Primary Structure of the *Thermoplasma* Proteasome and Its Implications for the Structure, Function, and Evolution of the Multicatalytic Proteinase†

Peter Zwickl,‡ Anja Grziwa,‡ Gabriela Pühler,‡ Burkhardt Dahlmann,§ Friedrich Lottspeich,‡ and Wolfgang Baumeister*,‡

Max-Planck-Institut für Biochemie, D-8033 Martinsried, FRG, and Diabetes-Forschungsinstitut an der Universität Düsseldorf, D-4000 Düsseldorf, FRG

Received September 25, 1991; Revised Manuscript Received November 14, 1991

ABSTRACT: The proteasome or multicatalytic proteinase is a high molecular mass multisubunit complex ubiquitous in eukaryotes but also found in the archaebacterium *Thermoplasma acidophilum*. While eukaryotic proteasomes contain 15–20 different subunits, the archaebacterial proteasome is made of two different subunits only, and yet the complexes are almost identical in size and shape. Cloning and sequencing the gene encoding the small (β) subunit of the *T. acidophilum* complex completes the primary structure of the archaebacterial proteasome. The similarity of the derived amino acid sequences of 233 (α) and 211 (β) residues, respectively, indicates that they arose from a common ancestral gene. All the sequences of proteasomal subunits from eukaryotes available to date can be related to either the α-subunit or β-subunit of the *T. acidophilum* "Urproteasome", and they can be distinguished by means of a highly conserved N-terminal extension, which is characteristic for α-type subunits. On the basis of circumstantial evidence we suggest that the α-subunits have regulatory and targeting functions, while the β-subunits carry the active sites.

The proteasome is a nonlysosomal high M_r multicatalytic proteolytic complex ubiquitous in eukaryotic cells [for recent reviews see Orlowski (1990) and Rivett (1989)]. It has been shown (Arrigo et al., 1988; Falkenburg et al., 1988) to be identical with the prosome, a molecular complex supposed to be involved in the regulation of gene expression because of its purported association with RNA (Schmid et al., 1984; Scherrer, 1990), and with cylindrin (Harris, 1968), a major component of erythrocytes. Characteristically, eukaryotic proteasomes contain 15–20 different polypeptides, all in the M_r range between 20000 and 35000. The currently available amino acid sequences of several proteasomal subunits from yeast (six subunits), *Drosophila* (three subunits), *Xenopus* (two subunits), mouse (one subunit), rat (five subunits), and man (eleven subunits) have revealed a high degree of conservation, making it likely that all proteasome subunits are encoded by members of one gene family (Fujiwara et al., 1990; Haass et al., 1990a). Although it is now widely agreed that proteasomes possess multiple proteolytic activities, the physiological role of proteasomes has yet to be unraveled. It has been proposed that a higher M_r complex (26S), with the proteasome or some of its subunits as the "catalytic core", degrades proteins labeled with ubiquitin in an ATP-dependent reaction (Hough et al., 1988; Matthews et al., 1989; Eytan et al., 1989; Driscoll & Goldberg, 1990). Heinemeyer et al. (1991) presented genetic evidence that proteasomes are involved in the degradation of ubiquitinated proteins. Disruption of genes encoding yeast proteasomal subunits revealed that some of them are essential for growth (Fujiwara et al., 1990; Emori et al., 1991; Heinemeyer et al., 1991). Recently Parham (1990) drew attention to similarities between the proteasome and the "low-molecular-weight proteins" (LMP)[1] (Monaco & McDevitt, 1982, 1984), an intracellular protein complex of unknown function with a M_r of 580000 and composed of 16 polypeptides (M_r 15000–30000), and he put forward the hypothesis that the proteasome-related LMP particle is a protease involved in antigen processing.

We have isolated proteasomes from the archaebacterium *Thermoplasma acidophilum* which are composed of two different polypeptides only (Dahlmann et al., 1989); the α-subunit has an apparent M_r of 27000 and the β-subunit 25000. In spite of their much simpler subunit composition, proteasomes from *T. acidophilum* and proteasomes of eukaryotic origin have the same basic architecture (Dahlmann et al., 1989; Baumeister et al., 1988). Electron micrographs show the barrel-shaped proteasome complex (Hegerl et al., 1991) in two basic projections: rectangular side views (15 nm long and 11 nm wide) with a characteristic pattern of four striations and ring-shaped end-on views (11 nm in diameter) exhibiting a weak, but significant 7-fold symmetry (Dahlmann et al., 1989; Zwickl et al., 1990). We have shown by immunoelectron microscopy that the α-subunits form the "disks" closing off the barrel at both ends, while the β-subunits constitute the two juxtaposed inner rings (Grziwa et al., 1991).

In this paper we present and compare the amino acid sequences of the two subunits of the *T. acidophilum* proteasome—the first complete primary structure of a proteasome. We show that the various subunits of eukaryotic proteasomes which have been sequenced to date can be related to either the α-subunit or β-subunit. Hence the *T. acidophilum* proteasome may be seen as an ancestral or "Urproteasome" from which the eukaryotic proteasomes developed

† The genetic sequence has been submitted to GenBank under Accession Number J05358.

* Address correspondence to this author at Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-8033 Martinsried, FRG.

‡ Max-Planck-Institut für Biochemie.

§ Diabetes-Forschungsinstitut.

[1] Abbreviations: ABTS, 2,2′-azinobis(3-ethylbenzothiazoline-sulfonate); BSA, bovine serum albumin; CNBr, cyanogen bromide; ELISA, enzyme-linked immunosorbent assay; LMP, low-molecular-weight proteins; NLS, nuclear localization signal; PAGE, polyacrylamide gel electrophoresis; PBS, 0.01 M sodium phosphate and 0.85% NaCl; PVDF, polyvinylidene difluoride; SDS, sodium dodecyl sulfate; 2-DE, two-dimensional gel electrophoresis.

0006-2960/92/0431-964$03.00/0

by multiple gene duplication events.

Materials and Methods

Purification of the Proteasome. The proteasomes from *T. acidophilum* were isolated and purified as described previously (Dahlmann et al., 1989).

Two-Dimensional Gel Electrophoresis, Amino Acid Analysis, Immunoblotting, and Glycan Detection. For amino acid analysis the subunits of the *T. acidophilum* proteasome were separated by two-dimensional gel electrophoresis (2-DE), electroblotted onto a siliconized glass-fiber sheet (Glassybond, Biometra, Göttingen, FRG), detected by Coomassie blue staining, excised, and subjected to amino acid analysis (Eckerskorn et al., 1988a). For immunoblotting the subunits of the *T. acidophilum* proteasome were separated by 2-DE, electroblotted onto a Immobilon PVDF membrane, and detected by subunit-specific antibodies, raised in rabbits (Grziwa et al., 1991). Glycan detection was performed with the DIG Glycan detection kit following the description provided with the kit, except for using Tricine–SDS–PAGE (Schägger & von Jagow, 1987) to separate the subunits.

Protein Sequencing. For N-terminal protein sequencing the two subunits of the *T. acidophilum* proteasome were separated by Tricine–SDS–PAGE (Schägger & von Jagow, 1987), electroblotted onto a siliconized glass-fiber sheet, detected by Coomassie blue staining, excised, and subjected to sequence degradation (Eckerskorn et al., 1988b). Tryptic and cyanogen bromide peptides from the β-subunit protein were generated, separated, and sequenced as described (Zwickl et al., 1991).

Gene Cloning and Sequencing. The gene encoding the small (β) subunit of the proteasome from *T. acidophilum* was cloned and sequenced as reported for the large (α) subunit (Zwickl et al., 1991).

