can be recognized on the basis of their repeat motifs and other features. These correspond to the x-type and y-type subunits defined initially on the basis of their mobility on SDS–PAGE. Thus, the x-type subunits contain tripeptide, hexapeptide, and nonapeptide motifs while the y-type subunits contain only hexa- and nonapeptides. These are discussed in detail in Section IV. The distribution of cysteine residues is important as this may determine the structures and properties of the glutenin polymers. These are mainly located in the N- and C-terminal domains, with three and five cysteines in the N-terminal

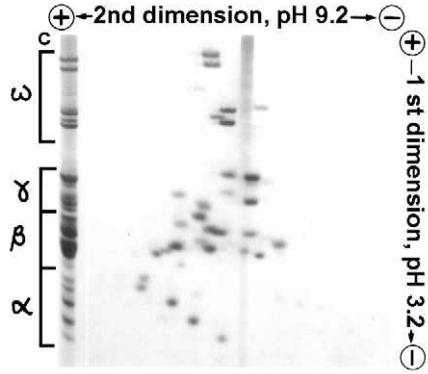


Figure 3.4. Two-dimensional electrophoresis of gliadins from wheat cv. Chinese Spring. The protein sample was applied at the center of the top of the gel and initially separated by electrophoresis in lactate buffer at pH 3.1. Electrophoresis was then carried out in Tris-glycine buffer, pH 9.2, containing 1.5 M dimethylformamide with the electric field at 90° to that used for the first dimension. Taken from Shewry (2003) with permission.

domains of x-type and y-type subunits, respectively, and single cysteines in the C-terminal domains of all subunits.

Despite their different properties, all of the S-rich prolamins (gliadins and LMW subunits of glutenin) have essentially similar structures, as summarized in Fig. 3.5. A short unique N-terminal sequence is followed by a repetitive domain based on one or more motifs that are rich in proline and glutamine residues. This is followed by a nonrepetitive C-terminal domain that, in the gliadins, contains all the cysteine residues. Six conserved cysteines are present in the α-type and γ-type gliadins (numbered 2 and 4–8 in Fig. 3.5), which form three interchain disulfide bonds, while the γ-gliadins contain an additional two cysteines (numbered 1 and 7 in Fig. 3.5), which form a fourth disulfide bond. There is more variation in the structures of the S-rich LMW subunits of glutenin, with some C-type subunits corresponding to α-type or γ-type gliadins that contain additional cysteine residues, which allow the formation of interchain disulfide bonds (see Shewry and Tatham, 1997; Anderson *et al*., 2001b). Figure 3.5 shows a "typical" B-type LMW subunit, which contains six of the conserved cysteines that are also present in the α/γ-type gliadins (numbered 1–5 and 7 in Fig. 3.5) and additional cysteines in the N-terminal sequence (*a* in Fig. 3.5) and C-terminal domain (*g* in Fig. 3.5), which may form interchain bonds (see Shewry and Tatham, 1997).

The ω-gliadins have been studied in less detail than the other gluten proteins, with only two full amino acid sequences being available (Hsia and

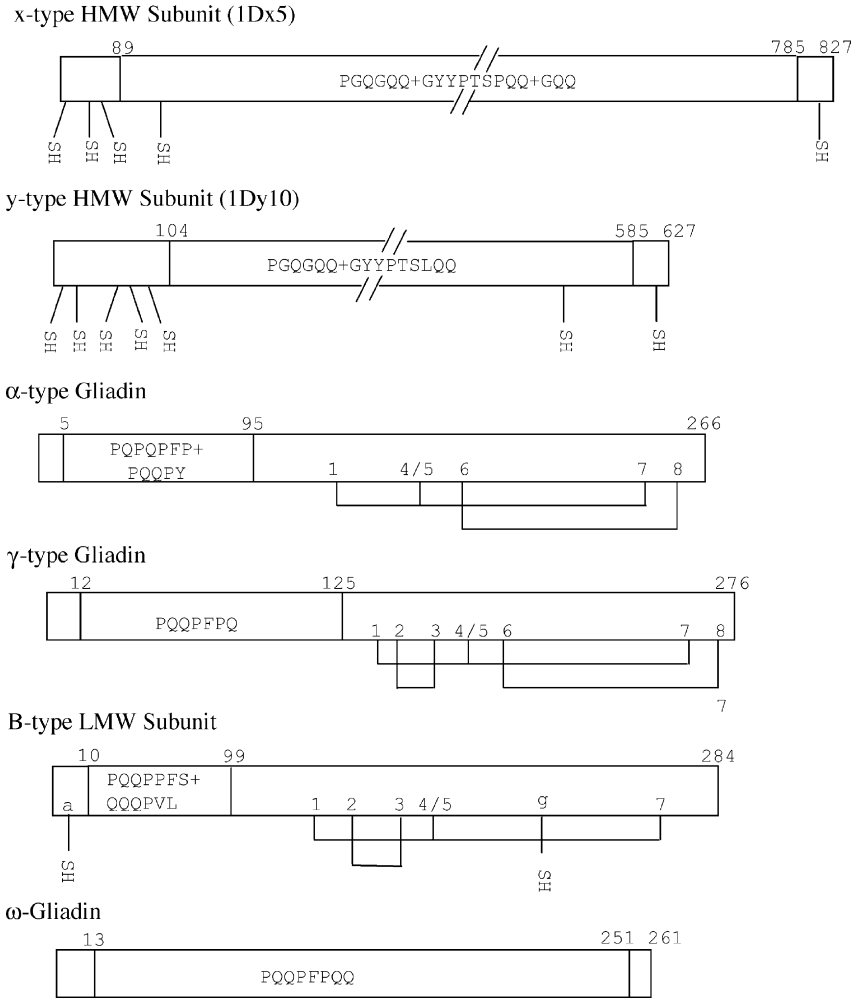


Figure 3.5. Summary of the sequences of "typical" wheat gluten proteins. The consensus repeat motifs in the repetitive domains are shown, using standard single-letter abbreviations for amino acids (G, glycine; F, phenylalanine; L, leucine; P, proline; Q, glutamine; S, serine; T, threonine, V, valine; Y, tyrosine). Taken from Shewry (2003) with permission.

Anderson, 2001). The example shown in Fig. 3.5 is encoded by the D genome of bread wheat and consists of 201 amino acids, with short N-terminal (13-residue) and C-terminal (10-residue) sequences flanking extensive tandem repeats based on a consensus octapeptide motif (PQQPFPQQ) related to the motifs present in the S-rich prolamins (Fig. 3.5).

Comparisons of the N-terminal and partial amino acid sequences of other ω-gliadins show that they have similar highly repetitive structures with short unique N-terminal domains (Kasarda *et al*., 1983; Tatham and Shewry, 1995), but that those encoded by the B genome have a different consensus repeat motif [(Q)(Q)QQXP] (DuPont *et al*., 2000).

Although typical ω-gliadins do not contain cysteine residues and hence do not form disulfide bonds, variant forms do occur in which the presence of single cysteine residues allows their incorporation into glutenin polymers (Masci *et al*., 1993, 1999). These are called the D-type LMW subunits of glutenin.

Anderson *et al*. (2001a) have reported the characterization of four novel types of low molecular mass proteins that are apparently related to the S-rich gluten proteins but lack extensive repeated sequences. However, they are only minor components and are not considered further.

III. CHROMOSOMAL ASSIGNMENT AND MAPPING OF GLUTEN PROTEIN GENES

A. Gliadins

The first reports on the chromosomal location of gluten protein genes date back to the late 1960s when Boyd and Lee (1967) and Shepherd (1968), using starch gel electrophoresis, analyzed the compensating nullisomic–tetrasomic and ditelosomic series developed by Sears (1954, 1966) in the bread wheat cv. Chinese Spring. These authors reported that gliadin proteins are controlled by genes present on the short arm of the homoeologous group 1 and 6 chromosomes. The poor resolution of the one-dimensional separations and the overlapping of many gliadin components meant that it was possible to assign only a few of them to specific chromosomes. Subsequently, Wrigley and Shepherd (1973) introduced a two-dimensional technique combining isoelectric focusing in the first dimension with starch gel electrophoresis at acidic pH in the second dimension. This resolved Chinese Spring gliadins into 46 components, allowing almost all of them to be assigned to the group 1 and 6 chromosomes. The development of more refined two-dimensional electrophoretic techniques, in which starch was replaced by polyacrylamide gels, resulted in increases in the resolution and number of components detected, but essentially confirmed the chromosomal assignments of gliadin coding-genes in Chinese Spring as well as in different bread wheat cultivars (Brown *et al*., 1981; Payne *et al*., 1982; Lafiandra *et al*., 1984). Similar studies of durum wheat (Joppa *et al*., 1983; Lafiandra *et al*., 1987) and wild relatives of wheat have confirmed these results (Lafiandra *et al*., 1993a) with the exception of *Dasypyrum villosum*, an

allogamous annual grass of the subtribe Triticinae (2*n* = 14, VV), where a third locus was detected on the long arm of chromosome 4V (Blanco *et al.*, 1991; Shewry *et al.*, 1991).

Extensive studies of the inheritance of gliadins in the progeny of specific crosses have indicated that the major gliadin genes occur in tightly linked clusters, termed blocks, with intrablock recombination being rare (Sozinov and Poperelya, 1980). It is now generally accepted that the ω- and γ-gliadins are controlled by clusters of tightly linked genes present at the *Gli-1* loci (*Gli-A1*, *Gli-B1*, and *Gli-D1*) on the short arms of the homoeologous group 1 chromosomes, whereas the α- and β-gliadins are controlled by the *Gli-2* loci (*Gli-A2*, *Gli-B2*, and *Gli-D2*) present on the short arms of the group 6 chromosomes (Payne, 1987). This spatial separation of gliadin genes on the group 1 and 6 chromosomes has been attributed to an ancient interchromosomal translocation, with the *Gli-2* locus originating from the translocation of a γ-type gene from chromosome 1 to chromosome 6, followed by divergence of the coding sequence to give rise to the α-type sequence (Shewry *et al.*, 1984). The *Gli-1* loci have been shown to be present on the distal parts of the group 1 chromosomes, showing independent or loose linkage with their respective centromeres (Shepherd, 1988). The *Gli-2* loci have been studied in less detail, but telocentric mapping showed 35% recombination between the *Gli-A2* locus and the centromere (Payne, 1987).

The individual *Gli-1* and *Gli-2* loci exhibit extensive polymorphism, as detected by electrophoretic techniques, with allelic blocks differing in the numbers, proportions, and mobilities of different components. This results in a great diversity of gliadin patterns, providing the basis for distinguishing different wheat cultivars (Metakovsky, 1991).

B. High molecular weight glutenin subunits

The use of reducing agents together with SDS–PAGE has allowed the genetic control of glutenin subunits to be determined. Using the Chinese Spring aneuploids described above, HMW subunit genes were located on the long arms of the homoeologous group 1 chromosome by Orth and Bushuk (1974) and Bietz *et al.* (1975). Subsequently, use of the discontinuous SDS–PAGE system of Laemmli (1970) to study different sets of intervarietal chromosome substitution lines allowed the results obtained in Chinese Spring to be extended to other bread wheat cultivars (Lawrence and Shepherd, 1980; Payne *et al.*, 1980; Galili and Feldman, 1985). These results provided firm evidence that in bread wheat the HMW subunits are encoded by genes at complex loci, designated *Glu-1*, present on the long arm of the homoeologous group 1 chromosomes (*Glu-A1*, *Glu-B1*, *Glu-D1*). Each locus contains two tightly linked genes (Harberd *et al.*, 1986) encoding subunits designated as x and y type based on their molecular

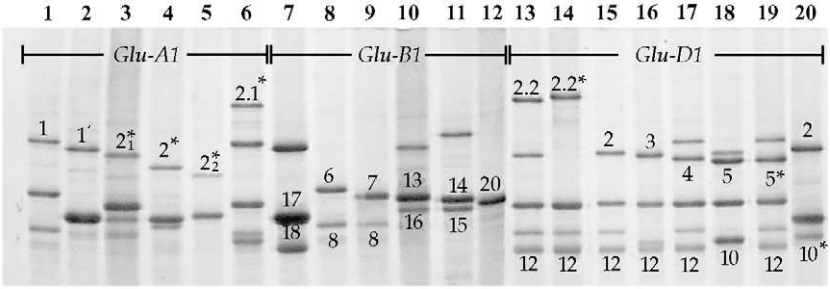


Figure 3.6. SDS–PAGE of allelic HMW glutenin subunits present at the three *Glu-1* loci. Taken from Shewry *et al.* (2003) with permission.

weights and biochemical characteristics (Payne *et al.*, 1981a). However, because of the silencing of some genes, only three to five HMW subunit genes are expressed in different bread wheat cultivars. In particular, two subunits are always expressed by the *Glu-D1* locus, two or one by the *Glu-B1*, locus and one or none (the null allele) by the *Glu-A1* locus. When only one subunit is expressed by the *Glu-B1* or *Glu-A1* loci, this is always x type. A similar situation occurs at the *Glu-A1* and *Glu-B1* loci in durum wheat. However, the y-type gene present at the *Glu-A1* locus may be expressed in cultivated and wild diploid wheats (*T. monococcum* subsp. *monococcum*, subsp. *boeoticum*, and *T. urartu*), in the wild tetraploid wheat *T. turgidum* subsp. *dicoccoides* (Waines and Payne, 1987; Levy *et al.*, 1988), and also in cultivated and wild forms of tetraploid wheats with the genomic formula AAGG (*T. timopheevii* subsp. *timopheevii* and subsp. *araraticum*) (Margiotta *et al.*, 1998).

Allelic variation has also been reported in the subunits encoded by each *Glu-1* locus in bread wheat cultivars (Lawrence and Shepherd, 1980; Payne and Lawrence, 1983) and a numbering system was introduced to identify different allelic subunits (Fig. 3.6).

Analyses of large collections of bread wheat landraces have identified rare alleles at the different *Glu-1* loci. Payne *et al.* (1983) and Lafiandra *et al.* (1997) have described the identification of subunits with high molecular weights encoded by the *Glu-D1* and *Glu-A1* loci (see lanes 6, 13, and 14 of Fig. 3.6). Null forms lacking x- and/or y-type subunits have also been detected, with Bietz *et al.* (1975) reporting the presence of seeds lacking both x- and y-type subunits encoded by the *Glu-D1* locus in the bread wheat landrace Nap Hal. Additional null types have since been reported by Payne *et al.* (1984a) and by Lafiandra *et al.* (1988). Combination of the different null *Glu-1* alleles has resulted in the production of genotypes with unusual HMW subunit compositions (Fig. 3.7). For example, Lawrence *et al.* (1988) have developed a set of bread wheat lines in which the number of subunits increased

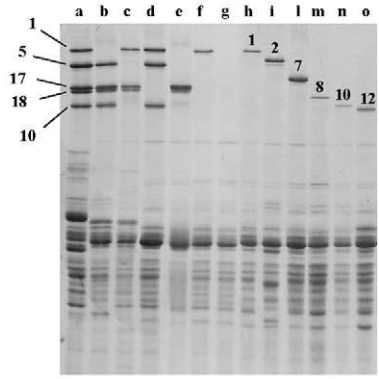


Figure 3.7. SDS–PAGE of HMW glutenin subunits present in bread wheat lines null at the *Glu-1* loci compared with those present in the bread wheat cultivar Yecora Rojo (lane a): *Glu-A1* null (lane b), *Glu-D1* null (lane c), *Glu-B1* null (lane d), *Glu-A1/Glu-D1* null (lane e), *Glu-B1/Glu-D1* null (lanes f and h) and *Glu-A1/Glu-B1/ Glu-D* null (lane g). Lanes i and l–o: Lines with a single Dx (lane i), Bx (lane 1), By (lane m), or Dy subunit (lanes n and o).

progressively from zero to five while Lafiandra *et al.* (2000b) have produced wheat lines with single x- or y-type subunits. This material has proved useful in determining the relative effects of individual HMW subunits on flour breadmaking properties and provided the possibility of developing wheats suitable for different end uses (Lawrence *et al.*, 1988; Payne and Seekings, 1996; Lafiandra *et al.*, 2000a).

Although the genes encoding x- and y-type subunits are tightly linked, rare cases of recombination have been reported to occur (Payne, 1987; Singh and Shepherd, 1988a). The *Glu-1* loci have been mapped relative to the centromeres by telocentric and translocation mapping, with the latter giving larger map distances (28–34 cM) than the former (7.6–10.1 cM). The identification of two chromosome mutants lacking about half of the long arms of chromosome 1B and 1D and also lacking the corresponding *Glu-B1-* and *Glu-D1*-encoded HMW subunits led Payne (1987) to conclude that the *Glu-1* loci are physically located on the distal halves of the chromosome arms. This was subsequently confirmed by Curtis and Lukaszewski (1991), who used C-banding patterns to study the distribution of recombination along chromosome 1B.

C. Low molecular weight glutenin subunits

The LMW glutenin subunits have proved more difficult to separate because of their complexity, heterogeneity, and similarity to each other and to some gliadin components. To overcome this, Jackson *et al.* (1983) fractionated total

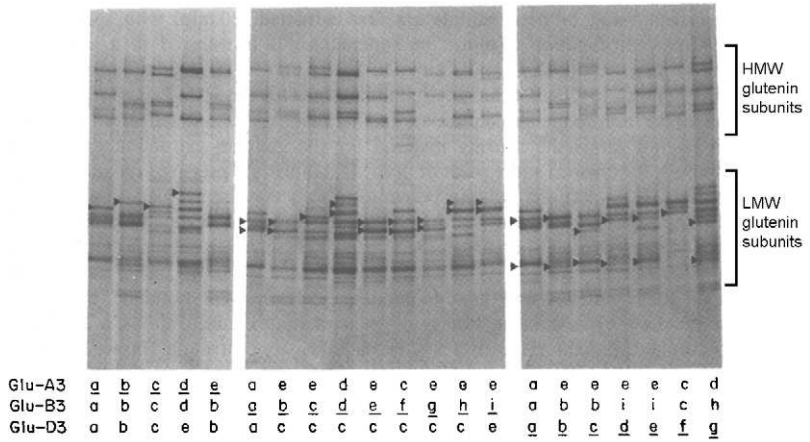


Figure 3.8. SDS–PAGE separation of HMW and LMW glutenin subunits present in different bread wheat cultivars. Designations for different alleles at the *Glu-3* loci are given at the bottom. Arrowheads indicate subunits that are useful in the identification of *Glu-3* alleles. Taken from Singh *et al.* (1991) with permission.

endosperm proteins by using two different two-dimensional systems that combined either IEF or nonequilibrium pH gradient electrophoresis (NEPHGE) in the first dimension, followed by SDS–PAGE in the second dimension (Fig. 3.3). Using these procedures to analyze the Chinese Spring aneuploids, LMW subunit genes were assigned to the short arms of the homoeologous group 1 chromosomes (Jackson *et al.*, 1983). Subsequent genetic analyses did not reveal any recombination between genes encoding LMW subunits and gliadins, so Payne *et al.* (1984b) concluded that the complex *Gli-1* loci contain three families of genes corresponding to the ω- and γ-gliadins and the LMW subunits. In contrast, Singh and Shepherd (1988b) reported evidence of recombination between genes for an ω-gliadin and B-type LMW subunit proteins. This has subsequently been confirmed by other research groups and consequently the symbol *Glu-3* was assigned to the loci encoding the B-type LMW subunits (Pogna *et al.*, 1990; Gupta and Shepherd, 1993). Recombination between B-type LMW glutenin subunits and gliadins encoded by the *Gli-A1* and *Gli-B1* loci on chromosomes 1A and 1B has also been observed in *T. durum*. Map distances of 1.3 cM between *Glu-A3* and *Gli-A1* and of 2.0 cM between *Glu-B3* and *Gli-B1* were reported by Ruiz and Carrillo (1993), confirming data reported by Pogna *et al.* (1990). The latter group also reported that *Glu-B3* was located between the centromere and *Gli-B1* on the short arm of chromosome 1B. Large allelic variation has been reported at the three *Glu-3* loci in bread wheat, with different alleles being designated by the name of the locus followed by letters (Fig. 3.8) (Gupta and Shepherd, 1988; Singh *et al.*, 1991).

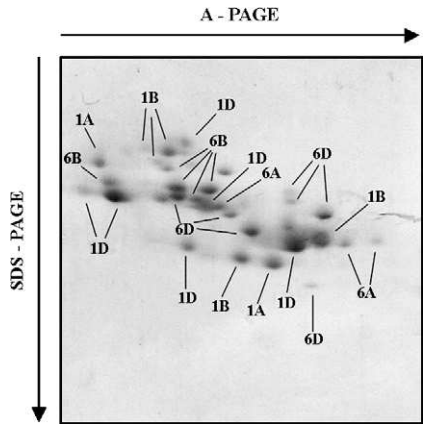


Figure 3.9. Two-dimensional electrophoresis of C-type LMW glutenin subunits of cv. Chinese spring and their chromosomal assignments.

The development of methods to purify fractions enriched in C-type LMW subunits (i.e., mutant forms of γ- and α-type gliadins) has allowed their genes to be mapped to the same group 1 and 6 chromosomes (Fig. 3.9), either tightly linked to or within the *Gli-1* and *Gli-2* loci (Masci *et al.*, 2002). This confirms earlier reports that the group 6 chromosomes contain genes encoding glutenin-type proteins (Gupta and Shepherd, 1993; Pogna *et al.*, 1995a; Felix *et al.*, 1996).

D. Minor gliadin and glutenin subunit loci

As discussed above, genes encoding most of the ω- and γ-gliadins are tightly clustered at the *Gli-A1*, *Gli-B1*, and *Gli-D1* loci on the distal ends of the short arms of chromosomes 1A, 1B, and 1D (Payne, 1987). However, a few gliadin components have been shown to be encoded by additional, dispersed genes (Fig. 3.10). An additional *Gli-B3* locus has been mapped on the short arm of chromosome 1B between the *Gli-B1* locus and the centromere (Galili and Feldman, 1984; Metakovsky *et al.*, 1986; Dachkevitch *et al.*, 1993). Furthermore, a gliadin locus was also mapped to the short arm of chromosome 1A, at a position comparable with that of *Gli-B3* (Sobko, 1984; Dachkevitch *et al.*, 1993), and therefore designated *Gli-A3*. The *Gli-B3* locus was subsequently found to be complex (Ruiz and Carrillo, 1993), containing tightly linked genes for an ω-gliadin and a B-type LMW subunit.

Studies of the inheritance of gliadin genes in the progeny of crosses between the bread wheat cultivar Salmone and six other Italian cultivars revealed an additional gliadin locus, *Gli-B5*, 1.4 cM from the *Gli-B1* locus on

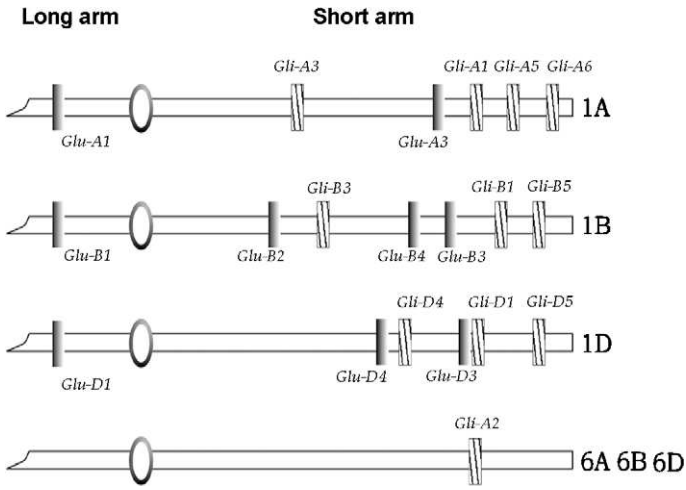


Figure 3.10. Chromosomal locations of major and minor gliadin and glutenin subunit loci in bread wheat.

chromosome 1B (Pogna *et al*., 1993). Similarly, a novel locus, *Gli-A5*, homoeologous to *Gli-B5*, was identified on chromosome 1A, 2 cM from *Gli-A1*, and evidence was obtained that a further locus homoeologous to *Gli-B5* also exists on chromosome 1D (Pogna *et al*., 1995b).

