Copyright © 1998 by the Genetics Society of America

Molecular Organization of the 20S Proteasome Gene Family from *Arabidopsis thaliana*

Hongyong Fu,* Jed H. Doelling,* Cassandra S. Arendt,† Mark Hochstrasser† and Richard D. Vierstra*

**Cellular and Molecular Biology Program and the Department of Horticulture, University of Wisconsin, Madison, Wisconsin 53706 and †Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637*

ABSTRACT

The 20S proteasome is the proteolytic complex in eukaryotes responsible for degrading short-lived and abnormal intracellular proteins, especially those targeted by ubiquitin conjugation. The 700-kD complex exists as a hollow cylinder comprising four stacked rings with the catalytic sites located in the lumen. The two outer rings and the two inner rings are composed of seven different α and β polypeptides, respectively, giving an α7/β7/β7/α7 symmetric organization. Here we describe the molecular organization of the 20S proteasome from the plant *Arabidopsis thaliana*. From an analysis of a collection of cDNA and genomic clones, we identified a superfamily of 23 genes encoding all 14 of the Arabidopsis proteasome subunits, designated *PAA-PAG* and *PBA-PBG* for *P*roteasome *A*lpha and *B*eta subunits *A–G*, respectively. Four of the subunits likely are encoded by single genes, and the remaining subunits are encoded by families of at least 2 genes. Expression of the α and β subunit genes appears to be coordinately regulated. Three of the nine Arabidopsis proteasome subunit genes tested, *PAC1* (α3), *PAE1* (α5) and *PBC2* (β3), could functionally replace their yeast orthologs, providing the first evidence for cross-species complementation of 20S subunit genes. Taken together, these results demonstrate that the 20S proteasome is structurally and functionally conserved among eukaryotes and suggest that the subunit arrangement of the Arabidopsis 20S proteasome is similar if not identical to that recently determined for the yeast complex.

PROTEIN degradation plays an integral role in cell physiology and development by removing abnormal proteins and important short-lived regulators. One abundant intracellular protease that has been implicated in various catabolic processes is the 20S proteasome, a 700-kD multisubunit protease with broad specificity (Coux *et al.* 1996; Hochstrasser 1996; Schmidt and Kloetzel 1997). The 20S proteasome (also called the prosome, multicatalytic proteinase, and macropain) is present in the cytoplasm and nucleus of eukaryotes and is closely related to protease complexes in some archaebacteria and eubacteria (Löwe *et al.* 1995; Maupin-Furlow and Ferry 1995; Tamura *et al.* 1995).

Whereas its exact function in prokaryotes is unclear, a number of critical functions have been ascribed to the 20S proteasome in eukaryotes. Most important is its role in ubiquitin-dependent proteolysis. Here, association of the 20S complex with a 19S regulatory complex creates the ATP-dependent 26S proteasome which degrades proteins covalently modified with one or more ubiquitins (Coux *et al.* 1996; Hochstrasser 1996; Vierstra 1996). The 20S proteasome also participates in the breakdown of certain proteins in the absence of ubiquitination, in proteolytic processing of precursor proteins, and in the production of antigenic peptides presented by the major histocompatibility class (MHC) I pathway in mammals and other metazoans (Coux *et al.* 1996; Hochstrasser 1996). In mammalian cells, the 20S complex can also associate with an additional ring-shaped complex called PA28 or the 11S regulator, which dramatically enhances its *in vitro* proteolytic activity (Knowlton *et al.* 1997). The importance of the 20S proteasome is reflected by the fact that all but one [the exception being PRE9 (α3) (Emori *et al.* 1991)] of the 14 yeast 20S proteasome subunits are essential (Coux *et al.* 1996; Hochstrasser 1996).

In recent years, the organization and structure of the archaeal, yeast, and mammalian 20S proteasomes have been examined in considerable detail (Löwe *et al.* 1995; Groll *et al.* 1997; Schmidt and Kloetzel 1997). The particle exists as a hollow cylinder ~148 Å in length and ~113 Å in width. It is created by the assembly of four stacked rings; the two peripheral rings are each composed of seven α subunits, whereas the two central rings are each composed of seven β subunits. The holocomplex contains three chambers with narrow entrances at each end restricting access. The central chamber is fashioned solely from β subunits and contains the catalytic sites of the protease complex. In the archaeon *Thermoplasma acidophilum*, the 20S proteasome has a simple subunit composition, containing only a single α-type and a single β-type subunit (Zwickl *et al.* 1992; Löwe *et al.* 1995). However, in the yeast *Saccharomyces cerevisiae* and in animals, the particle is more complex, containing seven distinct α-type and seven distinct β-type subunits

Corresponding author: Richard D. Vierstra, Department of Horticulture, 1575 Linden Drive, University of Wisconsin-Madison, Madison, WI 53706. E-mail: vierstra@facstaff.wisc.edu

in each particle (Chen and Hochstrasser 1995; Groll *et al.* 1997; Schmidt and Kloetzel 1997). The *T. acidophilum* β subunit and many of the yeast and human β subunits are synthesized as proproteins that are proteolytically processed during assembly of the particle. For the *T. acidophilum* β subunit and three of the yeast and human β subunits (β1, β2, and β5), this processing exposes a Thr residue at the N-terminus which forms part of a novel catalytic site that includes the N-terminal amino group (Seemüller *et al.* 1995; Löwe *et al.* 1995).

In contrast to our understanding of the particles from other kingdoms, little is known about the organization and structure of the plant 20S proteasome. The complex has been detected and purified from several sources including potato, tobacco, mungbean, pea, spinach and wheat, typically using its large size as one purification criterion (Kremp *et al.* 1986; Schliephacke *et al.* 1991; Skoda and Malek 1992; Ozaki *et al.* 1992; Fujinami *et al.* 1994; P. H. Hatfield and R. D. Vierstra, unpublished results). All appear to have a subunit composition similar to 20S proteasomes from animals and yeast, as determined by two-dimensional sodium dodecyl sulfate (SDS)-PAGE. Several of these subunits are antigenically related to those in mammals (Schliephacke *et al.* 1991; Ozaki *et al.* 1992). Likewise, electron microscopy of plant 20S proteasomes showed a similar barrel-shaped structure comprising four stacked rings. Genes encoding three subunits, two α and one β, have been described from *Arabidopsis thaliana*, with substantial amino acid sequence similarity to those from other eukaryotes (Genschik *et al.* 1992, 1994; Shirley and Goodman 1993). An Arabidopsis mutant bearing a deletion in one of these α subunits is phenotypically normal, suggesting that a second locus encoding this subunit exists (Shirley and Goodman 1993).

Given the importance of protein degradation to many phases of the plant life cycle, we expect the 20S proteasome has numerous essential functions in plants, especially with respect to its role in ubiquitin-dependent proteolysis (Vierstra 1996). As a first step in elucidating these functions, an extensive analysis of the molecular organization and structure of the 20S proteasome from *A. thaliana* was initiated. Here we describe a set of Arabidopsis genes that encodes the full complement of 20S proteasome subunits. Conservation of the polypeptide sequences indicates that the Arabidopsis 20S complex is organized similar to those from *T. acidophilum*, yeast, and animals. In fact, several of the Arabidopsis subunits can functionally replace their yeast orthologs *in vivo*.

MATERIALS AND METHODS

Identification of Arabidopsis genes encoding 20S proteasome subunits: A collection of Arabidopsis sequences encoding 20S proteasome subunits was identified from the *A. thaliana* ecotype Columbia databases (AtDB, Stanford University; http://genome-www.stanford.edu/Arabidopsis/) using the Blast program. The search first used peptide sequences of the 14 yeast *S. cerevisiae* 20S proteasome subunits as queries (Hochstrasser 1996). The nucleotide sequences of the Arabidopsis DNAs derived from this search were then used to explore the Arabidopsis expressed sequence tag (EST) and genomic nucleotide databases. After redundant entries were excluded, 106 Arabidopsis 20S proteasome DNA sequences were assembled: one genomic and three cDNA sequences had been published previously and corresponded to *PAD1* (TASG64; Genschik *et al.* 1992), *PAF1* (AtPSM30; Shirley and Goodman 1993), and *PBF1* (FAFP98 and TASG39.20; Genschik *et al.* 1994); 91 were ESTs; and 11 were genomic sequences identified from various *A. thaliana* chromosome sequencing projects.

Based on their relationship to the 14 yeast 20S proteasome subunits, the Arabidopsis sequences were classified into 14 groups (Table 1) corresponding to the seven α (α1–α7) and seven β (β1–β7) subunits. EST clones with the longest 5′ untranslated region (UTR) from each group (obtained from Arabidopsis Biological Resource Center (ABRC) at Ohio State University) were sequenced in their entirety. ESTs containing the entire presumptive coding region were available from each group, except *PAD2* and *PAF2*. Apparently full-length cDNA clones for these two were isolated from the Kieber and Ecker 1–2 kbp size-selected cDNA library (ARBC Stock Number CD4-14) by PCR using Pfu polymerase (Stratagene, La Jolla, CA). Sequence analyses were performed using programs from the University of Wisconsin Genetics Computer Group (UWGCG) software package (Deveraux *et al.* 1984). Accession numbers for the 21 Arabidopsis cDNAs encoding the 20S proteasome subunits described here are *PAA1* (AF043518), *PAA2* (AF043519), *PAB1* (AF043520), *PAC1* (AF043521), *PAD1* (AF043522), *PAD2* (AF043523), *PAE1* (AF043524), *PAE2* (AF043525), *PAF1* (AF043526), *PAF2* (AF043527), *PAG1* (AF043528), *PBA1* (AF043529), *PBB1* (AF043530), *PBB2* (AF043531), *PBC1* (AF043532), *PBC2* (AF043533), *PBD1* (AF043534), *PBD2* (AF043535), *PBE1* (AF043536), *PBF1* (AF043537), and *PBG1* (AF043538). Coding sequences for *PAB2* and *PAC2* were derived from genomic sequences (accession numbers: AC002986 and Z97338, respectively).

RNA and genomic DNA gel blot analysis: Total RNA was extracted from *A. thaliana* ecotype Wassilewskija (WS) and purified by LiCl precipitation (Rapp *et al.* 1992). Seedlings grown two weeks on Gamborg B-5 Media (GIBCO-BRL, Gaithersburg, MD) in 0.8% agar (GM agar) under continuous light or dark provided the plant material for green seedlings, roots (light-grown) and etiolated seedlings (dark-grown). Plants grown in soil under continuous light provided the stems, cauline leaves, siliques and flowers. RNA samples (7 μg) were size fractionated by electrophoreses in 1.2% agarose-formaldehyde gels and transferred onto Zeta-Probe membranes (Bio-Rad, Hercules, CA).

Total genomic DNA was isolated as described (Cone *et al.* 1989) from 2-wk-old *A. thaliana* ecotype WS grown on GM agar under continuous light. DNA samples (5 μg) were digested with *Bgl*II, *Eco*RI or *Eco*RV, size-fractionated by electrophoresis in 0.8% agarose gels, and transferred onto Zeta-Probe membranes. Specific probes to the various α and β subunit genes were amplified by PCR from the corresponding cDNAs using vector-specific primers, and the products were gel-purified. The PCR products were radiolabeled by incorporation of [α-^{32}P] dCTP during primer extension with random oligonucleotides. Hybridization of the probes to the membrane-bound DNA or RNA was performed at 65° in 0.5 m sodium phosphate, 7% SDS, 1 mm Na_4EDTA. High stringency wash conditions were 65° in 0.5× SSC and 0.1% SDS. Low stringency wash conditions were 65° in 3× SSC and 0.1% SDS (20× SSC = 3 m NaCl

and 0.3 m Na_3citrate). Following the washes, the blots were subjected to autoradiography.