Sequence Alignment and Motif Search. The N-terminal amino acid sequences of various subunits of eukaryotic proteasomes and the *T. acidophilum* α- and β-subunits were aligned with "CLUSTAL" (Higgins & Sharp, 1988). The two archaebacterial proteasomal proteins were compared with "PIRALIGN" (Orcutt et al., 1986). Search for protein motifs in the amino acid sequence of the *T. acidophilum* β-subunit was performed with "PROSITE" (Bairoch, 1990).

Preparation of Antisera and Affinity Purification of Antiproteasome Antibodies. Antisera against purified proteasomes were raised in rabbits. Preimmune sera were taken, and each animal was injected subcutaneously with an emulsion consisting of 1 volume of proteasome solution (150 μg in 500 μL of 0.01 M PBS, pH 7.2) and 1 volume of Freund's complete adjuvant. After 4 and 6 weeks booster injections of 100 μg of proteasome in Freund's incomplete adjuvant were given. Ten days later, the rabbits were bled completely and sera collected and stored at −20 °C. The antiproteasome antibodies were affinity-purified on a CNBr-activated Sepharose CL-4B column onto which the whole enzyme had been coupled (1 mg/mL swollen gel). In order to maintain the structure of the whole complex during the harsh antibody elution conditions, the Sepharose-coupled enzyme was cross-linked with 0.1% glutaraldehyde in PBS, pH 7.2, for 1 h. One milliliter of the antiserum was diluted 1:10 with PBS, pH 7.5, and circulated for 3 h (flow rate: 15 mL/h) on the column, which had been equilibrated with PBS. After unbound material had been washed out with PBS, the antibodies were released from the column with 3 M KSCN dissolved in PBS, pH 7.2, and the eluate was dialyzed extensively against PBS.

Synthesis of Pin-Bound Peptides and Epitope Mapping. Overlapping nonapeptides, starting from amino acid 1 through 233 of the α-subunit protein and 1 through 211 of the proteasome from *T. acidophilum*, were synthesized by the Pepscan method (Geysen et al., 1984), modified for Fmoc chemistry. The kit from Cambridge Research Biochemicals (U.K.) was used, following the manufacturer's instructions. The peptides are synthesized in covalent linkage to polypropylene rods having the format and the spacing of the wells in a microtiter plate. This allows subsequent ELISA to be rapidly performed. Control peptides known to be reactive with control sera were simultaneously synthesized and found to be immunologically reactive as expected. The support-coupled peptides were precoated with blocking buffer [1% bovine serum albumin (BSA) and 1% ovalbumin in PBS, pH 7.2] for 1 h at room temperature to block nonspecific adsorption of antibodies. The affinity-purified antibodies as well as the crude sera were diluted 1:200 to 1:500 in blocking buffer, depending on their titer against the whole enzyme, and incubated overnight at 4 °C. Three washes with 0.05% Tween 20 and PBS, pH 7.2, were followed by a 1-h incubation at room temperature with a 1:2000 dilution of goat-anti-rabbit IgG coupled to horseradish peroxidase in blocking buffer. Following this, three extensive washing steps removed excess conjugate. The presence of peptide-bound antibodies was detected by reaction for 20 min with a freshly prepared developing solution (0.5 mg/mL ABTS and 0.01% H_2O_2 in 0.1 M disodium hydrogen orthophosphate and 0.08 M citric acid, pH 4), and the color produced was read in a Titertek Multiskan photometer at 405 nm. Prior to retesting, bound antibody was removed from the peptides by ultrasonication of the polypropylene rods in a solution containing 1% SDS and 2 mM mercaptoethanol in 0.1 M sodium phosphate, pH 7.2, for 30 min at 60 °C.

Results

Two-Dimensional Gel Electrophoresis, Amino Acid Composition Analysis, and Immunoblotting of the Proteasomal Proteins. While in one-dimensional SDS–PAGE purified *T. acidophilum* proteasomes yield two bands with apparent M_r's of 27 000 and 25 000 (Dahlmann et al., 1989), 2-DE produces a slightly more complicated pattern. The appearance of the two-dimensional gels tends to be somewhat variable from one preparation to another. Sometimes two major spots with apparent M_r's of 27 000 and 25 000 (pI's of 5.6 and 6.8, respectively) and three minor spots with the same M_r as the larger of them (i.e., the α-subunit) but differing in their isoelectric points (pI's of 5.4, 5.8, and 6.2, respectively) are observed (see Figure 1A). In other preparations, three spots of approximately equal staining intensities appear, two at the higher M_r, but with different pI's (pI's of 5.6 and 5.8, respectively) (see Figure 1B). We have subjected these three protein spots from gels such as the one displayed in Figure 1B to amino acid composition analysis. Within the limits of experimental error, the two prominent high M_r spots turned out to have an identical amino acid composition distinctly different from the third one (Figure 1C). Thus, 2-DE is not in conflict with the proposition that the *T. acidophilum* proteasome is composed of two polypeptide species only; it rather indicates that the α-subunit is prone to modifications altering its pI. This conclusion is corroborated by immunoblotting with antibodies monospecific for the α- and β-subunits (data not shown); the anti-α antibody reacts equally with the two higher M_r spots, while the anti-β antibody reacts exclusively with the lower M_r spot. One possible modification is glycosylation, since glycan detection analysis of blots from Tricine–SDS–PAGE gels gave a positive result for the α-subunit (data not shown).

Cloning and Sequencing of the β-Subunit Gene. For Southern blot hybridization two oligonucleotide mixtures (5′-ATGGARAAYTTYATHATG-3′ and 5′-ATGA-

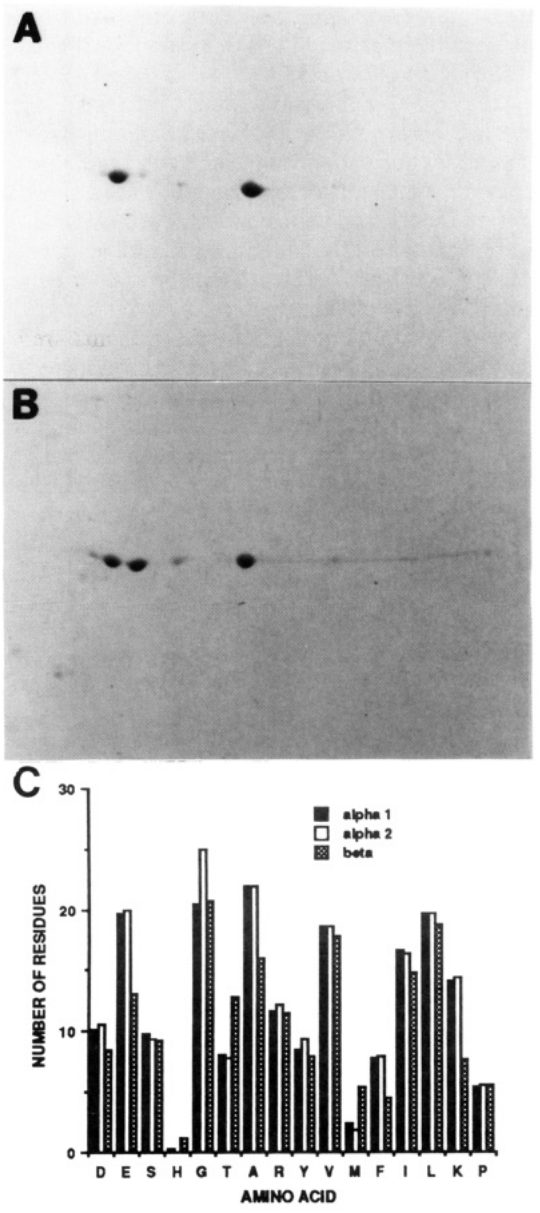

FIGURE 1: 2-DE of *T. acidophilum* proteasomes and amino acid composition analysis of the individual protein spots. Panels A and B show two-dimensional gels of two different proteasome preparations from *T. acidophilum*. Horizontal: isoelectric focusing, cathode left and anode right. Vertical: SDS–PAGE. A total of 10 μg of protein was applied. Panel C shows the result of an amino acid composition analysis of the three main spots visible on the gel shown in panel B. Numbers of the respective amino acid residues—indicated in the one-letter code on the x-axis—per protein spot are given on the y-axis. Within the limits of experimental error α_1 and α_2 have an identical composition, while β is distinctly different.