Evidence of a novel locus (*Gli-A6*), located distally to *Gli-A1* on chromosome 1A (with about 5% recombination), was reported by Metakovsky *et al*. (1996), while two new loci designated *Gli-D4* and *Gli-D5* have been described on the short arm of chromosome 1D, by Rodríguez-Quijano and Carrillo (1996a). *Gli-D4*, controlling a γ-gliadin, is present between the centromere and the *Gli-D1* locus at a map distance of 10.1 cM, whereas *Gli-D5*, encoding an ω-gliadin, was mapped at 3.7 cM from the *Gli-D1* locus on the distal part of the chromosome.

During the analysis of near-isogenic lines of cv. Alpe (Pogna *et al*., 1995a) and of mutant bread wheat lines lacking the *Glu-B3* locus (Redaelli *et al*., 1995), some B-type LMW subunits could not be assigned to any *Glu-3* locus, although all could be assigned to the short arms of the group 1 chromosomes (Gupta and Shepherd, 1990). However, the demonstration that a B-type LMW subunit is encoded by a novel locus loosely linked (about 20% recombination) with the *Glu-B3* locus on the short arm of chromosome 1B of durum wheat (Ruiz and Carrillo, 1993; Liu and Shepherd, 1995) suggested that this additional locus is also present on chromosome 1BS in bread wheat. This conclusion has been confirmed by Metakovsky *et al*. (1997a), who described a

B-type LMW subunit encoded by a gene recombining with *Gli-B1*/*Glu-B3* at a frequency of 24.8%. According to Liu and Shepherd (1995), this novel locus was designated *Glu-B2*. Furthermore, two polypeptides with relative molecular masses of 30,000 and 32,000 Da, which have N-terminal amino acid sequences homologous to those of B- and C-type LMW subunits, have been identified by Sreeramulu and Singh (1997). These proteins were assigned to the novel *Glu-D4* and *Glu-D5* loci on chromosomes 1D and 7D, respectively, with the *Glu-D4* locus being loosely linked to *Gli-D1*/*Glu-D3*.

IV. ANALYSIS OF GLUTEN PROTEIN MULTIGENE LOCI

Analysis of the structure and organization of the multigenic loci encoding wheat prolamins is complicated by two factors: the presence of related genes on the separate genomes and the high level of sequence homology between the proteins that they encode, leading to extensive cross-hybridization on Southern blotting.

A. The *Glu-1* loci

The simplest pattern is shown by the *Glu-1* loci encoding the high molecular weight subunits of glutenin. There is no cross-hybridization between the DNA sequences encoding these proteins and other gluten protein genes, although cross-hybridization between the individual *Glu-1* genes occurs. Southern blotting of DNA from hexaploid wheat varieties usually shows the presence of fragments corresponding to the *Glu-A1x*, *Glu-A1y*, *Glu-B1x*, *Glu-B1y*, *Glu-D1x*, and *Glu-D1y* genes, although some restriction enzymes may cut within the coding sequences, resulting in more than six fragments (Thompson *et al.*, 1983; Harberd *et al.*, 1986). The individual fragments can be assigned to the long arms of the group 1 chromosomes by analysis of nullisomic–tetrasomic and ditelosomic lines and to individual x-type or y-type proteins by comparison of their sizes with those derived from restriction digestion of cloned genomic fragments (Fig. 3.11). The HMW subunit genes generally occur as single copies. However, D'Ovidio *et al.* (1997) reported that the *Glu-B1x* gene encoding subunit 1Bx7 has been duplicated in the cultivar Red River 68, resulting in an increased amount of subunit protein and "overstrong" dough characteristics (D'Ovidio *et al.*, 1997; Lukow *et al.*, 1992).

In general, all hexaploid bread wheat varieties contain six HMW subunit genes, irrespectively of whether three, four, or five of these are expressed (J. Forde *et al.*, 1985; Margiotta *et al.*, 1993). However, Margiotta *et al.* (1993) showed that one variety (Nap Hal) had no fragments corresponding to chromosome 1D, indicating that it was a deletion mutant. They also showed

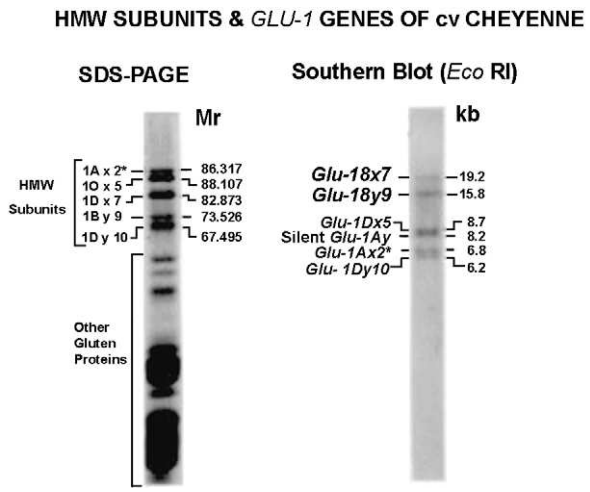


Figure 3.11. The HMW subunit proteins separated by SDS–PAGE and *Glu-1* genes determined by Southern blotting of bread wheat cv. Cheyenne. Taken from Shewry *et al.* (1989) with permission.

that several lines of bread wheat that contained unusually large 1Dx or 1Dy subunits showed correspondingly large hybridizing fragments, indicating that differences in fragment size can result from differences in the degree of reiteration of the repetitive DNA sequences that form the central part of the HMW subunits. More recently, D'Ovidio *et al.* (1996) have shown that the repetitive domain of the high M_r subunit 1Dx2.2* contains an insert of 187 amino acids (corresponding to 561 bases of DNA) and that this accounts for the size difference between restriction fragments generated from products amplified by polymerase chain reaction (PCR) from this gene and the 1Dx2-encoding gene. Polymorphism at *Glu-1* loci can also be used to identify cultivars and to follow alleles in breeding programmes as described by Reddy and Appels (1993).

Margiotta *et al.* (1993) also showed that all durum wheat cultivars gave restriction fragments corresponding to four *Glu-1* genes, although the existence of some deletion mutants cannot be ruled out. Similarly, extensive studies of diploid wheats have not been reported but J. Forde *et al.* (1985) reported the presence of two hybridizing fragments in an accession of the A genome diploid *T. monococcum*.

B. The *Gli-2* locus

The *Gli-2* loci encoding α-type gliadins are present on the group 6 chromosomes. No cross-hybridization between sequences encoding α-type gliadins and

other prolamins has been reported but extensive cross-hybridization within the *Gli-2* gene family occurs.

Harberd *et al.* (1985) reported that an α-gliadin cDNA hybridized to 12 bands in a *Hin*dIII restriction digest of Chinese Spring. These were estimated to contain between 1 and 5 gene copies each, giving a total copy number of 25–35 copies per haploid genome. All bands could be assigned to group 6 chromosomes (6A, 6B, 6D), using nullisomic–tetrasomic lines.

In contrast, much higher copy numbers have been reported by other workers, more than 100 copies (Okita *et al.*, 1985) and up to 150 copies (Anderson *et al.*, 1997) per haploid genome. Anderson *et al.* (1997) were able to assign most of the 20 *Hin*dIII fragments in the cultivars Cheyenne and Chinese Spring to individual chromosomes, using aneuploid and chromosome substitution lines. Furthermore, Anderson and Greene (1997) compared the sequences of the 27 cDNA and genomic sequences for α-type gliadins that were available then (20 of which were from cv. Cheyenne) and showed several gene subfamilies, some of which appeared to be specific for individual group 6 chromosomes.

The number of α-gliadin genes estimated by these studies is in excess of the total number of α-type gliadin proteins that can be separated by two-dimensional electrophoresis, implying that some may be pseudogenes. In fact, Anderson and Greene (1997) reported that half of the genomic α-gliadin gene sequences contained stop codons, indicating that they corresponded to pseudogenes. The presence of additional cysteine residues in α-type gliadin sequences reported for bread wheat (Okita *et al.*, 1985) and durum wheat (D'Ovidio *et al.*, 1995b) indicates that these may correspond to C-type LMW subunits of glutenin.

Vaccino and Metakovsky (1995) also used a cDNA probe encoding an α-type gliadin to study polymorphism at the *Gli-2* loci. This probe recognized several *Gli-A2* alleles but not *Gli-B2* or *Gli-D2* alleles. Similarly, D'Ovidio *et al.* (1992) used an α-gliadin sequence from durum wheat to study polymorphism at the *Gli-2* loci in a range of diploid, tetraploid, and hexaploid species of *Triticum* and *Aegilops*. Precise copy numbers were not determined but all loci were highly polymorphic and their degree of complexity broadly reflected the ploidy level of the species.

C. The *Gli-1/Glu-3* loci

Analysis of the gene families encoding the γ-gliadins and ω-gliadins (including the D group of LMW subunits) is difficult as sequences encoding these proteins may cross-hybridize. Thus, analyses carried out with a γ-gliadin sequence are likely to identify fragments encoding γ- and ω-gliadins. Sabelli and Shewry (1991) attempted to eliminte this problem, by comparing the hybridization

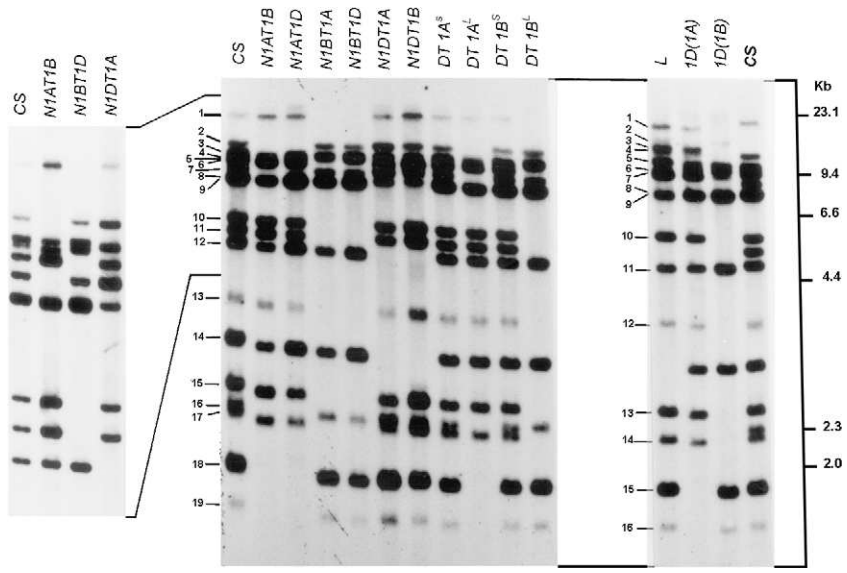


Figure 3.12. Assignment of restriction fragments hybridizing to a γ-gliadin cDNA probe to chromosomes in bread wheat cv. Chinese Spring and durum wheat cv. Langdon. DNA from cv. Chinese Spring (CS) and its nullisomic–tetrasomic and ditelocentric lines and from cv. Langdon (L) and the disomic substitution lines of chromosome 1D of cv. Chinese Spring into cv. Langdon [1D(1A), 1D(1B)] was digested with *Hin*dIII, separated on long (left) or normal (center and right) gels, and probed with a γ-gliadin cDNA labeled with ^{32}P. Hybridizing bands corresponding to γ- and ω-gliadins (1–19 in CS, 1–16 in L) are revealed by autoradiography. Taken from Sabelli and Shewry (1991) with permission.

patterns obtained with a γ-gliadin cDNA, a partial C hordein cDNA (both of which should hybridize to γ-type and ω-type genes), and the 3′ part of the γ-gliadin cDNA (which should be specific for γ-type genes). Comparison of the hybridization patterns obtained with these probes therefore allowed the copy numbers of γ-type + ω-type and γ-type genes to be estimated and of ω-type genes to be calculated by subtraction. Furthermore, the DNA fragments were separated on long and normal gels and the hybridization stringency was varied to give optimum separation and identification of the fragments. An example of the results obtained with the whole γ-gliadin cDNA probe is shown in (Fig. 3.12) and the results obtained for cv. Chinese Spring are summarized in Table 3.2. It can be seen that the total estimates vary between restriction digests, from about 32 copies when *Bam* HI was used to 55 with *Hin*dIII. The reason for this is not known. Sabelli and Shewry (1991) also compared the

Table 3.2. Numbers of Fragments and Gene Copies Corresponding to γ-Gliadins, ω-Gliadins, and Low Molecular Weight Subunits Revealed by Southern Blotting of DNA from Bread Wheat cv. Chinese Spring Digested with *Hin*dIII, *Eco*RI, and *Bam*HI

	*Hin*dIII		*Eco*RI		*Bam*HI	
	No. of fragments	Total copy no.	No. of fragments	Total copy no.	No. of fragments	Total copy no.
γ/ω Gliadin Sequences						
Total						
1A	7	16	5	11	4	10
1B	7	19	7	18	5	10
1D	5	20	4	19	3	12
1A + 1B + 1D	19	55	16	48	12	32
γ-type						
1A	3	7	2	5	3	6
1B	5	14	6	16	4	7
1D	4	18	3	9	2	4
1A + 1B + 1D	12	39	11	30	9	17
ω-type						
1A	4	9	3	6	1	4
1B	2	5	1	2	1	3
1D	1	2	1	10	1	8
1A + 1B + 1D	7	16	5	18	3	15
LMW Subunit Sequences						
1A	4	10	3	8	4	8
1B	4	12	5	11	2	3
1D	7	17	5	15	5	11
1A + 1B + 1D	15	39	13	34	11	22

patterns in cv. Chinese Spring with the durum wheat cv. Langdon. It was not possible to assign all of the hybridizing fragments in cv. Langdon to chromosomes and hence the results are not given in Table 3.2 However, the total numbers of fragments related to γ-gliadins and ω-gliadins were lower than in cv. Chinese Spring (24 and 12, respectively, compared with 39 and 16), which is consistent with the difference in ploidy. Furthermore, some of the fragments encoded by chromosomes 1A and 1B appeared to be conserved between cv. Chinese Spring and cv. Langdon in terms of their size, location, and copy number. As with the *Gli-2* loci, the γ-gliadin-related fragments identified by Sabelli and Shewry (1991) could be expected to include genes encoding C-type LMW subunits (D'Ovidio *et al.*, 1995a; Shewry and Tatham, 1997).

The same γ-gliadin cDNA probe was used by Vaccino and Metakovsky (1995). They compared the DNA restriction fragment patterns of a number of varieties of bread wheat, making no attempt to distinguish between γ- and ω-gliadins. They showed that all *Gli-B1* and *Gli-D1* alleles that differed in γ-gliadin proteins also differed in their restriction fragment length polymorphism (RFLP) patterns, with two distinct groups of *Gli-B1* alleles that could have originated from distinct genotypes of the B genome donor species. Polymorphism between some *Gli-A1* alleles was also demonstrated. The presence of some silent genes for γ-gliadins was demonstrated by Rafalski (1986), who isolated a genomic fragment encoding one expressed and one silent gene, and is also indicated by the results of Vaccino and Metokovsky (1995). Similarly, Hsia and Anderson (2001) isolated two genes encoding ω-gliadins, one of which contained an "in frame" stop codon and was therefore assumed to be a pseudogene.

Sabelli and Shewry (1991) used a similar approach to that discussed above to analyze the *Glu-3* loci encoding the B-type LMW subunits of glutenin. The probe used did not hybridize to the γ- and ω-gliadin fragments but, as pointed out by Cassidy *et al.* (1998), may not have hybridized with equal intensity to all the LMW subunit genes. Between 22 and 39 gene copies were estimated in cv. Chinese Spring depending on the restriction enzyme (Table 3.2), all of which were assigned to group 1 chromosomes. This is consistent with the estimate (Cassidy *et al.*, 1998) of 30–40 copies. Lower numbers of hybridizing fragments were reported in durum wheat cv. Langdon (six each with *Hin*dIII and *Bam*HI, five with *Eco*RI), only some of which could be assigned to group 1 chromosomes (Sabelli and Shewry, 1991).

Cassidy *et al.* (1998) also reviewed the available sequences for B-type LMW subunits (17 in total) but were unable to identify any pseudogenes. This contrasts with the situation with α-gliadin and γ-gliadin genes, as discussed above.

D. Analysis of null mutants at the *Gli-1/Glu-3* loci

Screening of germplasm collection by electrophoresis identified four bread and one durum wheat line in which specific ω- and γ-gliadin proteins were apparently absent (Lafiandra *et al.*, 1987; Sabelli *et al.*, 1992). Analysis of these by Southern blotting with a γ-gliadin cDNA probe, which would be expected to hybridize to γ-type and ω-gliadin genes, identified several hybridizing fragments that were absent from chromosome 1D of the bread wheat lines or chromosome 1B of the durum wheat line (Sabelli *et al.*, 1992). In addition, fragments that hybridized to an LMW subunit probe were also absent from the null lines, indicating that simultaneous deletion of genes encoding γ-gliadins, ω-gliadins, and LMW subunits had occurred.

V. ISOLATION, CHARACTERIZATION, AND EXPRESSION OF GLUTEN PROTEIN GENES

A. Structure and evolution of high molecular weight subunit genes

The first HMW glutenin subunit gene sequence to be reported was that of a partial cDNA isolated by J. Forde and co-workers in 1983 (J. Forde *et al.*, 1983). The first full-length HMW subunit gene sequences were reported in 1985. These were the silent *Glu-Aly* gene from cv. Cheyenne (J. Forde *et al.*, 1985), the *Glu-D1x* gene from cv. Yamhill (Sugiyama *et al.*, 1985) and the *Glu-Dly* gene encoding subunit 1Dy12 from cv. Chinese Spring (Thompson *et al.*, 1985). The remainder of the HMW glutenin subunit genes from cv. Cheyenne were subsequently isolated and sequenced. These were *Glu-A1x* encoding HMW subunit 1Ax2* (Anderson and Greene, 1989), *Glu-B1x* encoding HMW subunit 1Bx7 (Anderson and Greene, 1989), *Glu-Bly* encoding subunit 1By9 (Halford *et al.*, 1987), *Glu-D1x* encoding subunit 1Dx5 (Anderson *et al.*, 1989), and *Glu-Dly* encoding subunit 1Dy10 (Anderson *et al.*, 1989). The sequences of the *Glu-A1x* gene encoding subunit 1Ax1 of cv. Hope (Halford *et al.*, 1992) and two silent *Glu-A1x* alleles (Xin *et al.*, 1992; Bustos *et al.*, 2000) have also been reported.

Full-length sequences have also been obtained for the *Glu-A1x* and *Glu-A1y* genes of *T. timopheevii* (Wan *et al.*, 2002), the *Glu-C1y* gene and silent *Glu-D1x* allele of *A. cylindrica* (Wan *et al.*, 2002), the *Glu-D1y* genes of *T. tauschii* (Mackie *et al.*, 1996) and *A. cylindrica* (Wan *et al.*, 2002), and a silent *Glu-Dly* allele from *A. crassa*. The database accession numbers of all of these gene sequences are given in Table 3.3. Note that the x-type genes are sometimes referred to as *Glu-X1-1* instead of *Glu-X1x*, and the y-type genes as *Glu-X1-2*.

Despite the importance of HMW subunits to the processing properties of wheat grain, these are the only full-length gene sequences to have been reported. The length of HMW subunit transcripts (typically greater than 2 kb) and possibly their repetitive nature made full-length cDNAs difficult to synthesize in the early years of plant molecular biology, while the size of the wheat genome made many workers reluctant to perform lengthy library construction and screening. Full-length HMW subunit genes also proved difficult to amplify by polymerase chain reaction, again probably because of their length and repetitive structure. However, these difficulties have been overcome as more reliable and robust polymerases have become available (D'Ovidio *et al.*, 1995a; Wan *et al.*, 2002).

The HMW subunit amino acid sequences derived from their gene sequences enabled links to be made between primary amino acid sequence, predicted secondary and higher levels of protein structure, and the properties of the proteins and their role in gluten structure. It also enabled the identification

Table 3.3. Cloned High Molecular Weight Subunit Genes[a]

Gene	Species	Subunit	Accession numbers
Glu-A1x	*Triticum aestivum*	**1Ax1**	**X61009**
		1Ax2*	**M22208**
		Promoter only	AJ399893, Y10952, X98586
	Triticum dicoccum	Promoter only	X98712
	Triticum monococcum	Promoter only	Y12402
	Triticum spelta	Promoter only	AJ399886, AJ399880
	Triticum timopheevii	**1Ax**	**AJ306976**
		Promoter only	Y12406
	Triticum turgidum	Promoter only	AJ131815, Y10964
	Triticum urartu	Promoter only	Y12407
	Triticum sp. (archaeological)	Promoter only	X98589, Y10957, Y17993
Glu-A1x (silent)	*Triticum aestivum*	**Truncated 1Ax**	**AF145590, U19774**
Glu-A1y	*Triticum dicoccum*	Promoter only	X98711, X98715
	Triticum monococcum	Promoter only	Y12401
	Triticum spelta	Promoter only	AJ399881, AJ399887
	Triticum timopheevii	**1Ay**	**AJ306977**
	Triticum turgidum	Promoter only	AJ132190, Y10965
	Triticum urartu	Promoter only	Y12409
	Triticum sp. (archaeological)	Promoter only	X98592, Y10958, Y17994
Silent *Glu-A1y*	*Triticum aestivum*	**Truncated 1Ay**	**X03042**
		Promoter only	AJ399894, AY052153, X98587, Y10953
Silent *Glu-A1y*	*Triticum aestivum*	**Interrupted/ truncated 1Ay**	**X05995 (5′end), X05996 (3′ end)**
Glu-B1x	*Triticum aestivum*	**1Bx7**	**X13927, M22209**
		Promoter only	AJ399896, X98585, Y10954
	Triticum dicoccum	Promoter only	X98714
Glu-B1x	*Triticum spelta*	Promoter only	AJ399882, AJ399888
	Triticum turgidum	Promoter only	AJ131816, Y10966
	Triticum sp. (archaeological)	Promoter only	X98590, Y10959, Y10960, Y10961, Y17991, Y17996
Glu-B1y	*Triticum aestivum*	**1By9**	**X61026**
		Promoter only	AJ399895, Y10955, X98584
	Triticum dicoccum	Promoter only	X98713
	Triticum spelta	Promoter only	AJ399883, AJ399889
	Triticum turgidum	Promoter only	AJ131814, Y10967
	Triticum sp. (archaeological)	Promoter only	Y10962, Y17992, Y17997, X98591
Glu-C1y	*Aegilops cylindrica*	**1Cy**	**AJ306973**
Glu-D1x	*Triticum aestivum*	**1Dx2**	**X03346,**
		1Dx5	**X12928**
		Promoter only	AJ399897, Y10956, X98583
	Triticum spelta	Promoter only	AJ399884, AJ399890, AJ399891

Table 3.3. (*Continued*)

Gene	Species	Subunit	Accession numbers
	Triticum sp. (archaeological)	Promoter only	Y10963, Y17998, X98588
	Aegilops squarrosa	Promoter only	Y12408
Glu-D1x (silent)	*Aegilops cylindrica*	**Truncated 1Dx**	**AJ306975**
Glu-D1y	*Triticum aestivum*	**1Dy10**	**X12929,**
		1Dy12	**X03041**
		Promoter only	AJ399898, AY052151, AY052152
	Triticum spelta	Promoter only	AJ399885, AJ399892
	Triticum tauschii	**1Dy**	**U39229**
	Aegilops cylindrica	**1Dy**	**AJ306974**
	Aegilops squarrosa	Promoter only	Y12410
	Aegilops ventricosa	**y-type**	**AF226698**
Silent *Glu-D1y*	*Aegilops crassa*	**Truncated y-type**	**AF354289**
Glu-G1x	*Triticum timopheevii*	Promoter only	Y12404
Glu-G1y	*Triticum timopheevii*	Promoter only	Y12403, Y12405

[a]Entries containing coding sequences are in boldface.

of specific residues that were missing or present in certain HMW subunits that were associated with good or poor breadmaking quality. These are discussed in more detail in Section IV. The sequences also shed light on how the different genes had evolved and diverged, and led to experiments to characterize their promoters. The latter, which are described in detail in Section V. C, showed that the HMW subunit gene promoters were capable of driving high levels of expression of both homologous and heterologous gene sequences in wheat endosperm, making them important potential tools for wheat biotechnology (Altpeter *et al.*, 1996; Barro *et al.*, 1997; Lamacchia *et al.*, 2001). This may explain the plethora of promoter sequences that have been obtained (Table 3.3).

