FEBS Letters 349 (1994) 205–209

FEBS 14279

Existence of a molecular ruler in proteasomes suggested by analysis of degradation products

Thorsten Wenzel, Christoph Eckerskorn, Friedrich Lottspeich, Wolfgang Baumeister*

Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, 82152 Martinsried, Germany

Received 20 June 1994

Abstract

Analysis of the degradation products from two proteins, the insulin B-chain and human hemoglobin, generated by archaebacterial *Thermoplasma acidophilum* 20 S proteasomes, revealed an unexpectedly broad specificity. In spite of the vast number of different peptides found, they fell into a rather narrow size range. This suggests that a molecular ruler exists which determines the length of the cleavage products.

Key words: Proteasome; Molecular ruler; *Thermoplasma acidophilum;* Proteolytic degradation; Insulin B-chain; Hemoglobin

1. Introduction

Proteasomes, also known as the multicatalytic proteinase complex (EC 3.4.99.46), play a key role in the degradation of abnormal proteins and of short-lived regulatory proteins, both in the cytosol and in the nucleus [1–3]. Moreover, they are implicated in antigen presentation, generating peptides which associate with major histocompatibility complex (MHC) class I molecules [4–6]. The ATP-dependent degradation of ubiquitinated [7–9] and of certain non-ubiquitinated proteins [10] is linked to the 26 S proteasome (M_r ~2000 kDa) while the 20 S proteasome (M_r ~700 kDa), which constitutes the proteolytic core of the 26 S complex [11,12], degrades only unfolded proteins [13].

The eukaryotic 20 S proteasome is a ubiquitous, highly conserved barrel-shaped complex [14,15] comprising of up to 14 different subunits. Proteasomes indistinguishable on the quarternary structure level from eukaryotic proteasomes are found in the archaeon *Thermoplasma* [16]. They are made of only two different subunits, α and β, which occur in an $\alpha_{14}\beta_{14}$ stoichiometry; all hitherto available amino acid sequences of eukaryotic proteasomes can be related to either the α or β subunit of the *Thermoplasma* urproteasome [17]. Because of its relative simplicity the *Thermoplasma* proteasome has a pivotal role in advancing our understanding of the structure of this large protein complex [18–20].

In this communication we describe experiments using recombinant *Thermoplasma* proteasomes [21] for the degradation of two protein substrates, namely the insulin B-chain and human hemoglobin. Whereas oxidized insulin B-chain can be used as a substrate for the 20 S proteasome without pretreatment [22], native hemoglobin is not degraded by the *Thermoplasma* proteasome. At least partial unfolding is required, accomplished e.g. through phenylhydrazine treatment or hydrogen peroxide-mediated oxidation. Alternatively, pre-incubation with ubiquitin can be used to render hemoglobin susceptible to degradation [23]: ubiquitin appears to exert a chaotropic effect upon associating with proteins; it is unclear whether or not this effect has any physiological relevance. In the context of our experiments, the use of ubiquitin for conditioning the hemoglobin has the practical advantage that it does not chemically alter the constituent amino acids.

2. Materials and methods

2.1. Isolation, purification and activity assay of proteasomes

Recombinant *Thermoplasma acidophilum* proteasomes, expressed in *Escherichia coli* [21], were isolated and purified as described previously [19]. Routinely, activity assays of isolated proteasomes were performed by using Suc-Leu-Leu-Val-Tyr-NMec (Bachem Biochemica GmbH, Heidelberg, Germany), as described earlier [24].

2.2. Degradation of substrate proteins

Bovine insulin B-chain and human hemoglobin-A_0 (both from Sigma Chemie GmbH, Deisenhofen, Germany) were degraded with recombinant *Thermoplasma acidophilum* proteasomes, expressed in *Escherichia coli* and purified as described earlier [21]. Reaction mixtures contained in a volume of 1 ml: 1 mg oxidized bovine insulin B-chain, 50 mM Tris-HCl (pH 7.5) and 5 μg of recombinant *Thermoplasma acidophilum* proteasomes. Following incubation at room temperature for 24 h, the samples were analysed by LC/MS. For the degradation of human hemoglobin-A_0, a 1 mg sample was incubated with 1 mg ubiquitin from bovine red blood cells (Sigma Chemie GmbH, Deisenhofen) to render it degradable [23], 50 mM Tris-HCl (pH 7.5) and 5 μg of *Thermoplasma acidophilum* proteasomes. After an incubation period of 24 h at 60°C, the total peptide mixture was analyzed by LC/MS. After an incubation time of 24 h further addition of fresh proteasomes did not cause any further degradation, whereas the degradation of added insulin B-chain indicated the still extant activity of the proteasomes.