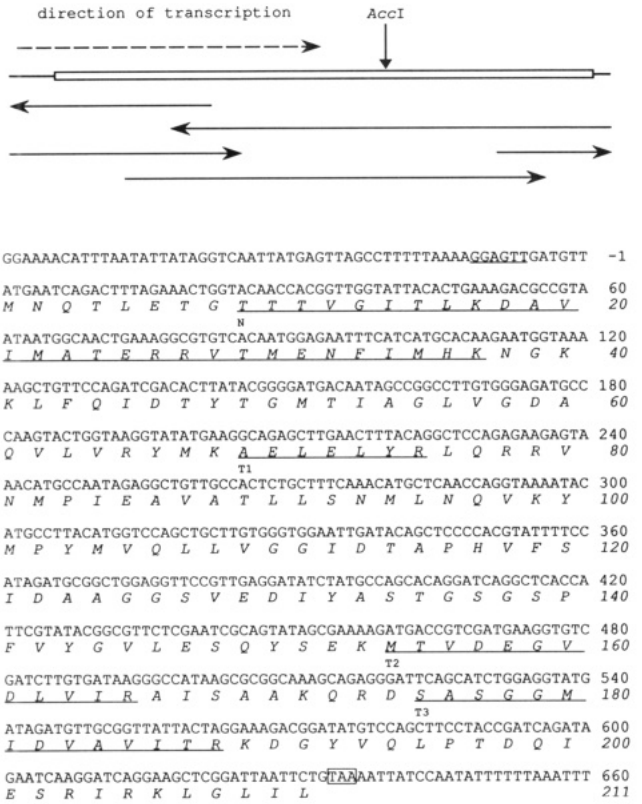

FIGURE 2: Restriction map, sequencing strategy, nucleotide sequence, and derived amino acid sequence of the DNA fragment containing the β-subunit gene. The upper panel shows the restriction endonuclease map of the cloned DNA fragment containing the β-subunit gene. The open box indicates the coding region of the gene. Continuous lines indicate 5′- and 3′-noncoding regions. The vertical arrow indicates a restriction endonuclease cleavage site. Horizontal arrows indicate the individual sequence runs. The dotted arrow shows the direction of transcription of the β-subunit gene. The lower panel shows the nucleotide sequence and the derived amino acid sequence of the DNA fragment containing the β-subunit gene. Nucleotides are numbered beginning with the first residue of the putative initiation codon ATG. The nucleotides upstream of the initiation codon are indicated by negative numbers. The underlined nucleic acids represent the putative Shine–Dalgarno sequence. The stop codon is boxed. The predicted amino acid sequence is shown below the nucleotide sequence in the one-letter code. Amino acid residues are numbered beginning with methionine, the putative first amino acid residue of the translation product. Underlined residues were determined by protein sequencing. T1, T2, and T3 mark the sequences of three tryptic peptides. N marks the sequence, which was obtained by N-terminal sequencing. Apparently the eight additional amino acids upstream of the N-terminal threonine are removed posttranslationally.

THGAYGTNGCNGT-3′) derived from the N-terminus and the tryptic peptide T2 of the β-subunit protein were synthesized (Figure 2). *Hin*dIII-generated DNA fragments of genomic DNA from *T. acidophilum*, giving strong signals with both probes, were ligated in pUC18, and the resulting recombinant plasmids were transformed into *E. coli* DH5α. Colony hybridization of the resulting 1100 clones revealed two positive ones.

The nucleotide sequence of the β-subunit gene and its 5′- and 3′-flanking regions was determined (Figure 2). The amino acid sequences of the N-terminus and of three tryptic peptides, generated from the purified β-subunit protein, determined by protein sequencing are identical with the derived amino acid sequence (Figure 2). The only in-frame initiation codon is located eight codons upstream from the triplet coding for the threonine, which was determined as the N-terminal amino acid residue of the mature protein (Figure 2). Upstream of the initiation codon a putative Shine–Dalgarno sequence (5′-GGAGTTGAT-3′) is identified, which is complementary to the 3′-end of the 16S rRNA (3′-CCUCCACUA-5′) of *T. acidophilum* (Ree et al., 1989). Probing Southern blots of *T. acidophilum* DNA fragments, generated by different independent restriction endonuclease digests, with α- and β-gene-specific oligonucleotides reveals that the genes encoding the *T. acidophilum* proteasomal proteins are not organized in an operon (data not shown).

Analysis of the Derived Amino Acid Sequence of the β-Subunit and Comparison with the α-Subunit. The nucleotide sequence of the β-subunit gene codes for a translation product of 211 amino acid residues, which is posttranslationally processed to a polypeptide of 203 amino acid residues (starting with threonine) and a M_r of 22 272 (Figure 2). Investigation

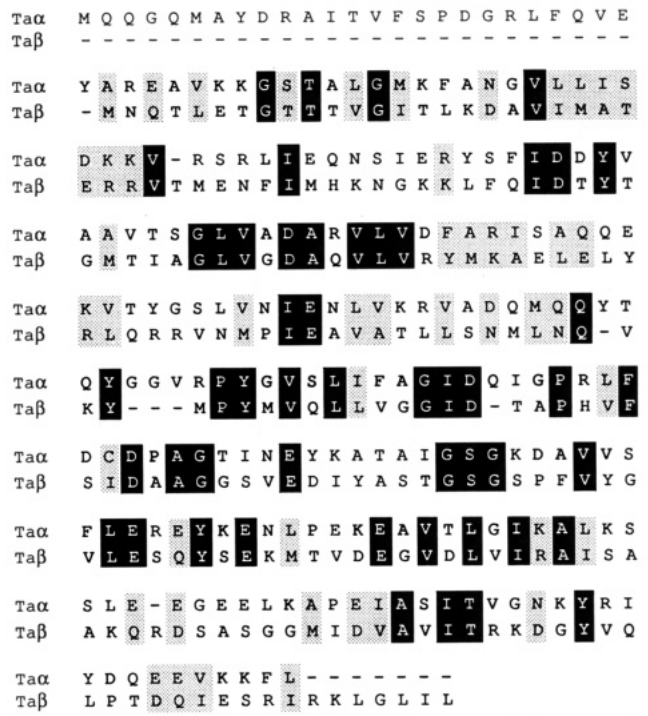

FIGURE 3: Alignment of the amino acid sequences of the *T. acidophilum* α- and β-subunits (Taα, Taβ). Dashes indicate gaps, introduced for better alignment. Identical residues are boxed in black; shaded boxes include conservative exchanges following the classification of Argos (1987): PG, ST, KR, EQND, FWHY, and AIVLMC.

of the derived amino acid sequence of the β-subunit with the computer-edited dictionary of protein sites and patterns PROSITE (Bairoch, 1990) identified no functional sites, in particular no established proteinase motif. Searching for related proteins revealed no sequence similarity of the β-subunit protein with any other protein, except for some eukaryotic proteasomal subunits.