The proteins encoded by all of the HMW subunit genes listed above have essentially similar structures. Each comprises a 21-amino acid signal peptide and short N- and C-terminal domains flanking a more extensive repetitive domain (Fig. 3.5). A dendrogram showing the evolutionary relationships of the different proteins is shown in Fig. 3.13. It was produced using only the N- and C-terminal domain sequences, since the highly repetitive nature of the central domain could distort the analysis. The N- and C-terminal domain sequences of HMW subunits 1Ax1, 1Bx7, 1Dx5, 1Ay (silent), 1By9, and 1Dx5 are also aligned in Fig. 3.14.

These analyses (see also Halford *et al.*, 1987; Shewry *et al.*, 1989; Wan *et al.*, 2002) confirm that the x- and y-type subunits classified on the basis of their mobility on SDS–PAGE and on genetic studies represent two different

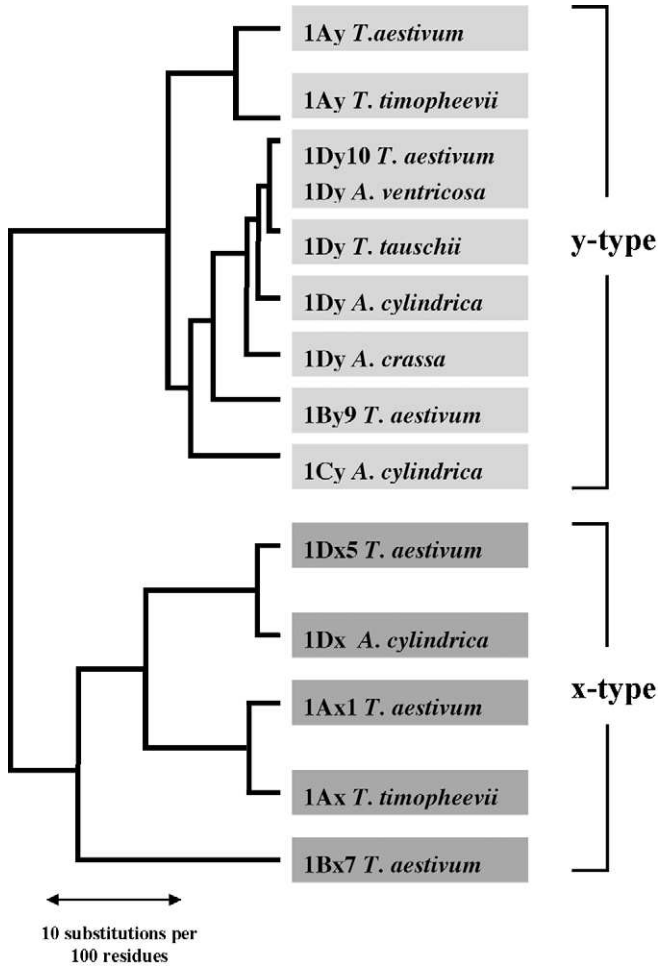


Figure 3.13. Dendrogram showing the relationships between HMW subunits encoded by genes on different genomes and from different wheat species. The tree was derived from multiple N- and C-terminal sequence alignments produced using the PILEUP program with a gap creation penalty of 3.0 and a gap extension penalty of 0.1. Evolutionary distances were calculated using the DISTANCES program, correcting for multiple substitutions at a single site, and displayed by the unweighted pair group method using arithmetic averages with the GROWTREE program (Wisconsin Package version 10; Genetics Computer Group, Madison, WI). Database accession numbers for the genes are given in Table 3.3.

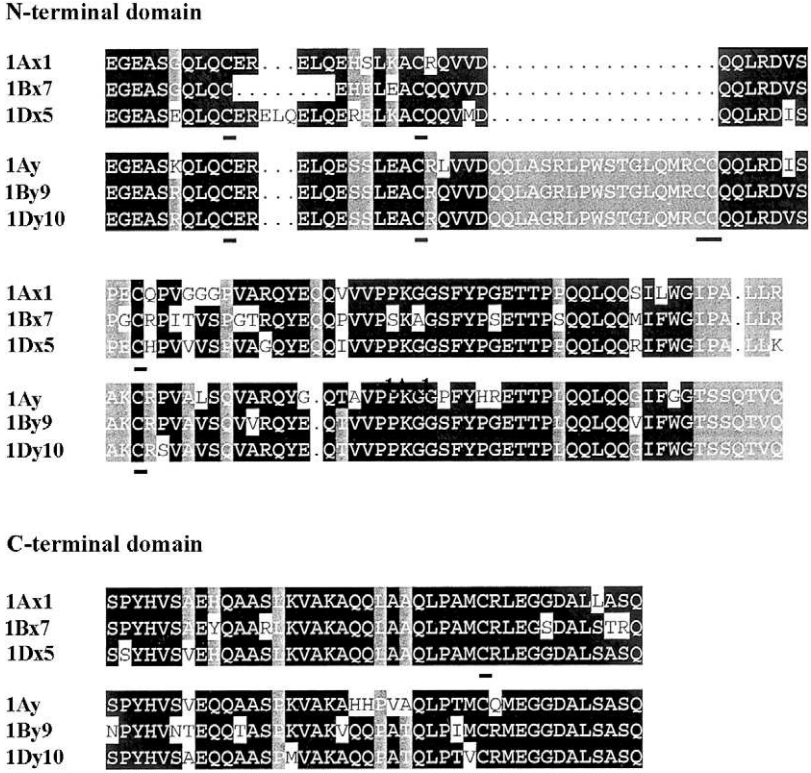


Figure 3.14. Alignments of the N- and C-terminal amino acid sequences of x- and y-type HMW subunits encoded by genes on different genomes. The alignment was produced using the PILEUP program with a gap creation penalty of 3.0 and a gap extension penalty of 0.1 (Wisconsin Package version 10; Genetics Computer Group, Madison, WI). Residues that are conserved in x- and y-type subunits are highlighted in black; those conserved in one type but not the other are highlighted in gray. Cysteines are indicated by solid bars. Database accession numbers for the genes are given in Table 3.3.

subclasses. This means that the first step in the evolutionary process that led to the diversity of HMW subunits in modern wheat and its relatives was a duplication of a single ancestral gene, *Glu-X1*, into two closely linked copies, *Glu-X1x* and *Glu-X1y*. These copies then diverged and their encoded proteins (x- and y-type HMW subunits) were already distinguishable when the speciation events occurred that eventually led to the ancestors of modern cultivated wheats and their wild relatives. Consequently, the *Glu-A1-*, *Glu-B1-*, *Glu-C1-*, and *Glu-D1-*encoded y-type HMW subunits from *T. aestivum*, *T. tauschii*, *A. cylindrica*, and *A. crassa*, for example, are more similar to each

other than they are to the x-type HMW subunits encoded by genes from the same genomes (Fig. 3.13).

This kind of analysis should also be useful in characterizing the HMW subunits of the wild relatives of wheat and determining the most likely origins of the genomes of bread wheat (Wan *et al*., 2002). However, it may not be possible to amplify by PCR the HMW subunit genes of more divergent genomes, using oligonucleotide primers based on the gene sequences of bread wheat (D'Ovidio *et al*., 1995b; Wan *et al*., 2002). For example, attempts to amplify the *Glu-C1x* gene of *A. cylindrica* and the *Glu-G1x* and *Glu-G1y* genes of *T. timopheevii*, using a variety of combinations of primers, were unsuccessful (Wan *et al*., 2002; Y. Wan, P. R. Shewry, and N. G. Halford, unpublished data).

The y-type HMW subunits contain five cysteine residues in the N-terminal domain and one in the C-terminal domain. Many, including subunits 1By9 and 1Dy10 (Fig. 3.15), also contain a single cysteine residue in the repetitive domain while the 1Dy subunit of *A. cylindrica* has two cysteines in the repetitive domain (Wan *et al*., 2002). The x-type subunits contain an 18-amino piece amino acid deletion in the N-terminal domain that includes two cysteine residues. However, subunit 1Dx5 also contains a cysteine residue at the N-terminal end of the repetitive domain (Fig. 3.15) that is not present in any other subunit, including its allelic form, 1Dx2. The number and distribution of cysteine residues in the HMW subunit proteins are relevant to their ability to form high M_r polymers stabilized by interchain disulfide bonds (Shewry and Tatham, 1997), which is discussed in more detail in Section V. Subunit 1Dx5 also contains a duplication of the Glu-Leu-Gln peptide at position 13, while 1Bx7 lacks this peptide and single residues immediately before and after it.

The size variation between different HMW subunits is attributable mostly to the insertion and duplication of blocks of repeats in the central repetitive domain (Fig. 3.15). In x-type subunits, the repetitive domain consists of three types of repeat (Halford *et al*., 1987; Shewry *et al*., 1989): a hexapeptide repeat with the consensus sequence Pro-Gly-Gln-Gly-Gln-Gln, a 9-amino acid repeat consisting of the hexapeptide plus a tripeptide with the consensus sequence Gly-Gln-Gln, and a 15-amino acid repeat consisting of the hexapeptide plus a nonapeptide with the consensus sequence Gly-Tyr-Tyr-Pro-Thr-Ser-Pro-Gln-Gln. In y-type subunits the repetitive domain contains only the hexapeptide and hexapeptide/nonapeptide repeats and in the nonapeptide leucine is more prevalent than proline at position 7 (Shewry *et al*., 1989). There are 23 amino acid residues at the N-terminal ends of the repetitive domains of both x- and y-type HMW subunits that show similarity with the other repeats but do not have the same structure, and single truncated hexapeptide/nonapeptide repeats at the C-terminal ends of the repetitive domain.

The processes of insertion and duplication of blocks of repeats appear to be quite rapid in some regions in some subunits, for example, the last stretch of hexapeptide/tripeptide repeats in subunit 1Ax1 consists of only 3 repeats, whereas the equivalent region in subunit 1Dx5 consists of 10. This region is missing entirely in subunit 1Bx7. In fact, subunit 1Bx7 lacks a number of other blocks of repeats present in the other x-type subunits while containing several that are not present in the others. It also appears to have diverged more rapidly in the N- and C-terminal domains (Fig. 3.14). Intriguingly, the same is not true of subunit 1By9, so this cannot be attributed simply to the B genome having diverged earlier than the A and D genomes.

There is considerable degeneracy within the repeat sequences, although the length of the repeats and the residues at certain positions (notably positions 3, 5, and 6 of the hexapeptide, positions 1–3 of the tripeptide, and positions 1, 6, and 8 of the nonapeptide) are tightly conserved (Shewry *et al.*, 1989; Feeney *et al.*, 2001). This is true not only for the HMW subunits from modern cultivated wheat varieties, but also for those from the primitive cultivated species, *T. timopheevii*, and the wild relative, *A. cylindrica* (Wan *et al.*, 2002). The significance of this for HMW subunit structure is discussed in Section VI. However, it is worth noting here that the repeat motifs of the HMW subunits from *T. timopheevii* and *A. cylindrica* are not significantly more or less degenerate than those of the HMW subunits from *T. aestivum* (Wan *et al.*, 2002). Also, no differences were observed between the sizes of the repetitive domains of the subunits from bread wheat, *T. timopheevii* and *A. cylindrica*. This indicates that any selection pressure affecting the conservation of residues at certain positions of the motifs or the length of the repetitive domain must operate similarly in all three species. We suggest, therefore, that the structure adopted by the proteins results from natural selection in relation to packaging of the proteins in the developing grain rather than artificial selection for processing quality.

The failure of selective breeding to alter the sequence, structure, or length of the HMW subunits may reflect the relatively low priority given to quality compared with yield in wheat breeding until recent times. Alternatively, it may indicate that changes in the expression levels or structure of individual subunits did not make enough difference to grain quality to be detected and selected for until the advent of modern genetic techniques and the analysis of large numbers of cultivars.

Mutations at variable positions of the repeat motifs appears to be random with no evidence for selection of mutations that give rise to conservative substitutions in the amino acid sequence. At position 1 of the hexapeptide, for example, the most commonly present codon is CCA, encoding proline (Fig. 3.16A). The next most common codons at this position are TCA (serine) and CTA (leucine), both differing from the consensus codon at only

A

```
1Ax1  RYYLSVTSPQQVSYYPGQASSQRPGQGQQ.....PGQGQQGEYYLTSPQQSGWMQQPGQGQARGYYPTSPQQSGQGQPGYYPTSPWQPEQLQ
1Bx7  RYYPSVTSSQQGSYYPGQASPQQSGQGQQPGQGQQHQQPGQGQQGYWPTSPQQPGQGQQGQGQPGYYPTS.QQ.....PGQKQ
1Dx5  RYYPSVT@PQQVSYYPGQASPQRPGQGQQ.....PGQGQQGYYPTSPQQPGQWQQPEQGQPRYYPTSPWQ.....SGQLQ

1Ax1  QPTGQGQRGQPGQGQQLEQGQGQGQQSGQGQPRYYPTSSQQPGQLQQLAQGQGQQPERGQGQQSGQGQQLGQGQGQQPGQKQQ.....SGQGQ
1Bx7  Q.....AGQGQQ.....SGQGQGSYYPTSPQQSGQGQQ.....PGQ
1Dx5  Q.....PAQGQQPGQGQGQQPGQGQGYYPTSSQLQPGQLQQPAQGQGQQPGQAQGQQPGQGQQPGQGQQPGQGQQPGQGQQPGQLGQGQ

1Ax1  QGYYPTSPQQLGQGQQ.....SGQGQLGYYPTSPQQ.....SGQGQSGYYPTSAQQPGQLQQSTQEQQ.....LGQ
1Bx7  PGYYPTSPQQSGQWQQPGQGQQPGQGQQSGQGQQGQQPGQGQQ.....PGQGQQGYYPISPQQPGQGQQPGQGQQGYYPTSLRQPGQWQQPGQGQQ.....PGQ
1Dx5  QGYYPTSLQQ.....SGQGQPGYYPTSLQQ.....LGQGQSGYYPTSPQQPGQEQQPRQLQQ.....PGQ

1Ax1  EQQPQQSGQGRQGQQSGQRQQDQQSGQGQQ.....KGQRQPGYYSTSPQQ.....LGQGQPRYYPTSPQQPGQEQQPRQLQQ.....
1Bx7  GQQGQQ.....PGQGQQ.....SGQGQGYYPTSLQQPGQGQQ.....LGQGQPGYYPTS.SEQGQQPGQGKQPGQGQGYYPTSPQQSGQGQQL
1Dx5  GQQCQQPGQGQQGQQ.....PGQGQQ.....SGQGQPGYYPTSPQQ.....S.SGQGPGYYPTSSQQPTQSQQ.....

1Ax1  .....PEQGQQGQQPEQGQGQQPGQGEQGQQPGQGQQGQQ.....PGQGQPGYYPTSPQQ.....SGQGQPGYYPTSPQQPGQLQQPA
1Bx7  GQGQPGYYPTSPQQ.....SGQGQQ.....SGQGQQGYYPTSPQQSGQGQQ.....SGQLQQPG
1Dx5  .....PGQGQQGQQVGQGQQPAGQ.....PGQGQQ.....PGQGQPGYYPTSPQQ.....SGQGQPGYYLTSPQQSGQGQQPG

1Ax1  QGQQPGQEQGQGQQ.....PGQGQQGQQPGQGQQ.....PGQGQPGYYPTSPQQSGQDQQLEQWQQGQGQPGHYTTSPLQPGQGQPGYYPTSPQQ
1Bx7  QGQQ.....SGQGQQCQQ.....SGQGQQGQQ.....SRQRQQAGQWQRGQGQPGYYPTSPQQ.....
1Dx5  QLQQSAQGQKGQQPGQGQQPGQGQQPGQGQQGQQ.....PGQGQPGYYPTSPQQSGQGQQPGQWQQRGQGQPGYYPTSPLQPGQGQQPGYDPTSPQQ

1Ax1  IGQGQQPGQLQQ.....PNQGQQGQQPGQGQQPGQGQQGQQ.....
1Bx7  PGQEQQSGQAQQSGQWQLVYYPTSPQQPGQLQQPAQGQQPAQGQQSAQEQQ.....
1Dx5  PGQGQQPGQLQQ.....PAQGQQGQQLAQGQQGQQPAQVQQGQRPAQGQQGQQPGQGQQGQQLGQG

1Ax1  .....PGQGQQPGQGQP..GYYPTSLQQSGQGQQPGWQQGQGLPGYYPTSSLQPEQGQQGYY
1Bx7  .....PGQAQQGQWQWQLVYYPTSPQQPGQLQQ.....PAQGQQGYYPTSPQQSGQGQQGYY
1Dx5  QQGQQPGQGQQGQQPAQGQQGQQPGQGQQGQQPGQGQQGQQPGQGQP.WYYPTSPQPSGQGQQPGQWQQGQGLPGYYLTSPLQLGQGQQGYY

1Ax1  PTSQQPGQGPQPGQWQQGQQGQQGYYPTSPQQSGQGQQPGQWLQGQGYYLTSPQQLGQGQQPROWLQPGQGQQGYYPTSPQQSGQGQQ.....
1Bx7  PTSPQQ.....SGQCQQGYYPTSPQQSGQGQQPGQGQQPRQSQQ.GYYPTSPQQSGQGQQ.....PGQGQQGYYPTSPQQSGQGQQPGHEQQ
1Dx5  PTSLQQPGQGQQPGQWQQGQCQHWYYPTSPQLSGQGQRPGQWLQPGQGQQ.GYYPTSPQQPGQGQQLGQWLQPGQGQQGYYPTSLQWTGQGQQ.....

1Ax1  .....LGQGQQGYYPTSPQQSGQGQQGYD
1Bx7  PGQWLQPGQGQQGYYPTSSQQSGQGHQPGQGQQGYA
1Dx5  .....SGQGQQGYY
```

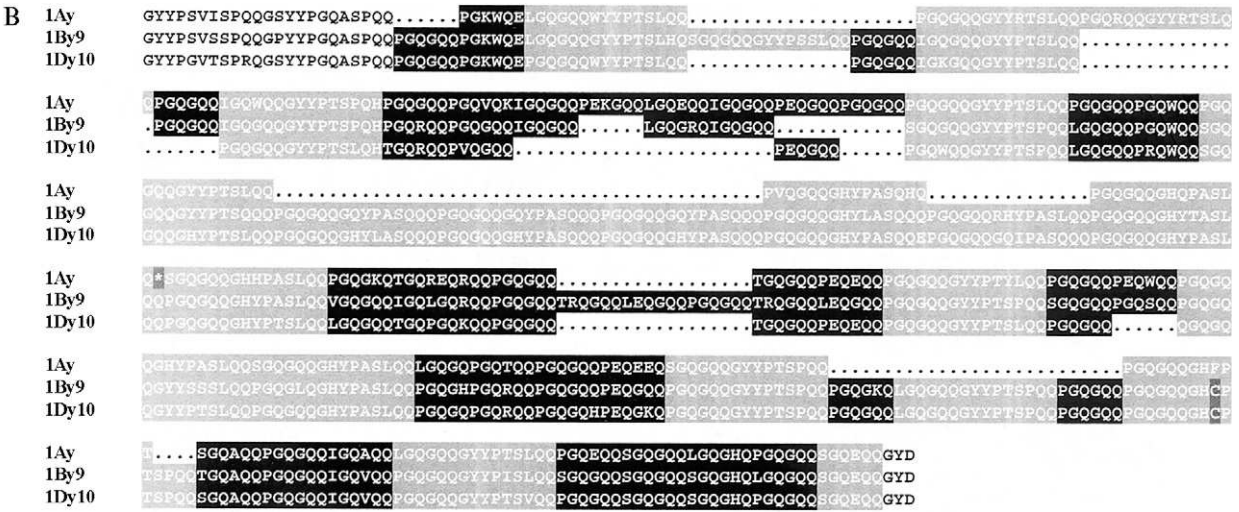


Figure 3.15. Multiple sequence alignments of the derived amino acid sequences of x-type (A) and y-type (B) HMW subunits encoded by genes on different genomes. The alignments were assembled by eye to demonstrate the repeat structure. Cysteine residues are highlighted, hexapeptide repeat motifs are highlighted in black, hexapeptide/nonapeptide 15-amino acid repeat motifs are highlighted in light gray, and hexapeptide/tripeptide 9-amino acid repeat motifs are highlighted in darker grey. Asterisks indicate stop codons. Database accession numbers for the genes are given in Table 3.3.

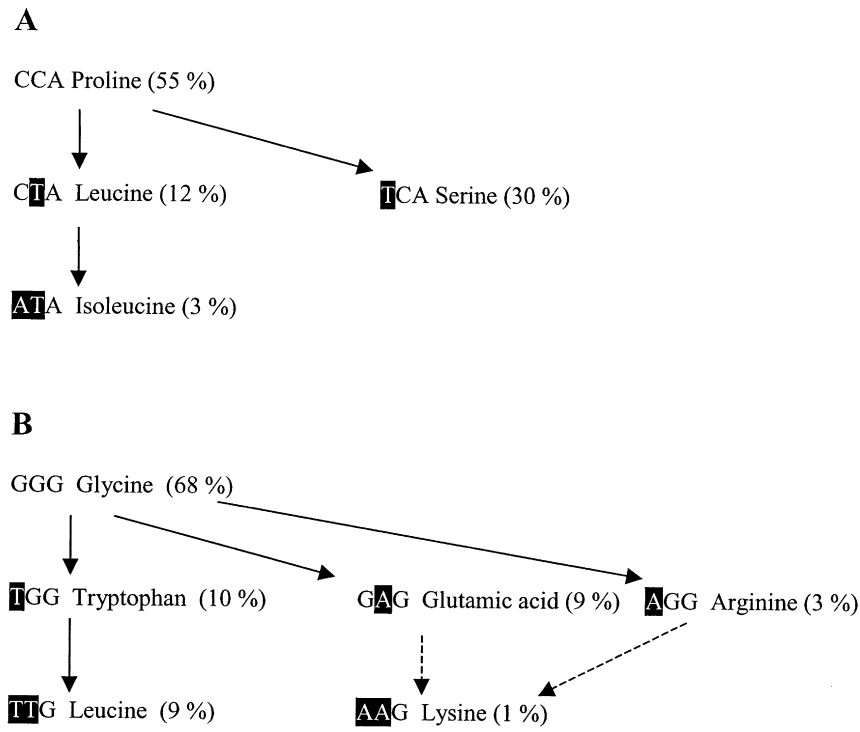


Figure 3.16. Codons present at position 1 (A) and position 4 (B) of the hexapeptide repeat motif of x-type HMW subunits. Figures in parentheses give the percentage of codons of the sequence shown at each position, and mutations are highlighted. Arrows indicate the simplest evolutionary route required to produce each variant codon.

one position, while the next most common is ATA (isoleucine), resulting from a further single mutation. Similarly, GGG (glycine) is the most common codon at position 4 of the hexapeptide (Fig. 3.16B), followed by TGG (tryptophan) and GAG (glutamic acid), both differing from the consensus at only one position, and then TTG (leucine), resulting from a second mutation. AGG (arginine) and AAG (lysine) are also present. Clearly, the most widely occurring variant codons are those produced by single mutations rather than those encoding amino acids similar to the consensus amino acids.

Where mutations do occur they are often repeated in adjacent motifs, sometimes in combination with other mutations. This gives a fascinating insight into how these and probably other repetitive sequences are evolving. A good example of this is a block of 15 amino acid repeats that vary from the consensus sequence (Fig. 3.17A) present in y-type subunits; these are shown for subunits 1By9 and 1Dy10 in Fig. 3.17B. A scheme of how the sequence of

subunit 1By9 could have evolved is shown in Fig. 3.17C. In both 1By9 and 1Dy10 this block comprises seven 15-amino acid repeats in which the tyrosine at position 2 of the nonapeptide has been replaced with histidine and the threonine at position 5 has been replaced with alanine. The same regions of 1Ay and 1Cy subunits have similar variant repeat motifs but only three are present, suggesting that the number of blocks is increasing through duplication.

The hypothesis that blocks of repeat motifs and therefore any mutations in them have been duplicated is supported by the fact that the probability of both the Tyr-His and Thr-Ala mutations occuring together by chance in all seven adjacent repeats but not elsewhere must be extremely low. An alternative explanation is that mutations are spread through adjacent repeats through gene conversion, and it is likely that both processes can occur. The low likelihood that random mutations are responsible is supported by the fact that the hexapeptides linked to these variant nonapeptides remain entirely unchanged apart from a GGG-to-CGG (Gly-to-Arg) mutation in one hexapeptide in subunit 1By9. In other words, the mutation rate in the hexapeptide motif would need to be many times lower than in the nonapeptide motif.

