Duplicated Proteasome Subunit Genes in *Drosophila melanogaster* Encoding Testes-Specific Isoforms

Xiaoqing Yuan,¹ Mary Miller and John M. Belote

Department of Biology, Syracuse University, Syracuse, New York 13244

Manuscript received February 5, 1996

Accepted for publication May 20, 1996

ABSTRACT

Using the previously cloned proteasome α -type subunit gene *Pros28.1*, we screened a *Drosophila melanogaster* genomic library using reduced stringency conditions to identify closely related genes. Two new genes, *Pros28.1A* (map position 92F) and *Pros28.1B* (map position 60D7), showing high sequence similarity to *Pros28.1*, were identified and characterized. *Pros28.1A* encodes a protein with 74% amino acid identity to *PROS28.1*, while the *Pros28.1B* gene product is 58% identical. The *Pros28.1B* gene has two introns, located in exactly analogous positions as the two introns in *Pros28.1*, while the *Pros28.1A* gene lacks introns. Northern blot analysis reveals that the two new genes are expressed only in males, during the pupal and adult stages. Tissue-specific patterns of expression were examined using transgenic flies carrying *lacZ*-fusion reporter genes. This analysis revealed that both genes are expressed in germline cells during spermatogenesis, although their expression patterns differed. *Pros28.1A* expression is first detected at the primary spermatocyte stage and persists into the spermatid elongation phase of spermiogenesis, while *Pros28.1B* expression is prominent only during spermatid elongation. These genes represent the most striking example of cell-type-specific proteasome gene expression reported to date in any system and support the notion that there is structural and functional heterogeneity among proteasomes in metazoans.

INTRACELLULAR proteolysis is an important mechanism by which the levels of various gene products can be modulated. Regulated protein turnover in eukaryotes is known to play key roles in a variety of processes, including cell differentiation, metabolic regulation, control of cell proliferation, and programmed cell death (for reviews, see GOLDBERG 1992; CALLIS 1995; HOCHSTRASSER 1995). Targeted protein degradation is also important for ridding cells of abnormal polypeptides that arise as the result of mutation or damage and that can potentially interfere with normal cell function. In eukaryotes, a major pathway for the breakdown of abnormal polypeptides and short-lived regulatory proteins is the ubiquitin-dependent proteolytic pathway (HERSHKO and CIECHANOVER 1992; JENTSCH 1992; CIECHANOVER 1994; JENTSCH and SCHLENKER 1995). In this system, proteins are targeted for destruction by their covalent attachment to the small, well-conserved polypeptide ubiquitin and are degraded by a large (26S, ~2000 kDa), multi-subunit complex that has as its core proteinase a structure known as the 20S proteasome (GOLDBERG and ROCK 1992; TANAKA *et al.* 1992; RECHSTEINER *et al.* 1993; RIVETT 1993; PETERS 1994). Among the specific substrates of the ubiquitin-proteasome system are the yeast transcriptional repressor MAT α 2 (HOCHSTRASSER *et al.* 1991), plant phyto-

chrome (SHANKLIN *et al.* 1987), the tumor suppressor, p53 (SCHEFFNER *et al.* 1993), the cystic fibrosis transmembrane conductance regulator (WARD *et al.* 1995), the c-Jun and c-Mos oncoproteins (NISHIZAWA *et al.* 1993; TREIER *et al.* 1994), I κ B α , an inhibitor of the Rel family of transcription factors (PALOMBELLA *et al.* 1994), and cyclin B, a cell cycle regulator (GLOTZER *et al.* 1991). Recent studies have also implicated this system in the maturation of the NF- κ B transcription factor, demonstrating that the proteasome is responsible for the proteolytic processing of a 105-kD precursor into its 50-kD active form (PALOMBELLA *et al.* 1994). Thus, in addition to a general "housekeeping" role, proteasomes may also have more specialized functions related to the turnover or processing of specific regulatory proteins.

Proteasomes are abundant and are found in the cytosol and nucleus of many cell types. They have been isolated from a variety of eukaryotes and from the archaeobacterium, *Thermoplasma acidophilum*, and their overall structures are well-conserved (TANAKA *et al.* 1992). Electron microscope (BAUMEISTER *et al.* 1988; PÜHLER *et al.* 1992; KOPP *et al.* 1993) and X-ray crystallography (LÖWE *et al.* 1995) studies indicate that the 20S proteasome is a hollow, barrel-shaped cylinder made up of four stacked rings of seven subunits each. The structure of the archaeon proteasome is relatively simple, with the outer rings containing a single type of subunit called α , and the inner rings containing a subunit called β . The subunit composition of eukaryotic proteasomes,

Corresponding author: John Belote, Department of Biology, 130 College Place, Syracuse, NY 13244. E-mail: jbelote@mailbox.syr.edu

¹ Present address: National Institutes of Health, Building 49, 49 Convent Dr., MSC 4480, Bethesda, MD 20892.

however, is more complex. In the yeast *Saccharomyces cerevisiae* 14 different genes, encoding seven α -like and seven β -like subunits, have been identified and cloned (HILT *et al.* 1994). Biochemical and molecular analyses of mammalian proteasomes suggest that their subunit compositions are similar to those of yeast, with each proteasome being comprised of seven different α -like and seven distinct β -like subunits (HENDIL *et al.* 1993; TANAHISHI *et al.* 1993). Within each of these two gene families, the different members are recognizably related, but the degree of amino acid identity between the different subunits is not high (*i.e.*, <30%). Proteasomes exhibit up to five endopeptidase activities. Mutational and inhibitor studies suggest that these are conferred by β subunits, with the active sites being situated within the proteasome's innermost chamber (HEINEMEYER *et al.* 1991; CHEN and HOCHSTRASSER 1995; LÖWE *et al.* 1995; SEEMÜLLER *et al.* 1995). None of the α subunits have been implicated in a direct hydrolytic function, and it is thought that they play structural or regulatory roles. Additional components comprising a 19S complex that caps the 20S proteasome at each end, yielding the 26S structure, are responsible for conferring ubiquitin-dependence, and are thought to play an important role in unfolding substrates and directing them into the proteasome's degradative tunnel (see JENTSCH and SCHLENKER 1995 for review).

The available evidence suggests that in yeast there is a homogeneous population of 20S proteasomes, each comprised of the same 14 subunits (CHEN and HOCHSTRASSER 1995). However, in metazoans a number of biochemical studies have indicated that there exist subpopulations of proteasomes that differ slightly in their subunit compositions (HAASS and KLOETZEL 1989; AHN *et al.* 1991). To further investigate the extent and nature of structural heterogeneity among proteasomes, and to begin to study its possible functional and biological significance, we sought to identify proteasome subunit isoforms that exhibit developmental or cell-type specificity. In this paper, we report the isolation and characterization of two proteasome subunit genes in *Drosophila* that exhibit spermatogenesis-specific expression patterns. These represent two of the most striking examples of developmentally regulated proteasome subunits reported to date in any system.

MATERIALS AND METHODS

General procedures: All standard techniques (*e.g.*, DNA extraction, restriction digestion, plasmid and phage DNA isolation, Southern blots, Northern blots, etc.) were done as described in SAMBROOK *et al.* (1989). The *Escherichia coli* strains used for propagating phage and plasmids were Q359 and DH5 α , respectively. Subcloning was done using the plasmid vector pGEM-3 blue (Promega), unless otherwise noted. Flies were cultured on standard *Drosophila* media at 25°.