Alignment of the derived amino acid sequences of the α- and β-subunits with the program PIRALIGN (Orcutt et al., 1986) revealed a significant overall similarity with 24% identical or 47% identical plus conserved residues (Figure 3). The highly conserved N-terminal region common to the *T. acidophilum* α-subunit and many eukaryotic proteasome subunits (Zwickl et al., 1991) is missing in the β-subunit.

Comparison of Thermoplasma and Eukaryotic Proteasomal Sequences. Comparing the two amino acid sequences of the *T. acidophilum* proteasome with sequences of the eukaryotic proteasomal subunits by means of the multiple sequence alignment program CLUSTAL (Higgins & Sharp, 1988) reveals significant sequence similarities between all of them (Table I). The similarity between the α- and the β-subunit of the *T. acidophilum* proteasome is in the same range as the similarities between the β-subunit and the various eukaryotic proteasomal proteins (ranging from 14% to 27% identity) (Table I). The β-subunit shows high identity with the yeast proteasomal protein PRE1 (22%), a subunit contributing to the "chymotrypsin-like" proteolytic activity of the proteinase yscE (Heinemeyer et al., 1991), as well as with the putative yeast proteasomal protein PUP1 (26%) (Haffter & Fox, 1991) and the human MHCII encoded putative proteasomal protein RING10 (27%) (Glynne et al., 1991).

Classification of Proteasomal Proteins by N-Terminal Protein Sequence Comparison. N-Terminal sequences of proteasomal proteins were aligned using the program CLUSTAL (Higgins & Sharp, 1988) (Figure 4). It turns out that the proteasomal proteins fall into two groups, either possessing or lacking an N-terminal sequence extension of about 26 residues, which, if present, is highly conserved. Those subunits which have this N-terminal extension share significant similarity with the *T. acidophilum* α-subunit, whereas those not having it are more similar to the *T. acidophilum* β-subunit. Therefore, the genes encoding the several subunits of eukaryotic proteasomes can all be related to either one or the other of the two *T. acidophilum* genes. The assignment of the human and rat C5 subunits is somewhat uncertain. These subunits have an unusually low level of identity (ranging from 12% to 20%) with other proteasomal proteins. Therefore, it remains to be established that the C5 subunits are intrinsic components of proteasomes rather than merely associated proteins.

Localization of Surface-Exposed Sequences by Epitope Mapping. Polyclonal antisera against the whole proteasome reacted immune positive in a conventional ELISA procedure. The affinity purification of the serum on Sepharose-coupled proteasomes selected for those antibodies binding to surface-exposed epitopes. The subsequently performed ELISA's on the overlapping nonapeptides of the α- and β-subunit sequences revealed for the α-subunit one quite outstanding signal for the peptides numbered 198–201, representing the amino acid sequence from residue 198 to residue 209 (Figure 5A). The amino acid residues present in all the four overlapping nonapeptides are 201-LEEGEE-206, constituting an acidic amino acid cluster, which is part of a region complementary to the NLS consensus sequence (Roberts, 1989).

ELISA's performed on nonapeptides of the β-subunit sequence did not reveal a single dominant epitope but rather a number of weaker, separate signals distributed over the entire sequence. The results for the β-subunit (Figure 5B) were obtained with the same affinity-purified serum being used for the α-subunit. With three other affinity-purified sera tested, a somewhat variable pattern was observed. Reacting peptides were defined as those represented by peaks which were clearly above the background level and which reacted with three different sera. According to this criterion, peptides with positive reactions were identified in three regions; these are residues 96–105, 186–196, and 200–208. All four epitopes correspond to antigenic-index maxima calculated according to the algorithm of Jameson and Wolf (1988).

DISCUSSION

A sequence identity of 24% and, taking into account also conservative replacements, an overall similarity of 47% are clear indications that the genes encoding the two subunits of the *Thermoplasma* proteasome arose from a common ancestor. The most conspicuous distinguishing feature of the derived amino acid sequences is the complete absence in the β-subunit of the 26 amino acid residues of the N-terminus of the α-subunit (Figure 3). When aligning the N-terminal sequences of proteasomal subunits from various eukaryotic species, it is striking to see that they fall into two classes, related to either the α-subunit or β-subunit of the *Thermoplasma* proteasome and distinguished by the same feature (Figure 4).

In evolutionary terms this means that an ancestral gene in an ancestral species has duplicated, creating the α- and β-subunits, which have been able to acquire different functions (see below) and hence are regarded as paralogous.[2] In the evolution of new species, further divergence of the two paralogous genes generated sets of subunits capable of performing similar functions but allowing for fine tuning, e.g., of substrate

[2] Paralogous genes result from the duplication and divergence of one gene within one genome.

Table I: Comparison of the Amino Acid Sequences of the Proteasomal Proteins from *T. acidophilum* and Various Eukaryotes[a]

	Taα	HsECζ	Dm28.1	RnLC3	XlC3	ScY7	RnLC9	Dm29	ScY13	RnLC8	ScC1	HsECι	ScC7α	HsECν	RnLC2	Dm35	HsLC5	RnLC5	Xlβ	ScPRE1	MmLMP2	HsRING12	HsECδ	ScPUP1	HsRING10	Taβ
Taα	-																									
HsECζ	36	-																								
Dm28.1	34	30	-																							
RnLC3	37	29	32	-																						
XlC3	38	29	31	95	-																					
ScY7	35	28	27	55	56	-																				
RnLC9	31	29	29	32	32	31	-																			
Dm29	29	28	26	30	31	31	68	-																		
ScY13	33	31	29	32	32	31	49	48	-																	
RnLC8	30	30	29	27	28	26	27	25	29	-																
ScC1	31	30	28	33	34	30	26	24	26	47	-															
HsECι	29	20	25	31	30	30	28	28	30	27	25	-														
ScC7α	28	22	24	26	27	24	23	23	23	27	27	51	-													
HsECν	29	25	30	29	29	25	23	24	24	22	26	20	22	-												
RnLC2	29	25	30	28	29	24	23	24	24	21	25	21	21	98	-											
Dm35	30	27	26	28	30	26	26	26	26	22	26	25	25	52	51	-										
HsLC5	12	15	14	13	13	15	18	17	15	12	13	17	17	12	12	13	-									
RnLC5	13	13	13	13	13	15	17	17	15	12	10	18	16	12	13	14	92	-								
Xlβ	14	14	13	12	12	14	14	13	13	13	15	12	12	13	13	13	17	17	-							
ScPRE1	15	18	16	15	15	14	16	15	16	13	15	14	14	16	16	15	20	19	16	-						
MmLMP2	16	15	14	10	13	14	13	12	15	13	12	13	11	12	11	13	16	16	17	14	-					
HsRING12	16	15	14	13	12	14	13	11	16	13	13	13	12	11	11	14	17	16	18	15	88	-				
HsECδ	14	14	14	14	14	14	14	14	16	12	11	14	13	13	13	12	19	19	17	15	60	57	-			
ScPUP1	18	17	14	14	16	14	16	16	18	13	18	12	13	13	13	15	13	13	17	20	26	25	25	-		
HsRING10	17	16	12	16	16	14	14	15	14	12	14	13	11	16	16	19	19	19	21	19	27	29	25	25	-	
Taβ	19	18	17	14	16	14	17	17	17	18	17	14	16	18	19	21	20	19	16	22	21	21	20	26	27	-