Comparison of the sequences of subunits 1By9 and 1Dy10 may provide direct evidence of gene conversion. The glutamine residue at position 2 of the nonapeptide occurs only once in this region of subunit 1Dy10, and is linked with an unusual residue (isoleucine) at position 3. The fact that subunits 1By9 and 1Dy10 have the same number of variant repeats in this region, but that glutamine occurs three times at position 2 in 1By9 but only once in 1Dy10 (and in a different repeat unit), suggests that the spread of this mutation in 1By9 occurred by gene conversion rather than repeat duplication.

B. Structure and evolution of low molecular weight subunit and gliadin genes

A complete α-gliadin sequence was first published in 1985 (Sumner-Smith *et al.*, 1985), a complete γ-gliadin sequence in 1986 (Sugiyama *et al.*, 1986), and an LMW subunit sequence in 1989 (Colot *et al.*, 1989). An ω-gliadin sequence was published in 2001 (Hsia and Anderson, 2001). Lists of LMW subunit and gliadin gene sequences and their accession numbers are given in Tables 3.4 and 3.5.

Analysis of these sequences is complicated by the fact that they originate from complex loci and most are derived from genomic clones and there is no proof that they are expressed. However, it confirms that the α- and γ-gliadins and LMW subunits contain highly repetitive central domains rich in proline and glutamine, flanked by nonrepetitive N- and C-terminal domains (see Fig. 3.5), while the ω-gliadins lack the C-terminal nonrepetitive domain.

A

P G Q G Q Q G Y Y P T S L Q Q
CCAGGACAAGGGCAACAAGGGTACTACCCAACTTCTCTGCA$^{A}/_{G}$CA$^{A}/_{G}$

B

1By9

P G Q G Q Q G Q Y P A S Q Q Q
CCAGGACAAGGGCAACAAGGGCAGTACCCAGCTTCTCAGCAGCAG

P G Q G Q Q G Q Y P A S Q Q Q
CCAGGACAAGGGCAACAAGGGCAGTACCCAGCTTCTCAGCAGCAG

P G Q G Q Q G Q Y P A S Q Q Q
CCAGGACAAGGGCAACAAGGGCAGTACCCAGCTTCTCAACAACAG

P G Q G Q Q G H Y L A S Q Q Q
CCAGGACAAGGGCAACAAGGGCACTACCTAGCTTCTCAGCAGCAG

P G Q G Q Q R H Y P A S L Q Q
CCAGGACAAGGGCAACAACGGCACTACCCAGCTTCTCTGCAGCAA

P G Q G Q Q G H Y T A S L Q Q
CCAGGACAAGGGCAACAAGGGCATTACACAGCTTCTCTGCAGCAA

P G Q G Q Q G H Y P A S L Q Q
CCAGGACAAGGGCAACAAGGGCATTACCCAGCTTCTCTGCAGCAG

1Dy10

P G Q G Q Q G H Y L A S Q Q Q
CCAGGACAAGGGCAACAAGGGCATTACCTAGCTTCTCAGCAGCAG

P G Q G Q Q G H Y P A S Q Q Q
CCAGGACAAGGGCAACAAGGGCACTACCCAGCTTCTCAGCAGCAG

P G Q G Q Q G H Y P A S Q Q Q
CCAGGACAAGGGCAACAAGGGCACTACCCAGCTTCTCAGCAGCAG

P G Q G Q Q G H Y P A S Q Q E
CCAGGACAAGGGCAACAAGGGCACTACCCAGCTTCTCAGCAAGAG

P G Q G Q Q G Q I P A S Q Q Q
CCAGGACAAGGGCAACAAGGGCAAATCCCAGCTTCTCAGCAGCAG

P G Q G Q Q G H Y P A S L Q Q
CCAGGACAAGGGCAACAAGGGCACTACCCAGCTTCTCTGCAGCAA

P G Q G Q Q G H Y P T S L Q Q
CCAGGACAAGGGCAACAAGGGCATTACCCAACTTCTCTACAGCAG

C

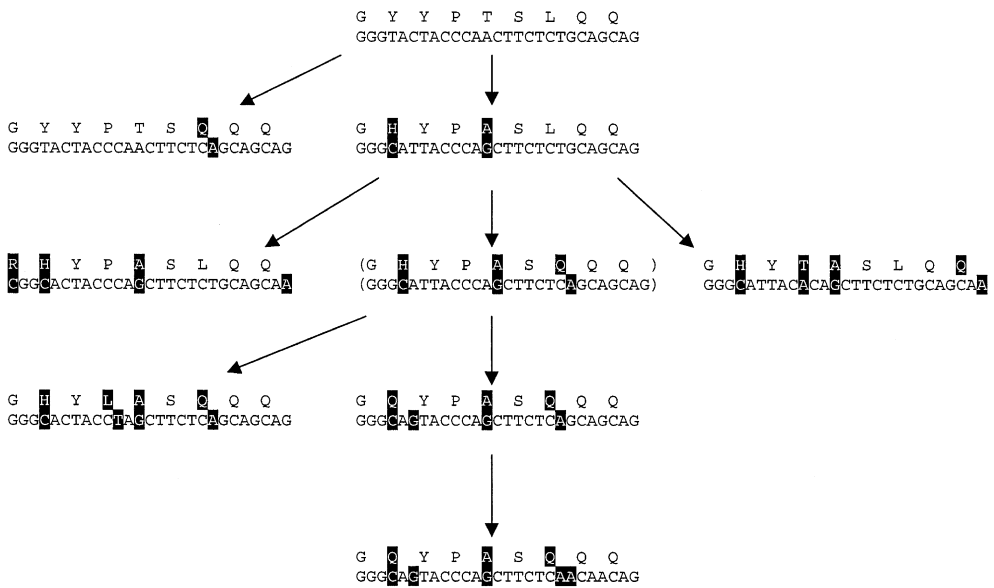


Figure 3.17. (A) Consensus sequence for 15-amino acid repeats comprising hexapeptide and nonapeptide motifs in y-type HMW subunits. (B) Blocks of variant 15-amino acid repeats in HMW subunits 1By9 and 1Dy10. Mutations and amino acid substitutions are highlighted. (C) Simplest evolutionary path for the production of the block of variant repeats shown in (B) for HMW subunit 1By9. The process may have involved mutation, the duplication of variant repeats, further mutation of some of the repeats in the block, followed by further duplication and mutation of some repeats. Alternatively, consensus repeats may have been duplicated to produce the block, and then mutations occurred and spread through adjacent repeats by gene conversion. It is likely that both processes have occurred.

Table 3.4. Cloned Low Molecular Weight Subunit Genes[a]

Locus	Species	Accession numbers
Glu-A3	*Triticum aestivum*	X84959
	Triticum durum	AJ293097
	Triticum durum	AJ293098
	Triticum durum	AJ293099
	Triticum monococcum	AF072898
Glu-A3 (silent)	*Triticum monococcum*	AF073525
Glu-B3	*Triticum aestivum*	X84960, Y17845
	Triticum durum	**Y14104**
Glu-D3	*Triticum aestivum*	**X13306**
		X84961
Unknown	*Triticum aestivum*	AB007763, AB007764, AB008497, **M11077**,[b] U86025, **U86026, U86027, U86028, U86029, U86030, X07747**, X84887
	Triticum durum	AJ007746, **X51759**, **X62588**, **Y18159**

[a]Entries containing full-length coding sequences are in boldface.
[b]This gene has been classified as a γ-gliadin gene (Okita *et al*., 1985), but undoubtedly encodes an LMW subunit with a Cys-to-Ser mutation at the C-terminal end. The encoded protein is likely to be monomeric.

Stretches of multiple glutamine residues characterize all the proteins and are present in the terminal as well as the repetitive domains. There are also similarities between the repetitive domains of all four classes of protein, for example, the sequence Pro-Tyr/Phe-Pro-Gln. This suggests that they are all related. The α- and γ-gliadins and the LMW subunits, however, are much more similar to each other than to the ω-gliadins.

The repeat structures of the gliadins and LMW subunits are not so well defined as that of the HMW subunits. While motifs such as that discussed above can be identified, they are not present in the ordered "stacks" evident in Fig. 3.15, so we have not attempted to discern the evolutionary processes that may have given rise to them. The presence of multiple glutamine residues in the nonrepetitive domains also complicates any analysis. However, an alignment of the C-terminal 53 residues of LMW subunits with corresponding regions from α- and γ-gliadins (Fig. 3.18) shows clearly how closely related they are (this region is not present in ω-gliadins). A prediction of the evolutionary relationships between and within the three classes is shown in Fig. 3.19. It indicates that the α- and γ-gliadins are slightly more closely related to each other than to the LMW subunits. The LMW subunits fall into four clusters and the γ-gliadins into three.

Table 3.5. Cloned Gliadin Genes[a]

Type	Species	Accession numbers
α-Gliadin	*Triticum aestivum*	AJ133602, AJ133603, AJ133604, AJ133605, AJ133606, AJ133607, AJ133608, AJ133609, AJ133610, AJ133611, AJ133612, **D84341**, K03074, K03075, K03076, M10092, M11073, M11074, M11075, M11076, **U08287**, **U50984**, **U51303**, **U51304**, **U51306**, **U51307**, X02538, X02539, X02540, X17361
	Triticum spelta	AJ130948
	Triticum urartu	M16496
Silent α-gliadin	*Triticum aestivum*	U51302, **U51305**, U51308, U51309, U51310, X54517, X54688, X54689
	Triticum urartu	M16497
γ-Gliadin	*Triticum aestivum*	**AF144104**, AF177532, AF234642, AF234643, AF234644, AF234645, AF234646, AF234647, AF234648, AF234649, AF234650, AF234651, **AJ133613**, AJ389694, AJ389693, AJ389692, AJ389691, AJ389690, AJ389689, AJ389668, AJ389667, AJ389666, **D78183**, J01309, M11335, M11336, **M13713**, M16060, **M16064**, **M36999**, X53412
	Triticum compactum	AJ389669, AJ389670
	Triticum dicoccoides	AJ389707, AJ389708
	Triticum dicoccum	AJ389705, AJ389706
	Triticum durum	AJ389699, AJ389700, AJ389701, AJ389702, AJ389703, AJ389704
	Triticum macha	AJ389671, AJ389672, AJ389673
	Triticum spelta	AF120267, AJ389674, AJ389695, AJ389696, AJ389697, AJ389698
	Triticum sphaerococcum	AJ389675, AJ389676
	Triticum timopheevii	AJ389709, AJ389710
	Triticum urartu	AJ389721, AJ389722
	Triticum vavilovii	AJ389677, AJ389678
	Triticum bicornis	AJ38711, AJ389712
	Triticum longissima	AJ389713, AJ389714
	Triticum searsii	AJ389715, AJ389716
	Triticum sharonensis	AJ389718, AJ389717
	Triticum speltoides	AJ389719, AJ389720
	Triticum squarrosa	AJ389679, AJ389680, AJ389681, AJ389682, AJ389683, AJ389684, AJ389685, AJ389686, AJ389687, AJ389688
Silent γ-gliadin	*Triticum aestivum*	**AF234648**, **M13712**
ω-Gliadin	*Triticum aestivum*	**AF280605**
Silent ω-gliadin	*Triticum aestivum*	**AF280606**

[a]Entries containing full-length coding sequences are in boldface.

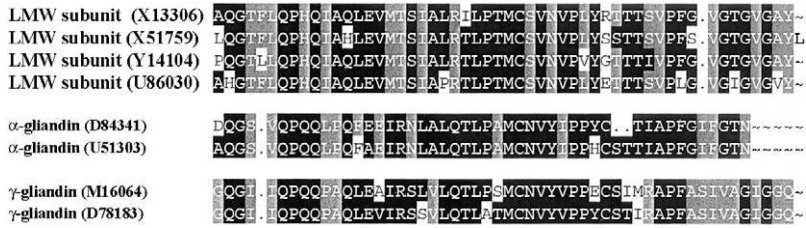


Figure 3.18. Alignment of the C-terminal amino acid sequences of LMW subunits and α- and γ-gliadins. The alignment was produced using the PILEUP program with a gap creation penalty of 3.0 and a gap extension penalty of 0.1 (Wisconsin Package version 10; Genetics Computer Group, Madison, WI). Residues that are conserved in all three gene families are highlighted in black; those conserved in one family but not the others are highlighted in gray. Database accession numbers for the genes are shown in parentheses.

It is important to note that the terms "gliadin" and "LMW subunit" are used according to whether a protein is present in polymers or as a monomer. Similarly, the names assigned to the proteins encoded by the genes that have been sequenced are based on predictions of whether the proteins exist in monomeric or polymeric forms. This depends on the number and distribution of cysteine residues. One of the first gene sequences identified as a γ-gliadin gene to be reported (Okita *et al.*, 1985; accession number M11077) actually encodes a protein that clusters with the LMW subunits in Fig. 3.19. However, it contains a Ser-to-Cys mutation that makes its cysteine distribution resemble that of the gliadins and it is, therefore, presumed to be monomeric. Similarly, some ω-gliadins form aggregates and are called D-type LMW subunits. While the classification of the proteins into gliadins and LMW glutenins is undoubtedly useful for protein chemists, it sits uneasily with the genetics.

C. Gene expression and regulatory sequences

The expression of prolamin genes is controlled primarily at the transcriptional level (Bartels and Thompson, 1986; Sørensen *et al.*, 1989). It is subject to tissue-specific and developmental regulation, the genes being expressed exclusively in the starchy endosperm during mid and late development, and nutritional regulation, responding sensitively to the availability of nitrogen and sulfur in the grain (Duffus and Cochrane, 1992; Giese and Hopp, 1984).

Since they show similar patterns of expression, it is to be expected that prolamin genes would have regulatory sequences in common and this is true to some extent. Schematics of gliadin, LMW subunit, and HMW subunit gene promoters including known and putative regulatory sequences are shown in Fig. 3.20.

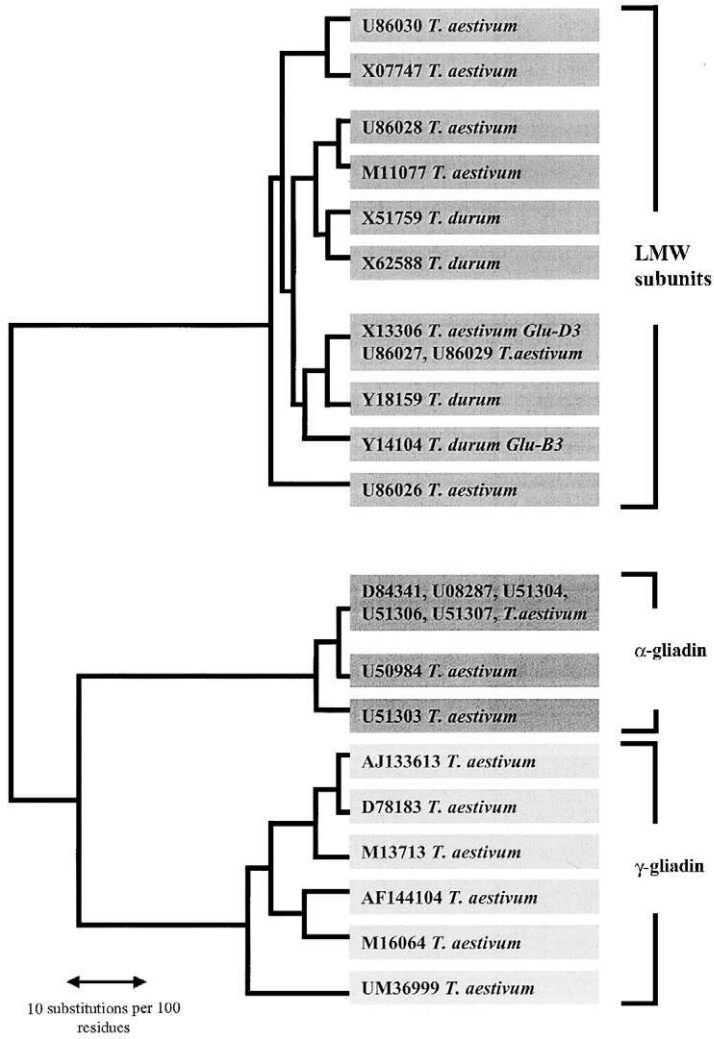


Figure 3.19. Dendrogram showing the relationships between LMW subunits and α- and γ-gliadins. The proteins are identified by the database accession numbers of the genes that encode them. The tree was derived from C-terminal sequence alignments produced using the PILEUP program with a gap creation penalty of 3.0 and a gap extension penalty of 0.1. Evolutionary distances were calculated using the DISTANCES program, correcting for multiple substitutions at a single site, and displayed by the unweighted pair group method using arithmetic averages with the GROWTREE program (Wisconsin Package version 10; Genetics Computer Group, Madison, WI).

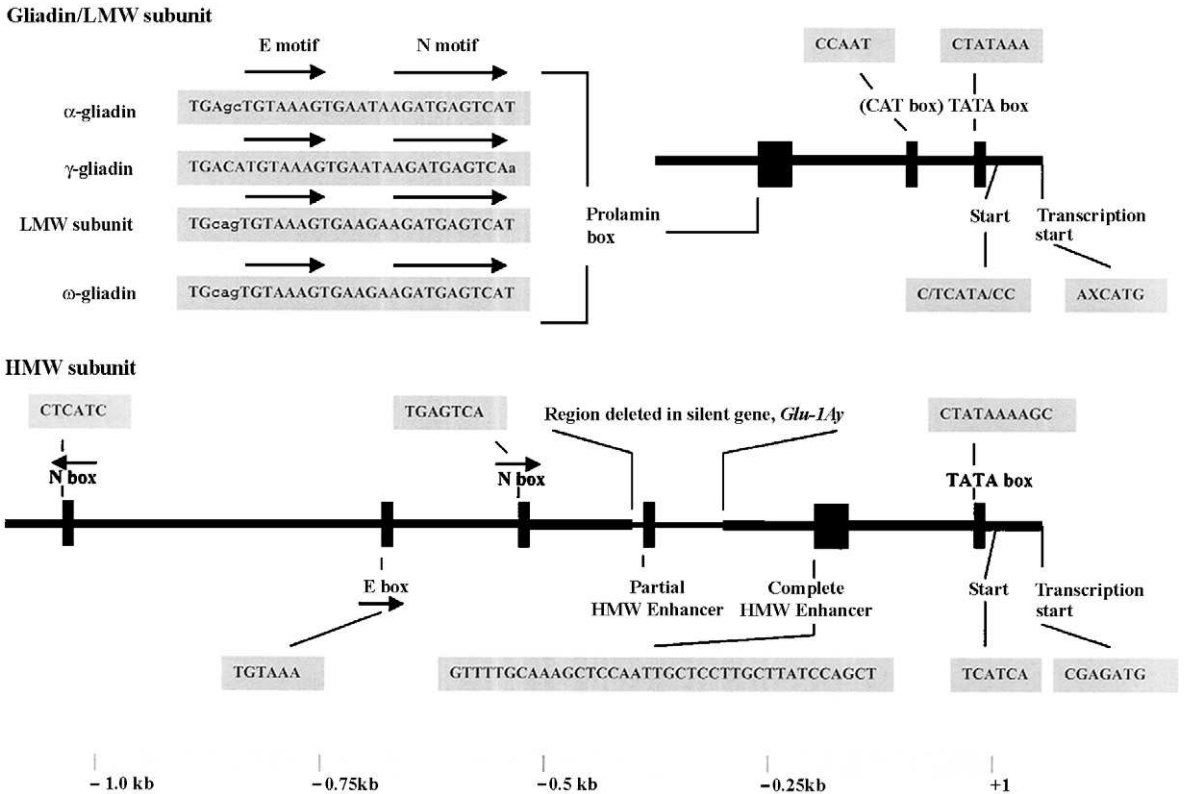


Figure 3.20. Regulatory sequences present in gliadin, LMW subunit, and HMW subunit gene promoters.

1. Proximal regulatory elements

The TATA box sequence of CTATAAA is absolutely conserved in all the prolamin genes. The context of the ATG translation start codon is conserved in the gliadin and LMW subunit genes, with a consensus of AXCATG, while that of the HMW subunit genes has a consensus sequence CGAGATG. The transcription start site has been mapped for one α-gliadin gene and the LMW subunit gene, while the sites shown for the γ- and ω-gliadin genes in Fig. 3.20 are predictions based on sequence comparisons. The consensus shown applies to the gliadin genes, the only LMW subunit gene for which the site has been mapped having a completely different sequence (AAGCAC). This gene contains a CCAAT box at position −124 and *in vivo* footprinting of its promoter showed the binding of a protein adjacent to this site, suggesting that the site has a regulatory function (Hammond-Kosack *et al.*, 1993). The α-gliadin gene contains a CCAAT sequence at position −68. Putative CAT sequences are also present in other prolamin genes, but there is no apparent conservation of sequence or position, so they are not included in Fig. 3.20.

Multiple stop codons and polyadenylation signals are present in many prolamin genes (all HMW subunit genes, for example, contain two translation stop signals in tandem, and almost all contain two polyadenylation signals). The reason for this is not clear, since there is no evidence that the presence of multiple signals increases the efficiency of RNA processing or translation.

2. The prolamin box

The search for regulatory sequences that could confer the tissue-specific, developmental, and nutritional regulation of prolamin gene expression began with a simple sequence comparison of the promoters of several wheat gliadin and barley hordein genes (B. G. Forde *et al.*, 1985). This revealed the presence of a conserved sequence, approximately 30 bp long and positioned around 300 bp upstream of the transcription start site. This was first called the −300 element and subsequently the prolamin box or endosperm element. Its sequence is shown in Fig. 3.20. It contains two conserved motifs, TGTAAAGT and G(A/G)TGAGTCAT, with a more variable region in between. The former has been called the endosperm motif (Hammond-Kosack *et al.*, 1993) (E motif in Fig. 3.20), the latter the GCN4-like motif (GLM), nitrogen element, or N motif (Hammond-Kosack *et al.*, 1993; Müller and Knudsen, 1993). The N motif is similar to the binding site of the GCN4 transcription factor, which is a component of the amino acid signaling pathway in yeast. The E motif is present in the promoters of storage protein genes from a wide range of other species, including some zein genes (Coleman and Larkins, 1999).

The position of the prolamin box is highly conserved, with the first T of the endosperm motif close to position −250. The α-gliadin and LMW subunit gene promoters contain additional complete or partial boxes further upstream in the promoter, but these do not appear to be required for promoter activity and are not present in the γ- or ω-gliadin promoters. The N motif is inverted in S-poor genes from barley (C hordein) and rye (ω-secalin) (Shewry *et al*., 1999) but not in the prolamin box of the ω-gliadin gene.

Functional analyses of a wheat prolamin promoter were first carried out by introducing LMW subunit gene promoter–chloramphenicol acetyltransferase (CAT) reporter gene constructs into transgenic tobacco (Colot *et al*., 1987). These established that the LMW subunit gene promoter conferred the same tissue-specific and developmental regulation of gene expression in tobacco as it did in wheat. The shortest promoter length tested in the study and found to be active was 326 bp. A regulatory role for the prolamin box was established by Müller and Knudsen (1993), using a barley C hordein gene promoter, and by Hammond-Kosack *et al*. (1993), using an LMW subunit promoter. Müller and Knudsen (1993) used a transient expression system involving particle bombardment of cultured barley endosperms with C hordein promoter–β-glucuronidase (GUS) constructs. These experiments also confirmed that the E and N motifs were separate elements, with the N motif acting as a negative element at low nitrogen levels and interacting with the E motif to give high expression when nitrogen levels were adequate.

Hammond-Kosack *et al*. (1993) used *in vivo* footprinting and gel retardation assays to show that the E motif within the prolamin box of an LMW subunit gene bound a putative transcription factor, ESBF-1, during early grain development. The N motif bound a second putative transcription factor, ESBF-II, before maximum expression of the gene. This study was followed by a functional analysis of the prolamin box of this gene in transgenic tobacco, which showed that both motifs were required for seed-specific expression, and the cloning of SPA, a bZIP transcriptional activator that recognized the N motif (Albani *et al*., 1997).