* Corresponding author. Fax: (49) (89) 8578 2641.

Abbreviations: HPLC, high-performance liquid chromatography; LC/MS, liquid chromatography/mass spectrometry; ATP, adenosine 5′-triphosphate; MHC, major histocompatibility complex; NMec, 7-amino-4-methylcoumarin; Suc, succinyl.

SSDI 0014-5793(94)00665-I

2.3. Liquid chromatography and mass spectrometry

The degradation products were separated by reversed-phase chromatography (System 1090, Hewlett Packard, Waldbronn, Germany) using a 2 mm column (C18, LiChroCart 125–2, Merck, Darmstadt, Germany). The solvent system used was 0.1% (v/v) trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in acetonitrile (solvent B). A gradient from 0 to 60% B was performed in 70 min and from 60% to 90% B in 15 min at a flow rate of 0.3 ml/min. In both cases the degradation of the substrate was almost complete. Ubiquitin remained undegraded. Mass spectrometry (B) was performed on-line (LC/MS) by coupling of chromatography (SMART-System, Pharmacia, Freiburg, Germany) to an atmospheric pressure ionization source (ion spray) fitted to tandem quadrupole instrument API III (Sciex, Thornhill, Ont., Canada). The column and the solvent system for the reversed phase separation was the same as described above. The gradient from 0% to 60% B was performed in 150 min and from 60% to 90% B in 50 min at a flow rate of 20 μl/min. Each scan was acquired over the range *m/z* 350–2000 using a step of 0.2 units, a dwell time of 0.5 ms. The molecular masses of peptides were obtained as single ions or were calculated from the *m/z* peaks in the charge distribution profiles of the multiply charged ions [25,26]. Amino acids or dipeptides could not be detected by this method; it is also impossible to calculate the exact amount of the derived peptides.

2.4. Sequence analysis of degradation products

For this purpose fractions were collected and sequence analysis of the derived peptides were performed using a 477 A sequencer model from Applied Biosystems (Foster City, CA).

2.5. Analysis of free amino acids

Ninhydrine-amino acid analysis was performed on a Beckman (München, Germany) 6300 Amino acid analyzer according to the instructions of the manufacturer. 6 nmol of the starting material were applied after 24 h.

3. Results and discussion

The degradation products were analysed by reversed-phase chromatography in conjunction with mass spectrometry and N-terminal sequencing. With both proteins a surprisingly large number of different peptides was generated. Most of the peaks in the reversed-phase chromatograms (Fig. 1) contained more than one peptide species. For the insulin B-chain we unambiguously identified 44 different peptides either by mass spectrometry alone or combined with N-terminal sequencing. In the case of human hemoglobin, the number of peptide species was >80. Of the 29 peptide bonds in the insulin B-chain only six (Phe^{1}–Val^{2}, Asn^{3}–Gln^{4}, Cys^{7}–Gly^{8}, Tyr^{16}–Leu^{17}, Glu^{21}–Arg^{22}, Pro^{28}–Lys^{29}) were not cleaved. This implies that contrary to conclusions drawn from experiments using synthetic peptide substrates, suggesting that the *Thermoplasma* proteasome is exclusively a 'chymotrypsin-like' proteinase [16], it is essentially a non-specific endopeptidase. Therefore, our results cast some doubt on the usefulness of assays based on small peptide substrates for characterising the various proteolytic activities associated with proteasomes [27], or for providing hints as to the nature of the active site(s), which continues to be an enigma.