[a] Scores give percentage of identical amino acid residues, obtained by pairwise comparison. Abbreviations are as defined in the legend to Figure 4.

```
Taα       MQQGQM----AYDRAITVFSPDGRLFQVEYAREAVKKGS-TALGMKFANGVLLISDKKVRSRLI-

HsECζ     MFLTRS----EYDRGVNTFSPEGRLFQVEYAIEAIKLGS-TAIGIQTSEGVCLAVEKRITSPLME
Dm28.1    MSS-------RYDRAVTIFSPDGHLLQVEYAQEAVRKGS-TAVGVRGANCVVLGVEKKSVAQLQ-
RnLC3     M--AER----GYSFSLTTFSPSGKLVQIEYALAAVAGGA-PSVGIKAANGVVLATEKKQKSILY-
XlC3      M--AER----GYSFSLTTFSPSGKLVQIEYALAAVAAGA-PSVGIKATNGVVLATEKK-------
ScY7      M--TDR-----YSFSLTTFSPSGKLGQIDYALTAVKQGV-TSLGIKATNGVVIATEKKSSSPLA-
RnLC9     M-S--R----RYDSRTTIFSPEGRLYQVEYAMEAIGHAG-TCLGILANDGVLLAAERRNIHKLL-
Dm29      M-A--R----RYDSRTTIFSPEGRLYQVEYAMEAISHAG-TCLGILAEDGILLAAECRSTNKLL-
ScY13     MGS--R----RYDSRTTIFSPEGRLYQVEYALESISHAG-TAIGIMASDGIVLAAERKVTSTLL-
RnLC8     MSSIGT----GYDLSASTFSPDGRVFQVEYAMKAVENSS-TAIGIRCKDGVVFGVEKLVLSKLY-
ScC1      MTSIGT----GYDLSNSVFSPDGRNFQVEYAVKAVENGT-TSIGIKCNDGVVFAVEKLITSKLL-
HsECι     ----------------HITIFSPEGRLYQVEYAFKAINQGGLTSVAVRGKDCAVIVTQKKVPDCLLD
ScC7α     MSGAAAASAAGYDRHITIFSPEGRLYQVEYAFKATNQTNINSLAVRGKDCTVVISQKKVPDKLL-
HsECν     MFRNQ------YDNDVTVWSPQGRIHQIEYAMEAVKQGS-ATVGLKSKTHAVLVALKRAQSELAA
RnLC2     MFRNQ------YDNDVTVWSPQGRIHQIEYAMEAVKQGS-ATVGLKSKTHAVLVALKRAQSELA-
Dm35      MFRNQ------YDSDVTVWSPQGRLHQVEYAMEAVKLGT-ATVGLKNKDYAVLVALCKPTSELS-

HsLC5     MLSST-AMYSAPGRDLGM-EPHRAAGPLQLRFSPYVFNGGTILAIAGEDFAIVASDTRLS-EGFS
RnLC5     MLS-T-AAYRDPDRELVM-GPQGSAGPVQMRFSPYAFNGGTVLAIAGEDFSIVASDTRLS-EGFS

Xlβ       ----------------------------------mvtgTSVLGVKFDGGVIIAADMLGSYGSLA
ScPRE1    -------------------------------------mDIILGIRVQDSVILASSKAVT-RGIS
MmLMP2    ------------------mlragaptagsfrteevhtgTTIMAVEFDGGVVVGSDSRVSAGTAV
HsRING12  ------------------mlragaptgdlpragevhtgTTIMAVEFDGGVVMGSDSRVSAGEAV
HsECδ     ------------------------------------XXIMAVQFDGGVVLGADSRTTTGSYI
ScPUP1    ----------maglsfdnyqrnnflaenshtqpkatstgTTIVGVKFNNGVVIAADTRSTQGPIV
HsRING10  ----------------------------------mahgTTTLAFKFQHGVIAAVDSRASAGSYI

Taβ       ----------------------------------------TTTVGITLKDAVIMATERRVTMENFI
```

FIGURE 4: Alignment of N-terminal sequences of proteasomal subunits. Amino acid residues are given in the one-letter code, X designating residues with uncertain assignment. Amino acid residues in lower-case letters are probably removed posttranslationally; this assumption is based on the comparison with the N-terminal protein sequence of the *T. acidophilum* β-subunit and of eukaryotic proteasomal subunits, where N-terminal protein sequences are available (Lee et al., 1990; Lilley et al., 1990). Dashes indicate gaps, introduced for better alignment. The N-terminal sequences of the α- and β-subunits of *T. acidophilum* (top and bottom) are shaded. The box marks a region of highly conserved residues, characteristic for α-type subunits. Abbreviations: Ta, *T. acidophilum*; Sc, *Saccharomyces cerevisiae* (yeast); Dm, *Drosophila melanogaster*; Xl, *Xenopus laevis*; Mm, *Mus musculus* (mouse); RnL, *Rattus norvegicus* (rat) liver; HsL, *Homo sapiens* (human) liver; HsE, *H. Sapiens* erythrocytes. Sequences are taken from the following sources: Taα (Zwickl et al., 1991); ScC1 and ScC7α (Fujiwara et al., 1990); ScY7 and ScY13 (Emori et al., 1991); ScPRE1 (Heinemeyer et al., 1991); ScPUP1 (Haffter & Fox, 1991); Dm28.1 (Haass et al., 1990a); Dm29 (Haass et al., 1990b); Dm35 (Haass et al., 1989); XlC3 (Fujii et al., 1991); Xlβ (van Riel & Martens, 1991); MmLMP2 (Martinez & Monaco, 1991); RnCL2 (Fujiwara et al., 1989); RnCL3 (Tanaka et al., 1990a); RnCL5 (Tamura et al., 1990); RnCL8 (Tanaka et al., 1990b); RnCL9 (Kumatori et al., 1990); HsCL5 (Tamura et al., 1991); HsECδ, -ι, -ν, -ζ (DeMartino et al., 1991); HsRING10 (Glynne et al., 1991); HsRING12 (Kelly et al., 1991).

specificities. Therefore, the *Thermoplasma* proteasome may be seen as an ancestral or Urproteasome. This does not necessarily imply that *Thermoplasma* itself is an ancestral species, it gives, however, as it appears, testimony of an ancestral situation. Such a scenario is not unusual; it has, in fact, all the traits discussed in the context of the evolutionary divergence of proteolytic enzymes and their inhibitors (Creighton & Darby, 1989). The hypervariability of their active sites appears to promote functional fine tuning, particularly with regard to substrate specificity, and it is reasonable to assume that this applies to proteasomes, too. It is an unusual feature of proteasomes, however, that despite this divergence the subunits remain physically associated, forming a structural and functional entity. The assembly in one complex offers advantages in performing catalytic reactions sequentially and it allows the subunits to share regulatory or targeting mechanisms.