To help interpret the fragmentation patterns from the DNA gel blots for *PAA1*, *PAC1*, *PAD1*, *PAE2*, *PAF1*, *PAG1*, *PBA1*, *PBB1*, *PBC1*, *PBD1*, *PBD2*, *PBE1* and *PBF1*, the presence or absence of *Bgl*II, *Eco*RI, and *Eco*RV restriction sites was determined. Specific gene sequences were PCR amplified from genomic DNA of *A. thaliana* ecotype WS using 5′ and 3′ gene-specific primers near the start and stop codons, respectively. The amplified fragments were digested with the three restriction enzymes individually and size-fractionated by electrophoresis in 0.8% agarose gels; their electrophoretic patterns were compared with those obtained with the uncut fragments.

Construction of *URA3*, *TRP1*, and *LEU2* plasmids bearing various wild-type yeast and Arabidopsis 20S subunit genes: The *URA3*-plasmids bearing yeast *PRE3* (β1), *PRE1* (β4), *PRS3* (β6), and *PRE4* (β7) for covering corresponding gene deletions in MHY1031, MHY1028, MHY1057, and MHY1029 were made as follows: To construct pRS316-PRE3, a 1.4-kbp *Bam*HI-*Xho*I fragment from p15E3 (from W. Heinemeyer, Universität Stuttgart, Germany) was subcloned into *Bam*HI/*Sal*I-digested pRS316 (Sikorski and Hieter 1989). The p15E3 plasmid carries a 1.4-kbp *Nru*I-*Sna*BI yeast genomic DNA fragment encompassing *PRE3* (Enenkel *et al.* 1994) subcloned between the *Bam*HI and *Xho*I sites of pRS315 (Sikorski and Hieter 1989). pDP83.PRE1-PRE4 was made by subcloning a 1.4-kbp *Hin*dIII/*Bam*HI *PRE4*-bearing fragment from p15E4 (from W. Heinemeyer) into the similarly digested pDP83.E1-L8 plasmid (Heinemeyer *et al.* 1991); pDP83.E1-L8 [*CEN14*, *URA3*] carries a genomic *PRE1* fragment. The *PRS3* gene was amplified from yeast genomic DNA using PCR with the following primers: AGCTTGGAGTAGGCATT and AGGTA CAGCACGGAAGAT. The 1.4-kbp PCR product was cloned into pCR2.1 (Invitrogen, Carlsbad, CA) and the insert was subsequently excised with *Eco*RI and subcloned into pRS316 (Sikorski and Hieter 1989) to generate pRS316-PRS3. Other *URA3*-based plasmids carrying yeast 20S proteasome genes were described previously (Chen and Hochstrasser 1995; Arendt and Hochstrasser 1997).

For controls in the 5-fluoroorotic acid (FOA) assays, genomic DNA fragments for *DOA5* (α5), *PRE3* (β1), *PUP1* (β2), *PUP3* (β3), *PRE1* (β4), *DOA3-His6* (β5), *PRS3* (β6), and *PRE4* (β7) were cloned into either the *LEU2*-plasmid YEplac181 or the *TRP1*-plasmid YCplac22 (Gietz and Sugino 1988) to generate YCplac22-DOA5, YCplac22-PRE3, YCplac22-PUP1, YEplac181-PUP3, YCplac22-PRE1, YCplac22-DOA3-His6, YCplac22-PRS3 and YEplac181-PRE4, respectively. The *PRE1* fragment was isolated as a 1.15-kbp *Eco*RI/*Hin*dIII fragment from pDP83.E1-L8. The *PRS3* fragment was isolated as an *Eco*RI fragment from pRS316-PRS3. The *PRE4*-bearing *Hin*dIII/*Bam*HI fragment was from p154E (see above).

The predicted full-length coding sequences including the 3′ UTR for the various *A. thaliana* 20S subunits were isolated from the corresponding cDNAs by PCR using Pfu polymerase. The 5′ primers were designed to add an *Nde*I site at the putative start codons and the 3′ primers were designed to add an *Eco*RI site to the 3′ UTR. The products were cloned into *Nde*I/*Eco*RI sites of a modified pRS424 plasmid described previously (Fu *et al.* 1998); it contains the 552-bp promoter sequence of yeast *MCB1*/*SUN1*, followed by *Nde*I and *Eco*RI cloning sites. All constructions were verified as correct by DNA sequence analysis.

Yeast strain construction: *S. cerevisiae* strains used in this study are listed in Table 3. All the strains are congenic with MHY501 (Chen *et al.* 1993). MHY787, MHY991, MHY996, and MHY784 bearing chromosomal deletions of *DOA5* (α5), *PUP1* (β2), *PUP3* (β3), and *DOA3* (β5), respectively, and harboring the corresponding wild-type gene on a *URA3*-plasmid were described previously (Chen and Hochstrasser 1995; Arendt and Hochstrasser 1997). The MHY1029 strain carrying a chromosomal deletion of *PRE4* (β7) was made as follows. Diploid MHY606 cells were transformed with the *pre4*Δ::*HIS3* allele described by Hilt *et al.* (1993), and histidine prototrophs were selected. Putative heterozygotes were sporulated and subjected to tetrad analysis. A strain that yielded tetrads displaying 2:2 segregation of viability (with all viable segregants being His−) was then transformed with pDP83. PRE1-PRE4, and uracil prototrophs were selected. Following sporulation and tetrad dissection, Ura^+/His^+ segregants were identified. To create a *pre1*::*LYS2* disruption allele, a 5.5-kbp *LYS2* fragment was excised with *Cla*I from pUB39 (Spence *et al.* 1995), the 5′-overhangs were filled in using T4 DNA polymerase, and the fragment was then ligated into pDP83.E1-L8 at the unique *Ecl*136II site within the *PRE1* sequence. In the resulting plasmid pDP83pre1::LYS2, the *LYS2* coding sequence is in the opposite orientation relative to the *PRE1* sequence. The *pre1*::*LYS2* allele was isolated from pDP83pre1::LYS2 on a 6.7-kbp *Nar*I/*Pst*I fragment that was then transformed into MHY606 cells. The resulting lysine prototrophs were screened for a single insertion in the *PRE1* locus using the same strategy as that described for *pre4*Δ/+ heterozygotes. A *pre1*::*LYS2*/+ heterozygote was transformed with pDP83.PRE1-PRE4, uracil prototrophs were selected, and, following sporulation and tetrad dissection, Ura^+ Lys^+ segregants were identified resulting in the strain MHY1028.

To produce a null allele of *PRE3*, the *HIS3* gene was PCR amplified with primers, TGTCATTTTCACTTTTCCACTCG CAACGGAATCCGGTGGCCTCTTGGCCTCCTCTAG and CT TGCTTACGAAATTCCCTTCTAGGATACTTTGTCGTTCTA ACAGTCGTTCAGAATGACACG. The primers were designed to amplify the entire *HIS3* gene and flanked by sequences complementary to the 5′ and 3′ coding region of *PRE3*; the amplified fragments were transformed into MHY606 cells. The resulting heterozygote (made by P. Chen), was transformed with pRS316-PRE3, uracil prototrophs were selected, and, following sporulation and tetrad dissection, Ura+/His+ segregants were identified, yielding MHY1031. An exactly analogous PCR-based deletion strategy was used to create the *prs3*-Δ2::*HIS3* and *pre9*-Δ2::*HIS3* null alleles. The primers used for making *prs3*-Δ2::*HIS3* were GAGAGTAGCAAGACTATTGAAC TATAAAGTTAAACAAAATATGGCTCTTGGCCTCCTCTAG and TTCTTTTTATACTATGATATGTATGCATTAATCTCTT TTTAGCTCTCGTTCAGAATGACACG. The MHY1057 strain was derived from a *prs3*Δ/+ heterozyote transformed with pRS316-PRS3. The primers used for amplifying *pre9*-Δ2::*HIS3* were CATGGGTTCCAGAAGATACGATTCCAGGACAACAA TTTTTCTCCCCCTCTTGGCCTCCTCTAG and GGCTCTC GAGCGATTCCGATCTTGAAATTTGCGCATCGTTCAGAAT GACACG. MHY606 transformants were screened for the deletion by colony PCR, and, following sporulation, tetrads were dissected. The tetrads all contained four viable spores, with two His^+ and two His^- segregants. The His^+ segregants (MHY1069) grew slightly slower at 30° than those His^-. The expected structure at the *PRE9* locus was confirmed by colony PCR of several complete tetrads. These data confirm the original observation that the *PRE9* gene is not required for viability (Emori *et al.* 1991).

Yeast media and techniques: Yeast rich (YPD) and minimal synthetic media (SD) were prepared as described previously (Kaiser *et al.* 1994). For complementation of amino acid-analog sensitivity, arginine and phenylalanine were omitted from the SD media and their analogs canavanine and *p*-fluorophenylalanine were added to 3 and 25 μg/ml, respectively. 5-fluoroorotic acid (FOA) was added to the SD media at a concentration of 1 g/l. Standard methods were used for genetic manipulation of yeast (Kaiser *et al.* 1994). Yeast transformation was carried out as described by Gietz *et al.* (1995), and cultures were incubated at 30°.

TABLE 1

Subunits of the Arabidopsis 20S proteasome

Systematic name[a]	Arabidopsis[b]	Yeast[c]	Human
α-Type Subunits			
α1	PAA1 (Prc1−At)	PRS2/C7/SCL1/Y8	IOTA/PROS27
	PAA2		
α2	PAB1 (Prc3−At)	PRE8/Y7	C3
	PAB2		
α3	PAC1 (Prc9−At)	PRE9/Y13	C9
	PAC2		
α4	PAD1 (Prc6a)	PRE6	XAPC7
	PAD2 (Prc6b)		
α5	PAE1	DOA5/PUP2	ZETA
	PAE2 (Prcz−At)		
α2	PAF1 (Prc2a−At)	PRE5	C2/PROS30
	PAF2 (Prc2b−At)		
α2	PAG1 (Prc8−At)	PRS1/PRE10/C1	C8
β-Type Subunits			
β1	PBA1 (Prcd−At)	PRE3	Y/Delta/LMP2[d]
β2	PBB1 (Prcfa−At)	PUP1	Z/MECL1[d]
	PBB2	(Prcfc−At)	
β3	PBC1 (Prct−At)	PUP3	C10
	PBC2		
β4	PBD1 (Prcgb−At)	PRE1/C11	C7
	PBD2 (Prcga−At)		
β5	PBE1 (Prce−At)	DOA3/PRE2/PRG1	X/MB1/ε/LMP7[d]
β6	PBF1 (Prc5−At)	PRS3/PRE7/C5	C5
β7	PBG1 (Prch−At)	PRE4	N3/Beta

[a] Groll *et al.* (1997).
[b] Names in parentheses are the corresponding designations of the Arabidopsis 20S proteasome subunits by Paramentier *et al.* (1997).
[c] All yeast subunits are essental except α3/PRE9.
[d] γ-interferon inducible.

FOA and amino acid analog sensitivity assays: Assays for complementation of essential genes using FOA resistance were performed as described by Arendt and Hochstrasser (1997). Yeast strains bearing chromosomal deletions for the various 20S subunit genes and harboring the corresponding wild-type yeast genes on a *URA3*-plasmid, were transformed with plasmids bearing yeast or Arabidopsis 20S subunits genes. Transformants were patched on solid YPD media, incubated overnight to allow random loss of either plasmid, and then streaked on solid SD media containing FOA.

Complementation of the amino acid analog sensitivity of the yeast *pre9*Δ (α3) strain was performed as previously described (Fu *et al.* 1998). Yeasts were grown in liquid SD media to late logarithmic phase. Cultures were washed once with SD medium (minus arginine and phenylalanine), resuspended in the same medium to A_{600} = 1.0, and spotted in a 10-fold dilution series onto solid SD medium supplemented with canavanine and *p*-fluorophenylalanine.