Genomic and cDNA library screening: A DNA fragment containing the *Pros28.1* gene was synthesized from genomic DNA using PCR. Oligonucleotide primers were designed

based on the published sequence (FRENTZEL *et al.* 1992) and were obtained from Genosys Biotechnologies, Inc. (The Woodlands, Texas). The sequences of the primers used were as follows: PROS28.1-5', GGGATCCGCTACGATCGTGCT and PROS28.1-3', GGGATCCGTTCTCCATGATGG. The amplification product was gel purified using the GeneClean Kit (BIO 101, Inc.) and digested with *Bam*HI to generate 377- and 415-bp fragments containing the 5' and 3' regions of the gene, respectively. These were subcloned into pGEM-3 and subsequently radiolabeled using the Random Primed Labelling Kit (Boehringer-Mannheim) for use as hybridization probes. A *D. melanogaster* genomic library, constructed in λ EMBL3 using DNA from a *th st tra cp in ri p^h/TM3* strain (J. M. BELOTE and M. MILLER, unpublished), was screened under reduced-stringency conditions using the following hybridization solution: 39% formamide, 5 $\times$ SSPE, 1 $\times$ Denhardt's solution, 0.1% SDS, and 0.2 mg/ml heat-denatured salmon sperm DNA. Following overnight hybridization at 42°, filters were washed three times in 0.1% SDS, 0.5 $\times$ SSPE at 50°, air-dried and subjected to autoradiography. Each positive plaque was plated at low plaque density and rescreened, and liquid phage stocks were established from single plaque isolates. Phage DNA was extracted and subjected to restriction mapping. Southern blot analyses were used to identify the restriction fragments containing the cross-hybridizing sequences, and these were subcloned into a plasmid vector for further analyses.

Appropriate restriction fragments from genomic clones were radiolabeled and used to screen late pupal and adult male cDNA libraries constructed in λ gt10 by L. KAUVAR (POOLE *et al.* 1985). Positive plaques were isolated, phage stocks prepared, and the *Eco*RI inserts subcloned into pGEM-3 for sequencing.

DNA sequencing: Double-stranded DNA templates were sequenced by the dideoxynucleotide chain termination method (SANGER *et al.* 1977) using ³⁵S-dATP (New England Nuclear) and Sequenase (U.S. Biochemicals) DNA polymerase. Primers for the sequencing reactions were either the T7 or SP6 promoter oligonucleotides purchased from Promega, or in some cases, gene-specific oligonucleotides from Genosys, Inc. or Gibco-BRL. The sequences reported here have been deposited in the GenBank database (accession numbers U46008 and U46009).

In situ hybridization: Salivary glands from late third instar larvae were dissected in 0.7% saline, incubated for 10 sec in 45% acetic acid, and placed in a drop of 1 part lactic acid:2 parts H₂O:1 part acetic acid for 5 min. A siliconized cover slip was then applied and the glands were squashed. After freezing the slide in liquid nitrogen, the cover slip was removed and the chromosomes dehydrated through an ethanol series. Hybridization of the chromosomes was carried out according to the procedure of PLILEY *et al.* (1986). Probes were prepared for *in situ* hybridization by random primed labeling of gel purified restriction fragments. The modified deoxynucleotide used was biotinylated dUTP from Gibco-BRL, and signal detection was done using a Detek-I-hrp Kit (ENZO) following the procedure of ASHBURNER (1989).

Northern blot analysis: PolyA⁺ RNA was isolated and electrophoresed through a formaldehyde agarose gel as described in SAMBROOK *et al.* (1989). The RNA was electroblotted onto GeneScreen (Dupont) nylon membrane and probed with random primed ³²P-labeled restriction fragments specific for each of the three *PROS28.1*-related genes. The membrane filter was initially probed with the *PROS28.1A*-specific probe. Following autoradiography, counts were removed from the filter by treating with boiling TE. The filter was then autoradiographed to confirm that there was no residual radioactivity remaining. The filter was then hybridized with the

PROS28.1B-specific probe, and following autoradiography and removal of the hybridized counts, the filter was probed with the *PROS28.1*-specific probe. Finally, the filter was hybridized with a probe specific for the ribosomal protein gene *rp49* (O'CONNELL and ROSBASH 1984) to confirm that all lanes contained similar amounts of intact RNA.

Construction of reporter genes: For the construction of the *Pros28.1B-lacZ* reporter gene, a 3.1-kb *EcoRI* fragment containing *Pros28.1B* was first subcloned into the plasmid vector p34H (TSANG *et al.* 1991) to yield p34H/P28.1B. *E. coli lacZ* gene sequences were cut out of the pOX1 vector with *BamHI* and *HindIII*. After subcloning this fragment into pEMBL8+, the *HindIII* site was filled in using Klenow polymerase and a *SaII* linker added. The 3.3-kb *BamHI/SaII* fragment was then isolated, treated with Klenow polymerase, and ligated into the *SmaI* site of p34E (TSANG *et al.* 1991) to yield plasmid p34E/*lacZB*. The *lacZ*-containing fragment of p34E/*lacZB* was released by digestion with *BamHI* and ligated to p34H/P28.1B, which had been cut with *BamHI* to release the internal *BamHI* fragment. The resulting subclones were checked by restriction digest analysis for those in which the *lacZ* sequences were in the proper orientation with respect to *Pros28.1B*. One such clone, p34H/P28.1B-lac, was digested with *EcoRI*, and the P28.1B-lac fragment was subcloned into the transformation vector pW8 (KLEMENZ *et al.* 1987) to yield pW8/P28.1B-lac.

For the construction of the *Pros28.1A-lacZ* reporter gene, a 5.7-kb *HindIII/SaII* fragment containing *Pros28.1A* was first subcloned into the plasmid vector pGEMEX1 (Promega) to give pGEMEX/P28.1A SH. This plasmid was linearized by digestion with *BamHI* and ligated to the 3.3-kb *lacZ*-containing *BamHI* fragment of p34E/*lacZB*. Subclones were tested by restriction digest analysis to identify those in which the *lacZ* gene was in the correct orientation with respect to *Pros28.1A* and one of these, pGEMEX/P28.1A-lac, was selected for further manipulation. The 9-kb *HindIII/SaII* fragment from pGEMEX/P28.1A-lac was isolated, treated with Klenow to fill in the ends, and ligated into the *SmaI* site of pGEM-3 to yield pGEM/P28.1A-lacSHK. The 9-kb *EcoRI/SaII* fragment was then isolated and ligated to *EcoRI/XhoI*-digested pW8, yielding pW8/P28.1A-lac, which was used for the transformation experiments.

P-element-mediated germline transformation: Microinjection of *Drosophila* embryos was done using standard methods (SPRADLING 1986; ASHBURNER 1989). The pW8/P28.1A-lac and pW8/P28.1B-lac plasmids were prepared using the CsCl density gradient centrifugation method. A mixture of reporter gene plasmid (300 µg/ml) and helper plasmid p π 25.7wc (100 µg/ml) was injected into preblastoderm embryos of the host strain, *w; P[ry⁺(Δ 2-3)](99B)* (ROBERTSON *et al.* 1988), which contains an additional source of *P*-element transposase. For each reporter gene, a minimum of four independent transformants were selected and, after replacement of the *P[ry⁺(Δ 2-3)](99B)*-bearing chromosome to stabilize the transposon, the chromosome containing the reporter gene was determined by linkage analysis.