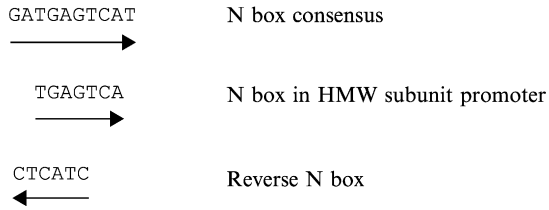
3. The high molecular weight subunit enhancer

Each active HMW subunit gene in bread wheat encodes a protein that accounts for approximately 2% of the total seed protein at maturity (Halford *et al*., 1992). This high level of endosperm-specific expression suggests that HMW subunit promoters have considerable potential for the seed-specific expression of transgenes. High levels of expression have been obtained from additional copies of HMW subunit genes introduced into transgenic wheat (see below and Altpeter *et al*., 1996; Barro *et al*., 1997). Furthermore, an HMW subunit

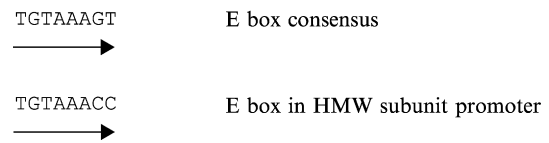
gene promoter has been shown to drive tight, readily detectable, endosperm-specific expression of a GUS reporter gene in transgenic wheat (Lamacchia *et al.*, 2001).

Apart from the high levels of activity, HMW subunit genes show the same patterns of gene expression as gliadin and LMW subunit genes. However, the prolamin box is not present in HMW subunit gene promoters and the HMW subunit enhancer was identified independently. Three studies showed that HMW subunit gene promoters were active in tobacco and gave high levels of endosperm-specific expression of both the HMW glutenin protein itself (Robert *et al.*, 1989) and of CAT and GUS reporter genes (Colot *et al.*, 1987; Halford *et al.*, 1989). Halford *et al.* (1989) found that the region of the promoter of a *Glu-1Dx* gene, from position −277 to +39 with respect to the transcription start site, was sufficient to drive high levels of GUS expression (Halford *et al.*, 1989). The major regulatory element was finally located in a 38-bp sequence, from position −186 to −148 of a *Glu-1Dy* gene, by Thomas and Flavell (1990), again using promoter–GUS reporter gene constructs in transgenic tobacco. The sequence of the element is shown in Fig. 3.20

The palindromic sequence TTTGCAAA within this element is repeated further upstream (shown as a partial HMW enhancer in Fig. 3.20). Interestingly, this sequence is within a region that is deleted in a silent HMW subunit gene, *Glu-1Ay*(J. Forde *et al.*, 1985). However, it is not required for activity of the *Glu-1Dx* promoter in tobacco (Halford *et al.*, 1989). Sequences corresponding to part of the N box and a reverse N box are present further upstream. These sequences align with the consensus N box sequence (Shewry *et al.*, 1999) as follows:



The promoter also contains a sequence that matches part of the E box:



These E and N motifs are not required for activity of the HMW subunit promoter in transgenic tobacco. In the experiments of Halford and co-workers,

higher levels of expression were obtained with a promoter that included the N box than with shorter promoters that contained the HMW enhancer but none of the other motifs, but a statistically significant difference could not be demonstrated (Halford *et al*., 1989).

D. Gene Silencing

Tables 3.3–3.5 include a number of entries from the gliadin and glutenin gene families that are described as silent. Many contain in-frame stop codons and the presence of multiple glutamine-encoding codons, CAA and CAG, probably makes gliadin and glutenin genes particularly prone to mutations that produce the stop codons TAA and TAG. The *Glu-D1x* gene from *A. cylindrica*, for example, contains a CAA (Gln) → TAA (Stop) mutation in codon 343 of the coding region (Wan *et al*., 2002). In-frame stop codons have also been found in *Glu-A1y* and *Glu-A1x* genes (J Forde *et al*., 1985; Bustos *et al*., 2000), resulting from similar CAG (Gln) → TAG (Stop) and CAA (Gln) → TAA (Stop) mutations, respectively, and are also present in silent gliadin and LMW subunit genes (Anderson and Greene, 1997; Rafalski, 1986). However, mutation is not the only way in which the coding regions of wheat prolamins have been disrupted, one *Glu-A1y* allele having been shown to contain a large transposon-like insertion (Harberd *et al*., 1987).

It is difficult to correlate the presence of a stop codon in a gliadin or LMW glutenin gene with the absence of a corresponding full-length protein because there are many proteins of a similar size that may not be separated by SDS–PAGE. The gliadin and LMW subunit proteins listed in Table 3.1 as silent are therefore assumed to be silent because of the presence of an in-frame stop codon in the coding region. This is not the case for HMW subunits, for which the number of expressed genes can be matched with the number of individual HMW subunit proteins identified by SDS–PAGE.

Disruption of the coding region of a gene does not necessarily prevent it from being expressed, particularly if the disruption is a minor one, such as mutation to produce a stop codon. The stop codon may be read through inefficiently or, more likely, a truncated protein may be produced. However, truncated HMW subunit proteins have never been detected in seed protein extracts. Western blot analysis of total proteins extracted from single seeds of *A. cylindrica*, for example, failed to detect any protein of the size predicted for a truncated HMW subunit protein encoded by the *Glu-D1x* gene (Wan *et al*., 2002).

The most likely explanation for this result is that promoters of silent HMW subunit genes are inactive and this has been shown for the promoter of a silent *Glu-A1y* gene linked to a reporter gene and introduced into tobacco

(Halford *et al.*, 1989). However, the basis for the inactivity has never been explained satisfactorily. Although the gene contains a deletion when compared with other HMW subunit gene promoters (Fig. 3.20), this is upstream of the minimal region required for activity of the *Glu-D1x* promoter in tobacco. It also contains all of the regulatory elements shown in Fig. 3.20 with 94% sequence identity to the *Glu-D1x* promoter in the region shown to be required for activity of the latter.

However, *Escherichia coli* cells containing a plasmid designed to express the truncated *Glu-D1x* gene from *A. cylindrica* failed to produce any 1Dx protein or to accumulate any RNA transcribed from the heterologous gene (Wan *et al.*, 2002). In contrast, four other HMW subunit genes expressed well in the same system using the same bacterial promoter. It is possible that the mRNA transcribed from the *Glu-D1x* gene was unstable and this could account for the silent nature of the *Glu-D1x* gene in *A. cylindrica* and the absence of a truncated 1Dx subunit in the grain. However, it would not explain the inability of a promoter from a silent HMW subunit gene to drive expression of a reporter gene.

VI. GLUTEN PROTEINS AND GRAIN QUALITY

A. Gliadins

Electrophoretic separations of gluten proteins allow the presence or absence of specific components to be correlated with functional properties, particularly when combined with the development and utilization of genetic material such as segregating populations, recombinant inbred lines, biotypes, intervarietal chromosome substitution lines, near-isogenic lines, and double haploids.

Sozinov and Poperelya (1980, 1982) analyzed progeny from many different crosses to demonstrate that allelic variation at the *Gli-1* and *Gli-2* loci was associated with variation in breadmaking properties, and also established a ranking order of the different blocks of gliadin components. The contributions of different gliadin components to variation in breadmaking performance has also been reported by others, including Pogna *et al.* (1982), Wrigley *et al.* (1982), Branlard and Dardevet (1985), and Metakovsky *et al.* (1997b,c), with allelic variation present at the *Gli-1* loci being more effective in influencing breadmaking properties compared with the *Gli-2* loci. In most of these studies correlations between gluten components and breadmaking properties of flour have been assessed by the SDS sedimentation test, an indirect test in which good breadmaking performance is associated with the formation of a large sediment when the flour is suspended in water containing SDS and lactic acid. Studies by Moonen *et al.* (1982) and Payne *et al.* (1987c) have demonstrated

that large sedimentation volumes result from the formation of an extensive gel, comprised exclusively of the larger and more insoluble glutenin polymers, leading to the conclusion, supported by genetic studies (Pogna *et al*., 1988, 1990), that the superior quality associated with gliadins present at the *Gli-1* loci is the result of linkage with genes encoding LMW subunits at the *Glu-3* loci.

Similarly, reports that differences in dough strength were correlated with the presence of certain *Gli-2* alleles (Metakovsky *et al*., 1997b; Branlard *et al*., 2001) could be related to the association with C-type LMW glutenin subunits, which have been reported to be closely linked to gliadin components encoded by genes on the group 6 chromosomes (Masci *et al*., 2002).

B. Glutenin Subunits

The establishment of allelic variation in HMW glutenin subunits as a major contributor to genetic variation in breadmaking quality was pioneered by Payne and co-workers with a series of studies starting at the end of the 1970s (Payne *et al*., 1979).

Payne *et al*. (1981b) studied random progeny of various crosses between lines with different HMW subunits and contrasting breadmaking properties, showing that allelic variation in subunits affected breadmaking properties as measured by the SDS sedimentation test. For example, the allelic pairs of HMW subunits 1Dx5 + 1Dy10 and 1Dx2 + 1Dy12 encoded by the *Glu-D1* locus were correlated with good and poor breadmaking properties, respectively. Similarly, differential effects of allelic subunits encoded by the *Glu-B1* and *Glu-A1* loci were also shown. Analyses of many different crosses led Payne *et al*. (1987a) to assign scores to the different HMW subunit alleles, according to their effects on breadmaking properties, with the overall quality score for a variety being calculated by simply adding the scores of the individual subunits present. Analyses of varieties grown in different countries have essentially confirmed the role of allelic difference at the three *Glu-1* loci in determining breadmaking properties (MacRitchie *et al*., 1990).

Subsequently, Payne *et al*. (1987b) used a set of near-isogenic lines developed in the bread wheat cultivar Sicco [HMW subunit composition 1Ax1 (*Glu-A1*), 1Bx7 + 1By9 (*Glu-B1*), 1Dx5 + 1Dy10 (*Glu-D1*)] to show effects associated with the replacement of subunits 1Dx5 + 1Dy10 with the allelic pair 1Dx2 + 1Dy12. SDS sedimentation, Farinograph (Brabender, Duisburg, Germany) and baking tests clearly showed the superior effects on breadmaking of the former pair compared with the latter. Furthermore, near-isogenic lines of Sicco, in which HMW subunits had been eliminated, demonstrated that the removal of subunits 1Dx5 + 1Dy10 had severe effects on the SDS sedimentation volume and loaf volume. These results were similar to those obtained by Lawrence *et al*. (1988), who crossed a mutant line of the cultivar Olympic,

which is null at the *Glu-B1* locus, with a line of the cultivar Gabo, which is null at the *Glu-A1* and *Glu-D1* loci, to generate a series of lines containing from zero to five subunits.

By the middle of the 1980s, the results of extensive genetic and biochemical studies had clearly demonstrated that differences in number and type of HMW glutenin subunits strongly affected the breadmaking properties through effects on the amount and size distribution of glutenin polymers.

Halford *et al.* (1992) carried out quantitative analyses of total protein extracts from 22 bread wheat cultivars and demonstrated that the presence of the allelic subunits 1Ax1 or 1Ax2*, when compared with the null allele, resulted in an increase in the proportion of total HMW subunits from about 8 to 10%. This led them to suggest that the increase in breadmaking quality associated with the presence of subunit 1Ax1 or 1Ax2* may result from an increase in the total proportion of HMW glutenin subunits, which can in turn result in a higher amount of large glutenin polymers. Popineau *et al.* (1994) used the near-isogenic lines developed in the bread wheat cultivar Sicco to show that elimination of HMW subunits associated with the *Glu-A1* and *Glu-D1* loci strongly reduced the amount of the large size glutenin polymers, as measured by size-exclusion chromatography (SE-HPLC), with a marked decrease in gluten viscoelasticity. The substitution of the pair 1Dx5 + 1Dy10, normally present in Sicco, with the pair 1Dx2 + 1Dy12 also negatively affected the size distribution of glutenin polymers and gluten viscoelasticity. These results were paralleled by studies performed by Gupta and MacRitchie (1994), who used recombinant inbred lines and biotypes differing in allelic composition at the *Glu-B1* (1Bx17 + 1By18 vs 1Bx20 + 1By20) or the *Glu-D1* locus (1Dx5 + 1Dy10 vs 1Dx2 + 1Dy12) to show that the allelic pairs 1Bx17 + 1By18 and 1Dx5 + 1Dy10 were associated with the production of larger amounts of large-sized glutenin polymers. No quantitative differences were found between the pairs of allelic combinations tested. However, subunit 1Dx5 contains an additional cysteine residue compared with subunit 1Bx2 while subunit 1Bx20 contains two cysteine residues compared with four in subunit 1Bx17. It has been suggested that these differences may be critical in determining the size distribution of glutenin polymers (Lafiandra *et al.*, 1993b, 1999; Gupta and MacRitchie, 1994).

More recently, doubled-haploid populations have been used to examine in detail the contributions of HMW subunits in influencing breadmaking properties (Kammholz *et al.*, 1998; Killermann and Zimmermann, 2000; Lukow, 2000; Cornish *et al.*, 2001). Killermann and Zimmermann (2000) used four doubled haploid populations differing in glutenin subunits associated at the *Glu-1* and *Glu-3* loci and demonstrated that the effects of the different alleles were strongly dependent on the genetic background of the material. In fact, comparison of the allelic pairs 1Dx5 + 1Dy10 versus 1Dx2 + 1Dy12, and of 1Bx7 + 1By9 versus 1Bx6 + 1By8, consistently showed superior effects of 1Dx5

+ 1Dy10 and 1Bx7 + 1By9, whereas comparison of the null allele with subunit 1 at the *Glu-A1* locus showed significant positive effects only in one population.

The development of new separation techniques for the LMW subunits was crucial in allowing comprehensive analyses to be made of allelic variation at the *Glu-3* loci (Fig. 3.8) of bread wheat, allowing them to be ranked in terms of their effects on dough strength and extensibility (Singh and Shepherd, 1988b; Gupta and MacRitchie, 1991; Gupta *et al*., 1991; Singh *et al*., 1991; Gupta and Shepherd, 1990; Jackson *et al*., 1996; Luo *et al*., 2001; Cornish *et al*., 2001).

As with HMW subunits, the development and wise use of particular genetic stocks has been extremely valuable in establishing correlations between LMW glutenin subunit composition and breadmaking properties. Payne *et al*. (1987c) stressed that there are several disadvantages to using random lines to relate individual gluten proteins to breadmaking properties, the major one being that crossing two parents generates considerable genetic variation and many progeny need to be evaluated. To avoid this they crossed cultivar Chinese Spring with the intervarietal chromosome substitution line CS (cultivar Hope 1A), the latter being identical to cultivar Chinese Spring except that its chromosome 1A had been replaced by the chromosome 1A of the bread wheat cultivar Hope. This particular combination was chosen because the two parents (cultivars Chinese Spring and Hope) differ in their glutenin composition at the *Glu-A1* (null vs 1Ax1) and *Glu-A3* (*Glu-A3a* vs *Glu-A3e*) loci. F_5 progeny derived from the cross by single seed descent were analyzed for protein composition by SDS–PAGE and for quality by SDS sedimentation. These analyses confirmed that progeny containing subunit 1Ax1 from cultivar Hope had much higher sedimentation volumes compared with those with the null allele inherited from cultivar Chinese Spring, agreeing with previous findings (Payne *et al*., 1979). In contrast, lines with the *Glu-A3a* allele from cultivar Chinese Spring had higher sedimentation volumes than those with the *Glu-A3e* allele derived from cultivar Hope. The former allele was shown to be associated with the presence of more LMW subunit protein, leading the authors to conclude that the relative amount of glutenin protein may be an important factor in determining allelic effects on breadmaking quality. In fact, the *Glu-A3e* allele is now known to be null and its deleterious effect on gluten properties has been confirmed by other research groups (Gupta *et al*., 1989). In particular, Gupta and MacRitchie (1994) also showed that the *Glu-A3e* allele in the bread wheat variety Halberd produced no major LMW subunits and that this negatively affected the size distribution of polymeric glutenin.

A set of genetic lines has been developed in the bread wheat cultivar Gabo, in which one, two, or all three short arms of the group 1 chromosomes have been replaced by the short arm of chromosome 1R of rye (Gupta and Shepherd, 1993). The short arm of chromosome 1R carries the *Sec-1* locus

controlling the synthesis of the monomeric secalins but not genes encoding LMW subunit-like proteins (Gupta and Shepherd, 1993). These lines, and in particular the triple translocation stock, served as a useful third parent in test crosses produced to study the inheritance and allelic relationships of LMW subunits (Gupta and Shepherd, 1993). These lines were also used by Gupta *et al.* (1995), together with the lines varying in number of HMW subunits produced by Lawrence *et al.* (1988), to assess the role of glutenin subunits in polymer formation. Results of the analyses of a complete set of genotypes carrying single, double, or triple translocations demonstrated that the effects of progressively eliminating LMW subunits strongly affected the size distribution of polymeric proteins and gluten quality, although the effects are less critical than when all the HMW subunits were deleted. These authors also concluded that although the *Glu-1* and *Glu-3* subunits alone are capable of forming polymers when present by themselves, they also interact to positively affect polymer formation, resulting in higher amounts of large polymers than predicted from their individual contributions. More recently, Cornish *et al.* (2001) analyzed a doubled haploid population and devised a glutenin strength score (GSS) based on different HMW and LMW glutenin subunit alleles present, identifying certain *Glu-3* alleles capable of conferring high extensibility to doughs.

The effects of the C and D groups of LMW subunits on quality have not been studied in detail, although they make a significant contribution to the glutenin polymers (Lew *et al.*, 1992). It has been suggested that, because some of these gliadin-related subunits have an odd number of cysteine residues with only one being likely to participate in intermolecular disulphide bonds, they would act as chain terminators and prevent elongation of developing glutenin polymers (Kasarda, 1989; Lafiandra *et al.*, 1999). As a consequence a negative effect on dough strength, through a tendency to decrease the average molecular weight of the glutenin polymer, is expected. This has proved to be the case for the chromosome 1D-encoded D-type LMW glutenin subunits (Masci *et al.*, 1993, 1999). The development of more appropriate material is required to elucidate the genetics of this group of proteins and their role in determining gluten functional properties.

C. Interactions Between *Glu-1* and *Glu-3* loci

It is clear that *Glu-1* and *Glu-3* loci also interact to affect breadmaking quality. Lorenzo *et al.* (1987) reported that the effects of individual *Glu-1* alleles on the SDS sedimentation test were affected by interactions with the *Glu-1* alleles encoded by the other two genomes. Similarly, Rousset *et al.* (1992) showed additive and epistatic effects of *Glu-1* loci in recombinant inbred lines from a

cross between two bread wheat cultivars that differed in their alleles at all three loci.

Additive and epistatic effects between HMW and LMW subunit loci have also been reported. Payne *et al*. (1987c) analyzed progeny derived from a cross between the cultivar Chinese Spring and the intervarietal chromosome substitution line Chinese Spring (Hope 1A). In addition to confirming the superiority of the *Glu-A1* allele from Hope and of the *Glu-A3* allele from Chinese Spring, the results also showed additive effects of the two superior alleles in some progeny.

Similarly, Gupta and MacRitchie (1994) analyzed a set of recombinant inbred lines for variation at the *Glu-1* and *Glu-3* loci, determining the effects of allelic variation on breadmaking properties. They showed additive and epistatic effects, concluding that interactions between different *Glu-1* loci or between *Glu-1* and *Glu-3* loci contributed to the reported inadequacy of the *Glu-1* score (MacRitchie *et al*., 1990). Rodríguez-Quijano and Carrillo (1996b) similarly analyzed crosses between different bread wheat varieties to show that interactions between *Glu-1* and *Glu-3* loci significantly contributed to variation in gluten strength.

VII. MANIPULATION OF GLUTEN PROTEIN COMPOSITION

There is little doubt that plant breeders have selected for specific combinations of gluten storage proteins over a long period, initially unconsciously as a consequence of selecting for improved processing quality. The identification over the last two decades of alleles at the *Glu-1* and *Glu-3* loci that are positively correlated with good dough properties and breadmaking performance (Payne *et al*., 1987a; Gupta and Shepherd, 1988; Cornish *et al*., 2001) now allows for these to be selected specifically, using electrophoresis to monitor the compositions of progeny from crosses. However, such improvement is still limited by the extent of variation in gluten protein composition of the bread wheat gene pool. Two approaches can be used to expand the range of variation beyond these limits.

A. Exploitation of other wheat species and wild relatives by chromosome engineering

High levels of variation are present in the gluten protein compositions of wild relatives of bread wheat and related diploid and tetraploid cultivated wheats, making them attractive sources of novel genes to improve the breadmaking quality of bread wheat or to confer new properties for novel end uses (Law and

Payne, 1983; Nevo and Payne, 1987; Waines and Payne, 1987; Lagudah and Halloran, 1988; Levy *et al.*, 1988; Ciaffi *et al.*, 1993, 1997; Saponaro *et al.*, 1995; Rodríguez-Quijano *et al.*, 1997; Lee *et al.*, 1999; Gianibelli *et al.*, 2001; Pfluger *et al.*, 2001).

The ease of exploitation of this genetic variation depends on the relationship of the wild relatives with cultivated wheats. Wild species of the sub-tribe Triticinae show different degrees of cytogenetic affinity and phylogenetic relationships between themselves and cultivated wheats, and can be classified into two main groups, depending on whether their genomes are homologous or homoeologous to those of cultivated wheats. The diploid donors of the A (*T. urartu*) and D (*T. tauschii*) genomes of cultivated wheats, and those polyploids that share one or two common genomes [e.g., *T. dicoccoides* (AABB) or *T. timopheevii* (AAGG)] with cultivated wheats, are included in the first group. The second group includes species that are more distantly related to cultivated wheats.

The low level of variation found for glutenins and gliadins encoded by the D genome of bread wheat could be extended by the introduction of allelic variants from its wild diploid progenitor, *T. tauschii* (Gianibelli *et al.*, 2001; Pfluger *et al.*, 2001). This species can be crossed directly with bread wheat, or crossed with durum wheat to generate synthetic hexaploids for indirect gene transfer to bread wheat (Mujeeb-Kazi *et al.*, 1996).

As discussed above, one approach to improving the breadmaking performance of cultivated wheats is to increase the dosage of genes actively expressing HMW glutenin subunits. One way of achieving this would to be replace existing alleles at the *Glu-A1* locus of bread wheat, which are null or express only an x-type subunit, with equivalent loci from *T. urartu*, *T. monococcum* (AA genome), *T. dicoccoides* (AABB) or *T. timopheevii* (subsp. *timopheevii* and subsp. *araraticum*) (AAGG), which express both x- and y-type subunits (Levy *et al.*, 1988; Ciaffi *et al.*, 1993; Margiotta *et al.*, 1998). Rogers *et al.* (1997) demonstrated the positive effect on gluten strength of introducing a *Glu-A1* allele encoding two subunits from *T. thaoudar*. Ciaffi *et al.* (1995) similarly introduced a *Glu-A1* allele expressing both x- and y-type subunits from *T. dicoccoides* into durum wheat, in an attempt to obtain lines with satisfactory breadmaking characteristics and good pasta quality. Some of the lines produced had exceptionally high dough strength and baking performance, as good as those of the bread wheat cultivars used as controls.

When the alien chromosome from wheat relatives is homoeologous and not homologous, other approaches can be exploited to induce recombination and the transfer of useful genes between different genomes (Jauhar and Chibbar, 1999). Thus, chromosome engineering, which involves the transfer of chromosomal segments between wheat and related Triticeae species through manipulation of the homoeologous pairing process (Ceoloni *et al.*, 1993), has

been successfully used to modify the protein composition of wheat. Using this approach, recombination between chromosome 1U of *A. umbellulata* and its wheat homoeologues, present in the bread wheat cultivar Chinese Spring, was achieved (Koebner and Shepherd 1987; Faridi, 1988), leading to replacement of the null allele present at the *Glu-A1* locus of Chinese Spring with the *Glu-U1* locus from *A. umbellulata* (expressing genes for two HMW subunits). This approach has also been explored for transferring homoeoloci present in other cultivated cereals; in particular, Lukaszewski and Curtis (1992, 1994) used chromosome engineering to transfer segments carrying the *Glu-D1* loci encoding the allelic pairs 1Dx5 + 1Dy10 and 1Dx2 + 1Dy12 onto chromosomes 1R and 1A of triticale to improve the breadmaking properties. Similarly, loci encoding subunits 1Dx5 + 1Dy10 and 1Dx2 + 1Dy12 have been introduced into several durum wheat cultivars, replacing the null allele at the *Glu-A1* locus (Ammar *et al*., 1997; Vitellozzi *et al*., 1997; Lafiandra *et al*., 2000b). This approach therefore presents a viable alternative to genetic engineering to manipulate the gluten protein composition to improve end use quality.