RESULTS

Identification of Arabidopsis genes encoding subunits of the 20S proteasome: As a first step in the analysis of the 20S proteasome from Arabidopsis, we initiated a detailed characterization of the corresponding genes. Using the 14 genes encoding the full complement of polypeptides from the yeast 20S proteasome (Coux *et al.* 1996; Hochstrasser 1996) as queries, we identified 23 related sequences in the various nucleic acid sequence databases of *A. thaliana*, ecotype Columbia. Three of the genes corresponded to the two α subunits and the one β subunit previously described by Genschik *et al.* (1992, 1994) and Shirley and Goodman (1993). The remaining twenty, which represented related but new loci, were 18 partially sequenced cDNAs obtained from expressed-sequence-tag (EST) databases and two genomic sequences deposited by the *A. thaliana* genome sequencing projects. [The search also detected related sequences in Genebank databases from other plant species such as rice, corn, soybean, spinach, alfalfa, tomato, and tobacco.] All but two of the cDNAs contained the entire coding region. For the two partial cDNAs [*PAD2* and *PAF2* (see Table 1)], the full coding sequences were obtained subsequently from an Arabidopsis cDNA library by PCR.

The cDNAs avaliable for the 21 proteasome genes (the exception are *PAB2* and *PAC2* which were identified only by genomic sequences) were sequenced in their entirety and used to derive the predicted full-

length amino acid sequence of the corresponding proteins. Coding regions of *PAB2* and *PAC2* were assembled by alignment with cDNAs of their respective Arabidopsis paralogs, *PAB1* and *PAC1*. The predicted start codon for most α and β subunit genes were identified as the 5′ most ATG codon adjacent to sequence conserved in orthologs from other species. In-frame stop codons were found immediately upstream of these presumptive start codons. For *PAA1*, *PAG1*, *PBD1*, the start codon was assigned as the 5′ most ATG beyond the conserved region. No in-frame stop codons were present upstream of these putative start codons. For *PBG1*, the predicted start codon was assigned based on a comparison of the coding sequence with that of the propeptide present within its yeast and mammalian orthologs (*e.g.*, Hilt *et al.* 1993).

Pairwise comparisons of derived amino acid sequences revealed that the 23 Arabidopsis 20S proteasome subunit genes cluster into two groups corresponding to the α and β families detected in other organisms (Coux *et al.* 1996; Hochstrasser 1996). The Arabidopsis α and β families are weakly but significantly related to each other (20–30% amino acid similarity over specific regions of the corresponding proteins) further supporting the proposal that the α and β families share a common progenitor (Zwickl *et al.* 1992; Coux *et al.* 1994). Phylograms of the sequence comparisons showed that one to two genes in each Arabidopsis family are closely related to each of the seven α and seven β subunit genes present in yeast (Figure 1). The phylograms adopted seven branches for both the α and β subunits; each branch contained a single yeast α or β subunit gene and one or more Arabidopsis genes. In all cases, the Arabidopsis α and β subunits in each branch were more similar to specific subunits in yeast and other organisms than to other α or β subunits in Arabidopsis (Figure 1 and data not shown). As examples, the α5 subunit (PAE1) is 66–78% similar to the α5 subunits from yeast, humans, *Drosophila melanogaster*, and *Caenorhabditis elegans* but only 48–57% similar to the other six Arabidopsis α subunits, whereas, the β3 subunit (PBC2) is 58–66% similar to the β3 subunits from yeast, humans, and *C. elegans* but only 29–45% similar to the other six Arabidopsis β subunits. From this analysis, we conclude that this Arabidopsis collection encodes the full complement of 20S proteasome subunits with clear matches to each of its putative yeast and human orthologs. It is remotely possible that the 20S proteasome from plants has additional subunits not found in the animal and yeast complexes. However, comparisons of the SDS-PAGE profiles of the plant and animal 20S proteasomes suggest a similar polypeptide composition (Kremp *et al.* 1986; Schliephacke *et al.* 1991; Skoda and Malek 1992; Ozaki *et al.* 1992; Fujinami *et al.* 1994; P. H. Hatfield and R. D. Vierstra, unpublished results).

On the basis of their homology with yeast subunits (Figure 1), the Arabidopsis collection was designated

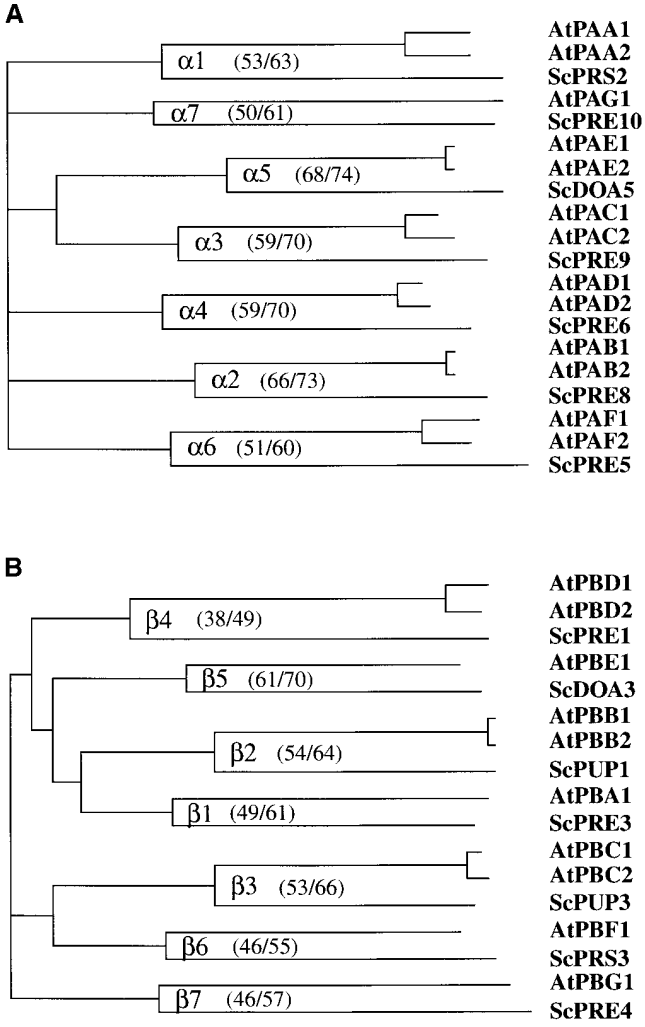

Figure 1.—Phylograms showing the amino acid sequence relationships of the complete collection of 20S proteasome subunits from *A. thaliana* with the 14 subunits that comprise the 20S proteasome from yeast (*S. cerevisiae*): (A) α subunits; (B) β subunits. Yeast subunits designations α1–α7 and β1–β7 were based on the position of the subunits within the crystal structure of the yeast 20S complex (Groll *et al.* 1997). Phylograms were created using the UW-GCG computer programs Paupsearch and Paupdisplay, using the bootstrap analysis option. Distance along the horizontal axis separating two sequences is proportional to the divergence between the sequences. Protein sequence comparisons (% identity/similarity) of the Arabidopsis subunits to their yeast orthologs are indicated to the right.

using the new systematic nomenclature for the yeast 20S proteasome which was organized from the crystal structure of the particle (Groll *et al.* 1997). The protein subunits, referred to as α1–α7 and β1–β7, likely assume the same positions in the particle as the equivalent yeast subunits based on our cross-species complementation data (see below). The corresponding genes were named using the protocol recommended for *A. thaliana* (Meinke and Koornneef 1997). They are referred to as *AtPAA-G* and *AtPBA-G* (for *A. thaliana P*roteasome *A*lpha subunits *A–G* and *P*roteasome *B*eta subunits *A–G*)

with numerical suffixes added to designate various members in each gene family. Table 1 relates the various Arabidopsis subunits with their presumptive orthologs from yeast and human. For five of the Arabidopsis subunits (α7, β1, and β5–β7), a single gene was identified. For each of the remaining nine (α1–α6 and β2–β4), two independent genes sharing substantial similarity were detected; the average identity among the nine related pairs is 86 and 95% at the nucleotide and amino acid sequence levels, respectively. As examples, *PAF1* and *PAF2*, encoding the α6 subunit, share 87 and 92% identity at the nucleotide and amino acid sequence levels, respectively, and *PBB1* and *PBB2* genes encoding the β2 subunit, shared 88 and 98% identity at the nucleotide and amino acid sequence levels, respectively. Significant divergence at the nucleotide level and subsequent genomic DNA blot analysis (see below) indicate that the members in each group of paralogs represent independent loci and not polymorphic alleles in *A. thaliana*.

20S proteasome α subunits: Thirteen of the Arabidopsis genes encode the seven α-type subunits. An amino acid sequence alignment and pairwise comparisons of sequence homology using a representative member of each of the seven α subunits are presented in Figure 2 and Table 2. Amino acid sequence identities/similarities within the family range from 34/44 to 47/56% with several regions exhibiting high conservation to corresponding subunits in other eukaryotes and to the *T. acidophilum* α subunit (Löwe *et al.* 1995; Groll *et al.* 1997). The most conserved region is near the N-terminus [residues 9–33 in α1/PAA1 (Figure 2)], which assumes an α-helical structure necessary for assembly and/or subsequent stabilization of appropriate α subunit/α subunit contacts in the *T. acidophilum* and yeast complexes (Löwe *et al.* 1995; Groll *et al.* 1997). The invariant Tyr residue at the N-terminal side of the α-helix, which plays a crucial role in this contact, is present in all the Arabidopsis α subunits [Tyr^{9} in PAA1 (α1) (Figure 2)].

The pore in the seven-membered α ring of the yeast complex has not been resolved (Groll *et al.* 1997). However, in *T. acidophilum*, a 13-Å opening is lined by bulky hydrophobic residues [Tyr^{126} (Löwe *et al.* 1995)]. Bulky aromatic hydrophobic residues are present at or near the same position in many of the Arabidopsis subunits (Tyr, Phe and/or Trp), suggesting that the pore in the plant complex is also hydrophobic (Figure 2). Several Arabidopsis α subunits (especially α6/PAF1 and PAF2), like their yeast orthologs, contain long C-terminal extensions. In the yeast 20S complex, these extensions project from the surface of the α-subunit disc, potentially providing important contact sites for accessory factors (Groll *et al.* 1997). A sequence potentially involved in nuclear localization of the yeast α subunits (Tanaka *et al.* 1990) is present in the Arabidopsis counterparts [KKVPDK (residues 54–59) in α1/PAA1 (Figure 2)], suggesting a similar role in nuclear import. A number of contextually conserved Gly residues have been detected in numerous α subunits from a variety of species (Coux *et al.* 1994). Most of these are also present in the Arabidopsis subunits [residues 20, 36, 41, 77, 135, 142, 158 and 169 in α1/PAA1 (Figure 2)].

20S proteasome β subunits: Ten of the Arabidopsis 20S proteasome genes encode the seven β-type subunits. An amino acid sequence alignment and pairwise comparisons of sequence homology using a representative member of each of the seven β subunits are presented in Figure 3 and Table 2. Consistent with results from other species (Coux *et al.* 1996), the β subunits are generally less related amongst themselves than are the α subunits with amino acid sequence identities/similarities ranging from 18/30 to 33/45%.

The single β subunit of *T. acidophilum* and five of the seven β subunits in yeast and humans are synthesized as proproteins that are proteolytically processed before generating the catalytically active holoenzyme complex (Seemüller *et al.* 1995; Chen and Hochstrasser 1996; Schmidtke *et al.* 1996; Arendt and Hochstrasser 1997). The β3 (PUP3/C10) and the β4 (PRE1/C7) subunits are not processed from their initial translation products (Groll *et al.* 1997; Schmidt and Kloetzel 1997). In yeast, prosequence cleavage is accomplished by the active β1, β2, and/or β5 subunits working autocatalytically or in *trans* to the other β subunits in the ring. N-terminal sequences related in length to the yeast and human prosequences are present in the Arabidopsis β1, β2, β5–β7 orthologs, suggesting that they are also synthesized as precursors that then require processing (Figure 3 and data not shown). The alignments predict that the Arabidopsis β1, β2, and β5 subunits are cleaved between a Gly/Thr pair [*e.g.*, Gly^{12}/Thr^{13} in PBA1 (β1)], thus liberating the Thr active site (see below). For PBF1 (β6), the predicted cleavage site is between Glu^{4} and His^{5} (instead of Asp/His in yeast β6) and for PBG1 (β7), it is between Arg^{23} and Thr^{24} [identical to its human counterpart and similar to the Asn/Thr cleavage site in yeast β7 (Chen and Hochstrasser 1995; Schmidtke *et al.* 1996; Groll *et al.* 1997)]. Assuming these cleavage sites are correct, prosequences of 12 (β1), 39 (β2), 57 (β5), 4 (β6), and 23 (β7) amino acids would be released.