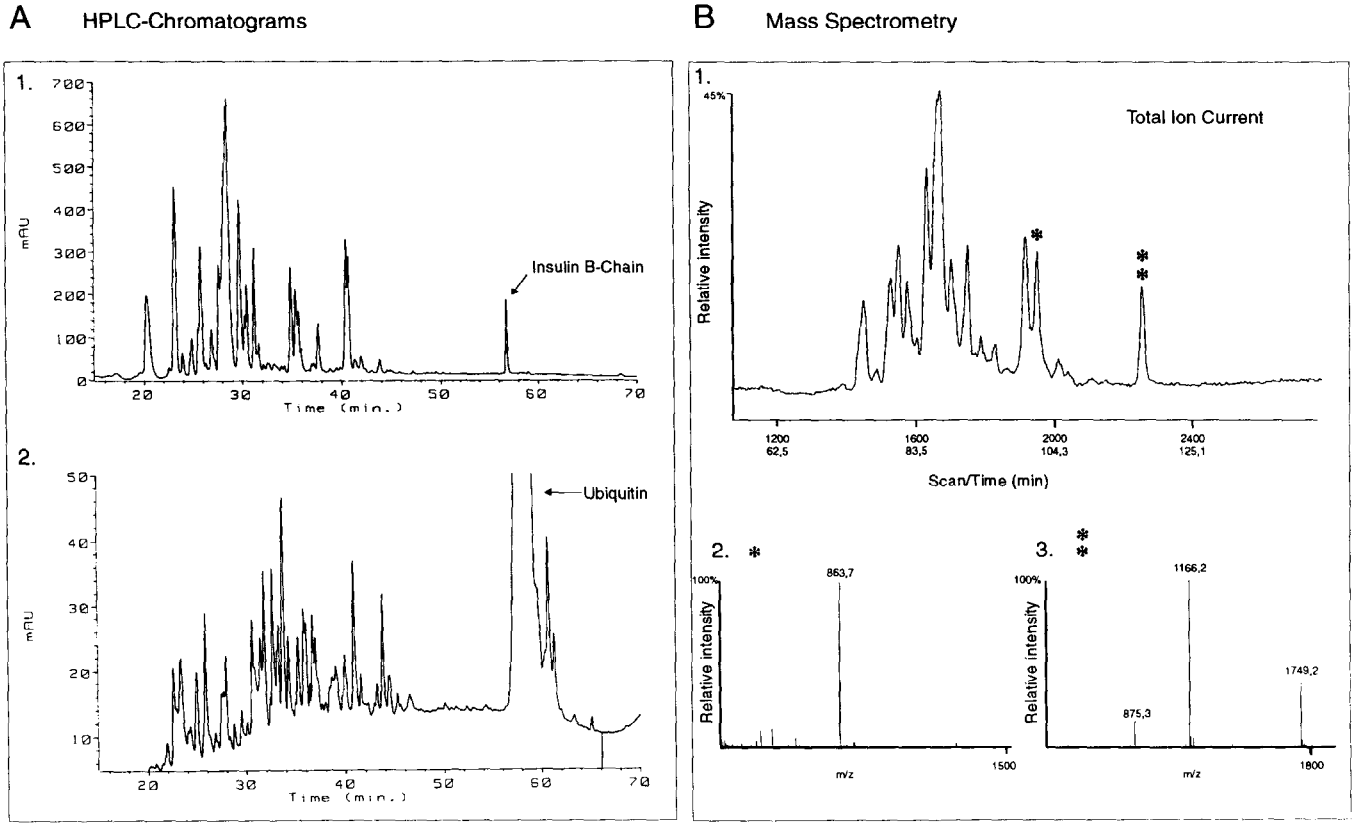


Fig. 1. Reverse-phase separation of the peptides derived from the oxidized bovine insulin B-chain (A1) and human hemoglobin-A_0 (A2) by digestion with *Thermoplasma* proteasomes, and corresponding mass spectrometric analysis (B). B1 shows a typical Total Ion Current of the derived peptides from the insulin B-chain and B2 and B3 two selected mass spectra of the peaks marked in B1 with asterisks.