B. Introduction of additional or novel genes by genetic engineering

1. Wheat transformation technology

Despite its immense economic importance, wheat was the last of the major cereals to be transformed and, indeed, among the last of the major crops. In fact, the first fertile transgenic wheat plants were reported only a decade ago (Vasil *et al*., 1992; Weeks *et al*., 1993) and wheat transformation is still carried out routinely in only a small number of laboratories worldwide.

One of the early limitations to wheat transformation was transgene delivery, as cereals are outside the natural host range of the soil bacterium *Agrobacterium* which has been widely used to mediate transformation of dicotyledonous species. Vasil *et al.* (1992) used particle bombardment, the DNA being coated on the outside of microscopic gold particles and fired into the cells. This is still the most widely used method for wheat transformation although transformation with *Agrobacterium*-based vectors has been reported (Cheng *et al*., 1997) following success with maize and rice (see Barcelo *et al*., 2001, for a review).

The regeneration of wheat plants from transformed cells or tissues also proved to be problematic, with Vasil *et al.* (1992) and several other groups using a single wheat cultivar, Bob White, which can be regenerated readily from immature scutella (cotyledons). However, improvements in regeneration conditions mean that it is now possible to transform a range of elite wheat cultivars by regeneration of bombarded immature scutella (Rasco-Gaunt and Barcello, 1999; Rasco-Gaunt *et al*., 2000; Pastori *et al*., 2001; Varshney and Altpeter,

2001). A similar protocol has been successfully applied to several cultivars of durum wheat (He *et al.*, 1999), which can also be transformed by bombardment and regeneration of small sections of immature inflorescences (Lamacchia *et al.*, 2001).

The final part of the technology for wheat transformation that must be considered is selectable marker genes. A selectable marker gene is required to allow the selection of transformed cells, by conferring resistance to a chemical that is toxic to untransformed cells or by allowing the utilization of a substrate that will not normally support cell growth. Regeneration without such selection would lead to chimaeric plants with rare transformed cells surrounded by untransformed cells. The breakthrough in wheat transformation was facilitated by the isolation of bacterial genes (*bar* and *pat*) that confer tolerance of two related herbicides, phosphinothricin and bialaphos (phosphinothricin-alanine-alanine). These genes allow more effective selection of transformed cereal tissues than previously used genes, which conferred resistance to antibiotics such as hygromycin.

Although herbicide tolerance genes may be more acceptable for consumers than those for antibiotic resistance, their use will ultimately affect the range of strategies available to farmers. Furthermore, there is concern from environmentalists that genes for herbicide resistance may ultimately spread into wild populations. Consequently, there is currently an active search for more benign marker genes.

2. Application of genetic engineering to the high molecular weight subunits of glutenin

The HMW subunits were among the first targets identified for manipulation in transgenic wheat. One reason for this is clearly their functional importance in determining dough strength. Furthermore, analyses of genotypes differing in the number of expressed genes have indicated a relationship between the number of expressed genes, the amount of HMW subunit, and dough elasticity. Thus, transformation with additional HMW subunit genes should lead to increased dough strength. However, the choice has also almost certainly been influenced by practical considerations. The HMW subunit genes were freely available, with the whole gene family from cv. Cheyenne having been cloned before 1990 (see Shewry *et al.*, 1992). In addition, the proteins encoded by the transgenes are readily identified in the grain, being clearly separated from other gluten proteins on SDS–PAGE (see Fig. 3.2).

The first reported expression of an HMW subunit in transgenic wheat was by Blechl and Anderson (1996), who transformed the cultivar Bob White with a chimaeric gene encoding a hybrid protein (comprising residues 1–124 of

subunit 1Dy10 + resides 130–848 of subunit 1Dx5) in order to resolve clearly the transgene product from the endogenous subunits 1Dx5 and 1Dy10 present in this cultivar. Subsequently, Altpeter *et al.* (1996) expressed the subunit 1Ax1-encoding gene from cv. Hope (Halford *et al.*, 1992), also in cv. Bob White. Twenty-one independent lines were selected, of which 9 expressed the 1Ax1-encoding transgene. The transgenic subunit was clearly visible by SDS–PAGE and densitometric analysis showed that it accounted for 0.6 to 2.3% of the total proteins. Both Altpeter *et al.* (1996) and Blechl and Anderson (1996) also showed that expression of the transgenes was generally stable up to at least three seed generations. Blechl and co-workers have since expressed authentic subunits 1Dx5 and 1Dy10 in the same cultivar (Anderson and Blechl, 2000).

The cultivar Bob White contains the 1BL/1RS chromosome translocation, which is frequently associated with poor processing quality due to the production of sticky dough. Consequently, it is not an ideal background to assess the impact of the HMW subunit transgenes on processing quality. This may account, at least in part, for the failure by Vasil *et al.* (2001) to observe major improvements in mixing and breadmaking properties when lines of cv. Bob White expressing the 1Ax1-encoding gene were grown in the field. Similarly, although Anderson and Blechl (2000) reported that transformation of Bob White to express increased amounts of subunits 1Dx5 and 1Dy10 (both of which are present in cv. Bob White itself) resulted in a substantial increase in the dough mixing time (from about 4 to 17 min), this was accompanied by a large decrease in the peak resistance. In contrast, Barro *et al.* (1997) used two normal (i.e., non-translocation) wheat lines that allowed the measurement of dough mixing properties.

The lines L88-6 and L88-31 form part of a near-isogenic series derived from crosses between mutants of the Australian spring wheat cultivars Olympic and Gabo. These mutants have null (silent or absent) alleles at the *Glu-A1*, *Glu-B1*, and *Glu-D1* loci, allowing the development of a series of lines expressing from zero (triple null line) to five (1Ax1, 1Dx5, 1Dy10, 1Bx17, and 1By18) subunits (Lawrence *et al.*, 1988). L88-6 expresses all five HMW subunits while L88-6 expresses only subunits 1Bx17 and 1By18.

Barro *et al.* (1997) produced several lines of L88-6 and L88-31 expressing transgenes encoding HMW subunits 1Ax1 and/or 1Dx5. Of particular interest was one line in L88-6, called B73-6-1, in which the presence of about 15 additional copies of the 1Dx5-encoding gene resulted in a 4-fold increase in the proportion of the corresponding proteins (Rooke *et al.*, 1999). Consequently, subunit 1Dx5 accounted for more than 10% of the total gluten proteins compared with about 2.7% in the control.

The mixing properties of B73-6-1 and other lines produced by Barro *et al.* (1997) were determined with a 12g Mixograph (National Manufacturing, Lincoln, NE) using material grown in pots and in field trials. The Mixograph

mixes dough, using a system of moving pins, and records the resistance. A range of parameters relating to dough mixing can be measured from the Mixograph curve, the most widely used being the maximum resistance (peak resistance, PR), the time to maximum resistance (mixing time, MT), and the rate of decrease in resistance after the peak (resistance breakdown, RBD). Analysis of flours from the field-grown material showed that the expression of subunit 1Ax1 in the L88-31 background (line B102-1-2) resulted in increased dough strength (Fig. 3.21) and this was associated with an increase in "soluble" polymers (i.e., those extractable by sonication with 2% SDS) (Popineau *et al.*, 2001). In contrast, expression of subunit 1Dx5 had similar but unexpected effects in both the L88-6 (B73-6-1) and L88-31 (B72-8-11b) backgrounds (Fig. 3.21). Flour from these lines failed to form a cohesive dough during hydration and this was associated with an increase in "insoluble" glutenin polymers (Popineau *et al.*, 2001).

Similar effects were observed when subunits 1Ax1 and 1Dx5 were expressed in transgenic durum wheat, with low expression levels of both subunits giving increased dough strength but high levels of 1Dx5 resulting in altered dough characteristics (He *et al.*, 1999).

The properties of transgenic bread and durum wheats expressing high levels of subunit 1Dx5 resembled, to some extent, those of "overstrong" bread wheats including the cultivars Glenlea and Red River, in which duplication of HMW subunit genes may have occurred (D'Ovidio *et al.*, 1997; Lukow *et al.*, 1992). Furthermore, blending of flour from either B73-6-1 or the transgenic durum line with flour from the bread wheat cultivar Banks resulted in increased mixing times up to about 50–60%, above which decreases were observed (He *et al.*, 1999; Roone *et al.*, 1999). In contrast, incorporation of as little as 10% of B73-6-1 resulted in a decrease in the loaf volume obtained with the UK cultivar Hereward in breadmaking tests and when incorporated at 50% the loaf volume was decreased by approximately 25% (unpublished results of P. R. Shewry *et al.*).

Subunit 1Dx5 is always present in bread wheat with subunit 1Dy10 and there is evidence that dimers of x-type and y-type subunits, including 1Dx5 + 1Dy10, occur as structural units in glutenin polymers (Lawrence and Payne, 1983; Werner *et al.*, 1992). Consequently, the unusual mixing properties resulting from the expression of high levels of subunit 1Dx5 without equivalent increases in subunit 1Dy10 could result from fundamental changes in glutenin structure that are required to accommodate the imbalance in subunit composition. In support of this, Butow *et al.* (2000) showed that the incorporation of additional subunit 1Dy10, purified from wheat, into dough of B73-6-1, using a reduction–re-oxidation procedure, resulted in increased mixing time when the flour was blended with that of a normal cultivar up to 100%.

A further factor that could contribute to the effects of high-level expression of subunit 1Dx5 is increased cross-linking due to the presence of an

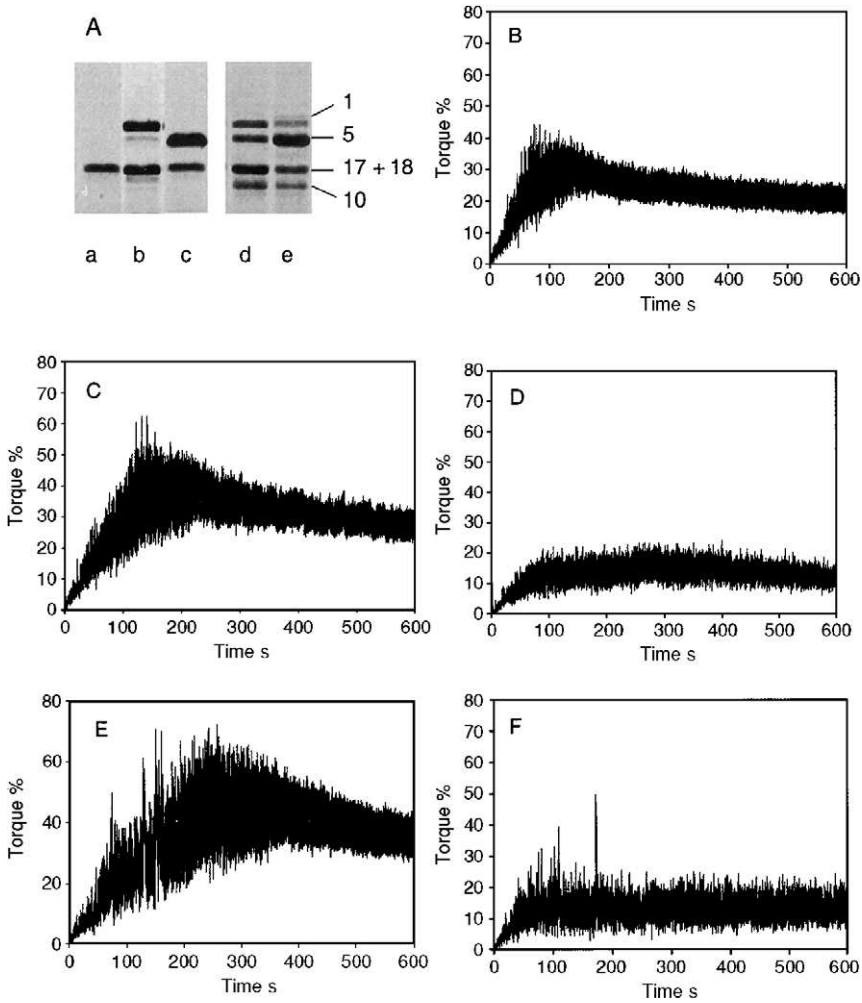


Figure 3.21. Analysis of the mixing properties of transgenic wheats expressing additional HMW subunits, using the 2g Mixograph. (**A**) SDS–PAGE of the HMW subunits from L88-31: control line (1A null, 1Bx17 + 1By18, 1D null) (lane a); B72-8-11b: transformed line expressing 1Dx5 subunit transgene in the L88-31 background (lane b); B102-1-2: transformed line expressing 1Ax1 subunit transgene in the L88-31 background (lane c); L88-6: control line (1Ax1, 1Bx17 + 1By18, 1Dx5 + 1Dy10) (lane d); B73-6-1: transformed line expressing 1Dx5 subunit transgene in the L88-6 background (lane e). (**B–F**) Mixographs of (**B**) L88-31; (**C**) B102-1-2; (**D**) B72-8-11b; (**E**) L88-6; and (**F**) B73-6-1. The resistance is given as torque percent and the mixing time is given in seconds. Reprinted with permission from Popineau *et al.* (2001).

additional cysteine residue within the repetitive domain of this protein compared with other HMW subunits. This explanation is consistent with studies by Popineau *et al.* (2001), who demonstrated that the rheological properties of gluten from B73-6-1 and B72-8-11b were similar to those of gluten fractions from the near-isogenic line Sicco, in which additional cross-links had been introduced by the enzyme transglutaminase (Larré *et al.*, 2000).

Alvarez *et al.* (2000) reported the first successful expression of HMW subunit transgenes (1Ax1 and 1Dx5) in a commercial cultivar (the Argentinean cv. Pro INTA Federal) and showed that one line expressing the 1Dx5 transgene had a twofold increase in mixing time (Alvarez *et al.*, 2001) while Pastori *et al.* (2000) reported the expression of the 1Ax1 transgene in the UK cultivars Cadenza and Canon.

Transgene silencing is a well-established phenomenon in plants that can result from transcriptional or posttranscriptional events (Vaucheret *et al.*, 1998). Blechl *et al.* (1998) and Alvarez *et al.* (2000) have both reported gene silencing related to transformation with HMW subunit genes. Blechl *et al.* (1998) reported that expression of the chimeric 1Dy10/1Dx5 transgene resulted in the silencing of endogenous HMW subunit genes while Alvarez *et al.* (2000) reported silencing of the endogenous 1Ax2* gene by the 1Ax1 transgene and silencing of all subunits by the 1Ax1 or 1Dx5 transgene.

3. Application of genetic engineering to other gluten proteins

The transformation of durum (pasta) wheat with additional genes for LMW subunits is of particular interest because of the correlation between quality and specific LMW subunit alleles (discussed above).

Tosi *et al.* (2000) reported the expression of mutant and wild-type LMW subunits in the Italian durum wheat cultivars Svevo and Ofanto, using PCR products amplified from the durum cv. Langdon with an HMW subunit gene promoter from bread wheat to drive expression. In addition, a C-terminal epitope tag was used to facilitate identification of the protein in total protein extracts (Fig. 3.22) and in the cells of the developing grain. The tagged subunit was incorporated into glutenin polymers, showing a distribution similar to the endogenous LMW subunits. However, analysis of two lines expressing the same tagged subunit showed different effects on the dough properties measured with a 2g Mixograph (P. Tosi, F. Békés, and P. R. Shewry, unpublished results). Whereas one line showed the expected increase in dough strength, reduced expression of endogenous subunits in the second line was associated with weakening of the dough.

D'Ovidio *et al.* (2000) also reported the identification of one line of bread wheat (cv. Bob White) expressing an LMW subunit genomic clone from

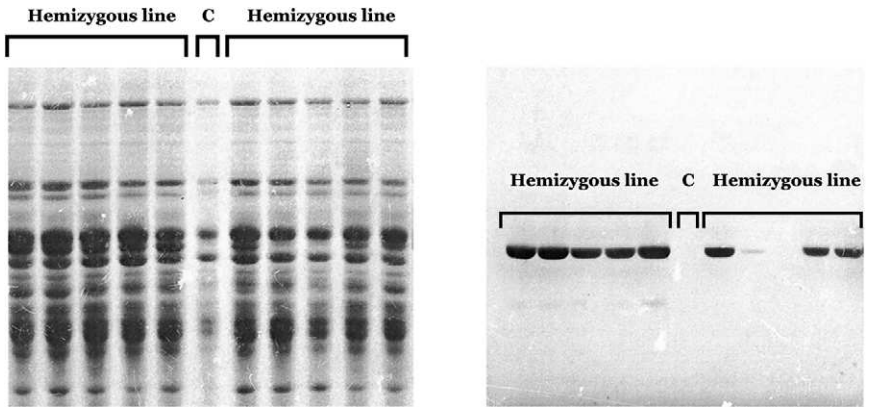


Figure 3.22. SDS–PAGE (A) and Western blot analysis (B) of T_2 seeds from two transgenic lines of cv. Ofanto expressing the LMW subunit gene pRDPT$_5$1B*. Reprinted with permission from Tosi *et al.* (2000).

bread wheat cv. Cheyenne but studies of the functional properties have not been reported.

VIII. CONCLUSIONS

The importance of wheat in food processing and human nutrition has resulted in a massive volume of research on the crop and grain. In particular, attention has focused on gluten proteins in order to explain their role in determining grain-processing properties and to facilitate improvement of end use quality. These studies have resulted in remarkable increases over the last two decades in our knowledge of wheat gluten protein genes and the proteins that they encode. Furthermore, the demonstration of "associations" between specific storage proteins and quality parameters has led to practical applications in crop improvement, exemplified by the routine selection of specific patterns of HMW subunits of glutenin in plant-breeding programs. Such applications do not require any detailed understanding of the molecular basis for the role of gluten proteins in determining end use quality, but the associations need to be rigorous.

A detailed understanding of the molecular basis for end use quality is, however, important for future improvement, in two respects. First, it will facilitate the identification of novel "quality-associated" proteins, and the genes that encode them, by defining important features such as protein size and ability to interact via covalent disulphide bonds or noncovalent hydrogen bonding. Second, it is now becoming possible to use this knowledge to design genes for "ideal" proteins and introduce them into wheat by genetic engineering technology.

This bringing together of basic and strategic studies is an excellent illustration of the importance of continuing to support research on the genetics of crop species, which runs counter to the current trend to focus resources on studies of model organisms.

Acknowledgments

Long Ashton Research Station receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

References

Albani, D., Hammond-Kosack, M. C. U., Smith, C., Conlan, S., Colot, V., Holdsworth, M., and Bevan, M. W. (1997). The wheat transcriptional activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell* **9,** 171–184.

Altpeter, F., Vasil, V., Srivastava, V., and Vasil, I. K. (1996). Integration and expression of the high-molecular-weight glutenin subunit 1Ax1 gene into wheat. *Nat. Biotechnol.* **14,** 1155–1159.

Alvarez, M. L., Guelman, S., Halford, N. G., Lustig, S., Reggiardo, M. I., Ryabushkina, N., Shewry, P. R., Stein, J., and Vallejos, R. H. (2000). Silencing of HMW glutenins in transgenic wheat expressing extra HMW subunits. *Theor. Appl. Genet.* **100,** 319–327.

Alvarez, M. L., Gómez, M., Carrillo, J. M., and Vallejos, R. H. (2001). Analysis of dough functionality of flours from transgenic wheat. *Mol. Breeding.* **8,** 103–108.

Ammar, K., Lukaszewski, A. J., and Banowetz, G. M. (1997). Effect of *Glu-D1*(5+10) on gluten strength and polymeric protein composition in durum wheat. *Cereal Foods World* **42,** 610.

Anderson, O. D., and Blechl, A. E. (2000). Transgenic wheat: Challenges and opportunities. *In* "Transgenic Cereals" (L. O'Brien and R. Henry, eds.), pp. 1–27. American Association of Cereals Chemists St. Paul, MN.

Anderson, O. D., and Greene, F. C. (1989). The characterization and comparative analysis of HMW glutenin genes from genomes A and B of hexaploid wheat. *Theor. Appl. Genet.* **77,** 689–700.

Anderson, O. D., and Greene, F. C. (1997). The α-gliadin gene family. II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes. *Theor. Appl. Genet.* **95,** 59–65.

Anderson, O. D., Greene, F. C., Yip, R. E., Halford, N. G., Shewry, P. R., and Malpica-Romero, J.-M. (1989). Nucleotide sequences of the two high molecular weight glutenin genes from the D-genome of a hexaploid bread wheat, *T. aestivum* L. cv. Cheyenne. *Nucleic Acids Res.* **17,** 461–462.

Anderson, O. D., Litts, J. C., and Greene, F. C. (1997). The α-gliadin gene family. I. Characterization of ten new wheat α-gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and Southern analysis of the gene family. *Theor. Appl. Genet.* **95,** 50–58.

Anderson, O. D., Hsia, C. C., Adalsteins, A. E., Lew, E. J.-L., and Kasarda, D. D. (2001a). Identification of several new classes of low-molecular-weight wheat gliadin related proteins and genes. *Theor. Appl. Genet.* **103,** 307–315.

Anderson, O. D., Hsia, C. C., and Torres, V. (2001b). The wheat γ-gliadin genes: Characterization of ten new sequences and further understanding of γ-gliadin gene family structure. *Theor. Appl. Genet.* **103,** 323–330.

Barcelo, P., Rasco-Gaunt, S., Thorpe, C., and Lazzeri, P. A. (2001). Transformation and gene expression. *In* "Advances in Botanical Research," pp. 59–126. Academic Press, London.

Barnes, P. J. (1989). Wheat in milling and baking. *In* "Cereal Science and Technology" (G. H. Palmer, ed.), pp. 367–412. Aberdeen University Press, Aberdeen, UK.

Barro, F., Rooke, L., Békés, F., Gras, P., Tatham, A. S., Fido, R., Lazzeri, P. A., Shewry, P. R., and Barceló, P. (1997). Transformation of wheat with high molecular weight subunit genes results in improved functional properties. *Nat. Biotechnol.* **15,** 1295–1299.

Bartels, D., and Thompson, R. D. (1986). Synthesis of messenger-RNAs coding for abundant endosperm proteins during wheat grain development. *Plant Sci.* **46,** 117–125.

Bean, S. R., and Lookhart, G. L. (2000). Electrophoresis of cereal storage proteins. *J. Chromatogr. A* **881,** 23–36.

Bean, S. R., Bietz, J. A., and Lookhart, G. L. (1998). High-performance capillary electrophoresis of cereal proteins. *J. Chromatogr. A.* **814,** 25–41.

Beccari, J. B. (1745). De Frumento. *In* "De Bononiensi Scientiarum et Artium atque Academia Commentarii," Tomi Secundi, Part I, pp. 122–127.

Bietz, J. A., Shepherd, K. W., and Wall, J. S. (1975). Single-kernel analysis of glutenin: Use in wheat genetics and breeding. *Cereal Chem.* **52,** 513–532.

Blanco, A., Resta, P., Simeone, R., Parmar, S., Shewry, P. R., Sabelli, P., and Lafiandra, D. (1991). Chromosomal location of seed storage protein genes in the genome of *Dasypirum villosum* (L.) Candargy. *Theor. Appl. Genet.* **82,** 358–362.

Blechl, A. E., and Anderson, O. D. (1996). Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. *Nat. Biotechnol.* **14,** 875–879.

Blechl, A. E., Le, H. Q., and Anderson, O. D. (1998). Engineering changes in wheat flour by genetic transformation. *J. Plant Physiol.* **152,** 703–707.

Boyd, W. J. R., and Lee, J. W. (1967). The control of wheat gluten synthesis at the genome and chromosome levels. *Experientia* **23,** 332–333.

Branlard, G., and Darvedet, M. (1985). Diversity of grain proteins and bread wheat quality. I. Correlations between gliadin bands and flour quality characteristics. *J. Cereal Sci.* **3,** 329–343.

Branlard, G., Darvedet, M., Saccomano, R., Lagoutte, F., and Gourdon, J. (2001). Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica* **119,** 59–67.

Brown, J. W. S., Law, C. N., Worland, A. J., and Flavell, R. B. (1981). Genetic variation in wheat endosperm proteins: An analysis by two-dimensional electrophoresis using intervarietal chromosomal substitution lines. *Theor. Appl. Genet.* **59,** 361–371.

Bustos, A. D., Rubio, P., and Jouve, N. (2000). Molecular characterisation of the inactive allele of the gene *Glu-A1* and the development of a set of AS-PCR markers for HMW glutenins of wheat. *Theor. Appl. Genet.* **100,** 1085–1094.