Processing of the *T. acidophilum* β subunit and the catalytically active yeast and human subunits β1 (PRE3/Y), β2 (PUP1/Z), and β5 (DOA3/X) is necessary to expose a Thr^{1} residue at their N-termini, which then becomes organized into a novel active site involving the free α-amino group (Seemüller *et al.* 1995; Chen and Hochstrasser 1996). In the *T. acidophilum* complex, this amino group works in a catalytic tetrad that also contains the Thr^{1} hydroxyl group, Glu^{17}, and the ε amino group of a proximal Lys^{33}. Glu^{17} appears to maintain the correct orientation of Lys^{33} relative to Thr^{1}. Three residues adjacent to Thr^{1} (Ser^{129}, Asp^{166}, and Ser^{169}) in the three-dimensional complex also may be involved in forming the catalytic site (Seemüller *et al.* 1996). From

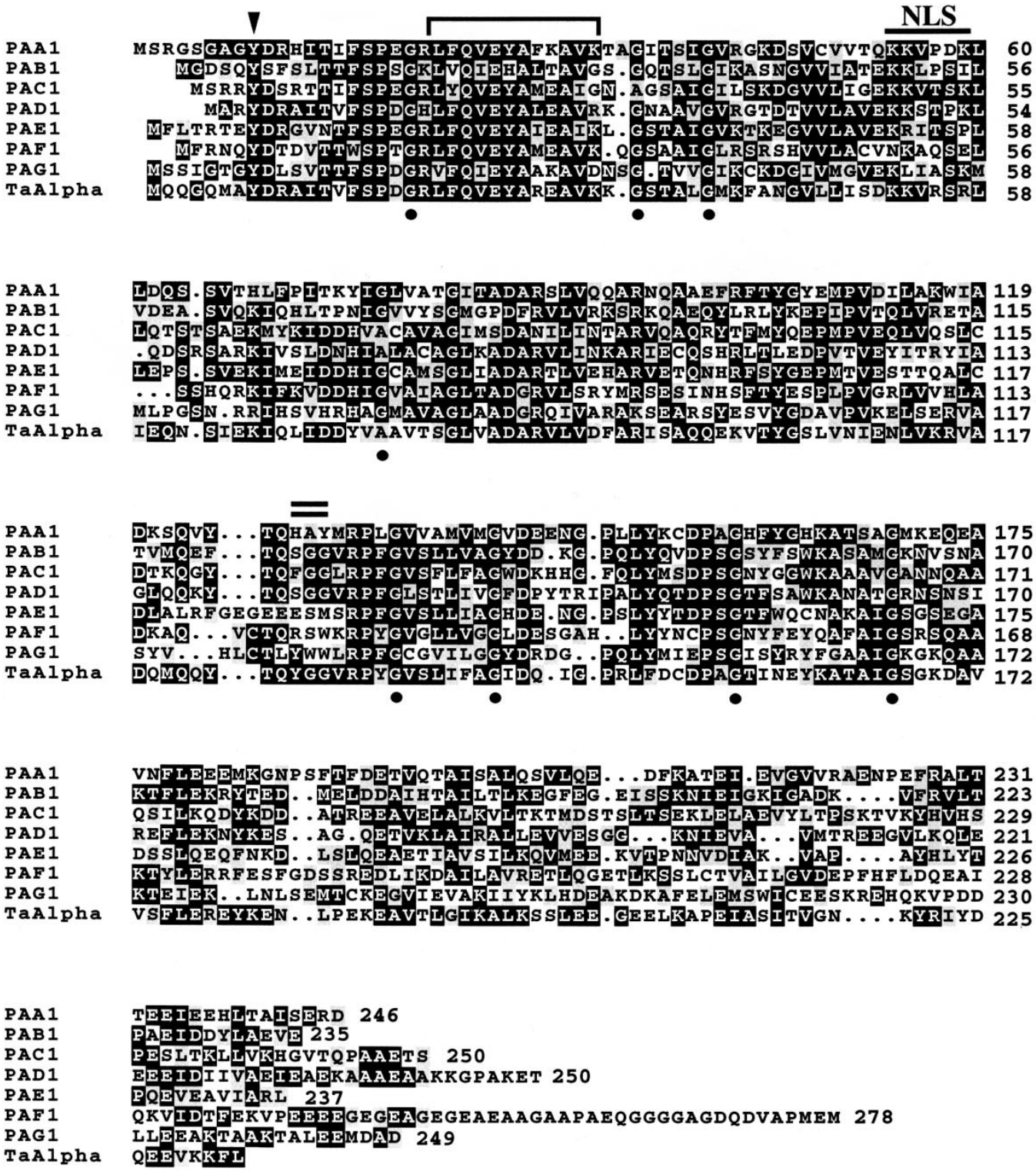

Figure 2.—Derived amino acid sequence alignment of the representative members of the α-subunit gene family (*PAA-PAG*) from *A. thaliana.* The alignment was created using the computer program BoxShade 2.7 (Netserve@embl-heidelberg.de; Bioinformatics Group, Lausanne, Switzerland). Identical and similar residues are in reverse type and shaded boxes, respectively. The bracket indicates the N-terminal α-helical region required in the corresponding yeast subunits for interactions among the α subunits; the line shows the putative NLS sequence; and the double line identifies the amino acids that border the pore of the *T. acidophilum* α-subunit ring (Löwe *et al.* 1995; Groll *et al.* 1997). The Tyr residue essential for assembly of the α ring is indicated by the arrowhead. (●) marks a collection of conserved Gly residues present in many α and β subunits (Coux *et al.* 1994). Sequences encoding subunit α4/PAD1 and α6/PAF1 were first described by Genschik *et al.* (1992) and Shirley and Goodman (1993) and previously designated TASG64 and AtPSM30, respectively. The amino acids sequence for the α subunit from *T. acidophilum* (TaAlpha) is included for comparison (Zwickl *et al.* 1992).

TABLE 2

Amino acid sequence homology (% identity/similarity) of the Arabidopsis 20S α (α1/PAA-α7/PAG) and β (β1/PBA-β7/PBG) subunits[a,b]

α Subunits	PAB1	PAC1	PAD1	PAE1	PAF1	PAG1
PAA1	36/50 (234)	39/50 (207)	39/48 (231)	38/52 (231)	34/44 (229)	32/50 (176)
PAB1		37/50 (212)	40/52 (235)	36/52 (234)	37/48 (235)	34/47 (183)
PAC1			44/56 (200)	44/56 (214)	39/50 (212)	33/49 (209)
PAD1				44/57 (200)	39/47 (198)	35/49 (173)
PAE1					47/56 (179)	33/48 (201)
PAF1						39/49 (171)
β Subunits	**PBB1**	**PBC2**	**PBD1**	**PBE1**	**PBF1**	**PBG1**
PBA1	33/45 (180)	18/30 (194)	25/34 (160)	33/42 (191)	25/34 (198)	21/31 (193)
PBB1		19/33 (187)	22/33 (160)	33/39 (172)	26/36 (146)	20/32 (186)
PBC2			19/29 (162)	26/35 (185)	33/45 (129)	20/32 (186)
PBD1				29/38 (153)	21/32 (142)	21/33 (151)
PBE1					26/34 (146)	20/32 (201)
PBF1						23/33 (147)

[a] Determined by UW-GCG program Bestfit.
[b] Numbers in parentheses are best-matched peptide lengths for each pairwise α and β subunit comparison.

sequence alignments with the comparable *T. acidophilum* and yeast subunits, we propose that a similar protease active site is present in the β1, β2, and β5 subunits of Arabidopsis. For instance in PBA1 (β1), all six of the contextually essential residues are evident, including the Thr and Lys residues at positions 13 and 45, an invariant Asp at position 29 (which could substitute for the Glu residue), and the Ser/Asp/Ser sequence at positions 142, 179 and 182 (Figure 3). The β3 (PBC1 and 2), β4 (PBD1 and 2), and β6 (PBF1) subunits lack the N-terminal Thr residue whereas subunit β7 (PBG1) does have the conserved Thr but lacks the essential Lys, suggesting that these subunits, like their yeast and human counterparts (Schmidt and Kloetzel 1997), are not proteolytically active.

For the yeast 20S proteasome, mutagenesis studies indicate that subunit PRE3 (β1) carries the peptidylglutamyl peptide hydrolyzing (acidic) activity, DOA3 (β5) carries the chymotrypsin-like (hydrophobic) activity, and PUP1 (β2) carries the trypsin-like (basic) activity (Chen and Hochstrasser 1996; Arendt and Hochstrasser 1997; Heinemeyer *et al.* 1993, 1997). Based on sequence similarity, we expect that the Arabidopsis counterparts, PBA1 (β1), PBB1 and 2 (β2), and PBE1 (β5), will have similar peptidase specificities. In each, the predicted substrate specificity (S1) pocket [residues 32, 43, 47, 57, 61 and 65 in PBA1 (β1)] is nearly identical to the pocket identified in the corresponding yeast ortholog (Groll *et al.* 1997) (Figure 3). Covalent inhibition of the yeast β5 subunit by lactacystin involves hydrogen bonding with Thr^{1}, Thr^{21}, Met^{45}, and Gly^{47} residues in the S1 pocket (Groll *et al.* 1997). Identically situated residues are present in Arabidopsis β5 subunit PBE1 (except for Thr^{21}, which is replaced by a Ser (Figure 3)], suggesting that lactacystin inhibits the Arabidopsis 20S proteasome as well. Like the α-ring pore, the pores of the β rings in *T. acidophilum* and yeast are surrounded by one or more bulky hydrophobic residues (Löwe *et al.* 1995; Groll *et al.* 1997). A similar hydrophobic patch is present in all the Arabidopsis β subunits (Figure 3). Many of the Arabidopsis β subunits also contain a number of contextually conserved Gly residues [residues 17, 23, 59, 114, 120, 130, 141 and 143 in PBA1 (β1) (Figure 3)] present in the β subunits from a variety of other species (Coux *et al.* 1994).

Genomic DNA gel blot analysis: To estimate the number of genes encoding each 20S proteasome subunit, genomic DNA from the *A. thaliana* ecotype WS was digested with one of three restriction endonucleases (*Bgl*II, *Eco*RI or *Eco*RV) and subjected to DNA gel blot analysis using probes specific for each of the 14 proteasome subunit gene families [*PAA-PAG* (α1–α7) and *PBA-PBG* (β1–β7)]. Following hybridization, the blots were washed at low stringency to enable detection of all closely related fragments (Figure 4, A and B) and then at high stringency to identify fragments specific to the probe (data not shown). To help assign genomic fragments to individual genes, specific DNAs were PCR amplified from *A. thaliana* genomic DNA and similarly digested with the same set of restriction endonucleases (data not shown).