X-Gal staining: Flies were dissected in *Drosophila* Ringer's or PBS, fixed in 1% glutaraldehyde in 50 mM sodium cacodylate, and stained for β -galactosidase expression using the method of GÖNCZY *et al.* (1992). The tissues were rinsed with staining buffer (10 mM NaH₂PO₄/Na₂HPO₄ pH 7.2, 1 mM MgCl₂, 150 mM NaCl, 5 mM K₄[Fe^{III}(CN)₆] and 5 mM K₃[Fe^{III}(CN)₆]), and then incubated for 30 min at room temperature in staining buffer. Five hundred microliters of staining buffer was prewarmed to 37°, and 10 µl of 10% 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal) in dimethylformamide was added. The tissues were incubated at 37° in this solution for 2–12 hr and then washed in PBS plus 1 mM EDTA to stop the staining reaction. The tissues were

dehydrated through an ethanol series, mounted in a 90% glycerol:10% PBS solution under a bridged coverslip, and examined under either bright field or Nomarski optics using a Zeiss Axioplan microscope.

RESULTS

Isolation of *Pros28.1*-related genes of *D. melanogaster*:

A *Drosophila* cDNA clone encoding a 28-kD proteasome α -like subunit was previously isolated from a λ gt11 expression library by HAASS *et al.* (1990), using a proteasome-specific antibody probe, and the corresponding gene, called *Pros28.1*, was subsequently cloned from a genomic library by FRENTZEL *et al.* (1992). Since affinity-purified antisera against PROS28.1 protein exhibited weak cross-reaction with more than one 28-kD subunit in proteasome preparations, and since, in Southern blots, the *Pros28.1* probe showed additional, though weak, hybridization signals even under high-stringency conditions, it was proposed that this gene is a member of a gene family encoding closely related proteasome subunits (HAASS *et al.* 1990). To investigate the nature of these *Pros28.1*-related sequences, we isolated them by screening a genomic library with a *Pros28.1* probe under reduced stringency hybridization conditions. In this experiment, eight recombinant phage were isolated, out of 75,000 screened. Restriction mapping analysis placed these into three classes (Figure 1). The first is defined by a phage that strongly hybridizes to the probe and that has a restriction map consistent with that previously described for the *Pros28.1* gene itself (FRENTZEL *et al.* 1992). The second class is comprised of six overlapping phage that contain a region that cross-hybridizes with moderate strength with the probe and that have restriction maps indicating that they identify a locus distinct from *Pros28.1*. The third class of recombinant phage has a single member, and, like those of Class II, it hybridizes to the probe with moderate strength. The restriction map of this phage indicates that it represents a third locus containing *Pros28.1*-related sequences.

The three genomic regions identified by the recombinant phage were localized on the chromosome map by *in situ* hybridization to larval salivary gland chromosomes. The results reveal that the three loci are unlinked. As expected, the phage from class I hybridizes to region 14B4, the site on the X chromosome previously reported to be the cytogenetic position of *Pros28.1* (FRENTZEL *et al.* 1992). Class II phage hybridize to the right arm of chromosome 3, at position 92F. This locus will be referred to as *Pros28.1A*. The class III phage hybridizes to the right arm of chromosome 2, at 60D7, and this locus will be referred to as *Pros28.1B*. When additional *in situ* hybridizations were carried out, using *Pros28.1* sequences as probes and using slightly reduced stringency hybridization conditions, it was noticed that, in addition to the strong signal at 14B4, there were also weak hybridization signals seen at 92F and 60D7. No

CLASS 1

λ Dm4.1

CLASS 2

λ Dm13.1

CLASS 3

λ Dm12.1

FIGURE 1.—Recombinant λ phage containing *Pros28.1*-related sequences. The open and closed rectangles represent the λ phage arms, and the lines correspond to the *D. melanogaster* insert DNA. The bold lines indicate restriction fragments that hybridize to the *Pros28.1* probe. E, *EcoRI*; B, *BamHI*; H, *HindIII*; S, *SalI*.

other sites showing weak hybridization signals were evident.

DNA sequence analyses of the *Pros28.1* gene family:

For each of the three classes of phage, the region that hybridized to the probe was identified by Southern blot analysis, and relevant restriction fragments subcloned into a plasmid vector to facilitate sequencing. In addition, cDNA clones corresponding to *Pros28.1* and *Pros28.1B* were isolated as detailed in MATERIALS AND METHODS. Despite numerous attempts using several different libraries, we failed to isolate any cDNA clones corresponding to *Pros28.1A*. The reason for this is not known, since Northern blot analysis (see below) indicated that this gene does produce polyadenylated transcripts in pupae and adult males, albeit at a low level of abundance.

Class I clones were partially sequenced to confirm that they really did correspond to *Pros28.1*, and no discrepancies were found between this and the previously published sequence (HAASS *et al.* 1990; FRENTZEL *et al.* 1992).

Genomic clones of *Pros28.1A* were sequenced, and the results depicted in Figure 2A. There is an open reading frame (ORF) of 249 codons, the exact size of the *Pros28.1* ORF. One notable feature of the *Pros28.1A* gene, when compared to *Pros28.1*, is its lack of introns. While most reported examples of intronless gene duplications involve nonfunctional pseudogenes, DNA sequence comparisons between *Pros28.1* and *Pros28.1A*, taken together with the observation that *Pros28.1A* is

transcribed and translated (see below), strongly suggest that this is a functional gene.

Comparison of the deduced protein products of *Pros28.1* and *Pros28.1A*, shown in Figure 3, shows a high degree of amino acid identity (74%), with the amino-terminal two-thirds being particularly well-conserved. Comparison of the DNA sequences within the coding regions show 74% identity at the nucleotide level. The extensive differences in the nucleotide sequences in the 5' and 3' noncoding regions (34 and 33% identity, respectively), and the numerous instances of synonymous substitutions (*i.e.*, silent, third position changes) within the coding region indicate that this gene duplication is not extremely recent and further suggest that natural selection has acted to maintain a complete ORF.

Genomic and cDNA clones of *Pros28.1B* were also sequenced, and the results are depicted in Figure 2B. This gene also potentially encodes a protein of 249 amino acids, with strong sequence similarity to PROS28.1 (Figure 3). Within the coding regions, *Pros28.1* and *Pros28.1B* share ~72% nucleotide identity while noncoding regions are poorly conserved (<40% identity). Unlike *Pros28.1A*, the *Pros28.1B* gene structure resembles that of *Pros28.1*, with the two introns being positioned in exactly analogous positions, after codon 34 and within codon 111. Again, the amino-terminal two-thirds is particularly well-conserved, although the overall similarity (58% identity) is less than that observed between PROS28.1 and PROS28.1A. As is the case with *Pros28.1A*, the maintenance of a 249-amino acid ORF and the high