Butow, B. J., Rooke, L., Tatham, A. S., Shewry, P. R., Savage, A. W. J., Darlington, H., Larroque, O., Murray, D., and Békés, F. (2000). The effect of HMW-GS subunit *x*/*y* ratios on the functional properties of doughs from transgenic and non-transgenic wheat. *In* "Proceedings of the 11th ICC Cereal and Bread Congress and of the 50th Australian Cereal Chemistry Conference" (M. Wootton, I. L. Batey and C. W. Wrigley, eds.), pp. 391–395. Royal Australian Chemical Institute, Melbourne, Australia.

Cassidy, B. G., Dvorak, J., and Anderson, O. D. (1998). The wheat low-molecular-weight glutenin genes: Characterization of six new genes and progress in understanding gene family structure. *Theor. Appl. Genet.* **96,** 743–750.

Ceoloni, C., Ciaffi, M., Lafiandra, D., and Giorgi, B. (1993). Chromosome engineering as a means of transferring 1D storage protein genes from common to durum wheat. *In* "Proceedings of the 8th International Wheat Genetics Symposium" (Z. S. Li and Z. Y. Xin, eds.), pp. 159–163. China Agriculture Scientech Press, Beijing, China.

Cheng, M., Fry, J. E., Pang, S., Zhou, H., Hironaka, C. M., Duncan, D. R., Conner, T. W., and Wan, Y. (1997). Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. *Plant Physiol.* **115,** 971–980.

Ciaffi, M., Lafiandra, D., Porceddu, E., and Benedettelli, S. (1993). Storage protein variation in wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) from Jordan and Turkey. I. Electrophoretic characterization of genotypes. *Theor. Appl. Genet.* **86,** 474–480.

Ciaffi, M., Lafiandra, D., Turchetta, T., Ravaglia, S., Bariana, H., Gupta, R., and MacRitchie, F. (1995). Breadmaking potential of durum wheat lines expressing both *x* and *y* type subunits at the *Glu-A1* locus. *Cereal Chem.* **72,** 465–469.

Ciaffi, M., Dominici, L., and Lafiandra, D. (1997). Gliadin polymorphism in wild and cultivated einkorn wheats. *Theor. Appl. Genet.* **94,** 68–74.

Coleman, C. E., and Larkins, B. A. (1999). The prolamins of maize. *In* "Seed Proteins" (R. Casey and P. R. Shewry, eds.), pp. 109–140. Kluwer Academic Press, Dordrecht, The Netherlands.

Colot, V., Robert, L. S., Kavanagh, T. A., Bevan, M. W., and Thompson, R. D. (1987). Localization of sequences in wheat endosperm protein genes which confer tissue-specific expression in tobacco. *EMBO J.* **6,** 3559–3564.

Colot, V., Bartels, D., Thompson, R., and Flavell, R. (1989). Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin genes. *Mol. Gen. Genet.* **216,** 81–90.

Cornish, G. B., Békés, F., Allen, H. M., and Martin, J. M. (2001). Flour proteins linked to quality traits in an Australian doubled haploid wheat population. *Aust. J. Agric. Res.* **52,** 1339–1348.

Curtis, C. A., and Lukaszewski, A. J. (1991). Genetic linkage between C-bands and storage protein genes in chromosome 1B of tetraploid wheat. *Theor. Appl. Genet.* **81,** 245–252.

Dachkevitch, T., Redaelli, R., Biancardi, A. M., Metakovsky, E. V., and Pogna, N. E. (1993). Genetics of gliadins coded by the group 1 chromosomes in the high quality bread wheat cultivar Neepawa. *Theor. Appl. Genet.* **86,** 389–399.

D'Ovidio, R., Tanzarella, O. A., and Porceddu, E. (1992). Isolation of an α-type gliadin gene from *Triticum durum* Desf. and genetic polymorphism at the *Gli-2* loci. *J. Genet. Breeding.* **46,** 41–48.

D'Ovidio, R., Masci, S., and Porceddu, E. (1995a). Development of a set of oligonucleotide primers specific for genes at the *Glu-1* complex loci of wheat. *Theor. Appl. Genet.* **91,** 189–194.

D'Ovidio, R., Simeone, M., Masci, S., Porceddu, E., and Kasarda, D. D. (1995b). Nucleotide sequence of a γ-type glutenin gene from durum wheat: Correlation with a γ-type glutenin subunit from the same biotype. *Cereal Chem.* **72,** 443–449.

D'Ovidio, R., Lafiandra, D., and Porceddu, E. (1996). Identification and molecular characterization of a large insertion within the repetitive domain of a high-molecular-weight glutenin subunit gene from hexaploid wheat. *Theor. Appl. Genet.* **93,** 1048–1053.

D'Ovidio, R., Masci, S., Porceddu, E., and Kasarda, D. D. (1997). Duplication of the *Bx7* high-molecular-weight glutenin subunit gene in bread wheat (*Triticum aestivum* L.) cultivar 'Red River 68'. *Plant Breeding* **116,** 525–531.

D'Ovidio, R., Fabbri, R., Patacchini, C., Masci, S., Lafiandra, D., Porceddu, E., Blechl, A. E., and Anderson, O. D. (2000). Production of transgenic bread wheat lines over-expressing a LMW glutenin subunit. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 101–104. Royal Society of Chemistry, Cambridge.

Duffus, C. M., and Cochrane, M. P. (1992). Grain structure and composition. *In* "Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology" (P. R. Shewry, ed.), pp. 291–317. CAB International, Wallingford, UK.

DuPont, F. M., Vensel, W. H., Chan, R., and Kasarda, D. D. (2000). Characterization of the 1B-type ω-gliadins from *Triticum aestivum* cultivar Butte. *Cereal Chem.* **77,** 607–614.

Dworschak, R. G., Werner, E., Standing, K. G., Preston, K. R., Marchylo, B. A., Nightingale, M. J., Stevenson, S. G., and Hatcher, D. W. (1998). Analysis of wheat gluten proteins by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **33,** 429–435.

Faridi, I. M. N. (1988). "Genetical Studies of Grain Protein and Developmental Characters in Wheat." Ph.D. thesis, University of Cambridge, Cambridge.

Feeney, K. A., Tatham, A. S., Gilbert, S. M., Fido, R. J., Halford, N. G., and Shewry, P. R. (2001). Synthesis, expression and characterization of peptides comprised of perfect repeat motifs based on a wheat seed storage protein. *Biochim. Biophys. Acta.* **1546,** 346–353.

Feldman, M., Lupton, F. G. H., and Miller, T. E. (1995). Wheat *Triticum* spp. (Gramineae-Triticinae). *In* "Evolution of Crop Plants" (J. Smartt and N. W. Simmonds, eds.), pp. 184–192. Longman Scientific & Technical, Harlow, UK.

Felix, I., Martinant, J. P., Bernard, M., Bernard, S., and Branlard, G. (1996). Genetic characterisation of storage proteins in a set of F_1-derived haploid lines in bread wheat. *Theor. Appl. Genet.* **92,** 340–346.

Food and Agriculture Organization (1999). Production. *In* "FAO Yearbook." Food and Agriculture Organization of the United Nations, Rome.

Forde, B. G., Heyworth, A., Pywell, J., and Kreis, M. (1985). Nucleotide sequence of a B1 hordein gene and the identification of possible upstream regulatory elements in endosperm storage protein genes from barley, wheat and maize. *Nucleic Acids Res.* **13,** 7327–7339.

Forde, J., Forde, B. G., Fry, R. P., Kreis, M., Shewry, P. R., and Miflin, B. J. (1983). Identification of barley and wheat cDNA clones related to the high-M_r polypeptides of wheat gluten. *FEBS Lett.* **162,** 360–366.

Forde, J., Malpica, J.-M., Halford, N. G., Shewry, P. R., Anderson, O. D., Greene, F. C., and Miflin, B. J. (1985). The nucleotide sequence of a HMW glutenin subunit gene located on chromosome 1A of wheat (*Triticum aestivum* L.). *Nucleic Acids Res.* **13,** 6817–6831.

Galili, G., and Feldman, M. (1984). Mapping of glutenin and gliadin genes located on chromosome 1B of common wheat. *Mol. Gen. Genet.* **193,** 293–298.

Galili, G., and Feldman, M. (1985). Genetic control of endosperm proteins in wheat. 3. Allocation to chromosomes and differential expression of high molecular weight glutenin and gliadin genes in intervarietal substitution lines of common wheat. *Theor. Appl. Genet.* **69,** 583–589.

Gianibelli, M. C., Gupta, R. B., Lafiandra, D., Margiotta, B., and MacRitchie, F. (2001). Polymorphism of high M_r glutenin subunits in *Triticum tauschii*: Characterisation by chromatography and electrophoretic methods. *J. Cereal Sci.* **33,** 39–52.

Giese, H., and Hopp, E. (1984). Influence of nitrogen nutrition on the amount of hordein, protein Z and β-amylase messenger RNA in developing endosperms of barley. *Carlsberg Res. Commun.* **49,** 365–383.

Gooding, M. J., and Davies, W. P. (1997). "Wheat Production and Utilization: Systems, Quality and the Environment." CAB International, Wallington, UK.

Gupta, R. B., and MacRitchie, F. (1991). A rapid one-step one-dimensional SDS-PAGE procedure for analysis of subunit composition of glutenin in wheat. *J. Cereal Sci.* **14,** 105–109.

Gupta, R. B., and MacRitchie, F. (1994). Allelic variation at glutenin subunit and gliadin loci, *Glu-1*, *Glu-3* and *Gli-1* of common wheats. II. Biochemical basis of the allelic effects on dough properties. *J. Cereal Sci.* **19,** 19–29.

Gupta, R. B., and Shepherd, K. W. (1988). Low molecular weight glutenin subunits in wheat: Their variation, inheritance and association with breadmaking quality. *In* "Proceedings of the 7th International Wheat Genetics Symposium," pp. 943–949. Institute of Plant Science Research, Cambridge.

Gupta, R. B., and Shepherd, K. W. (1990). Two-step one-dimensional SDS-PAGE analysis of LMW subunits of glutenin. I. Variation and genetic control of the subunits in hexaploid wheats. *Theor. Appl. Genet.* **80,** 65–74.

Gupta, R. B., and Shepherd, K. W. (1993). Production of multiple wheat-rye 1RS translocation stocks and genetic analysis of LMW subunits of glutenin and gliadins in wheats using these stocks. *Theor. Appl. Genet.* **85,** 719–728.

Gupta, R. B., Singh, N. K., and Shepherd, N. K. (1989). The cumulative effect of allelic variation in LMW and HMW glutenin subunits on dough properties in progeny of two bread wheats. *Theor. Appl. Genet.* **77,** 57–64.

Gupta, R. B., Békés, F., and Wrigley, C. W. (1991). Prediction of physical dough properties from glutenin subunit composition in bread wheats: Correlation study. *Cereal Chem.* **68,** 328–333.

Gupta, R. B., Popineau, Y., Lefebvre, J., Cornec, M., Lawrence, G. J., and MacRitchie, F. (1995). Biochemical basis of flour properties in bread wheats. II. Changes in polymeric protein formation and dough/gluten properties associated with the loss of low M_r or high M_r glutenin subunits. *J. Cereal Sci.* **21,** 103–116.

Halford, N. G., Forde, J., Anderson, O. D., Greene, F. C., and Shewry, P. R. (1987). The nucleotide and deduced amino acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*T. aestivum* L.) and comparison with those of genes from chromosomes 1A and 1D. *Theor. Appl. Genet.* **75,** 117–126.

Halford, N. G., Forde, J., Shewry, P. R., and Kreis, M. (1989). Functional analysis of the upstream regions of a silent and an expressed member of a family of wheat seed protein genes in transgenic tobacco. *Plant Sci.* **62,** 207–216.

Halford, N. G., Field, J. M., Blair, H., Urwin, P., Moore, K., Robert, L., Thompson, R., Flavell, R. B., Tatham, A. S., and Shewry, P. R. (1992). Analysis of HMW glutenin subunits encoded by chromosome 1A of bread wheat (*Triticum aestivum* L.) indicates quantitative effects on grain quality. *Theor. Appl. Genet.* **83,** 373–378.

Hammond-Kosack, M. C. U., Holdsworth, M. J., and Bevan, M. W. (1993). *In vivo* footprinting of a low molecular weight glutenin gene (*LMWG-1D1*) in wheat endosperm. *EMBO J* **12,** 545–554.

Harberd, N. P., Bartels, D., and Thompson, R. D. (1985). Analysis of the gliadin multigene loci in bread wheat using nullisomic-tetrasomic lines. *Mol. Gen. Genet.* **198,** 234–242.

Harberd, N. P., Bartels, D., and Thompson, R. D. (1986). DNA restriction-fragment variation in the gene family encoding high molecular weight (HMW) glutenin subunits of wheat. *Biochem. Genet.* **24,** 579–596.

Harberd, N. P., Flavell, R. B., and Thompson, R. D. (1987). Identification of a transposon-like insertion in a *Glu-1* allele of wheat. *Mol. Gen. Genet.* **209,** 326–332.

He, G., Rooke, L., Steele, S., Békés, F., Gras, P., Tatham, A. S., Fido, R., Barcelo, P., Shewry, P. R., and Lazzeri, P. (1999). Transformation of pasta wheat (*Triticum durum* L.) with HMW glutenin subunit genes and modification of dough functionality. *Mol. Breeding* **5,** 377–386.

Hsia, C. C., and Anderson, O. D. (2001). Isolation and characterization of wheat ω-gliadin genes. *Theor. Appl. Genet.* **103,** 37–44.

Jackson, E. A., Holt, L. M., and Payne, P. I. (1983). Characterization of high molecular weight gliadin and low molecular weight glutenin subunits of wheat endosperm by two dimensional electrophoresis and the chromosomal localization of their controlling genes. *Theor. Appl. Genet.* **66,** 29–37.

Jackson, E. A., Morel, M. H., Sontag-Strohm, T., Branlard, G., Metakovsky, E. V., and Redaelli, R. (1996). Proposal for combining the classification systems of alleles of *Gli-1* and *Glu-3* loci in bread wheat (*T. aestivum* L.). *J. Genet. Breeding* **50,** 321–336.

Jauhar, P. P., and Chibbar, R. N. (1999). Chromosome-mediated and direct gene transfers in wheat. *Genome* **42,** 570–583.

Joppa, L. R., Khan, K., and Williams, N. D. (1983). Chromosomal location of genes for gliadin polypeptides in durum wheat *Triticum turgidum* L. *Theor. Appl. Genet.* **64,** 289–293.

Kammholz, S. J., Grams, R. A., Banks, P. M., and Sutherland, M. W. (1998). Segregation of glutenins in wheat × maize-derived doubled haploid wheat populations. *Aus. J. Agric. Res.* **49,** 1253–1259.

Kasarda, D. D. (1989). Glutenin structure in relation to wheat quality. *In* "Wheat Is Unique" (Y. Pomeranz, ed.), pp. 277–302. American Association of Cereal Chemists, St. Paul, MN.

Kasarda, D. D., Autran, J.-C., Lew, E. J.-L., Nimmo, C. C., and Shewry, P. R. (1983). N-terminal amino acid sequences of ω-gliadins and ω-secalins: Implications for the evolution of prolamin genes. *Biochim. Biophys. Acta.* **747,** 138–150.

Killermann, B., and Zimmermann, G. (2000). Relationship betwen allelic variation of *Glu-1, Glu-3* and *Gli-1* prolamin loci and baking quality in doubled haploid wheat populations. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 66–70. Royal Society of Chemistry, Cambridge.

Koebner, R. M. D., and Shepherd, K. W. (1987). Allosyndetic recombination between a chromosome of *Aegilops umbellulata* and wheat chromosome. *Heredity* **59,** 33–45.

Konzak, C. F. (1977). Genetic control of the content, amino acid composition, and processing properties of proteins in wheat. *In* "Advances in Genetics" (E. W. Caspari, ed.), pp. 407–582. Academic Press, New York.

Kreis, M., Forde, B. G., Rahman, S., Miflin, B. J., and Shewry, P. R. (1985). Molecular evolution of the seed storage proteins of barley, rye and wheat. *J. Mol. Biol.* **183,** 499–502.

Kruger, J. E., and Bietz, J. A. (1994). "High-Performance Liquid Chromatography of Cereal and Legume Proteins." American Association of Cereal Chemists, St. Paul, MN.

Laemmli, U. K. (1970). Cleavage of structural proteins durino the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.

Lafiandra, D., and Kasarda, D. D. (1985). One- and two-dimensional (two-pH) polyacrylamide gel electrophoresis in a single gel: Separation of wheat proteins. *Cereal Chem.* **62,** 314–319.

Lafiandra, D., Morris, R., and Kasarda, D. D. (1984). Chromosomal assignment of genes coding for the wheat gliadin protein components of the cultivars Cheyenne and Chinese Spring by two-dimensional (two-pH) electrophoresis. *Theor. Appl. Genet.* **68,** 531–539.

Lafiandra, D., Colaprico, G., Kasarda, D. D., and Porceddu, E. (1987). Null alleles for gliadin blocks in bread and durum wheat cultivars. *Theor. Appl. Genet.* **74,** 610–616.

Lafiandra, D., Benedettelli, S., and Porceddu, E. (1988). Null forms for storage proteins in bread wheat and durum. *In* "Proceedings of the 7th International Wheat Genetics Symposium," pp. 963–967. Institute of Plant Science Research, Cambridge.

Lafiandra, D., Ciaffi, M., and Benedettelli, S. (1993a). Seed storage proteins of wild wheat progenitors. *In* "Biodiversity and Wheat Improvement" (A. B. Damania, ed.), pp. 329–340. John Wiley & Sons, New York.

Lafiandra, D., D'Ovidio, R., Porceddu, E., Margiotta, B., and Colaprico, G. (1993b). New data supporting high M_r glutenin subunit 5 as the determinant of quality differences among the pairs 5+10 vs. 2+12. *J. Cereal Sci.* **18,** 197–205.

Lafiandra, D., Tucci, G. F., Pavoni, A., Turchetta, T., and Margiotta, B. (1997). PCR analysis of x- and y-type genes present at the complex *Glu-A1* locus in durum and bread wheat. *Theor. Appl. Genet.* **94,** 235–240.

Lafiandra, D., Masci, S., Blumenthal, C., and Wrigley, C. W. (1999). The formation of glutenin polymer in practice. *Cereal Foods World* **44,** 572–578.

Lafiandra, D., Masci, S., D'Ovidio, R., and Margiotta, B. (2000). The genetics of wheat gluten proteins: An overview. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 3–10. Royal Society of Chemistry, Cambridge.

Lafiandra, D., Margiotta, B., Colaprico, G., Masci, S., Roth, M. R., and MacRitchie, F. (2000). Introduction of the D-genome related high- and low-M_r glutenin subunits into durum wheat and their effect on technological properties. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 51–54. Royal Society of Chemistry, Cambridge.

Lagudah, E. S., and Halloran, G. M. (1988). Phylogenetic relationships of *Triticum tauschii*, the D genome donor to hexaploid wheat. I. Variation in HMW subunits of glutenin and gliadins. *Theor. Appl. Genet.* **75,** 592–598.

Lamacchia, C., Shewry, P. R., Di Fonzo, N., Forsyth, J. L., Harris, N., Lazzeri, P. A., Napier, J. A., Halford, N. G., and Barcelo, P. (2001). Endosperm-specific activity of a storage protein gene promoter in transgenic wheat seed. *J. Exp. Bot.* **52,** 243–250.

Larré, C., Denery-Papini, S., Popineau, Y., Deshayes, G., Desserme, C., and Lefebvre, J. (2000). Biochemical analysis and rheological properties of gluten modified by transglutaminase. *Cereal Chem.* **77,** 121–127.

Law, C. N., and Payne, P. I. (1983). Genetical aspect of breeding for improved grain protein content and type in wheat. *J. Cereal Sci.* **1,** 79–93.

Lawrence, G. J., and Payne, P. I. (1983). Detection by gel electrophoresis of oligomers formed by the association of high-molecular-weight glutenin protein subunits of wheat endosperm. *J. Exp. Bot.* **34,** 254–267.

Lawrence, G. J., and Shepherd, K. W. (1980). Variation in glutenin protein subunits of wheat. *Aust. J. Biol. Sci.* **33,** 221–233.

Lawrence, G. J., MacRitchie, F., and Wrigley, C. W. (1988). Dough and baking quality of wheat lines deficient in glutenin subunits controlled by the *Glu-A1*, *Glu-B1* and *Glu-D1* loci. *J. Cereal Sci.* **7,** 109–112.

Lee, Y. J., Békés, F., Gupta, R., Appels, R., and Morell, M. K. (1999). The low-molecular-weight glutenin subunits proteins of primitive wheats. I. Variation in A-genome species. *Theor. Appl. Genet.* **98,** 119–125.

Levy, A. A., Galili, G., and Feldman, M. (1988). Polymorphism and genetic control of high molecular weight glutenin subunits in wild tetraploid wheat *Triticum turgidum* var. *dicoccoides*. *Heredity* **61,** 63–72.

Lew, E. J. L., Kuzmicky, D. D., and Kasarda, D. D. (1992). Characterization of low molecular weight glutenin subunits by reversed-phase high performance liquid chromatography, sodium dodecyl sulfate polyacrylamide gel electrophoresis and N-terminal amino acid sequencing. *Cereal Chem.* **69,** 508–515.

Liu, C. Y., and Shepherd, K. W. (1995). Inheritance of B subunits of glutenin and ω- and γ-gliadins in tetraploid wheats. *Theor. Appl. Genet.* **90,** 1149–1157.

Lorenzo, A., Kronstad, W. E., and Vieira, L. G. E. (1987). Relationships between high molecular weight glutenin subunits and loaf volume in wheat as measured by sodium dodecyl sulphate sedimentation test. *Crop Sci.* **27,** 253–257.

Lukaszewski, A. J., and Curtis, C. A. (1992). Transfer of the *Glu-D1* gene from chromosome 1D of bread wheat to chromosome 1R in hexaploid triticale. *Plant Breeding* **109,** 203–211.

Lukaszewski, A. J., and Curtis, C. A. (1994). Transfer of the *Glu-D1* gene from chromosome 1D to chromosome 1A in triticale. *Plant Breeding* **112,** 177–182.

Lukow, O. M. (2000). Genetic analysis of dough strength using doubled haploid lines. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 61–65. Royal Society of Chemistry, Cambridge.

Lukow, O. M., Forsyth, S. A., and Payne, P. I. (1992). Over-production of HMW glutenin subunits coded on chromosome 1B in common wheat, *Triticum aestivum*. *J. Genet. Breeding* **46,** 187–192.

Luo, C., Griffin, W. B., Branlard, G., and McNeil, D. L. (2001). Comparison of low- and high molecular weight glutenin allele effects on flour quality. *Theor. Appl. Genet.* **102,** 1088–1098.

Mackie, A. M., Lagudah, E. S., Sharp, P. J., and Lafiandra, D. (1996). Molecular and biochemical characterisation of HMW glutenin subunits from *T. tauschii* and the D genome of hexaploid wheat. *J. Cereal Sci.* **23,** 213–225.

MacRitchie, F., duCros, D. L., and Wrigley, C. W. (1990). Flour polypeptides related to wheat quality. *Adv. Cereal Sci. Technol.* **10,** 79–145.

Margiotta, B., Colaprico, G., D'Ovidio, R., and Lafiandra, D. (1993). Characterization of high M_r subunits of glutenin by combined chromatographic (RP-HPLC) and electrophoretic separations and restriction fragment length polymorphism (RFLP) analyses of their encoding genes. *J. Cereal Sci.* **17,** 221–236.

Margiotta, B., Urbano, M., Colaprico, G., Turchetta, T., and Lafiandra, D. (1998). Variation of high molecular weight glutenin subunits in tetraploid wheats of genomic formula AAGG. *In* "Proceedings of the 9th International wheat Genetics Symposium" (A. E. Slinkard, ed.), pp. 195–197. University Extension Press, University of Saskatchewan, Regina, SK, Canada.

Masci, S., Lafiandra, D., Porceddu, E., Lew, E. J.-L., Tao, H. P., and Kasarda, D. D. (1993). D-glutenin subunits: N-terminal sequences and evidence for the presence of cysteine. *Cereal Chem.* **70,** 581–585.