This analysis was consistent with the amino acid sequence alignments, indicating that many of the proteasome subunits are encoded by multiple genes in Arabidopsis. For the 6 α-subunits and the three β subunits that are encoded by two identified cDNAs (Figure 1 and Table 1), DNA gel blot patterns agreed with the presence of two or more genes in the WS ecotype (Fig-

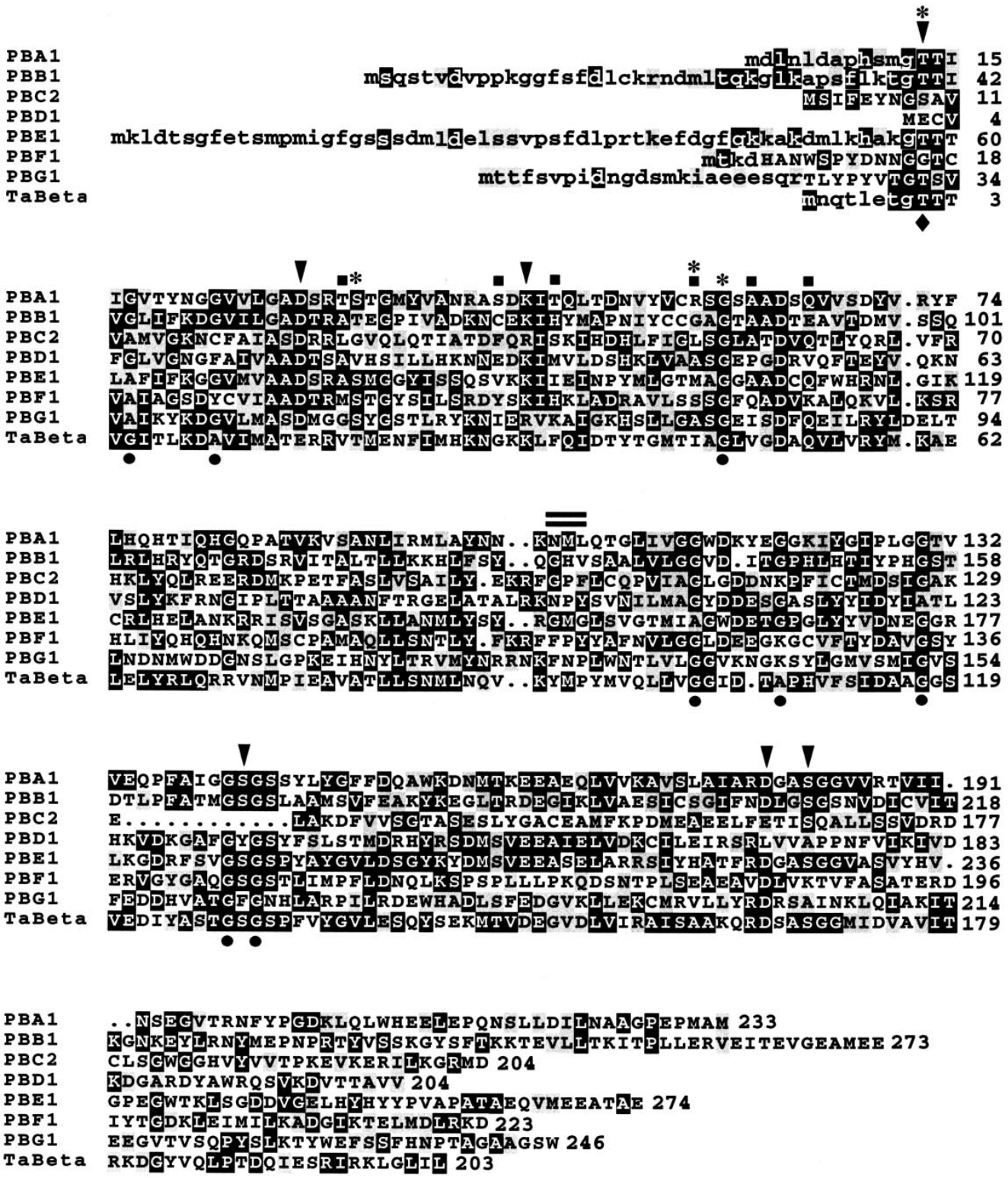

Figure 3.—Derived amino acid sequence alignment of the representative members of the β-subunit gene family (*PBA-PBG*) from *A. thaliana.* The alignment was created using the computer program BoxShade 2.7. Identical and similar residues are in reverse type and shaded boxes, respectively. The N-terminal sequences predicted to be removed during the maturation of the subunits, PBA1 (β1), PBB1 (β2), PBE1-PBG1 (β5–β7) are indicated in lower case letters. The residues predicted to comprise part of the protease active site for subunits PBA1 (β1), PBB1 (β2), and PBE1 (β5) are indicated by arrowheads; the double line identifies the amino acids that border the pore of the *T. acidophilum* β-subunit ring (Löwe *et al.* 1995; Groll *et al.* 1997); ■, positions of amino acid residues that correspond to the S1 substrate specificity pocket in yeast PRE3 (β1), PUP1 (β2), and DOA3 (β5) subunits; *, residues in subunit β5 that possibly interact with the inhibitor lactacystin; ●, a collection of conserved Gly residues present in many α and β subunits (Coux *et al.* 1994). Sequences encoding subunit PBF1 (β6) were previously identified by Genschik *et al.* (1994) and formly designated FAFP98. The amino acids sequence for the β subunit from *T. acidophilum* (TaBeta) is included for comparison (Zwickl *et al.* 1992). The numbering of TaBeta starts with the internal Thr (Thr[1]; ◆) released after removal of the 8-amino-acid prosequence.

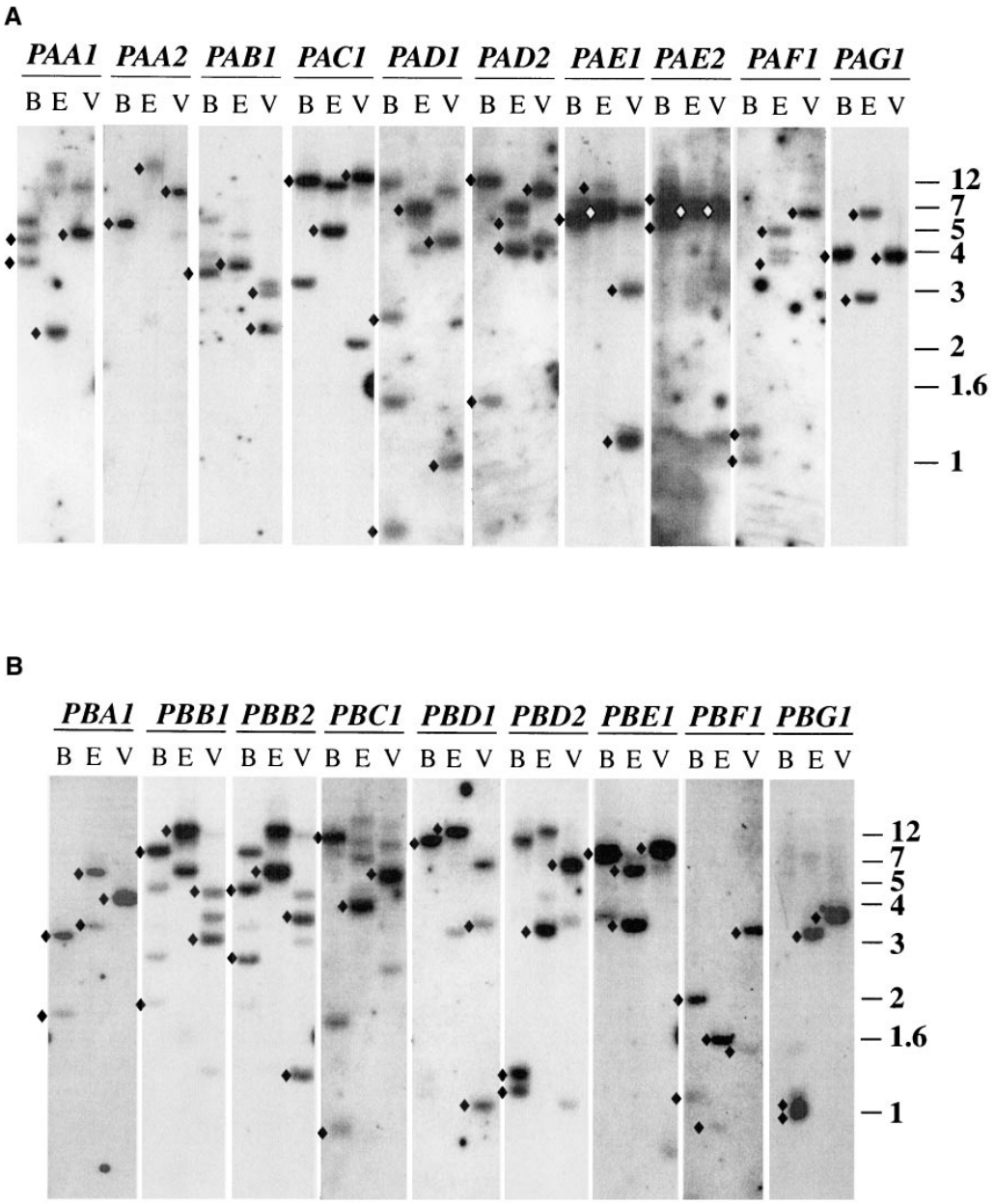

Figure 4.—Genomic DNA gel blot analysis of genes encoding various Arabidopsis 20S proteasome subunits. DNA from *A. thaliana* ecotype WS was isolated, digested singly with *Bgl* II (B), *Eco*RI (E) or *Eco*RV (V), and size-fractionated by electrophoresis in 0.8% agarose gels. The DNA was transferred to Zeta-Probe membranes and hybridized to the gene-specific probes encoding various α and β subunits as indicated in materials and methods. (A and B) Hybridization patterns following low stringency washes using α-subunit and β-subunit specific probes, respectively. Each band marked by a diamond represents a genomic DNA fragment that corresponds to the gene-specific probe used in that panel.

ure 4, A and B). For the *PAA-PAF* (α1–α6) subunit families, the DNA gel blots were consistent with only two genes. For the remaining 3 subunits encoded by the identified cDNAs [*PBB* (β2), *PBC* (β3) and *PBD* (β4)], additional bands were detected that could not be unequivocally attributed to the two identified genes, suggesting the possibility of a third gene. For four of the five subunits where only a single cDNA was identified [*PAG1* (α7), *PBA1* (β1), *PBF1* (β6), and *PBG1* (β7) (Figure 1)], the patterns were compatible with the presence of a single gene (Figure 4, A and B). For the other subunit represented by a single cDNA, *PBE1* (β5), additional bands were detected in the DNA gel blots suggesting that a close relative of *PBE1* exists but has not yet been isolated.

RNA gel blot analysis of 20S proteasome subunits: RNA gel blot analysis examined the expression patterns of the α and β subunit genes in Arabidopsis. Blots of total RNA were probed with cDNAs for three of the α-subunits [*PAA1* (α1), *PAC1* (α3), and *PAG1* (α7)] and

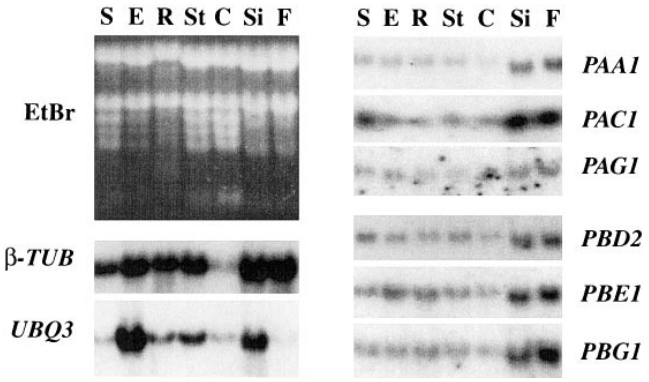

Figure 5.—RNA gel blot analysis of Arabidopsis 20S proteasome subunit gene expression. Total RNA isolated from light-grown seedlings (S), dark grown seedlings (E), roots (R), stems (St), cauline leaves (C), siliques (Si) and flowers (F) was size-fractionated on 1.2% agarose-formaldehyde gels, transferred to Zeta-Probe membrane and hybridized to the gene-specific probe as indicated. EtBR, duplicate RNA samples visualized by ethidium bromide staining. DNA probes include the following: *PAA1*, *PAC1*, and *PAG1*, which encode α1, α3 and α7 subunits, respectively; *PBD2*, *PBE1* and *PBG1*, which encode β4, β5 and β7 subunits, respectively; β-*TUB* that encodes β-tubulin; and *UBQ3* that encodes a tetrameric polyubiquitin.

three of the β-subunits [*PBD2* (β4), *PBE1* (β5), and *PBG1* (β7)] and washed at high stringency (0.5× SSC and 65°) to restrict hybridization to closely related mRNAs (as determined by DNA gel blot analysis). Hybridization to a β-tubulin cDNA and the polyubiquitin gene *UBQ3* was also tested to confirm the integrity of the mRNAs (Sun and Callis 1997). As shown in Figure 5, all seven RNA preparations contained intact rRNAs as detected by ethidium bromide staining. With the exception of the RNA from cauline leaves, near equal levels of β-tubulin mRNA were apparent in all samples. mRNA for *UBQ3* also was present in all tissues with a distribution that matched that reported previously by Sun and Callis (1997); the levels were highest in etiolated seedlings and siliques, intermediate in roots and stems, and lowest in light-grown seedlings, cauline leaves, and flowers. In addition to detecting the UBQ3 mRNA, the UBQ3 probe also weakly detected mRNAs of different sizes that likely represent other *A. thaliana* ubiquitin genes (Callis *et al.* 1995).