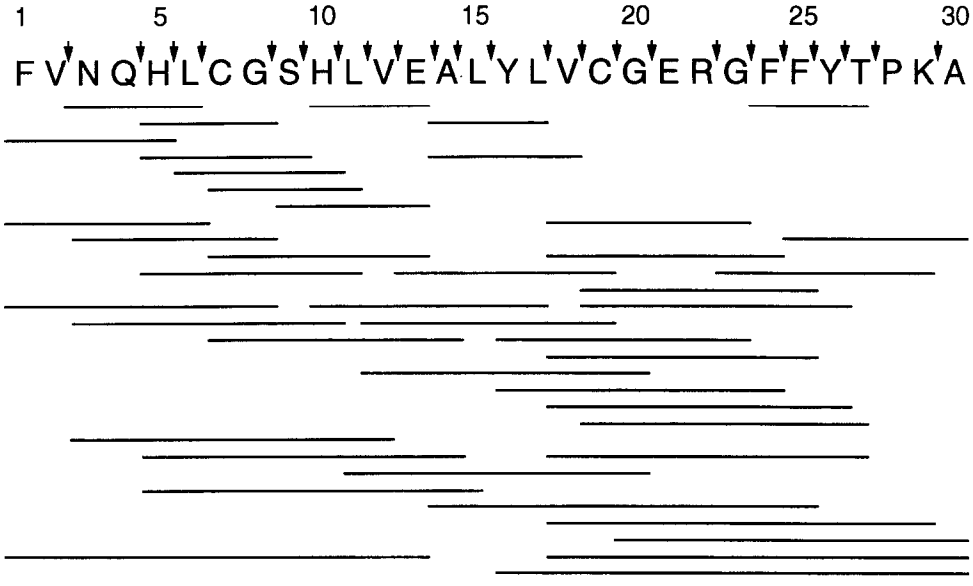
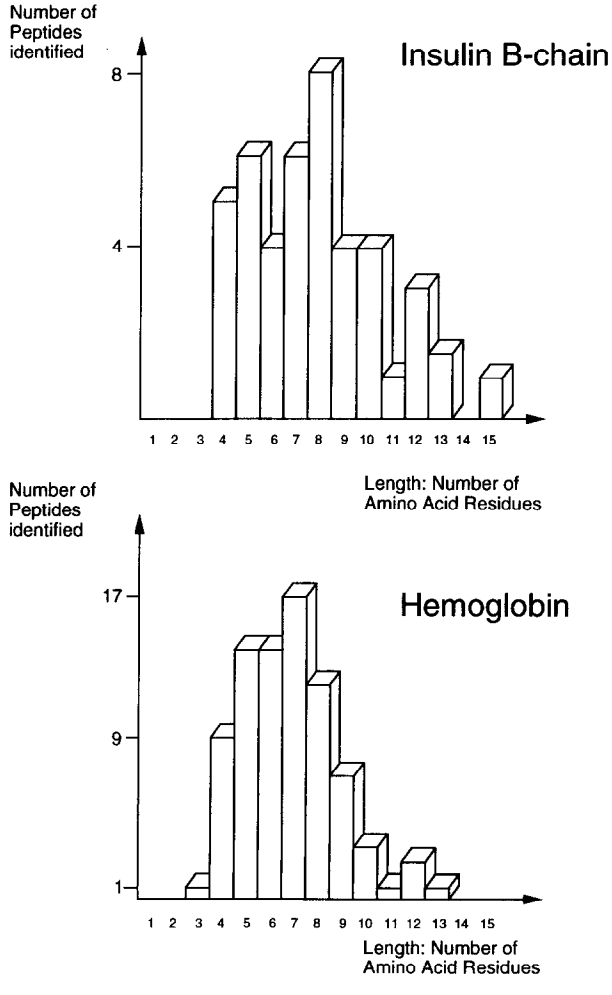


Fig. 2. Degradation of the oxidized bovine insulin B-chain by *Thermoplasma acidophilum* proteasomes. The sequence is represented in single letter code and all the cleavage sites detected are marked by arrows. Lines represent, regardless of their relative abundance, the peptides identified by mass spectrometry (Fig. 1), complemented, when necessary, by N-terminal sequencing.

In terms of the evolution of the proteasome, the present data make it unlikely that with the diversification of the α- and β-type subunits, several distinct proteolytic activities have been added to a chymotryptic activity of the 'urproteasome'. The data rather suggest that the molecular antecedant was a non-specific proteinase and that the diversification which followed multiple gene duplication events allowed for some functional fine tuning, either by changing the active site or the substrate binding region. Subunit substitutions, such as the incorporation of the MHC-encoded LMP 2 and LMP 7 subunits upon γ-interferon stimulation, allow further modulations of the activity and thus can favour the generation of more specific peptide species, as observed [28,29].

Notwithstanding the broad specificity of the *Thermoplasma* proteasome, the peptides generated fell into a surprisingly narrow size range. With the insulin B-chain, hepta-, octa- and nonapeptides constituted 46% of the total peptides and with hemoglobin, 45%. Only 14% of the insulin B-chain and only 5% of the hemoglobin derived peptides were longer than decapeptides. If peptides of a certain length are cut out of an unfolded polypeptide chain in a more or less random fashion, inevitably some 'odds and ends' will be produced. Only small amounts of amino acids have been detected. With 6 nmol starting material 360 pmol of free amino acids were found,



→

Fig. 3. Histogram of the length distribution of all identified peptides derived from the oxidized bovine insulin B-chain and from human hemoglobin-A_o. The graphs do not take into account the differences in the relative abundance of a given peptide.