A

```

taatcattacactatgtgtagaacaaaggatttgacttcaaacttcaaattgaagaaattcatattgtttcaagtaataacttgtcacaagcaaatgtg 100
      M S S R Y G R A L T I F S P D G H L L Q V E Y
tgttttgttttgcctgctaactaaccacaaaggATGTCTTCACGCTACGGTCGAGCTCTAACCATATTTTCGCCGGATGGCCACCTGTTGCAGGTGGAGTA 200
A Q E A V R K G S T A V G V R G A N C V V L G V E K S S V S E M Q
CGGCCAGGAAGCTGTTGAAAAGGGCTCCACAGCGGTTGGCGTTCGTGGAGCCAATTGTGTGGTGCCTGGGTGTGAGAAGAGCTCGGTGTCCGAGATGCAG 300
R D R T V R K I S M L D R H V A L A F A G L T A D A R I L I N R G Q
GAGGATCGGACGGTGCACAAGATCAGCATGCTCGATCGCCACGTGGCGTTGGCATTGTCCGGTCTCACGGCTGATGCACGGATCCTTTATAAATCGAGGGC 400
V E C Q S H R L N F E N Q V T L E Y I T R Y L A Q L K Q K Y T Q V
AGGTCGAGTGCCAGAGCCATCGACTCAACTTCGAAAACAGGTGACCTCGAGTACATTACAGATACCTTGCCAGCTCAAGCAGAAGTACACGCAGGT 500
N G R P S F G I S C L I G G I D A D G S A R L F A T E P S G I F H
CAATGGACGTCGCTTTTCGGTATATCGTGTCTAATAGGAGGCATTGACGCCGATGGCTCTGCACGCTCTCTTCATACCGAACCATCCGGATATTTCCAC 600
E Y K A T A T G R W A N T V R E F F E K A Y S D H E V T T K C D A I
GAGTACAAGGCGACCGCAACTGGGCGTGGGCAAAATCTGTGAGGGAGTTTGTGAGAAGCCCTATAGCGATCATGAAGTGACAACCAATGTGATGCCA 700
K L A M R A L L E V T Q M S Q M R L E V A V L E N G K P M K M L D
TCAAATAGCCATCGCGGCCCTTCGCAAGTAACCCAGATGAGCCAGATCGCCCTGGAGGTGGCCGATTTGGAGAATGGCAAGCCAATGAAAATGCTGGA 800
S V V I S E I V K I V Q N E K E L Q A K A H K M K R *
CAGTGTAGTAATTTTCGAAAATAGTCAAGATCGTTCAAGAACGAAAGAGCTCCAAGCGAAGGCACATAAAATGAAGAGATgaattgtcacctacagat 900
cgttgattttttatgttcatcaaatgttattagatattcatgttatattgttttgaggctttgtttaaatgtttgtttgtttgtttttaaatta 1000
attaatatagcagtttaaacatgttttgatgtttgtgaaggattttatctttgtataatgcagtaattaattaagattacgactttaaaattataacttt 1100
tgaatttcaaatatgcagatataaagcactcactctaacgccacgaactgtcgtgcaggtcaccaggatagcaggtaca 1182

```

B

```

atggattcagattatgtagatttgcctttgtagcggattacctcattatataaaaaatgatttaattgcataacagtttgagaggacactaatagaaaaa 100
taaaatataaaataaagctagtaaaaagtacattttatgataaacgaattgaagaaaaaatgcaatcaaaacaatatgtcttcaaaaggggtaattatgt 200
tttcataaaatttggttaaaagtggatttccagtagcgcactagcagggcgttttagaagtagcagagataactaaaatttcgatcttatacacaaagtacact 300
      M A H G Y D R A V T I Y S P D G H L L Q V E Y
ccatagattttcaataagtagccccacttcaATGGCTCACGGATACGATCGCGCGGTGACCATTCTACACCGACGGTCACCTGCTGCAGGTGGAGTA 400
A Q E A V R R G S T V
CGCCACGAGGCGCTCCGAAGGGTTCCACTGTGtaagtctacctggaggccaccgcacttgccattagtttaactctcaataactaaccacaaagAT 500
G L R T N N A I V I G V E K R S V G D L Q E E R M V R K I C M L D
GGGTCTGCGCACCAACACGCCATTGTAAATCGGCGTGGAGAAGCGCTCCGTGGCGGACCTGTCAGGAGGAGCGCATGGTGCAGCAAGATTGCAATGCTCGAC 600
D H V V M T F S G L T A D A R I L V S R A Q M E A Q S H R L N F E K
GACCACGTCGTGATGACTTTTCGGGTCTCACCGCGGACGCACGGATCTGTGTAAGCCGCGCCCAATGGAGGCCAGAGCCATCGTCTGAATTTTCGAGA 700
P T T V E Y I T R
AGCCACGACCGTGGAGTACATAACCGgtgaggaaggagcactgcctacacagggaagaatgagcagtgatgagctgaggaattgcaaaccaaaaat 800
actcatgcatacttttctgcctttcgaaaaagtttaagaactgtctctcagtgcatccccctaggcaaatgttgatttaccttttatgtctcttccg 900
Y I A Q L K Q N Y T Q S N G R R P F G L S C L V G G F D E
cttccatgcagATACATAGCCAGCTGAAACAGAACTACACGCAGAGCAATGGGAGCGTCCATTGGATTGTCTGTCTGGTTGGGGGCTTCGACGAG 1000
D G T P H L F Q T D P S G G I F Y E W R A N T T G R S S Q P V R D Y M
GATGGCATCCGACCTGTTCACAAACGGATCCCTCGGGCATCTTCTACGAGTGGAGGCCAACACACCGGCGCTCAAGCCAGCCGGTTCGCGATTACA 1100
E K H A D E I L T I A D E R G I K H I V R T L V S V S S L N H T Q
TGGAGAAGCACCGGAGAGATCTGACCATCGCCGACGAGCGCGCATCAAGCACATAGTCCGCACCTGCTCAGCGTGTCTCGCTCAACCACTCA 1200
M E V A V L K Y R Q P L R M I D H Q V L A D L E R T V R R E I E D
AATGAGGTGGTGTGCTCAAGTACCGCAGCCACTGCGTATGATTGATCACCAGGTCTGTGGCGATCTGGAGCGAACCCTTCGACGGGAAATTGAAGAC 1300
E P S E S A R A P *
GAGCCGAGCGAGTCGGCCCCGTCTCCCTagttttggcggccaccctgacccattctgttgcgaatggccctgagttcatgaccaatcaaaagttgata 1400
accggtcgtgctctcttttctggcaattcaatttcggccggggagttggagaaaagtaatcagcagggaataatgtgtaagcgggcaaatcacttaatt 1500
tgtctccaggcgctatttgtttactttatccctctcccgctagcgtacgctcttcagccaaccaagttaaaaagttttaagtgaaaatgcgaattgaa 1600
ttgtagaggaaaaataaagaatgcgccacacatgcagaaaatgtttgtcaatcgagatggtttcacttgattgggaataaaatcctagatgccaaggaat 1700
tgctccttgatcctgatgaaagaacattgttttctgaaggcatctttacattaaa
.....

```

FIGURE 2.—DNA and amino acid sequences of two new *Pros28.1*-related genes. (A) Sequence of *Pros28.1A*. (B) Sequence of *Pros28.1B*. Nucleotides written in lowercase represent noncoding regions, while uppercase letters indicate coding sequences. The dots lie under the polyadenylation site of *Pros28.1B*, as determined from the sequence of the cDNA. The underlined sequence is the consensus polyadenylation signal sequence. The numbers refer to the nucleotide positions at the end of each line.

ratio of synonomous to nonsynonomous substitutions when *Pros28.1* and *Pros28.1B* are compared, together with the finding that *Pros28.1B* is transcribed and translated (see below), argue that it, too, is not a pseudogene.