When probed with each of the six 20S proteasome subunit genes, corresponding mRNAs of the expected size were detected in the RNA prepared from each tissue (Figure 5). However, compared to those for β-tubulin and *UBQ3*, the signal intensities obtained with the proteasome probes were substantially lower, suggesting lower levels of mRNA (Figure 5 and data not shown). The mRNA levels for each subunit were roughly similar in all tissues except siliques and flowers where slightly higher levels were seen relative to those of β-tubulin. Despite the low levels of β-tubulin and *UBQ3* mRNA, mRNA for the six proteasome subunit genes could be detected in the RNA isolated from cauline leaves.

Arabidopsis α3, α5, and β3 subunits can functionally replace their yeast orthologs: Amino acid sequence conservation between the Arabidopsis 20S proteasome subunits and their counterparts in other organisms suggested that the Arabidopsis subunit families assume identical locations in the particle and perform similar functions (Figure 1). To examine whether this primary sequence homology could be translated into functional homology, we tested in yeast whether nine of the *A. thaliana* subunits (α3, α5, and β1–β7) could functionally replace their corresponding yeast orthologs (Table 1). For the eight subunits tested that are essential for yeast [α5/DOA5, β1/PRE3, β2/PUP1, β3/PUP3, β4/PRE1, β5/DOA3, β6/PRS3, and β7/PRE4 (Hochstrasser 1996)], yeast strains were constructed which bear a chromosomal deletion of the respective proteasome subunit gene but express the corresponding wild-type gene from a *URA3*-plasmid (Table 3). High copy *TRP1*-plasmids bearing the various Arabidopsis 20S proteasome subunits were then introduced into these haploid yeast deletion strains by transformation, and the presence of the plasmid-born copy of the yeast gene was selected against by plating on 5-fluoroorotic acid (FOA) which is toxic to cells expressing Ura3. Streaked transformants that formed colonies on FOA plates were those in which the Arabidopsis 20S proteasome gene was able to rescue the loss of the yeast subunit gene.

All eight tested yeast strains bearing the 20S proteasome subunit genes on a *URA3*-plasmid could grow on FOA-containing media when transformed with a *TRP1*- or *LEU2*-plasmid containing the corresponding yeast subunit but not with a corresponding empty plasmid (Figure 6A). When the eight corresponding Arabidopsis subunits genes were tested similarly, only plasmids bearing *PAE1* (α5) or *PBC2* (β3) rescued colony growth on FOA-containing media for yeast strains bearing deletions in the orthologous genes; *PAE1* (α5) rescued the loss of *DOA5* (α5) and *PBC2* rescued the loss of *PUP3* (β3) (Figure 6B). However, the colony growth rate of the yeast deletion strains harboring either Arabidopsis *PAE1* or *PBC1* was slower than the rate of those containing the corresponding yeast genes, suggesting that the Arabidopsis α5 and β3 subunits only partially restored yeast 20S proteasome function (Figure 6B and data not shown). In contrast, *TRP1*-plasmids bearing *PBA1* (β1), *PBB1* (β2), *PBD1* (β4), *PBE1* (β5), *PBF1* (β6), and *PBG1* (β7) failed to restore viability to yeast strains missing the orthologous yeast subunits (*PRE3*, *PUP1*, *PRE1*, *DOA3*, *PRS3*, and *PRE4*, respectively) (Figure 6B).

To demonstrate that successful complementation of *doa5*Δ and *pup3*Δ by *PAE1* and *PBC2* was specific for these Arabidopsis genes, we attempted to rescue the two yeast strains with other Arabidopsis subunit genes. Using the same FOA selection, none of the six other Arabidopsis α and β subunits genes restored viability to the yeast *doa5*Δ and *pup3*Δ strains. *PAC1* encoding an

TABLE 3

Yeast Strains

Strain	Genotype	Chromosomal deletion
MHY501[a]	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1*	
MHY606[b]	*MAT*α/**a** *his3-*Δ*200/his3-*Δ*200 leu2-3, 112/leu2-3, 112 ura3-52/ura3-52 lys2-801/lys2-801 trp1-1/trp1-1*	
MHY784[b]	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 doa3*Δ*::HIS3* [YCp50-DOA3]	DOA3 (β5)
MHY787[b]	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 doa5*Δ*::HIS3* [pYCp50-DOA5]	DOA5 (α5)
MHY991[c]	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 pup1*Δ*::LEU2* [pRS316-PUP1]	PUP1 (β2)
MHY996[c]	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 pup3-*Δ*1::HIS3* [pRS316-PUP3]	PUP3 (β3)
MHY1028	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 pre1*Δ*::LYS2* [pDP83.PRE1-PRE4]	PRE1 (β4)
MHY1029	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 pre4*Δ*::HIS3* [pDP83.PRE1-PRE4]	PRE4 (β7)
MHY1031	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 pre3*Δ*::HIS3* [pRS316-PRE3]	PRE3 (β1)
MHY1057	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 prs3-*Δ*2::HIS3* [pRS316-PRS3]	PRS3 (β6)
MHY1069	*MAT*α *his3-*Δ*200 leu2-3, 112 ura3-52 lys2-801 trp1-1 pre9*Δ*::HIS3*	PRE9 (α3)

[a] Chen *et al.* 1993.
[b] Chen and Hochstrasser 1995.
[c] Arendt and Hochstrasser 1997.

Arabidopsis α3 subunit was unable to complement the loss of *DOA5* (α5) (Figure 7A). Likewise, *PBA1*, *PBB1*, *PBD1*, *PBE1*, and *PBF1* encoding the Arabidopsis β1, β2, β4, β5, and β6 subunits, respectively, were unable to complement the loss of *PUP3* (β3) (Figure 7B). To examine the possibility that the Arabidopsis β3 subunit could complement deletions of other yeast β subunits besides PUP3 (β3), we attempted to rescue viability of the yeast *pup1*Δ (β2), *pre1*Δ (β4), *doa3*Δ (β5), and *prs3*Δ (β6) strains with *PBC2*. None produced viable colonies on FOA-containing media, indicating that Arabidopsis PBC2 (β3) can exchange only with its yeast ortholog (data not shown).

Although the yeast α3 subunit encoded by *PRE9* is not essential, a haploid yeast strain containing a chromosomal deletion of *PRE9* (Table 3) is sensitive to the amino acid analogs canavanine and *p*-fluorophenylalanine (Emori *et al.* 1991 and Figure 6C). Resistance to the analogs could be restored by extrachromosomal expression of Arabidopsis *PAC1* (α3) on a centromeric low copy plasmid (Figure 6C). When adjusting for cell number before plating, the growth rate of the yeast *pre9*Δ strain containing *PAC1* was indistinguishable from that of wild-type yeast. Complementation of the analog sensitivity of *pre9*Δ was subunit-specific since introduction of another Arabidopsis α subunit gene *PAE1* (α5) did not restore growth as compared to Arabidopsis *PAC1* (Figure 6C). In fact, expression of *PAE1* appeared to slightly reduce growth beyond that seen for *pre9*Δ.

DISCUSSION

Here we describe a collection of 23 genes encoding the full complement of subunits within the *A. thaliana* 20S proteasome. As in other eukaryotes, these genes can be organized into the α and β subunit families encoding the seven distinct α subunits and seven distinct β subunits of the 20S particle. Sequence alignments with orthologs from *T. acidophilum*, yeast, and animals indicate substantial amino acid conservation of the subunits occurs across kingdoms. Like others (Zwickl *et al.* 1992), we found that the various α and β subunits from Arabidopsis are more similar in amino acid sequence to corresponding subunits in other organisms than to different Arabidopsis subunits. These data further support the notion that the α and β families were derived from a common ancestral gene and that the multiple subunits in each family subsequently evolved prior to the divergence of the plant, fungal, and animal kingdoms. ESTs were available for most of the 20S subunit genes (the exceptions were *PAB2* and *PAC2*, which were derived from genomic sequences), indicating that a majority of the genes represent functional loci. The expression patterns of six subunits were coincident, suggesting that the genes are coordinately regulated. Enhanced expression was observed in flower and silique tissue, implicating the 20S proteasome in flower development and seed maturation.

During the preparation of this manuscript, Parmentier *et al.* (1997) reported the sequence of a subset of the Arabidopsis 20S proteasome subunit genes described here. In general, their study agrees well with ours except that: (i) several of the cDNAs they reported encoded only part of the full amino-acid sequence (*PAA1*, *PAE1*, *PAD2*, *PAF2*, *PBB1*, *PBG1*); (ii) one gene in their collection likely contains a 2-bp deletion which truncates the reading frame (*PBC1*); (iii) two genes they characterized as unique may be duplicates of each other (homologs of *PAD2*); and (iv) we identified additional family members not found in their collection (*PAA2*, *PAB2*, *PAC2*, *PAE2*, *PBC2*). In contrast to their designations, our nomenclature matches more closely those recommended for Arabidopsis gene assignments (Meinke