The conservation of the proteasome quaternary structure from *Thermoplasma* (Dahlmann et al., 1989) to higher eukaryotes, such as rat (Baumeister et al., 1988), is quite remarkable. It clearly imposes strong structural constraints to subunit divergence. For the *Thermoplasma* proteasome we have shown by means of immunoelectron microscopy that the α-subunits form the 7-fold symmetric disks located at the opposite ends of the barrel-shaped complex, while the β-subunits form the two central rings sandwiched between them (Grziwa et al., 1991). It seems reasonable to assume that in eukaryotic proteasomes the various subunits related to the α- or β-subunit, respectively, occupy equivalent positions. Moreover, it is tempting to speculate that the highly conserved N-terminal extension characteristic for the α-type subunits plays a crucial role in the assembly of the complex and in maintaining the pattern of interactions in the quaternary structure in spite of the subunit divergence. It could be involved either in making the intersubunit contacts within the α-subunit disks or in the disks made of α-type subunits in eukaryotic proteasomes or in maintaining the interaction between the α-disks and the β-type subunits. We regard such a structural role of the N-terminus of the α-subunit as more likely than its direct involvement in the biological function of the protein, the more since the epitope scanning experiments provided no evidence for the N-terminus, or nearby regions of the polypeptide, to be surface exposed.

An issue of utmost importance is, of course, the identification of the functional roles of the α- and β-subunits in the *Thermoplasma* proteasome. Although an unambiguous assignment of functions is not yet possible, we put forward a working hypothesis and we present circumstantial evidence in favor of it. We propose that it is the β-subunit which carries the active site, while the α-subunit serves a regulatory and a targeting function. Although we cannot rule out the possibility that the active site is shared between the α- and β-subunits, there are two arguments in favor of attributing the catalytic activity to the β-subunit: (I) The yeast PRE1 protein, which must be classified as a β-type subunit, has been shown by mutant studies to contribute a chymotrypsin-like activity (Heinemeyer et al., 1991), and this type of activity is also characteristic for the *Thermoplasma* proteasome (Dahlmann

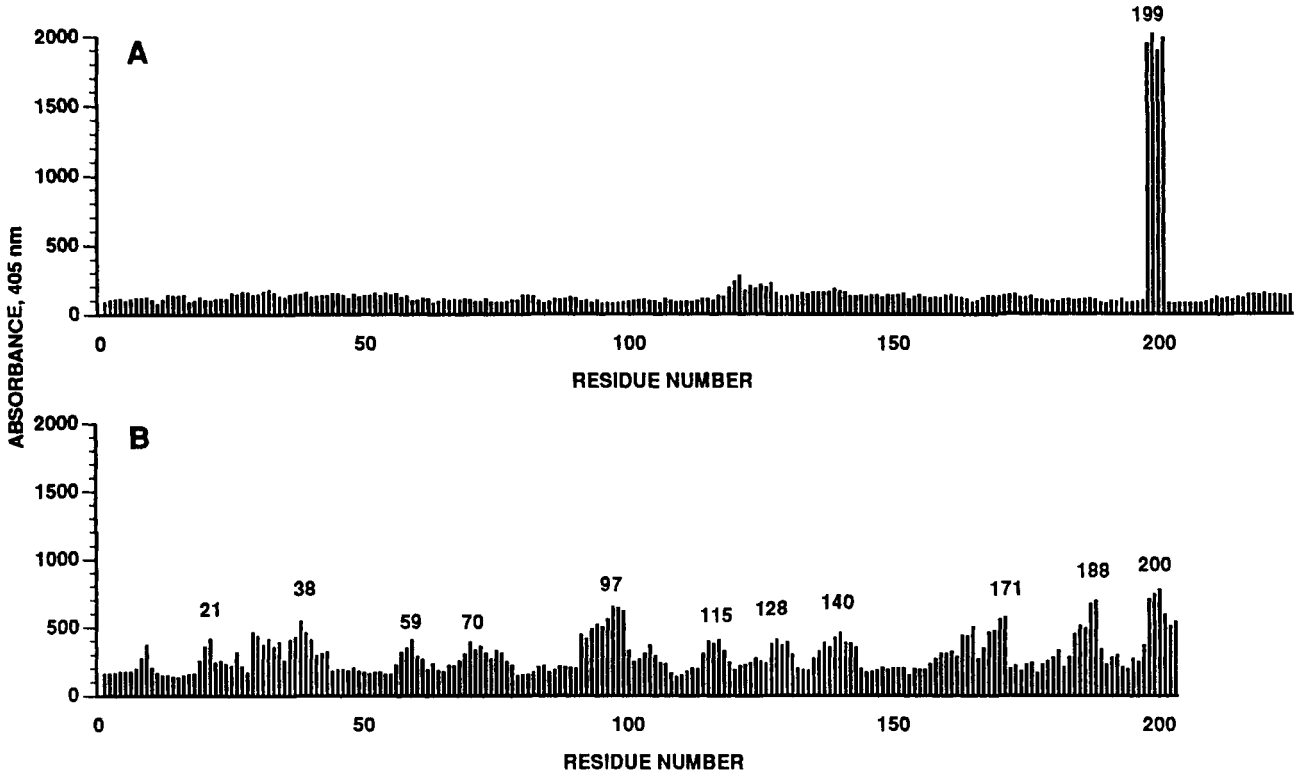

FIGURE 5: Scans of all possible overlapping nonapeptides based on the sequence of the α-subunit (A) and the β-subunit (B) protein with an affinity-purified antiproteasome serum. Peptides were reacted with a 1:200 dilution of the affinity-purified serum, and antibody binding was ascertained by ELISA (Geysen et al., 1987). Each vertical bar represents the absorbance of a single nonapeptide obtained in the ELISA test. Residue numbers start at the amino terminus. The antibody-binding peptides that gave a significantly higher absorbance at 405 nm compared with the background were identified as surface-exposed epitopes.

et al., 1989). (II) Although the β-subunit—and likewise all hitherto sequenced subunits of eukaryotic proteasomes—shows no sequence motif characteristic for serine proteases, it contains all the essential amino acid residues forming the catalytic triad, or the "charge relay system" (Ser, Asp, His), whereas the α-subunit contains no histidine at all. Moreover, labeling and band-shift experiments with the serine proteinase specific inhibitor diisopropyl fluorophosphate indicate that, if any, it is the β-subunit which interacts with this inhibitor. In relation to mammalian serine proteinases, such as the trypsin or chymotrypsin families, the evolution of the proteasomal subunits should probably be regarded as convergent, explaining why there is no sequence homology. This does not rule out, of course, that proteasomes have acquired a similar catalytic mechanism with a similar charge relay system; the proper spatial disposition of the residues involved might or might not be based on a similar folding pattern of the polypeptide chain.

Without any intention to dispute the proteolytic activities of proteasomes, it might be of some heuristic value to take a critical position. The assignment of activities such as "trypsin-like" or chymotrypsin-like is almost entirely based on experiments with small synthetic peptides, and the basic activities toward these substrates tend to be rather low; activating agents either are often nonphysiological (SDS or high temperature) per se or confer activation at nonphysiological concentrations (>500 mM $CaCl_2$). Similarly, the use of inhibitors as classification criteria must be met with skepticism if they become potent only at concentrations bearing the risk of perturbing the structure of the proteasome. It should also be recalled that active sites reminiscent of those found in serine proteinases are used for the chemically similar hydrolysis of ester bonds, e.g., in lipases (Brenner, 1988; Brzozowski et al., 1991), and it would not be a surprise to find similar motifs in other enzymes interacting with peptide bonds such as isomerases or transpeptidases. All this should be kept in mind in searching for the in vivo substrate(s) of proteasomes.