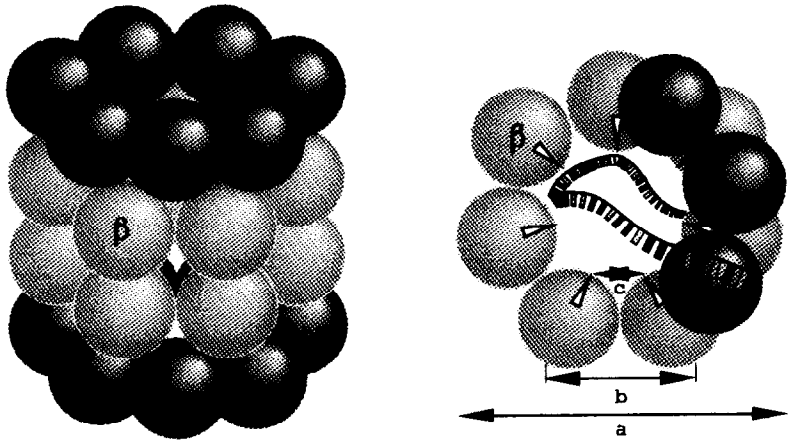


Fig. 4. Schematic illustration of the molecular ruler hypothesis. Left: model showing the subunit arrangement in the *Thermoplasma* proteasome. Four seven-membered rings collectively form a barrel-shaped complex. The α-subunits (dark) form the two outer rings or discs, whereas the β-subunits (light) make up the inner pair of rings. The β-rings are rotated by about 25° with respect to the α-rings [14,15]. Eukaryotic proteasomes have the same quarternary structure and the various α- and β-type subunits occupy positions equivalent to the α- and β-subunits of the *Thermoplasma* proteasome. Right: the *Thermoplasma* proteasome seen 'end-on'; for the sake of clarity some α-subunits have been removed. At present, it is neither known which amino acids are involved in the catalytic mechanism nor where the catalytic sites are located in the structure. There is circumstantial evidence that the β-subunits carry the active sites (e.g. LMP 2 and LMP 7 are β-type subunits); on the other hand only fully assembled proteasomes appear to have proteolytic activity indicating that the interface of the α- and β-subunits has a critical role in function. Regardless of these uncertainties the model allows to obtain an estimate of the minimum and maximum distance between the multiple active sites, symbolized by arrow-heads. The maximum distance $b = 4.2$ nm places them at the periphery, the minimum distance $c = 2$ nm places them near the surface of the inner cavity. The diameter of the proteasome a is approximately 11.5 nm. The molecular ruler hypothesis suggests that either two active sites act in concert, thus excising peptides of a certain length from a polypeptide chain, or that peptides must have a certain length in order to bind strongly enough to an extended binding groove; this too could define a lower length limit for the degradation of peptides.

mainly alanine. This rules out that the *Thermoplasma* proteasome has an exopeptidase activity.

Closer inspection of the peptides derived from the insulin B-chain (Fig. 2) revealed that most of the shorter peptides (<7mers) originated from the N-terminal region, while the longer peptides (>10mers) corresponded predominantly to C-terminal sequences. Such 'edge effects' may be taken as an indication that regions being located upstream or downstream of the scissile bond affect substrate binding. In agreement with this, the larger of the two proteins, hemoglobin, yielded a length distribution histogram (Fig. 3) with a smaller half-width.

The size distribution of the peptides generated by the *Thermoplasma* proteasome suggests that some kind of a 'molecular ruler' exists which determines their length. The provision of a molecular ruler in a multisubunit complex is not without precedent: for the fatty acid synthetase it was inferred from an analysis of fatty acid chain lengths that an intrinsic ruler controls chain termination [30]. With the proteasome it is tempting to put forward the hypothesis that it is the distance between multiple active sites which determines the length of the degradation products (Fig. 4). This hypothesis assumes that a stretch of the unfolded polypeptide chain of the substrate protein binds to an extended groove. Since only the assembled proteasome complex appears to have enzymatic activity it is conceivable that this binding groove is located at the interface of the α- and β-subunits. With two active sites acting in concert, peptides of a defined length could be excised from a polypeptide chain stretched out along the binding groove.