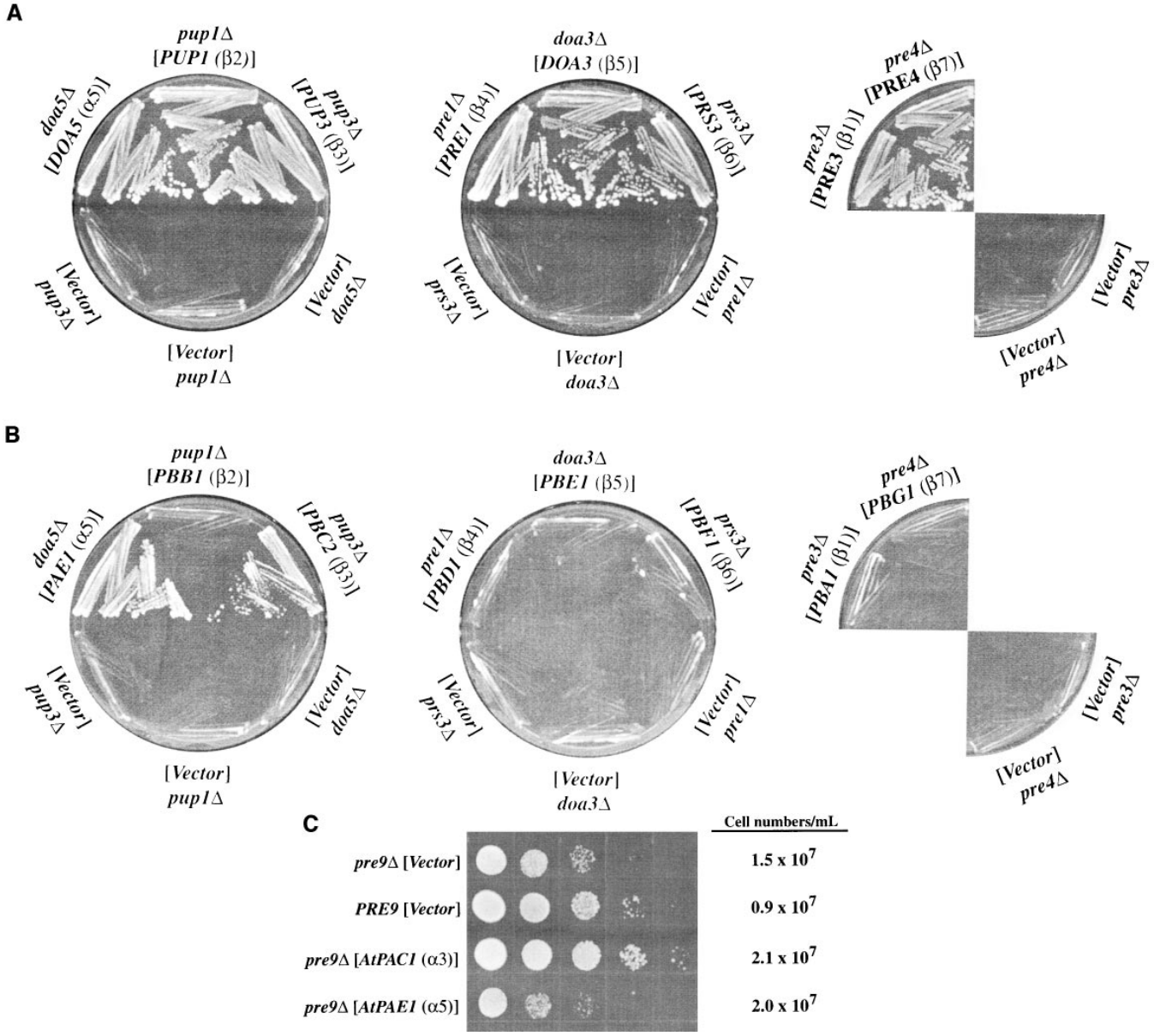

Figure 6.—Functional complementation of yeast 20S proteasome subunit genes with orthologs from *A. thaliana*. For the essential yeast subunit genes *DOA5* (α5), *PRE3* (β1), *PUP1* (β2), *PUP3* (β3), *PRE1* (β4), *DOA3* (β5), *PRS3* (β6), and *PRE4* (β7), tests were performed with yeast strains bearing a chromosomal deletion for the corresponding 20S subunit gene and containing the wild-type gene on a *URA3*-plasmid. These yeast strains were transformed with an empty *TRP1*- or *LEU2*-plasmid (Vector Laboratories, Burlingame, CA) or one expressing the corresponding yeast (A) or Arabidopsis (B) subunit gene. The transformants were patched on YPD medium, incubated overnight, then streaked on FOA-containing media, and grown for 6 days at 30°. (C) Complementation of the amino-acid-analog growth sensitivity of a yeast strain bearing a chromosomal deletion of *PRE9* (α3) with its Arabidopsis ortholog *PAC1*. Wild-type (*PRE9*) yeast or the *pre9*Δ strain was transformed with either an empty *TRP1*-plasmid (Vector) or one expressing Arabidopsis *PAC1* (α3) or Arabidopsis *PAE1* (α5). Cultures from the various transformants were resuspended to similar cell densities and spotted in a 10-fold dilution series on SD medium containing canavanine and *p*-fluorophenylalanine. The cells were grown for 7 days at 30°. Cell densities of the resuspended cultures prior to plating on the analog-containing media are shown on the right.

and Koornneef 1997) and thus were retained here (see Table 1). Parmentier *et al.* (1997) and the various Arabidopsis genome sequencing projects have mapped many of the subunit genes on the *A. thaliana* chromosomes and found them to be randomly distributed.

Based on overall amino acid sequence conservation, we expect that the families of Arabidopsis α and β subunits assume three dimensional structures similar to their *T. acidophilum* and yeast orthologs (Löwe *et al.* 1995; Groll *et al.* 1997). This is especially true for the β subunits, β1, β2 and β5, where almost all of the residues essential for catalysis, substrate specificity, and inhibitor

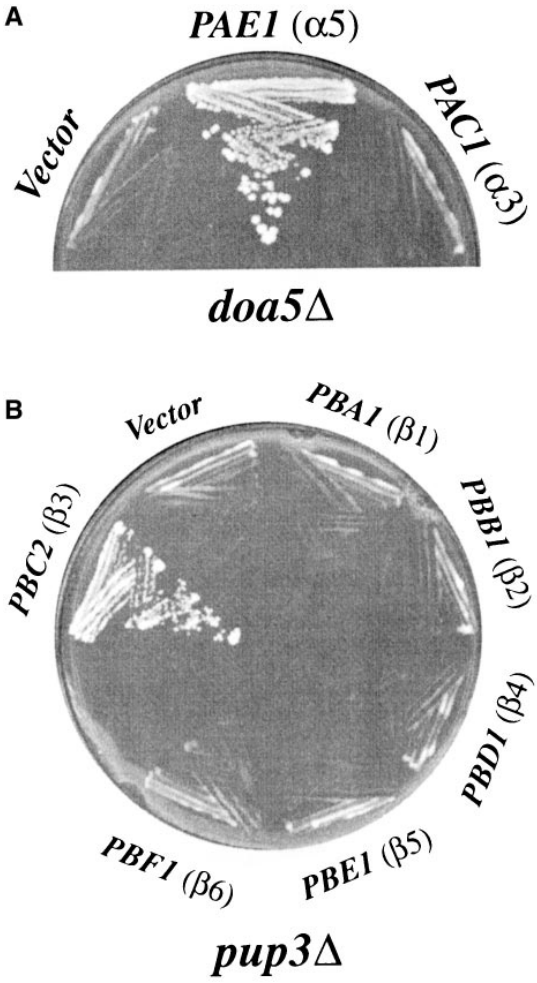

Figure 7.—Specificity of functional complementation of 20S subunit genes between Arabidopsis and yeast. To determine whether functional replacement of yeast *DOA5* (α5) and *PUP3* (β3) is specific to the corresponding *A. thaliana* subunit genes, *PAE1* and *PBC2*, respectively, complementation of the various subunit deletion strains was tested with other Arabidopsis α and β subunit genes. (A) Complementation of *doa5*Δ containing wild-type *DOA5* on a *URA3*-plasmid, with empty *TRP1*-plasmid (Vector) or one expressing Arabidopsis *PAC1* (α3) or *PAE1* (α5). (B) Complementation of *pup3*Δ containing wild-type *PUP3* on a *URA3*-plasmid, with an empty *TRP1*-plasmid (Vector) or one expressing either Arabidopsis *PBA1* (β1), *PBB1* (β2), *PBC2* (β3), *PBD1* (β4), *PBE1* (β5), or *PBF1* (β6). In all tests, the transformants were patched on YPD medium, incubated overnight, then streaked on FOA-containing media, and grown for six days at 30°.

binding are contextually conserved (Figure 3). Moreover, N-terminal extensions similar to the prosequences found in their yeast and mammalian counterparts likely indicate that the Arabidopsis β1, β2, β5, β6, and β7 subunits are synthesized as longer precursors which then require proteolytic processing to release the mature subunits that assemble into the 20S particle (Hochstrasser 1996; Schmidt and Kloetzel 1997). For β1, β2, and β5, the exposed Thr residue at the N- terminus then would become part of the catalytic site of the subunit.

For three of the seven Arabidopsis subunits tested, this primary sequence similarity could be translated into functional homology. Genes encoding the α subunits, α3/PAC1 and α5/PAE1, and the β subunit, β3/PBC2, were able to specifically complement deletions of their yeast counterparts. *PAE1* (α5) and *PBC2* (β3) rescued yeast deletions of the essential subunits *DOA5* (α5) and *PUP3* (β3), respectively. *PAC1* (α3) restored amino acid-analog resistance to a yeast strain missing *PRE9* (α3). Previously, Seelig *et al.* (1993) showed by stable transfection of mouse cultured cells that a Drosophila α2 subunit could structurally replace its ortholog in the mouse 20S proteasome. Here we extend this conservation by showing for the first time that proteasome subunits from one species can functionally replace their orthologs in another species. This cross-species complementation is remarkable given the extent of sequence divergence between the yeast and Arabidopsis α3, α5, and β3 subunits (Figure 1) and the multiple structural constraints that must exist for each subunit to allow its proper assembly into and function within the 20S complex (Groll *et al.* 1997).

Those Arabidopsis subunit genes that failed to complement their yeast orthologs were *PBA1* (β1), *PBB1* (β2), *PBD1* (β4), *PBE1* (β5), *PBF1* (β6), and *PBG1* (β7). Whether this failure represents an inability of the Arabidopsis subunits to exchange with the yeast subunits in the complex or simply inadequate expression of functional Arabidopsis proteins is unknown. Interestingly, both the Arabidopsis α subunits tested successfully replaced their yeast counterparts but only one of the seven β subunits tested did. This success rate could be related to the higher sequence similarity of the α subunits between yeast and Arabidopsis (Figure 1). Alternatively, it could imply that the assembly and/or function of the α ring is more tolerant of sequence variations or that the position of the α subunits on the periphery of the 20S cylinder confers greater structural freedom. These possibilities are supported by the observation that the only non-essential polypeptide in yeast is an α subunit [α3/PRE9 (Emori *et al.* 1991)]; presumably in a *pre9*Δ strain, another yeast α subunit must replace the missing PRE9 polypeptide to maintain the seven-membered α ring. With respect to the β1, β2, β5, β6, and β7 subunits, it is also conceivable that the Arabidopsis prosequences necessary for 20S particle assembly are cleaved poorly in yeast (Schmidt and Kloetzel 1997).

From DNA sequence alignments and DNA gel blot analyses, we can generate a nearly complete set of the *A. thaliana* genes encoding 20S proteasome subunits. In the α family, the majority of the subunits [*PAA-PAF* (α1–α6)] appear to be encoded by two genes, all of which are available in the EST and genomic databases. Only α7 appears to be encoded by a single gene (*PAG1*). In the β family, β1, β6, and β7 appear to be encoded by a single gene (*PBA1*, *PBF1*, and *PBG1*, respectively) and β5 appears to be encoded by only two genes. β2,

β3, and β4 synthesis may involve a third gene in addition to the two reported here. Taken together, the data suggest that the collection of proteasome sequences described here is missing at most only four more genes.

Whereas each 20S proteasome subunit is encoded by a single gene in yeast, many of the Arabidopsis subunits are encoded by two or more. What is the reason for this genetic redundancy? One simple explanation is that the duplications provide backup copies and/or insure that sufficient amounts of the corresponding subunits are produced. Similar duplications have been observed for many other proteins in the ubiquitin/26S proteasome pathway besides those in the 20S complex (Vierstra 1996). Genetic analysis also supports this explanation. Shirley and Goodman (1993) previously isolated a radiation-induced deletion of the α6 subunit gene *PAF1*; the mutant was phenotypically normal suggesting that the *PAF2* paralog is sufficient for proteasome function. Likewise, we have recently identified a T-DNA insertion mutant of the α5 subunit gene *PAE2*; this mutant is also viable, suggesting that its relative *PAE1* provides adequate α5 function (J. H. Doelling and R. D. Vierstra, unpublished results).

Alternatively, paralogs may be differentially regulated to allow synthesis of the subunit across a wider range of developmental states and environmental conditions. Similar situations have been observed in animals. For example, up-regulation of one or more 20S subunit genes has been reported in malignant hematopoietic cells and in mitotically-activated mononuclear cells of humans (Kumatori *et al.* 1990), during programmed cell death of *Manduca sexta* intersegmental muscles (Jones *et al.* 1995), in oocyte maturation of the starfish *Asterina pectinifera* (Sawada *et al.* 1992), and following γ-interferon induction in humans (Coux *et al.* 1996). In Arabidopsis, Genschik *et al.* (1994) found that the levels of mRNA from *PAD1* and *PBF1*, encoding the α4 and β6 subunits, respectively, were increased in proliferating cultured cells, implicating these gene products in cell division events.