With respect to the α-subunit of the *Thermoplasma* proteasome, there are several lines of evidence indicating that it might indeed have a regulatory and a targeting function: (I) In 2-DE the α-subunit gives rise to multiple spots with different p*I*'s, indicating that it is subject to modifications. (II) There is evidence for the α-subunit to be glycosylated. (III) The amino acid sequence contains potential phosphorylation sites, including a tyrosine autophosphorylation site at Tyr 124. (IV) It contains the consensus sequence of nuclear localization signals (amino acid residues 50–55) as well as a sequence complementary to it (residues 202–211). (V) Finally, the α-subunits are located at the top and the bottom of the *Thermoplasma* proteasome, a position quite suitable for targeting function(s). The notion of a nuclear localization signal (NLS) should not too easily be dismissed on the grounds of this relatively indistinct signature or because of its enigmatic role in a prokaryotic cell, not possessing a nucleus. The occurrence of an NLS sequence was recently also reported for an IgA protease associated protein from a pathogenic *Neisseria*, and it was shown to be capable of functioning as a nuclear transport signal in an in vitro assay (Pohlner et al., 1990). It is also noteworthy that the RnC3 subunit of the rat proteasome, which is related to the α-subunit (see Figure 4), has an NLS sequence in a corresponding position of its sequence (Tanaka et al., 1990c), and also sequences complementary to the NLS sequence have been found in several eukaryotic proteasomal subunits, of the α-type. It has been shown repeatedly that proteasomes occur in nuclei (Arrigo et al., 1988; Tanaka et al., 1989), and there has been some discussion of the possible mechanism of the nuclear translocation of proteasomes and the manner in which the NLS sequence and the sequence complementary to it might interact

(Tanaka et al., 1990c). It is remarkable in this context that the sequence complementary to the NLS sequence is by far the predominant epitope on the *Thermoplasma* α-subunit (Figure 5); this proves that it is surface exposed, and it indicates that it has some flexibility as it is required for any interaction with the NLS sequence, possibly in a phosphorylation-regulated manner. The occurrence of an NLS sequence in the absence of a nucleus is less puzzling when one considers that the primary target or the "receptor" must not reside in the nuclear envelope or, more specifically, the nuclear pore complex.

ACKNOWLEDGMENTS

We thank W. Zillig (Martinsried) for providing *T. acidophilum* cells and H.-P. Klenk (Martinsried) for providing genomic DNA from *T. acidophilum*.

REFERENCES

Argos, P. (1987) *J. Mol. Biol. 193*, 385–396.
Arrigo, A. P., Tanaka, K., Goldberg, A. L., & Welch, W. J. (1988) *Nature 331*, 192–194.
Bairoch, A. (1990) *PROSITE: a Dictionary of Protein Sites and Patterns*, Fifth Release, University of Geneva.
Baumeister, W., Dahlmann, B., Hegerl, R., Kopp, F., Kuehn, L., & Pfeifer, G. (1988) *FEBS Lett. 241*, 239–245.
Brenner, S. (1988) *Nature 334*, 528–530.
Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Hugu-Jensen, B., Patkar, S. A., & Thim, L. (1991) *Nature 351*, 491–494.
Creighton, T. E., & Darby, N. J. (1989) *Trends Biochem. Sci. 14*, 319–324.
Dahlmann, B., Kopp, F., Kuehn, L., Niedel, B., Pfeifer, G., Hegerl, R., & Baumeister, W. (1989) *FEBS Lett. 251*, 125–131.
DeMartino, G. N., Orth, K., McCullough, M. L., Lee, L. W., Munn, T. Z., Moomaw, C. R., Dawson, P. A., & Slaughter, C. A. (1991) *Biochim. Biophys. Acta 1079*, 29–38.
Driscoll, J., & Goldberg, A. L. (1990) *J. Biol. Chem. 265*, 4789–4792.
Eckerskorn, C., Jungblut, P., Mewes, W., Klose, J., & Lottspeich, F. (1988a) *Electrophoresis 9*, 830–838.
Eckerskorn, C., Mewes, W., Goretzki, H., & Lottspeich, F. (1988b) *Eur. J. Biochem. 176*, 509–519.
Emori, Y., Tsukahara, T., Kawasaki, H., Ishiura, S., Sugita, H., & Suzuki, K. (1991) *Mol. Cell. Biol. 11*, 344–353.
Eytan, E., Ganoth, D., Armon, T., & Hershko, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7751–7755.
Falkenburg, P.-E., Haass, C., Kloetzel, P.-M., Niedel, B., Kopp, F., Kuehn, L., & Dahlmann, B. (1988) *Nature 331*, 190–192.
Fujii, G., Tashiro, K., Emori, Y., Saigo, K., Tanaka, K., & Shiokawa, K. (1991) *Biochem. Biophys. Res. Commun. 178*, 1233–1239.
Fujiwara, T., Tanaka, K., Kumatori, A., Shin, S., Yoshimura, T., Ichihara, A., Tokunaga, F., Aruga, R., Iwanaga, S., Kakizuka, A., & Nakanishi, S. (1989) *Biochemistry 28*, 7332–7340.
Fujiwara, T., Tanaka, K., Orino, E., Yoshimura, T., Kumatori, A., Tamura, T., Chung, C. H., Nakai, T., Yamaguchi, K., Shin, S., Kakizuka, A., Nakanishi, S., & Ichihara, A. (1990) *J. Biol. Chem. 265*, 16604–16613.
Geysen, H. M., Meloen, R. H., & Barteling, S. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3998–4002.
Geysen, H. M., Rodda, S. J., Mason, T. J., Tribbick, G., & Schoofs, P. G. (1987) *J. Immunol. Methods 102*, 259–274.
Glynne, R., Powis, S. H., Beck, S., Kelly, A., Kerr, L.-A., & Trowsdale, J. (1991) *Nature 353*, 357–360.
Grziwa, A., Baumeister, W., Dahlmann, B., & Kopp, F. (1991) *FEBS Lett. 290*, 186–190.
Haass, C., Pesold-Hurt, B., Multhaup, G., Beyreuther, K., & Kloetzel, P.-M. (1989) *EMBO J. 8*, 2373–2379.
Haass, C., Pesold-Hurt, B., Multhaup, G., Beyreuther, K., & Kloetzel, P.-M. (1990a) *Gene 90*, 235–241.
Haass, C., Pesold-Hurt, B., & Kloetzel, P. M. (1990b) *Nucleic Acids Res. 18*, 4018.
Haffter, P., & Fox, T. D. (1991) *Nucleic Acids Res. 19*, 5075.
Harris, J. R. (1968) *Biochim. Biophys. Acta 150*, 534–537.
Hegerl, R., Pfeifer, G., Pühler, G., Dahlmann, B., & Baumeister, W. (1991) *FEBS Lett. 283*, 117–121.
Heinemeyer, W., Kleinschmidt, J. A., Saidowsky, J., Escher, C., & Wolf, D. H. (1991) *EMBO J. 10*, 555–562.
Higgins, D. G., & Sharp, P. M. (1988) *Gene 73*, 237–244.
Hough, R., Pratt, G., & Rechsteiner, M. (1988) in *Ubiquitin* (Rechsteiner, M., Ed.) pp 101–134, Plenum, New York.
Jameson, B. A., & Wolf, H. (1988) *Cabios 4*, 181–186.
Kelly, A., Powis, S. H., Glynne, R., Radley, E., Beck, S., & Trowsdale, J. (1991) *Nature 353*, 667–668.
Kumatori, A., Tanaka, K., Tamura, T., Fujiwara, T., Ichihara, A., Tokunaga, F., Onikura, A., & Iwanga, S. (1990) *FEBS Lett. 264*, 279–282.
Lee, L. W., Moomaw, C. R., Orth, K., McGuire, M. J., DeMartino, G. N., & Slaughter, C. A. (1990) *Biochim. Biophys. Acta 1037*, 178–185.
Lilley, K. S., Davison, M. D., & Rivett, A. J. (1990) *FEBS Lett. 262*, 327–329.
Martinez, C. K., & Monaco, J. J. (1991) *Nature 353*, 664–667.
Matthews, W., Driscoll, J., Tanaka, K., Ichihara, A., & Goldberg A. L. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2597–2601.
Monaco, J. J., & McDevitt, H. O. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3001–3005.
Monaco, J. J., & McDevitt, H. O. (1984) *Nature 309*, 797–799.
Orcutt, B. C., Dayhoff, M. O., George, D. G., & Barker, W. C. (1986) *User's guide to software systems of protein identification resource, alignment score program*, National Biomedical Research Foundation, Washington, DC.
Orlowski, M. (1990) *Biochemistry 29*, 10289–10297.
Parham, P. (1990) *Nature 348*, 674–675.
Pohlner, J., Langenberg, U., & Meyer, T. F. (1990) *Zentralbl. Bakteriol., Suppl. 19*, 461–469.
Ree, H. K., Cao, K., Thurlow, D. L., & Zimmermann, R. A. (1989) *Can. J. Microbiol. 35*, 124–133.
Rivett, J. A. (1989) *Arch. Biochem. Biophys. 268*, 1–8.
Roberts, B. (1989) *Biochim. Biophys. Acta 1008*, 263–280.
Schägger, H., & von Jagow, G. (1987) *Anal. Biochem. 166*, 368–379.
Scherrer, K. (1990) *Mol. Biol. Rep. 14*, 1–9.
Schmid, H. P., Akhayat, O., Martins de Sa, C., Puvion, F., Köhler, K., & Scherrer, K. (1984) *EMBO J. 3*, 29–34.
Tamura, T., Tanaka, K., Kumatori, A., Yamada, F., Tsurumi, C., Fujiwara, T., Ichihara, A., Tokunaga, F., Aruga, R., & Iwanaga, S. (1990) *FEBS Lett. 264*, 91–94.
Tamura, T., Lee, D. H., Osaka, F., Fujiwara, T., Shin, S., Chung, C. H., Tanaka, K., & Ichihara, A. (1991) *Biochim. Biophys. Acta 1089*, 95–102.
Tanaka, K., Kumatori, A., Ii, K., & Ichihara, A. (1989) *J. Cell. Physiol. 139*, 34–41.