Expression of the *Pros28.1*-related genes during development: The above results demonstrate the existence of two new *Pros28.1*-like genes that appear to encode isoforms of PROS28.1, and that are not closely

P28.1	MSSRYDRAVTIFSPDGHLLQVEYAQEA VRKGS TAVGVRGANCVVLGV EKK	50
P28.1A	MSSRYGRALTIFSPDGHLLQVEYAQEA VRKGS TAVGVRGANCVVLGV EKS	50
P28.1B	MAHG YDRAVTIYSPDGHLLQVEYAQEA VR RGS T VMGLRTNNAIVIGVEKKR	50
	SVAQLQEDRKVRKICMLDNHVVMAFAGLTADARIMINRAQVE CQSHRLNV	100
	SVSEM QEDRTVRKISMLDRHVALAFAGLTADARILINRGQVE CQSHRLNF	100
	SVGDLQEE RMVRKICMLDDHVVMTFSGLTADARILVSR AQME AQSHRLNF	100
	EDPVTLEYITRFIAQLKQKYTQSNGR RPF GISC LIGGF DADGSAHLFQTE	150
	ENQVTLEYITRYLAQLKQKYTQVNGRPSFGISC LIGGIDADG SARLFATE	150
	EKPTTVEYITRYIAQLKQNYTQSNGR RPF GLSCLVGGFDE DGT PHLFQTD	150
	PSGIFYEYKANATGRSAKV VREFF EKSYREEEVANEHGAVKLAI RALLEV	200
	PSGIFHEYKATATGRWANTVREFF EKAYS DHEVTTKCDAIKLAMRALLEV	200
	PSGIFYEWRANTTGRSSQPVRDYMEKHADEILTIAD ERGIKHIVRTLVS	200
	AQSGQNNLEVAIMENGKPLKMLD TDVITDYVKIIEKEKEE ELEKKKQKK *	249
	TQMSQMRLEVAVL ENGKPMKMLDSVISEIVKIVQNEKELQAKAHKMKR *	249
	SSLNHTQMEVAVLKYRQPLRMIDHQLADLERTVRREIEDEPSESARAP *	249

FIGURE 3.—Comparison of the deduced amino acid sequences of the PROS28.1, PROS28.1A, and PROS28.1B proteins. Shaded areas represent amino acid identities between two or more of the sequences. The sequence of PROS28.1 is taken from FRENTZEL *et al.* (1992) and is available from the GenBank database, accession number X62286.

related to any of the other α -like subunits that make up the eukaryotic proteasome. One major question that arises is “Are these isoforms expressed in a stage- or cell-type-specific manner that might be related to some specialized function of the proteasome?” As the first step toward addressing this question, we carried out Northern blot analysis on RNA extracted from individuals of different developmental stages.

In these experiments, *Pros28.1* transcripts were detected throughout most of development, with their steady-state levels being very high during the first few hours of embryogenesis, and then showing moderate fluctuation during the larval, pupal and adult stages. The observed pattern of expression is consistent with the previously reported Northern blot results of HAASS *et al.* (1990) and is similar to what is seen with another *Drosophila* proteasome subunit gene, *l(3)73Ai* encoding a β -type subunit, that has been examined by developmental Northern analysis (SAVILLE 1991; SAVILLE and BELOTE 1993). The expression patterns of *Pros28.1A* and *Pros28.1B*, however, are strikingly different, with transcripts being detected only during the pupal and adult stages. Examination of the sex-specificity of this pupal and adult expression gave the surprising result that both genes are expressed in a male-specific manner (Figure 4), while *Pros28.1* RNA is found in both sexes. The levels of abundance of these male-specific RNAs are about an order of magnitude lower than that of *PROS28.1*, as estimated by the exposure times required to give similar signals on the autoradiographs. The estimated sizes of the *PROS28.1* and *PROS28.1A* transcripts are similar, ~ 1.1 kb, while that of *PROS28.1B* is detectably larger, ~ 1.3 kb, consistent with the observed size of the *PROS28.1B* cDNA described above.

Testes-specific expression of *Pros28.1A* and *Pros28.1B*

reporter genes: To examine in more detail this unusual male-limited expression of these *Pros28.1*-related genes during the pupal and adult stages, we constructed reporter genes that fused the *E. coli lacZ* coding region in-frame with the *Pros28.1A* or *Pros28.1B* protein coding regions, and, using *Pelement*-mediated germline trans-

FIGURE 4.—Northern blot analysis of the *Pros28.1*-related genes. Poly-A⁺ RNA was isolated from male and female pupae and adults, and subjected to Northern blot analysis as described in MATERIALS AND METHODS. The same filter was used for all three blots, in the following order of probeds: *Pros28.1A*, *Pros28.1B*, then *Pros28.1*.

FIGURE 5.—Testes-specific expression of *Pros28.1A* and *Pros28.1B* as seen by X-gal staining of transgenic flies carrying *lacZ* reporter genes. (A) Dissected reproductive system from a transgenic male carrying a *Pros28.1A-lacZ* reporter gene stained with X-gal. Reporter gene expression is seen as blue staining. (B) Expression of the *Pros28.1A-lacZ* reporter gene in primary spermatocytes at the 16-cell cyst stage. Elongating spermatid bundles also show X-gal staining. (C) Dissected reproductive system from a transgenic male carrying a *Pros28.1B-lacZ* reporter gene stained with X-gal. (D) Expression of the *Pros28.1B-lacZ* reporter gene as seen by X-gal staining of elongating spermatid bundles. Earlier stage spermatocytes do not appear to stain. ag, accessory gland; eb, ejaculatory bulb; sc, spermatocyte cyst; st, elongating spermatid bundles; t, testis.

formation, produced transgenic flies carrying each reporter construct. Tissue-specific localization of the reporter gene activities were detected by staining dissected tissues with the β -galactosidase chromogenic substrate, X-gal. In both cases, expression of the reporter genes appeared to be limited to the testes.

The *Pros28.1A-lacZ* reporter gene had ~ 3.5 kb of DNA upstream of the transcribed region and ~ 2 kb of DNA downstream of the coding region. Four transformed lines were produced and stained for β -galactosidase activity using X-gal. All of the lines carrying *Pros28.1A-lacZ* showed similar patterns of reporter gene expression (Figure 5, A and B). The X-gal staining first appears during the stage at which the primary spermatocytes exist in cysts of 16 cells. The stain persists during the following two meiotic divisions and begins to fade as the spermatid bundles elongate. As spermatids individualize, a cystic bulge forms and passes down the spermatid bundle, collecting discarded cytoplasm as it progresses. These cystic bulge regions were sometimes

noticeably blue, but there was no detectable staining of mature individualized sperm.

The *Pros28.1B-lacZ* reporter gene had ~ 1.2 kb of DNA upstream, and ~ 1.4 kb of DNA downstream of the polyadenylation site. Five different transformed lines were generated, and the expression of the reporter was determined by staining dissected flies with X-gal. While *Pros28.1B-lacZ* was also expressed in the testes, the staining pattern was different than that seen with *Pros28.1A-lacZ*. As shown in Figure 5, C and D, flies carrying the *Pros28.1B-lacZ* reporter showed very intense X-gal staining in many of the elongating sperm bundles within the testes. The X-gal staining was uniform along the length of the spermatid bundles, although not all spermatid bundles were stained. These unstained spermatid bundles might represent a more advanced stage, since there was no staining of mature sperm found at the base of the testes and in the vas deferens. Unlike what was seen with *Pros28.1A-lacZ*, in this case there did not appear to be any appreciable staining of preelongation

spermatocytes. This late spermiogenesis-specific staining pattern was seen in four of the five transformed lines examined. A fifth transformed line failed to show any X-gal staining, even though the expression of the w^+ eye color marker encoded by the pW8 vector was quite strong. It is assumed that the lack of reporter gene expression in this line is either due to a mutation in the transgene or to some type of chromosomal position effect that is overriding the regulatory elements carried on the transposon.

That the observed staining patterns of *Pros28.1A-lacZ* and *Pros28.1B-lacZ* are due to expression of the reporter transgenes and are not the result of endogenous β -galactosidase activity is shown by the fact that control flies, lacking a transgene, fail to show any testes staining. In addition, the observed staining patterns were confirmed by immunostaining testes with antibody specific for the *E. coli* β -galactosidase enzyme (not shown).