Another intriguing possibility is that members of each α-type and β-type family encode proteins that impart distinct proteolytic specificities and/or functions to the 20S complex (*e.g.*, assembly with the 19S regulatory complex to form the 26S proteasome). The former strategy appears to be important for the production of appropriate antigenic peptides for the MHC class I immune response pathway in humans (Coux *et al.* 1996). Here, a unique form of the 20S proteasome is assembled by replacing the catalytic β subunits, β1/Y, β2/Z, and β5/X with related subunits, LMP2, MECL1, and LMP7. *LMP2* and *LMP7* genes are located within the MHC gene cluster and are transcribed upon γ-interferon treatment. Integration of LMP2 and LMP7 into the 20S particle appears to change the profile of peptide digestion products toward those favored by MHC class I presentation (Coux *et al.* 1996). Conceivably, a similar strategy could be operating in plants using the different α and β family members to alter the substrate specificity and/or peptidase activity of the 20S complex. However, because the paralogs of the Arabidopsis subunit families show greater similarity between each other than do the inducible human LMP2, MECL1 and LMP7 subunits with respect to their constitutive counterparts (90–99% versus 68–77% similarity), the integration of different subunits into the Arabidopsis complex may have more subtle effects.

We thank Dr. Steven van Nocker for assistance with the yeast complementation, Drs. Peggy Hatfield and Sharon Thoma for preliminary sequence determinations, Dr. Mark Gosink for help with the sequence analysis, and the *Arabidopsis* DNA Stock Center at Ohio State University for making available the EST clones analyzed in this study. This work was supported by the U.S. Department of Agriculture-National Research Initiative Competitive Grants Program (97-35301-4218) and the Research Division of the University of Wisconsin College of Agriculture and Life Sciences (Hatch Grant 142-3936) to R.D.V., the National Institutes of Health Grant (GM-46904) to M.H., a National Institutes of Health Postdoctoral Fellowship (1F32 GM19136-01) to J.H.D., and a Howard Hughes Predoctoral Fellowship to C.S.A.

LITERATURE CITED

Arendt, C. S., and M. Hochstrasser, 1997 Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation. Proc. Natl. Acad. Sci. USA **94:** 7156–7161.

Callis, J. A., T. B. Carpenter, C.-W. Sun and R. D. Vierstra, 1995 Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. Genetics **139:** 921–939.

Chen, P., and M. Hochstrasser, 1995 Biogenesis, structure, and function of the yeast 20S proteasome. EMBO J. **14:** 2620–2630.

Chen, P., and M. Hochstrasser, 1996 Autocatalytic subunit processing couples active site formation in the 20S proteasome to completion of assembly. Cell **86:** 961–972.

Chen, P., P. Johnson, T. Sommer, S. Jentsch and M. Hochstrasser, 1993 Multiple ubiquitin-conjugating enzymes participate in the *in vivo* degradation of the yeast MATα2 repressor. Cell **74:** 357–369.

Cone, K. C., E. B. Frisch and T. E. Phillips, 1989 *dek1* interferes with aleurone differentiation. Maize Genetics Coop. Newslett. **63:** 67–68.

Coux, O., H. G. Nothwang, I. Silva Pereira, F. Recillas Targa, F. Bey *et al.*, 1994 Phylogenic relationships of the amino acid sequences of prosome (proteasome, MCP) subunits. Mol. Gen. Genet. **245:** 769–780.

Coux, O., K. Tanaka and A. L. Goldberg, 1996 Structure and functions of the 20S and 26S proteasomes. Annu. Rev. Biochem. **65:** 801–847.

Deveraux, J., P. Haberberli and O. Smithies, 1984 A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:** 387–395.

Emori, Y., T. Tsukahara, H. Kawasaki, S. Ishiura, H. Sugita *et al.*, 1991 Molecular cloning and functional analysis of three subunits of the yeast proteasome. Mol. Cell. Biol. **11:** 344–353.

Enenkel, C., H. Lehmann, J. Kipper, R. Gückel, W. Hilt *et al.* 1994 *PRE3*, highly homologous to the human major histocompatibility comples-linked LMP2 (RING12) genes, codes for a yeast proteasome subunit necessary for the peptidylglutamyl- peptide hydrolyzing activity. FEBS Lett. **341:** 193–196.

Fu, H., S. Sadis, D. M. Rubin, M. Glickman, S. van Nocker *et al.*, 1998 Multiubiquitin chain binding and protein degradation are mediated by distinct domains within the 26 S proteasome subunit Mcb1. J. Biol. Chem. **273:** 1970–1981.

Fujinami, K., N. Tanahashi, K. Tanaka, A. Ichihara, Z. Cejka *et al.*,

1994 Purification and characterization of the 26 S proteasome from spinach leaves. J. Biol. Chem. **269:** 25905–25910.

Genschik, P., G. Phillips, C. Gigot and J. Fleck, 1992 Cloning and sequence analysis of a cDNA clone from *Arabidopsis thaliana* homologous to a proteasome α subunit from *Drosophila*. FEBS Lett. **309:** 311–315.

Genschik, P., E. Jamet, G. Phillips, Y. Parmentier, C. Gigot *et al.*, 1994 Molecular characterization of a β-type proteasome subunit from *Arabidopsis thaliana* co-expressed at a high level with an α-type proteasome subunit early in the cell cycle. Plant J. **6:** 537–546.

Gietz, R. D., and A. Sugino, 1988 New yeast-*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six-base pair restriction sites. Gene **74:** 527–534.

Gietz, R. D., R. H. Schiestl, A. R. Willems and R. A. Woods, 1995 Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast **11:** 355–360.

Groll, M., L. Ditzel, J. Löwe, D. Stock, M. Bochtler *et al.*, 1997 Structure of 20S proteasome from yeast at 2.4 Å resolution. Nature **386:** 463–471.

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.

Heinemeyer. W., A. Gruhler, V. Möhrle, Y. Mahé and D. H. Wolf, 1993 PRE2, highly homologous to the human major histocompatibility complex-linked *Ring10* gene, codes for a yeast proteasome unit necessary for chymotryptic activity and degradation of ubiquitinated proteins. J. Biol. Chem. **268:** 5115–5120.

Heinemeyer, W., M. Fischer, T. Krimmer, U. Stachon and D. H. Wolf, 1997 The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing. J. Biol. Chem. **272:** 25200–25209.

Hilt, W., C. Enenkel, A. Gruhler, T. Singer and D. H. Wolf, 1993 The *PRE4* gene codes for a subunit of the yeast proteasome necessary for peptidylglutamyl-peptide-hydrolyzing activity: mutations link the proteasome to stress- and ubiquitin-dependent proteolysis. J. Biol. Chem **268:** 3479–3486.

Hochstrasser, M., 1996 Ubiquitin-dependent protein degradation. Annu. Rev. Genet **30:** 405–439.

Jones, M. E., M. F. Haire, P.-M. Kloetzel, D. L. Mykles and L. M. Schwartz, 1995 Changes in the structure and function of the multicatalytic proteinase (proteasome) during programmed cell death in the intersegmental muscles of the hawkmoth, Manduca sexta. Dev. Biol. **169:** 436–447.

Kaiser, C., S. Michaelis and A. Mitchell, 1994 *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Knowlton, J. R., S. C. Johnston, F. C. Whitby, C. Realini, Z. Zhang *et al.*, 1997 Structure of the proteasome activator REGα (PA28α). Nature **390:** 639–643.

Kremp, A., M. Schliephacke, U. Kull and H.-P. Schmid, 1986 Prosomes exist in plant cells too. Exp. Cell Res. **166:** 553–557.

Kumatori, A., K. Tanaka, N. Inamura, S. Slone, T. Ogura *et al.*, 1990 Abnormally high expression of proteasomes in human leukemic cells. Proc. Natl. Acad. Sci. USA **87:** 7071–7075.

Löwe, J., D. Stock, F. Jap, P. Zwickl, W. Baumeister *et al.*, 1995 Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4Å resolution. Science **268:** 533–539.

Maupin-Furlow, J. A., and J. G. Ferry, 1995 A proteasome from the methanogenic archaeon, *Methanosarcina thermophila*. J. Biol. Chem. **270:** 28617–28622.

Meinke, D., and M. Koornneef, 1997 Community standards for Arabidopsis genetics. Plant J. **12:** 247–253.

Ozaki, M., K. Fujinami, D. Tanaka, Y. Ameiniya, T. Sato *et al.*, 1992 Purification and initial characterization of the proteasome from the higher plant *Spinacia oleracea*. J. Biol. Chem. **267:** 21678–21684.

Parmentier, Y., D. Bouchez, J. Fleck and P. Genschik, 1997 The proteasome gene family in *Arabidopsis thaliana*. FEBS Lett. **416:** 281–285.

Rapp, J. C., B. J. Baumgartner and J. E. Mullet, 1992 Quantitative analysis of transcription and RNA levels of 15 barley chloroplast genes. J. Biol. Chem. **267:** 21404–21414.

Sawada, M. T., T Someno, M. Hoshi and H. Sawada, 1992 Participation of 650-kDa protease (20S proteasome) in starfish oocyte maturation. Dev. Biol. **150:** 414–418.

Schliephacke, M., A. Kremp, H.-P. Schmid, K. Kohler and U. Kull, 1991 Prosomes (proteasomes) of higher plants. Eur. J. Cell Biol. **55:** 114–121.

Schmidt, M., and P. M. Kloetzel, 1997 Biogenesis of eukaryotic 20S proteasomes: the complex maturation pathway of a complex enzyme. FASEB J. **11:** 1235–1243.

Schmidtke, G., R. Kraft, S. Kostka, P. Henklein, C. Frömmel *et al.*, 1996 Analysis of mammalian 20S proteasome biogenesis: the maturation of β subunits is an ordered two-step mechanism involving autocatalysis. EMBO J. **15:** 6887–6898.

Seelig, A., G. Multhaup, B. Pesold-Hult, K. Beyreuther and P.-M. Kloetzel, 1993 Drosophila proteasome Dm25 subunit substitute the mouse MC3 subunit in hybrid proteasomes. J. Biol. Chem. **268:** 25561–15567.

Seemüller, E., A. Lupas, D. Stock, J. Löwe, R. Huber *et al.*, 1995 Proteasome from *Thermoplasma acidophilum*: a threonine protease. Science **268:** 579–582.

Seemüller, E., A. Lupas and W. Baumeister, 1996 Autocatalytic processing of the 20S proteasome. Nature **382:** 468–471.

Shirley, B. W., and H. M. Goodman, 1993 An *Arabidopsis* gene homologous to mammalian and insect genes encoding the largest proteasome subunit. Mol. Gen. Genet. **241:** 586–594.

Sikorski, R. S., and P. Hieter, 1989 A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics **122:** 19–27.

Skoda, B., and L. Malek, 1992 Dry pea seed proteasome. Plant Physiol. **99:** 1515–1519.

Spence, J., S. Sadis, A. L. Haas and D. Finley, 1995 A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. Mol. Cell. Biol. **15:** 1265–1273.

Sun, C.-W., and J. Callis, 1997 Independent modulation of *Arabidopsis thaliana* polyubiquitin mRNAs in different organs and in response to environmental changes. Plant J. **11:** 1017–1027.

Tamura, T., I. Nagy, A. Lupas, F. Lottspeich, Z. Cejka *et al.*, 1995 The first characterization of a eubacterial proteasome: the 20S complex of *Rhodococcus*. Curr. Biol. **5:** 766–774.

Tanaka, K., T. Yoshimura, T. Tamura, T. Fujiwara, A. Kumatori *et al.*, 1990 Possible mechanism of nuclear translocation of proteasomes. FEBS Lett. **271:** 41–46.

Vierstra, R. D., 1996 Proteolysis in plants: mechanisms and functions. Plant Mol. Biol. **32:** 275–302.

Zwickl, P., A. Grizwa, G. Pühler, B. Dahlmann, F. Lottspeich *et al.*, 1992 Primary structure of the Thermoplasma proteasome and its implications for the structure, function, and evolution of the multicatalytic proteinase. Biochemistry **31:** 964–972.

Communicating editor: D. Preuss