Masci, S., Egorov, T. A., Ronchi, C., Kuzmicky, D., Kasarda, D. D., and Lafiandra, D. (1999). Evidence for the presence of only one cysteine residue in the D-type low molecular weight subunits of wheat glutenin. *J. Cereal Sci.* **29,** 17–25.

Masci, S., Rovelli, L., Kasarda, D. D., Vensel, W. H., and Lafiandra, D. (2002). Characterisation and chromosomal localisation of C-type low-molecular-weight glutenin subunits in the bread wheat cultivar Chinese Spring. *Theor. Appl. Genet.* **104,** 422–428.

Metakovsky, E. V. (1991). Gliadin allele identification in common wheat. II. Catalogue of gliadin alleles in common wheat. *J. Genet. Breeding* **45,** 325–344.

Metakovsky, E. V., Akhmedov, M. G., and Sozinov, A. A. (1986). Genetic analysis of gliadin-encoding genes reveals gene clusters as well as single remoted genes. *Theor. Appl. Genet.* **73,** 278–285.

Metakovsky, E., Chernacov, V. M., Upelniek, U. P., Redaelli, R., Dardevet, N., Branlard, G., and Pogna, N. E. (1996). Recombination mapping of minor ω-gliadin coding loci on chromosome 1A in common wheat: A revision. *J. Genet. Breeding* **50,** 277–286.

Metakovsky, E. V., Branlard, G., Chernakov, V. M., Upelniek, V. P., Redaelli, R., and Pogna, N. E. (1997a). Recombination mapping of some 1A-, 1B-, 1D- and 6D-controlled gliadins and low molecular weight glutenin subunits in common wheat. *Theor. Appl. Genet.* **94,** 788–795.

Metakovsky, E. V., Annichiarico, P., Boggini, G., and Pogna, N. E. (1997b). Relationships between gliadin alleles and dough strength in Italian common wheats. *J. Cereal Sci.* **25,** 229–236.

Metakovsky, E. V., Felix, I., and Branlard, G. (1997c). Association between dough quality (*W* value) and certain gliadin alleles in French common wheat cultivars. *J. Cereal Sci.* **26,** 371–373.

Moonen, J. H. E., Scheepstra, A., and Graveland, A. (1982). Use of SDS-sedimentation test and SDS-polyacrylamide gel electrophoresis for screening breeders' samples of wheat for breadmaking quality. *Euphytica* **31,** 677–690.

Mujeeb-Kazi, A., Rosas, V., and Roldan, S. (1996). Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats (*T. turgidum* L. s. lat. $\times$ *T. tauschii*; $2n = 6x = 42$, AABBDD) and its potential utilization for wheat improvement. *Genet. Resources Crop Evol.* **43,** 129–134.

Müller, M., and Knudsen, S. (1993). The nitrogen response of a barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box. *Plant J.* **4,** 343–355.

Nevo, E., and Payne, P. I. (1987). Wheat storage proteins: Diversity of HMW glutenin alleles in wild emmer from Israel. I. Geographical patterns and ecological predictability. *Theor. Appl. Genet.* **74,** 827–836.

Okita, T. W., Cheesbrough, V., and Reeves, C. D. (1985). Evolution and heterogeneity of the α-/β-type and γ-type gliadin DNA sequences. *J. Biol. Chem.* **260,** 8203–8213.

Orth, R. A., and Bushuk, W. (1974). Studies of glutenin. VI. Chromosomal location for subunits of glutenin of common wheat. *Cereal Chem.* **51,** 118–126.

Orth, R. A., and Shellenberger, J. A. (1988). Origin, production, and utilization of wheat. *In* "Wheat Chemistry and Technology" (Y. Pomeranz, ed.), pp. 1–14. American Association of Cereal Chemists, St. Paul, MN.

Osborne, T. B. (1924). "The Vegetable Proteins," pp. 154. Longmans, Green & Co., London.

Pastori, G. M., Steele, S. H., Jones, H. D., and Shewry, P. R. (2000). Transformation of commercial wheat varieties with high molecular weight glutenin subunit genes. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 88–92. Royal Society of Chemistry, Cambridge.

Pastori, G. M., Wilkinson, M. D., Steele, S. H., Sparks, C. A., Jones, H. D., and Parry, M. A. J. (2001). Age-dependent transformation frequency in elite wheat varieties. *J. Exp. Bot.* **52,** 857–863.

Payne, P. I. (1987). Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annu. Rev. Plant Physiol.* **38,** 141–153.

Payne, P. I., and Lawrence, G. J. (1983). Catalogue of alleles for the complex gene loci, *Glu-A1*, *Glu-B1* and *Glu-D1* which code for high-molecular-weight subunits of glutenin in hexaploid wheat. *Cereal Res. Commun.* **11,** 29–35.

Payne, P. I., and Seekings, J. A. (1996). Characterisation of Galahad-6, Galahad-7 and Galahad-8, isogenic lines that contain only one HMW glutenin subunit. *In* Proceedings of the 7th International Gluten Workshop" (C. W. Wrigley, ed.), pp. 14–17.

Payne, P. I., Corfield, K. G., and Blackman, J. A. (1979). Identification of a high-molecular weight subunit of glutenin whose presence correlates with breadmaking quality in wheats of related pedigree. *Theor. Appl. Genet.* **60,** 229–236.

Payne, P. I., Law, C. N., and Mudd, E. E. (1980). Control by homoeologous group 1 chromosomes of the high molecular weight subunits of glutenin, a major protein of the wheat endosperm. *Theor. Appl. Genet.* **58,** 113–120.

Payne, P. I., Holt, L. M., and Law, C. N. (1981a). Structural and genetical studies of the high molecular subunits of wheat glutenin. I. Allelic variation in subunits amongst varieties of wheat. *Theor. Appl. Genet.* **60,** 229–236.

Payne, P. I., Corfield, K. G., Holt, L. M., and Blackman, J. A. (1981b). Correlations between the inheritance of certain high-molecular weight subunits of glutenin and breadmaking quality in progenies of six crosses of bread wheat. *J. Sci. Food Agric.* **32,** 51–60.

Payne, P. I., Holt, L. M., Lawrence, G. J., and Law, C. N. (1982). The genetics of gliadin and glutenin, the major storage proteins of the wheat endosperm. *Qualitas Plantarum Plant Foods Hum. Nutr.* **31,** 229–241.

Payne, P. I., Holt, L. M., and Lawrence, G. D. (1983). Detection of a novel high-molecular-weight subunit of glutenin in some Japanese hexaploid wheats. *J. Cereal Sci.* **1,** 3–8.

Payne, P. I., Holt, L. M., Jackson, E. A., and Law, C. N. (1984a). Wheat storage proteins: Their genetics and their potential for manipulation by plant breeding. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **304,** 359–371.

Payne, P. I., Jackson, E. A., Holt, L. M., and Law, C. N. (1984b). The genetic linkage between endosperm protein genes on each of the short arms of chromosomes 1A and 1B in wheat. *Theor. Appl. Genet.* **67,** 235–243.

Payne, P. I., Holt, L. M., Jarvis, M. G., and Jackson, E. A. (1985). Two dimensional frationation of the endosperm proteins of bread wheat (*Triticum aestivum*): Biochemical and genetical studies. *Cereal Chem.* **62,** 319–326.

Payne, P. I., Nightingale, M. A., Krattiger, A. F., and Holt, L. M. (1987a). The relationship between the HMW glutenin subunit composition and the breadmaking quality of British-grown wheat varieties. *J. Sci. Food Agric.* **40,** 51–65.

Payne, P. I., Holt, L. M., Harinder, K., McCartney, D. P., and Lawrence, G. J. (1987b). The use of near-isogenic lines with different HMW glutenin subunit composition in studying breadmaking quality and glutenin structure. *In* "Proceedings of the 3rd International Gluten Workshop" (R. D. Lasztity and F. Békés, eds.), pp. 216–226. .

Payne, P. I., Seekings, J. A., Worland, A. J., Jarvis, M. G., and Holt, L. M. (1987c). Allelic variation of glutenin subunits and gliadins and its effects on breadmaking quality in wheat: Analysis of F_5 progeny from "Chinese Spring" × "Chinese Spring (Hope 1A)." *J. Cereal Sci.* **6,** 103–118.

Pfluger, L. A., D'Ovidio, R., Margiotta, B., Peña, R., Mujeeb-Kazi, A., and Lafiandra, D. (2001). Characterisation of high- and low-molecular weight glutenin subunits associated to the D genome of *Aegilops tauschii* in a collection of synthetic hexaploid wheats. *Theor. Appl. Genet.* **103,** 1293–1301.

Pogna, N. E., Boggini, G., Corbellini, M., Cattaneo, M., and Dal Belin Peruffo, A. (1982). Association between gliadin electrophoretic bands and quality in common wheat. *Can. J. Plant Sci.* **4,** 913–918.

Pogna, N. E., Lafiandra, D., Feillet, P., and Autran, J.-C. (1988). Evidence for direct causal effect of low molecular weight subunits of glutenins on gluten viscoelasticity in durum wheats. *J. Cereal Sci.* **7,** 211–214.

Pogna, N. E., Autran, J. C., Mellini, F., Lafiandra, D., and Feillet, P. (1990). Chromosome 1B encoded gliadins and glutenin subunits in durum wheat: Genetics and relationship to gluten strength. *J. Cereal Sci.* **11,** 15–34.

Pogna, N. E., Metakovsky, E. V., Redaelli, R., Raineri, F., and Dachkevitch, T. (1993). Recombination mapping of *Gli-5*, a new gliadin-coding locus on chromosomes 1A and 1B in common wheat. *Theor. Appl. Genet.* **87,** 113–121.

Pogna, N. E., Redaelli, R., Vaccino, P., Peruffo, A. D. B., Curioni, A., Metakovsky, E. V., and Pagliaricci, S. (1995a). Production and genetic characterization of near-isogenic lines in the bread wheat cultivar Alpe. *Theor. Appl. Genet.* **90,** 650–658.

Pogna, N. E., Metakovsky, E. V., Redaelli, R., Dachkevitch, T., and Cernakov, V. M. (1995b). The group 1 chromosomes of wheat contain several loci coding for gliadins. *In* "Proceedings of the 8th International Wheat Genetics Symposium" (Z. S. Li and Z. Y. Xin, eds.), pp. 137–140. China Agriculture Scientech Press, Beijing, China.

Popineau, Y., Cornec, M., Lefebre, J., and Marchylo, B. (1994). Influence of high M_r glutenin subunits on glutenin polymers and rheological properties of gluten and gluten sufractions of near isogenic lines of wheat Sicco. *J. Cereal Sci.* **19,** 231–242.

Popineau, Y., Deshayes, G., Lefebvre, J., Fido, R., Tatham, A. S., and Shewry, P. R. (2001). Prolamin aggregation, gluten viscoelasticity, and mixing properties of transgenic wheat lines expressing 1Ax and 1Dx high molecular weight glutenin subunit transgenes. *J. Agric. Food Chem.* **49,** 395–401.

Rafalski, J. A. (1986). Structure of wheat γ-gliadin genes. *Gene* **43,** 221–229.

Rasco-Gaunt, S., and Barcelo, P. (1999). Immature inflorescence culture of cereals. *In* "Methods in Molecular Biology: Plant Cell Culture Methods" (R. D. Hall, ed.), pp. 71–81. Humana Press, Totowa, NJ.

Rasco-Gaunt, S., Riley, A., Lazzeri, P., and Barcelo, P. (2001). Procedures allowing the transformation of a range of European elite wheat (*Triticum aestivum* L.) varieties via particle bombardment. *J. Exp. Bot.* **52,** 865–874.

Redaelli, R., Morel, M.-H., Autran, J.-C., and Pogna, N. E. (1995). Genetic analysis of low M_r glutenin subunits fractionated by two-dimensional electrophoresis A-PAGE × SDS-PAGE. *J. Cereal Sci.* **21,** 5–13.

Reddy, P., and Appels, R. (1993). Analysis of a genomic DNA segment carrying the wheat high-molecular-weight (HMW) glutenin Bx17 subunit and its use as an RFLP marker. *Theor. Appl. Genet.* **85,** 616–624.

Riley, R., and Chapman, V. (1958). Genetic control of the cytologically diploid behavior of hexaploid wheat. *Nature* **182,** 713–715.

Robert, L. S., Thompson, R. D., and Flavell, R. B. (1989). Tissue-specific expression of a wheat high molecular weight glutenin gene in transgenic tobacco. *Plant Cell* **1,** 569–578.

Rodríguez-Quijano, M., and Carrillo, J. M. (1996a). Linkage map of prolamin loci *Gli-D4* and *Gli-D5* in hexaploid wheat. *Plant Breeding* **114,** 40–44.

Rodríguez-Quijano, M., and Carrillo, J. M. (1996b). Relationship between allelic variation of *Glu-1* and *Gli-1/Glu-3* prolamin loci and gluten strength in hexaploid wheat. *Euphytica* **91,** 141–148.

Rodríguez-Quijano, M., Nieto-Taladriz, M. T., and Carrillo, J. M. (1997). Variation in B-LMW glutenin subunits in Einkorn wheats. *Genet. Resources Crop Evol.* **44,** 539–543.

Rogers, W. J., Miller, T. E., Payne, P. I., Seekings, J. A., Holt, L. M., and Law, C. N. (1997). Introduction to bread wheat (*Triticum aestivum* L.) and assessment for breadmaking quality of alleles from *T. boeoticum* Boiss ssp *thaoudar* at *Glu-A1* encoding two high molecular weight subunits of glutenin. *Euphytica* **93,** 19–29.

Rooke, L., Békés, F., Fido, R., Barro, F., Gras, P., Tatham, A. S., Barcelo, P., Lazzeri, P., and Shewry, P. R. (1999). Overexpression of a gluten protein in transgenic wheat results in greatly increased dough strength. *J. Cereal Sci.* **30,** 115–120.

Rousset, M., Carrillo, J. M., Qualset, C. O., and Kasarda, D. D. (1992). Use of recombinant inbred lines of wheat for study of associations of high-molecular-weight glutenin subunit alleles to quantitative traits. 2. Milling and bread-baking quality. *Theor. Appl. Genet.* **83,** 403–412.

Ruiz, M., and Carrillo, J. M. (1993). Linkage relationships between prolamins genes on chromosomes 1A and 1B of durum wheat. *Theor. Appl. Genet.* **87,** 353–360.

Sabelli, P., and Shewry, P. R. (1991). Characterization and organisation of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis. *Theor. Appl. Genet.* **83,** 209–216.

Sabelli, P., Lafiandra, D., and Shewry, P. R. (1992). Restriction fragment analysis of null forms at the *Gli-1* loci of bread and durum wheat. *Theor. Appl. Genet.* **83,** 428–434.

Saponaro, C., Pogna, N. E., Castagna, R., Pasquini, M., Cacciatori, P., and Redaelli, R. (1995). Allelic variation at the *Gli-A1*^{m}, *Gli- A2*^{m} and *Glu-A1*^{m} loci and breadmaking quality in diploid wheat *Triticum monococcum*. *Genet. Res. Cambridge* **66,** 127–137.

Sears, E. R. (1954). The aneuploids of common wheat. *Missouri Agric. Exp. Station Res. Bull.* **572,** 1–58.

Sears, E. R. (1966). Nullisomic–tetrasomic combinations in hexaploid wheat. *In* "Chromosome Manipulations and Plant Genetics" (R. Riley and K. R. Lewis, eds.), pp. 29–45. Oliver and Boyd, London.

Sears, E. R. (1972). Chromosome engineering in wheat. *Stadler Symp.* **4,** 23–38.

Shepherd, K. W. (1968). Chomosomal control of wheat endosperm proteins in wheat and rye. *In* "Proceedings of the 3rd International Wheat Genetics Symposium" (K. W. Finlay and K. W. Shepherd, eds.), pp. 86–96. Australian Academy of Science, Canberra.

Shepherd, K. W. (1988). Genetics of wheat endosperm proteins in retrospect and prospect. *In* "Proceedings of the 7th International Wheat Genetics Symposium," pp. 943–949. Institute of Plant Science Research, Cambridge.

Shewry, P. R. (1995). Plant storage proteins. *Biol. Rev.* **70,** 375–426.

Shewry, P. R. (2003). Wheat gluten proteins. *In* "Wheat Gluten Protein Analysis" (P. R. Shewry and G. Lookhart, eds.), pp. 1–17. AACC, St. Paul, MN.

Shewry, P. R., and Tatham, A. S. (1997). Disulphide bonds in wheat gluten proteins. *J. Cereal Sci.* **25,** 207–227.

Shewry, P. R., and Tatham, A. S. (1999). The characteristics, structures and evolutionary relationships of prolamins. *In* "Seed Proteins" (P. R. Shewry and R. Casey, eds.), pp. 11–36. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Shewry, P. R., Miflin, B. J., and Kasarda, D. D. (1984). The structural and evolutionary relationships of the prolamin storage proteins of barley, rye and wheat. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **304,** 333–339.

Shewry, P. R., Halford, N. G., and Tatham, A. S. (1989). The high molecular weight subunits of wheat, barley and rye: genetics, molecular biology, chemistry and role in wheat gluten structure and functionality. *Oxford Surv. Plant Mol. Cell Biol.* **6,** 163–219.

Shewry, P. R., Sabelli, P. A., Parmar, S., and Lafiandra, D. (1991). α-type prolamins are encoded by genes on chromosomes 4Ha and 6Ha of *Haynaldia villosa* Schur (syn. Dasypyrum villosum L.). *Biochem. Genet.* **29,** 207–211.

Shewry, P. R., Halford, N. G., and Tatham, A. S. (1992). The high molecular weight subunits of wheat glutenin. *J. Cereal Sci.* **15,** 105–120.

Shewry, P. R., Tatham, A. S., and Halford, N. G. (1999). The prolamins of the Triticeae. *In* "Seed Proteins" (P. R. Shewry and R. Casey, eds.), pp. 37–84. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Shewry, P. R., Halford, N. G., Tatham, A. S., Popineau, Y., Lafiandra, D., and Belton, P. S. (2003). The high molecular weight subunits of wheat glutenin and their role in determining wheat processing properties. *Adv. Food Nutri Res.* **45,** 219–302.

Singh, N. K., and Shepherd, K. W. (1988a). Linkage mapping of the genes controlling endosperm storage proteins in wheat. II. Genes on the long arms of group-1 chromosomes. *Theor. Appl. Genet.* **75,** 642–650.

Singh, N. K., and Shepherd, K. W. (1988b). Linkage mapping of the genes controlling endosperm storage proteins in wheat. I. Genes on the short arms of group-1 chromosomes. *Theor. Appl. Genet.* **75,** 628–641.

Singh, N. K., Shepherd, K. W., and Cornish, G. B. (1991). A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. *J. Cereal Sci.* **14,** 203–208.

Sobko, T. I. (1984). Identification of a new locus which controls the synthesis of alcohol-soluble endosperm proteins in winter bread wheat. *J. Agric. Sci. (Kiev)* **320,** 78–80.

Sørensen, M. B., Cameron-Mills, V., and Brandt, A. (1989). Transcriptional and post-transcriptional regulation of gene expression in developing barley endosperm. *Mol. Gen. Genet.* **217,** 195–201.

Sozinov, A. A., and Poperelya, F. A. (1980). Genetic classification of prolamines and its use for plant breeding. *Ann. Technol. Agric.* (*Paris*) **29,** 229–245.

Sozinov, A. A., and Poperelya, F. A. (1982). Polymorphism of prolamines and variability of grain quality. *Qualitas Plantarum Plant Foods Hum. Nutr.* **31,** 243–249.

Sreeramulu, G., and Singh, N. K. (1997). Genetic and biochemical characterization of novel low molecular weight glutenin subunits in wheat (*Triticum aestivum* L.). *Genome* **40,** 41–48.

Sugiyama, T., Rafalski, A., Peterson, D., and Soll, D. (1985). A wheat HMW glutenin subunit gene reveals a highly repeated structure. *Nucleic Acids Res.* **13,** 8729–8737.

Sugiyama, T., Rafalski, A., and Soll, D. (1986). The nucleotide sequence of a wheat γ-gliadin genomic clone. *Plant Sci.* **44,** 205–209.

Sumner-Smith, M., Rafalski, J. A., Sugiyama, T., Soll, M., and Soll, D. (1985). Conservation and variability of wheat α/β gliadin genes. *Nucleic Acids Res.* **11,** 3905–3916.

Tatham, A. S., and Shewry, P. R. (1995). The S-poor prolamins of wheat, barley and rye. *J. Cereal Sci.* **22,** 1–16.

Thomas, M. S., and Flavell, R. B. (1990). Identification of an enhancer element for the endopserm-specific expression of high molecular weight glutenin. *Plant Cell* **2,** 1171–1180.

Thompson, R. D., Bartels, D., Harberd, N. P., and Flavell, R. B. (1983). Characterization of the multigene family coding for HMW glutenin subunits in wheat using cDNA clones. *Theor. Appl. Genet.* **67,** 87–96.

Thompson, R. D., Bartels, D., and Harberd, N. P. (1985). Nucleotide sequence of a gene from chromosome 1D of wheat encoding a HMW-glutenin subunit. *Nucleic Acids Res.* **13,** 6833–6846.

Tosi, P., Napier, J. A., D'Ovidio, R. D., Jones, H. D., and Shewry, P. R. (2000). Modification of the LMW glutenin subunit composition of durum wheat by microprojectile-mediated transformation. *In* "Wheat Gluten" (P. R. Shewry and A. S. Tatham, eds.), pp. 93–96. Royal Society of Chemistry, Cambridge.

Vaccino, P., and Metakovsky, E. V. (1995). RFLP patterns of gliadin alleles in *Triticum aestivum* L.: Implications for analysis of the organization and evolution of complex loci. *Theor. Appl. Genet.* **90,** 173–181.

Varshney, A., and Altpeter, F. (2001). Stable transformation and tissue culture response in current European winter wheats (*Triticum aestivum* L.). *Mol. Breeding* **8,** 295–309.

Vasil, I. K., Bean, S., Zhao, J., McCluskey, P., Lookhart, G., Zhao, H.-P., Altpeter, F., and Vasil, V. (2001). Evaluation of baking properties and gluten protien composition of field grown transgenic wheat lines expressing high molecular weight glutenin gene 1Ax1. *J. Plant Physiol.* **158,** 521–528.

Vasil, V., Castillo, A., Fromm, M., and Vasil, I. (1992). Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Bio/Technology* **10,** 667–674.

Vaucheret, H., Béclin, C., Elmayan, T., Feuerbach, F., Gordon, C., Morel, J.-B., Mourrain, P., Palauqui, J.-C., and Vernhettes, S. (1998). Transgene-induced gene silencing in plants. *Plant J.* **16,** 651–659.

Vitellozzi, F., Ciaffi, M., Dominaci, L., and Ceoloni, C. (1997). Isolation of a chromosomally engineered durum wheat line carrying the common wheat *Glu-D1d* allele. *Agronomie* **17,** 413–419.

Waines, J. G., and Payne, P. I. (1987). Electrophoretic analysis of the high-molecular-weight glutenin subunits of *Triticum monococcum*, *T. urartu*, and the A genome of bread wheat (*T. aestivum*). *Theor. Appl. Genet.* **74,** 71–76.

Wan, Y., Wang, D., Shewry, P. R., and Halford, N. G. (2002). Isolation and characterization of five novel high molecular weight subunit of glutenin genes from *Triticum timopheevi* and *Aegilops cylindrica*. *Theor. Appl. Genet.* **104,** 828–839.

Weeks, J., Anderson, O., and Blechl, A. (1993). Rapid production of multiple independent lines of fertile transgenic wheat *Triticum aestivum*. *Plant Physiol.* **102,** 1077–1084.

Werner, W. E., Adalstain, A. E., and Kasarda, D. D. (1992). Composition of high-molecular-weight glutenin subunit dimers formed by partial reduction of residue glutenin. *Cereal Chem.* **69,** 535–541.

Wrigley, C. W., and Shepherd, K. W. (1973). Electrofocusing of grain proteins from wheat genotypes. *Ann. N. Y. Acad. Sci.* **209,** 154–162.

Wrigley, C. W., Robinson, P. J., and Williams, W. T. (1982). Association between individual gliadin proteins and quality, agronomic and morphological attributes of wheat cultivars. *Aust. J. Agric. Res.* **33,** 409–418.

Xin, H., Jiang, Y., Huang, W., Gu, Q., and Sun, C (1992). Nucleotide sequence of a high molecular-weight glutenin gene in wheat. *Acta Bot. Sin.* **34,** 729–735.

Index

E