Testes are comprised of two general types of cells: germline cells that are present as different staged spermatocytes and elongating spermatid bundles, and somatic cells, including the cyst cells that encompass the groups, or cysts, of developing germ cells (see LINDSLEY and TOKUYASU 1980 and FULLER 1993 for reviews of *Drosophila* spermatogenesis). While it is apparent that both *Pros28.1A-lacZ* and *Pros28.1B-lacZ* reporter genes are expressed in germline cells (spermatocytes and spermatids), it is unclear if there is any expression of these genes in any somatic component of the testes. To investigate this possibility, germlineless flies were generated using the maternal effect mutant, *tudor* (*tud*) (BOSWELL and MAHOWALD 1985). Females homozygous for *tud* were crossed to males carrying either *Pros28.1A-lacZ* or *Pros28.1B-lacZ* to generate transgenic males that developed rudimentary testes lacking germ cells. The reproductive systems from such males were dissected and stained with X-gal to see if any of the somatic components of the testes were expressing these reporter genes. In neither case was there any X-gal staining in these agametic testes, indicating that the reporter gene expression in normal testes is germline-dependent.

DISCUSSION

One major question in proteasome research concerns the extent and nature of proteasome heterogeneity. That is, are proteasomes homogeneous in their subunit composition, or do there exist proteasome subtypes that carry special subunits that modify the functional properties of proteasomes in certain cells? In the yeast *S. cerevisiae* the available evidence suggests that all 20S proteasomes contain the same 14 subunits, at least in vegetatively growing cells. For example, only 14 different proteasome subunit genes have been identified and cloned from yeast, and interspecific sequence comparisons indicate that they encode seven distinct α -type and seven distinct β -type subunits (HILT *et al.*

1994). If yeast proteasomes existed in multiple forms with varying subunit compositions, then one might expect to find more genes encoding additional subunits. Other evidence suggesting that yeast proteasomes are uniform in their subunit structure comes from studies of strains carrying proteasome subunit mutants. CHEN and HOCHSTRASSER (1995) analyzed the biochemical properties of yeast proteasomes from strains carrying two different proteasome subunit mutants and found that the chromatographic properties of all detectable proteasomes in these strains were drastically altered when compared to wild type. Moreover, further analysis of the proteasomes displaying the aberrant anionic elution properties showed that they were comprised of the same 14 known proteasome subunits found in wild-type cells. If different populations of proteasomes did exist in yeast cells, then one would not expect all to have identically altered chromatographic profiles in both mutant strains.

But while yeast proteasomes seem to be homogeneous in their structure, the situation in metazoans is more complex. Evidence that proteasome heterogeneity exists comes primarily from biochemical studies. For example, in *Drosophila* proteasomes isolated from different developmental stages exhibit detectable differences in their subunit compositions, as seen in 2D gels (HAASS and KLOETZEL 1989). Similarly, in chick embryonic muscle cells the proteolytic activities of proteasomes change during development, and this is correlated with a developmental- and tissue-specific alteration in the pattern of proteasome subunits (AHN *et al.* 1991). In cultured mouse cells, it has been found that different subpopulations of proteasomes can be isolated on the basis of their chromatographic properties, and enzymatic analyses of the different subtypes indicate that they have altered peptide-hydrolyzing activities (SEELIG *et al.* 1994). In none of these cases, however, is the exact nature of these differences known. They could represent posttranslational modifications of the typical 20S subunits, or they might represent expression of subunit isoforms that are encoded by additional genes. It is also not known what the biological significance of these differences might be, if any.

There is, however, one well-documented example of a specialized proteasome subtype in mammals and that concerns the role that proteasomes play in antigen presentation (reviewed in GACZYNSKA *et al.* 1994). That is, in response to the antiviral cytokine γ -interferon, two special β -type subunits, encoded by the genes *LMP2* and *LMP7* of the major histocompatibility complex (MHC), are synthesized. These subunits replace two constitutive proteasomal subunits to yield a proteasome with altered proteolytic properties (DRISCOLL *et al.* 1993; GACZYNSKA *et al.* 1993; MICHALEK *et al.* 1993). These so-called "immunoproteasomes" are responsible for the hydrolysis of intracellular (viral) antigens into small peptides, which are then bound to MHC class-I

molecules before their presentation to cytolytic T lymphocytes. It has been postulated that these reconfigured proteasomes are better suited for producing peptides of the appropriate size and ends for more efficient transport and presentation to the cell surface during the antiviral immune response (GACZYNSKA *et al.* 1994).

Of interest here is whether there might exist other examples of subunit heterogeneity among proteasomes that might underlie functional differences. The present study was aimed at examining the extent and nature of proteasome heterogeneity in *Drosophila* by isolating additional genes that potentially encode proteasome subunit isoforms. The *Pros28.1A* and *Pros28.1B* loci described here represent two genes that encode proteins closely related to the α -type subunit PROS28.1, but that exhibit testes-specific expression. The maintenance of long and complete ORFs that are well-conserved in their deduced amino acid sequences over the entire protein-coding regions, taken together with the finding that both genes are transcribed and translated, argues strongly that these represent functional genes. Moreover, the finding that both genes, which appear to have arisen through independent gene duplication mechanisms (see below), exhibit testes-specific expression patterns suggests that proteasome function during spermatogenesis might have special requirements. Of extreme interest, of course, is what functional significance do these testes-specific genes have. During spermatogenesis gonial precursor cells undergo dramatic changes in cell shape, with mitochondrial and nuclear remodelling and elimination of much of the cytoplasm, en route to becoming mature spermatozoa. It is likely that these extensive morphological changes are accompanied by a large amount of protein degradation. The induction of these testes-specific proteasome components during spermatogenesis might be related to a phenomenon that has been observed in mammalian systems, where another component of the ubiquitin/proteasome pathway, the ubiquitin-conjugating (E2) enzymes, have been found to be highly expressed in rat testes (WING and JAIN 1995). A similar case of induction of ubiquitin-conjugating enzymes in terminally differentiating erythroid cells has also been reported (WEFES *et al.* 1995), and it has been proposed that in both of these cases this induction leads to alterations in the level or specificity of proteolysis, which plays an important role in bringing about the dramatic changes in cell morphology that accompany cell differentiation of spermatids or reticulocytes, respectively.

While this is yet to be tested, it is possible that synthesis of PROS28.1A and PROS28.1B results in restructured proteasomes that are particularly efficient at carrying out the degree and type of proteolytic activity characteristic of spermatogenesis. A more specific form of this hypothesis is that there might be specific proteasome substrates that must be degraded in a timely fashion during spermatogenesis, and that proteasomes con-

taining these testes-specific subunits are better suited for this purpose than the usual somatic proteasomes.

In the case of the "immunoproteasomes" discussed above, the subunits that are being replaced are β subunits that act directly in the proteolytic function of the 20S proteasome. Since the exact roles of the α subunits within the proteasome are not known, it is hard to predict what mechanistic changes, if any, these testes-specific subunits would confer upon proteasomes that contain them. It is possible that 20S proteasomes containing these α subunits would differ in their efficiency or rate of proteolysis, due to differential interactions with components of the 19S cap complex or other regulators, or because of slight alterations in the 20S proteasome's architecture. However, since the amino acid differences between the somatically-expressed PROS28.1 subunit and the two testes-specific isoforms are scattered throughout the protein, it is not possible to pin down a single domain of the 20S proteasome's predicted quaternary structure that might be critically affected by substitution of these α subunits.

Whatever the functional significance of these testes-specific subunits might be, this question can be best addressed by future mutational studies. For instance, it would be of interest to define the phenotypic effects of *Pros28.1A* and *Pros28.1B* mutants. (For example, are they male sterile mutants? If so, is the phenotypic effect on spermatogenesis informative? Do such mutants cause changes in the 2D-gel profile of testes-specific proteins that might identify putative substrates for proteasome-mediated proteolysis during spermatogenesis?). One question that also must await further investigation concerns how widespread the occurrence of these testes-specific α -like subunit isoforms is among other metazoans. An additional topic of interest is whether the existence of testes-specific proteasome subunits is limited to this particular α -like subunit or whether any of the other subunits also exist in alternative forms.