Tanaka, K., Fujiwara, T., Kumatori, A., Shin, S., Yoshimura, T., Ichihara, A., Tokunaga, F., Aruga, R., Iwanaga, S., Kakizuka, A., & Nakanishi, S. (1990a) *Biochemistry 29*, 3777-3785.

Tanaka, K., Kanayama, H., Tamura, T., Lee, D. H., Kumatori, A., Fujiwara, T., Ichihara, A., Tokunaga, F., Aruga, R., & Iwanaga, S. (1990b) *Biochem. Biophys. Res. Commun. 171*, 676-683.

Tanaka, K., Yoshimura, T., Tamura, T., Fujiwara, T., Kumatori, A., & Ichihara, A. (1990c) *FEBS Lett. 271*, 41-46.

van Riel, M. C. H. M., & Martens, G. J. M. (1991) *FEBS Lett. 291*, 37-40.

Zwickl, P., Pfeifer, G., Lottspeich, F., Kopp, F., Dahlmann, B., & Baumeister, W. (1990) *J. Struct. Biol. 103*, 197-203.

Zwickl, P., Lottspeich, F., Dahlmann, B., & Baumeister, W. (1991) *FEBS Lett. 278*, 217-221.

Articles

Recognition of Tertiary Structure in tRNAs by $Rh(phen)_2phi^{3+}$, a New Reagent for RNA Structure-Function Mapping†

Christine S. Chow,‡ Linda S. Behlen,§ Olke C. Uhlenbeck,§ and Jacqueline K. Barton*,‡

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309

Received September 5, 1991; Revised Manuscript Received October 18, 1991

ABSTRACT: With photoactivation $Rh(phen)_2phi^{3+}$ promotes strand cleavage at sites of tertiary interaction in tRNA. The rhodium complex, which binds double-helical DNA by intercalation in the major groove, yields no cleavage in double-helical regions of the RNA or in unstructured single-stranded regions. Instead, $Rh(phen)_2phi^{3+}$ appears to target regions which are structured so that the major groove is open and accessible for stacking with the complex, as occurs where bases are triply bonded. So as to examine the specificity of this novel reagent and to evaluate its use in probing structural changes in RNAs, cleavage studies have been conducted on two structurally characterized tRNAs, $tRNA^{Phe}$ and $tRNA^{Asp}$ from yeast, the unmodified yeast $tRNA^{Phe}$ transcript, and a chemically modified $tRNA^{Phe}$, as well as on a series of $tRNA^{Phe}$ mutants. On $tRNA^{Phe}$ strong cleavage is observed at residues G22, G45, U47, Ψ55, and U59; weaker cleavage is observed at A44, m^7G46, and C48. On $tRNA^{Asp}$ cleavage is found at residues A21 through G26, Ψ32, and U48, with minor cleavage apparent at A44, G45, A46, Ψ55, U59, and U60. There is a striking similarity in cleavage observed on these tRNAs, and the sites of cleavage mark regions of tertiary folding. Cleavage on the unmodified $tRNA^{Phe}$ transcript resembles closely that found on native yeast $tRNA^{Phe}$, but additional sites, primarily in the anticodon loop and stem, are evident. The results indicate that globally the structures containing or lacking the modified bases appear to be the same; the differences in cleavage observed may reflect a loosening or alteration in the structure due to the absence of the modified bases. Cleavage results on mutants of $tRNA^{Phe}$ illustrate $Rh(phen)_2phi^{3+}$ as a sensitive probe in characterizing tRNA tertiary structure. Results are consistent with other assays for structural or functional changes. Uniquely, $Rh(phen)_2phi^{3+}$ appears to target directly sites of tertiary interaction. Cleavage results on mutants which involve base changes within the triply bonded region of the molecule indicate that it is the structure of the triply bonded array rather than the individual nucleotides which are being targeted. Chemical modification to promote selective depurination of the third base (m^7G46) involved in the triple in the folded, native tRNA leads to the reduction of cleavage by the metal complex; this result shows directly the importance of the stacked triple base structure for recognition by the metal complex. The cleavage results are consistent with the notion that $Rh(phen)_2phi^{3+}$ preferentially targets regions of tertiary structure in the tRNA because these regions are structured so that the major grooves are open and accessible to stacking by the complex. Since sites cleaved by the rhodium complex mark a range of tertiary structures, $Rh(phen)_2phi^{3+}$ appears to be a powerful and unique probe in characterizing the folded structures of RNAs.

Increasing evidence suggests that many of the diverse biological functions of RNA require the molecule to maintain a precise three-dimensional structure (Yanofsky, 1981; Altman, 1984; Cech, 1987; Guerrier-Takada et al., 1989; Rould et al., 1989). A popular approach toward understanding the structure-function relationships of RNA involves extensive site-directed or random mutagenesis of the molecule and assay of the mutants in vitro. In these studies it is important to distinguish whether a decrease in activity of a mutant is the result of a change in an essential nucleotide or the less interesting consequence of a more general alteration of the overall structure. Thus it is important to develop methods to probe rapidly the subtle changes in the configuration of mutant RNA.

† We are grateful to the National Institutes of Health for their financial support. In addition, C.S.C. thanks the Parsons Foundation for their fellowship support.

* To whom correspondence should be addressed.

‡ California Institute of Technology.

§ University of Colorado.

0006-2960/92/0431-972$03.00/0