Another question raised by this study concerns how these testes-specific genes arose during evolution. Since *Pros28.1B* shows a similar degree of divergence from both *Pros28.1* and *Pros28.1A* (58 and 53% amino acid identity, respectively), with *Pros28.1A* being significantly more similar to *Pros28.1* (74%), it is most reasonable to assume that the *Pros28.1B* gene arose first as a duplication of *Pros28.1*, and that *Pros28.1A* arose later by an independent duplication event involving *Pros28.1*. As mentioned earlier, one notable feature of *Pros28.1A* is that the two introns seen in *Pros28.1* and *Pros28.1B* are absent in this gene. This suggests that *Pros28.1A* might have arisen via a mechanism involving an mRNA intermediate. In such a case, reverse transcription from a spliced mRNA in a germline cell would produce an intronless cDNA that could become inserted into an ectopic position in the genome. Examples of such processed "retrogenes" are not uncommon in vertebrates

(see WEINER *et al.* 1986 for a review), however only a few examples have been documented in *Drosophila* (JEFFS and ASHBURNER 1991; MARFANY and GONZALEZ-DUARTE 1992; CURRIE and SULLIVAN 1994). Moreover, the great majority of intronless gene duplications represent nonfunctional "pseudogenes", due to the fact that mRNA, and the corresponding cDNA, typically contains no transcriptional control sequences. If the *Pros28.1A* gene did arise through a retrotransposition event, it would represent one of the few examples of a transcribed and translated gene duplication that arose through this type of mechanism. Other examples of such transcribed retrogenes are thought to either have fortuitously integrated downstream of a functional promoter element, or to have been derived from an aberrant transcript that happened to initiate upstream of the normal promoter. Other features usually associated with retrogenes, *e.g.*, remnants of a poly-A tract at the 3' end and flanking direct repeats, are not present in *PROS28.1A*, however such sequences would be rapidly lost as the gene sequence diverged due to neutral drift. Of course, it cannot be ruled out that *Pros28.1A* arose by a more conventional gene duplication mechanism and that the two introns were subsequently precisely removed by deletion.

We are grateful to YUJING LIU, TONY HOWELLS and DIETMAR ZAISS for encouragement during the course of this work and for helpful comments on the manuscript. We thank TOM STARMER for his assistance with the sequence comparisons and MARIANA WOLFNER for providing the *tudor* strain. This material is based upon work supported by the National Science Foundation under grant MCB-9506885. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

LITERATURE CITED

- AHN, J. Y., S. O. HONG, K. B. KWAK, S. S. KANG, K. TANAKA *et al.*, 1991 Developmental regulation of proteolytic activities and subunit pattern of 20S proteasome in chick embryonic muscle. *J. Biol. Chem.* **266**: 15746–15749.
- ASHBURNER, M., 1989 *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- BAUMEISTER, W., B. DAHLMANN, R. HEGERL, F. KOPP, L. KUEHN *et al.*, 1988 Electron microscopy and image analysis of the multicatalytic proteinase. *FEBS Lett.* **241**: 239–245.
- BOSWELL, R. E., and A. P. MAHOWALD, 1985 *tudor*, a gene required for the assembly of the germ plasm in *Drosophila melanogaster*. *Cell* **43**: 97–104.
- CALLIS, J., 1995 Regulation of protein degradation. *Plant Cell* **7**: 845–857.
- CHEN, P., and M. HOCHSTRASSER, 1995 Biogenesis, structure and function of the yeast 20S proteasome. *EMBO J.* **14**: 2620–2630.
- CIECHANOVER, A., 1994 The ubiquitin-proteasome pathway. *Cell* **79**: 13–22.
- CURRIE, P. D., and D. T. SULLIVAN, 1994 Structure, expression, and duplication of genes which encode phosphoglyceromutase of *Drosophila melanogaster*. *Genetics* **138**: 353–363.
- DRISCOLL, J., M. G. BROWN, D. FINLEY and J. J. MONACO, 1993 MHC-linked *LMP* gene products specifically alter peptidase activities of the proteasome. *Nature* **365**: 262–264.
- FRENTZEL, S., M. TROXELL, C. HAASS, B. PESOLD-HURT, K. H. GLÄTZER *et al.*, 1992 Molecular characterization of the genomic regions of the *Drosophila* α -type subunit proteasome genes *PROS-Dm28.1* and *PROS-Dm35*. *Eur. J. Biochem.* **205**: 1043–1051.
- FULLER, M. T., 1993 Spermatogenesis, pp. 71–148 in *The Development of Drosophila melanogaster*, vol. I, edited by M. BATE and A. MARTINEZ ARIAS. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- GACZYNSKA, M., K. ROCK and A. L. GOLDBERG, 1993 γ -Interferon and expression of major histocompatibility complex (MHC) genes regulate peptide hydrolysis by proteasomes. *Nature* **365**: 264–267. (Erratum. *Nature* **374**: 290.)
- GACZYNSKA, M., K. ROCK and A. L. GOLDBERG, 1994 Role of proteasomes in antigen presentation. *Enzyme Protein* **47**: 354–369.
- GLOTZER, M., A. W. MURRAY and M. W. KIRSCHNER, 1991 Cyclin is degraded by the ubiquitin pathway. *Nature* **349**: 132–138.
- GOLDBERG, A. L., 1992 The mechanism and functions of ATP-dependent proteases in bacterial and animal cells. *Eur. J. Biochem.* **203**: 9–23.
- GOLDBERG, A. L., and K. ROCK, 1992 Proteolysis, proteasomes and antigen presentation. *Nature* **357**: 375–379.
- GÓNCZY, P., S. VISWANATHAN and S. DINARDO, 1992 Probing spermatogenesis in *Drosophila* with P-element enhancer detectors. *Development* **114**: 89–98.
- HAASS, C., and P.-M. KLOETZEL, 1989 The *Drosophila* proteasome undergoes changes in its subunit pattern during development. *Exp. Cell Res.* **180**: 243–252.
- HAASS, C. B., G. PESOLD-HURT MULTHAUP, K. BEYREUTHER and P.-M. KLOETZEL, 1990 The *Drosophila* *PROS-28.1* gene is a member of the proteasome gene family. *Gene* **90**: 235–241.
- HEINEMEYER, W., J. A. KLEINSCHMIDT, J. SAIDOWSKY, C. ESCHER and D. H. WOLF, 1991 Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase: mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival. *EMBO J.* **10**: 555–562.
- HENDIL, K. B., K. G. WELINDER, D. PEDERSEN, W. UERKVITZ and P. KRISTENSEN, 1994 Subunit stoichiometry of human proteasomes. *Enzymes Protein* **47**: 232–240.
- HERSHKO, A., and A. CIECHANOVER, 1992 The ubiquitin system for protein degradation. *Annu. Rev. Biochem.* **61**: 761–807.
- HILT, W., W. HEINEMEYER and D. H. WOLF, 1994 Studies on the yeast proteasome uncover its basic structural features and multiple *in vivo* functions. *Enzyme Protein* **47**: 189–201.
- HOCHSTRASSER, M., 1995 Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. *Curr. Opin. Cell Biol.* **7**: 215–223.
- HOCHSTRASSER, M., M. J. ELLISON, V. CHAU and A. VARSHAVSKY, 1991 The short-lived MAT α 2 transcriptional regulator is ubiquitinated *in vivo*. *Proc. Natl. Acad. Sci. USA* **88**: 4606–4610.
- JEFFS, P., and M. ASHBURNER, 1991 Processed pseudogenes in *Drosophila*. *Proc. R. Soc. Lond. B* **244**: 151–159.
- JENTSCH, S., 1992 Ubiquitin-dependent protein degradation: a cellular perspective. *Trends Cell Biol.* **2**: 98–103.
- JENTSCH, S., and S. SCHLENKER, 1995 Selective protein degradation: a journey's end within the proteasome. *Cell* **82**: 881–884.
- KLEMENZ, R., U. WEBER and W. J. GEHRING, 1987 The white gene as a marker in a new P-element vector for gene transfer in *Drosophila*. *Nucleic Acids Res.* **15**: 3947–3959.
- KOPP, F., B. DAHLMANN and K. B. HENDIL, 1993 Evidence indicating that the human proteasome is a complex dimer. *J. Mol. Biol.* **229**: 14–19.
- LINDSLEY, D. L., and K. T. TOKUYASU, 1980 Spermatogenesis, pp. 225–294 in *The Genetics and Biology of Drosophila*, Vol. 2d, edited by M. ASHBURNER and T. R. F. WRIGHT. Academic Press, New York.
- LÖWE, J., D. STOCK, B. JAP, P. ZWICKL, W. BAUMEISTER *et al.*, 1995 Crystal structure of the 20S proteasome from the archeon *T. acidophilum* at 3.4 Å resolution. *Science* **268**: 533–539.
- MARFANY, G., and R. GONZALEZ-DUARTE, 1992 Evidence for retrotranscription of protein-coding genes in the *Drosophila subobscura* genome. *J. Mol. Evol.* **35**: 492–501.
- MICHALEK, T. M., P. E. GRANT, C. GRAMM, A. GOLDBERG and L. K. ROCK, 1993 A role for the ubiquitin-dependent proteolytic pathway in MHC class I-restricted antigen presentation. *Nature* **363**: 552–554.
- NISHIZAWA, M., N. FURANO, K. OKAZAKI, H. TANAKA, Y. OGAWA *et al.*, 1993 Degradation of Mos by the N-terminal proline (Pro²)-dependent ubiquitin pathway on fertilization of *Xenopus* eggs: possible significance of natural selection for Pro² in Mos. *EMBO J.* **12**: 4021–4027.

- O'CONNELL, P., and M. ROSBASH, 1984 Sequence, structure, and codon preference of the *Drosophila* ribosomal protein 49 gene. *Nucleic Acids Res.* **12**: 5495–5513.
- PALOMBELLA, V. J., O. J. RANDO, A. L. GOLDBERG and T. MANIATIS, 1994 The ubiquitin-proteasome pathway is required for processing the NF- κ B1 precursor protein and the activation of NF- κ B. *Cell* **78**: 773–785.
- PETERS, J.-M., 1994 Proteasomes: protein degradation machines of the cell. *Trends Biochem. Sci.* **19**: 377–382.
- PLILEY, R., J. L. FARMER and D. E. JEFFERY, 1986 *In situ* hybridization of biotinylated DNA probes to polytene salivary chromosomes of *Drosophila* species. *Dros. Inform. Serv.* **63**: 147–149.
- POOLE, S. J., L. KAUVAR, B. DREES and T. KORNBERG, 1985 The *engrailed* locus of *Drosophila*: structural analysis of an embryonic transcript. *Cell* **40**: 37–43.
- PÜHLER, G., S. WEINKAUF, L. BACHMANN, S. MÜLLER, A. ENGEL *et al.*, 1992 Subunit stoichiometry and three-dimensional arrangement in proteasomes from *Thermoplasma acidophilum*. *EMBO J.* **11**: 1607–1616.
- RECHSTEINER, M., L. HOFFMAN and W. DUBIEL, 1993 The multicatalytic and 26S proteases. *J. Biol. Chem.* **268**: 6065–6068.
- RIVETT, A. J., 1993 Proteasomes: multicatalytic proteinase complexes. *Biochem. J.* **291**: 1–10.
- ROBERTSON, H. M., C. R. PRESTON, R. W. PHILLIS, D. M. JOHNSON-SCHLITZ, W. K. BENZ *et al.*, 1988 A stable genomic source of *P* element transposase in *Drosophila melanogaster*. *Genetics* **118**: 461–470.
- SAMBROOK, J., F. E. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463–5467.
- SAVILLE, K. J., 1992 The genetic and molecular characterization of the *DTS-5* locus of *Drosophila melanogaster*. Ph.D. thesis, Syracuse University, Syracuse, NY.
- SAVILLE, K. J., and J. M. BELOTE, 1993 Identification of an essential gene, *l(3)73A α* , with a dominant temperature-sensitive lethal allele, encoding a *Drosophila* proteasome subunit. *Proc. Natl. Acad. Sci. USA* **90**: 8842–8846.
- SCHEFFNER, M., J. M. HUIBREGTSE, R. D. VIERSTRA and P. M. HOWLEY, 1993 The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**: 495–505.
- SEELIG, A., B. BOES and P.-M. KLOETZEL, 1994 Characterization of mouse proteasome subunit MC3 and identification of proteasome subtypes with different cleavage characteristics. *Enzyme Protein* **47**: 330–342.
- SEEMÜLLER, E., A. LUPAS, D. STOCK, J. LÖWE, R. HUBER *et al.*, 1995 Proteasome from *Thermoplasma acidophilum*: a threonine protease. *Science* **268**: 579–582.
- SHANKLIN, J., M. JABBen and R. D. VIERSTRA, 1987 Red light-induced formation of ubiquitin-phytochrome conjugates: identification of possible intermediates of phytochrome degradation. *Proc. Natl. Acad. Sci. USA* **84**: 359–363.
- SPRADLING, A. C., 1986 *P* element-mediated transformation, pp. 175–198 in *Drosophila: A Practical Approach*, edited by D. B. ROBERTS. IRL Press, Oxford.
- TANAHASHI, N., C. TSURUMI, T. TAMURA and K. TANAKA, 1994 Molecular structures of 20S and 26S proteasomes. *Enzyme Protein* **47**: 241–251.
- TANAKA, K., T. TAMURA, T. YOSHIMURA and A. ICHIHARA, 1992 Proteasomes: protein and gene structures. *New Biologist* **4**: 173–187.
- TREIER, M., L. M. STASZEWSKI and D. BOHMANN, 1994 Ubiquitin-dependent c-Jun degradation *in vivo* is mediated by the δ domain. *Cell* **78**: 787–798.
- TSANG, T., V. COPELAND and G. T. BOWDEN, 1991 A set of cloning vectors for rapid and versatile adaptation of restriction fragments. *BioTechniques* **10**: 330.
- WARD, C. L., S. OMURA and R. R. KOPITO, 1995 Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* **83**: 121–127.
- WEFES, I., L. D. MASTRANDREA, M. HALDEMAN, S. T. KOURY, J. TAMBURLIN *et al.*, 1995 Induction of ubiquitin-conjugating enzymes during terminal erythroid differentiation. *Proc. Natl. Acad. Sci. USA* **92**: 4982–4986.
- WEINER, A. M., P. L. DEININGER and A. EFSTADIADIS, 1986 Nonviral retroposons: genes, pseudogenes and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* **55**: 631–661.
- WING, S. S., and P. JAIN, 1995 Molecular cloning, expression and characterization of a ubiquitin conjugation enzyme (E2_{17kD}) highly expressed in rat testis. *Biochem. J.* **305**: 125–132.

Communicating editor: V. G. FINNERTY