It is unclear at present precisely where the active sites are located. It is clear however, from the structure and simple geometric considerations that the largest possible distance between active sites is 4.2 nm. Assuming that the subunits are ideal spheres, the smallest distance is approximately 2 nm, placing the subunits close to the inner surface of a hollow cylinder. With more wedge-shaped subunits the active sites would move closer towards the cylinder axis and the distance between them would then decrease further. A distance of 2 nm is rather close to the length of a hepta- or octapeptide in an extended conformation. If we allow for some bulging out of the polypeptide chain, comparable to polypeptides bound to MHC molecules [31] somewhat larger stretches could be accomodated. This could account for the heterogeneity in peptide length noted.

In light of our results the existence of a molecular ruler in proteasomes seems likely. However, it may not necessarily be that the distance between active sites provides it. One could also envisage that properties of the binding groove determine the average length of the cleaved peptides. That is assuming that the binding strength per unit length is relatively low, the interaction of a relatively long peptide with an extended binding groove would be required for sufficiently high affinity to effect cleavage.

Peptides that are too short upstream or downstream of the scissile bond might have a small chance of becoming arrested; this could explain why cleavage does not proceed beyond the size of the predominantly found peptides.

References

[1] Rivett, A.J. (1993) Biochem. J. 291, 1–10.
[2] Goldberg, A.L. and Rock, K.L. (1992) Nature 357, 375–379.
[3] Tanaka, K., Tamura, T., Yoshimura, T. and Ichihara, A. (1992) New Biologist 4, 173–187.
[4] Brown, M.G., Driscoll, J. and Monaco, J.J. (1991) Nature 353, 355–357.
[5] Glynne, R., Powis, S.H., Beck, S., Kelly, A., Kerr, L.A. and Trowsdale, J. (1991) Nature 353, 357–360.
[6] Driscoll, J. and Finley, D. (1992) Cell 68, 823–825.
[7] Hershko, A. and Ciechanover, A. (1992) Annu. Rev. Biochem. 61, 761–807.
[8] Hilt, W. and Wolf, D.H. (1992) Mol. Microbiol. 6, 2437–2442.
[9] Seufert, W. and Jentsch, S. (1992) EMBO J. 11, 3077–3080.
[10] Murakami, Y., Matsufuji, S., Kameji, T., Hayashi, S., Igarashi, K., Tamura, T., Tanaka, K. and Ichihara, A. (1992) Nature 360, 597–599.
[11] Rechsteiner, M., Hoffman, L. and Dubiel, W. (1993) J. Biol. Chem. 268, 6065–6068.
[12] Peters, J.M., Cejka, Z., Harris, J.R., Kleinschmidt, J.A. and Baumeister, W. (1993) J. Mol. Biol. 234, 932–937.
[13] Fagan J.M. and Waxman, L. (1992) J. Biol. Chem. 267, 23015–23022.
[14] Baumeister, W., Dahlmann, B., Hegerl, R., Kopp, F., Kuehn, L. and Pfeifer, G. (1988) FEBS Lett. 241, 239–245.
[15] Schauer, T., Nesper, M., Kehl, M., Lottspeich, F., Müller-Taubenberger, A., Gerisch, G. and Baumeister, W. (1993) J. Struct. Biol. 111, 135–147.
[16] Dahlmann, B., Kopp, F., Kuehn, L., Niedel, B., Pfeifer, G., Hegerl, R. and Baumeister, W. (1989) FEBS Lett. 251, 125–131.
[17] Zwickl, P., Grziwa, A., Pühler, G., Dahlmann, B., Lottspeich, F. and Baumeister, W. (1992) Biochemistry 31, 964–972.
[18] Hegerl, R., Pfeifer, G., Pühler, G., Dahlmann, B. and Baumeister, W. (1991) FEBS Lett. 283, 117–121.
[19] Pühler, G., Weinkauf, S., Bachmann, L., Müller, S., Engel, A., Hegerl, R. and Baumeister, W. (1992) EMBO J. 11, 1607–1616.
[20] Jap, B., Pühler, G., Lücke, H., Typke, D., Löwe, J., Stock, D., Huber, R. and Baumeister, W. (1993) J. Mol. Biol. 234, 881–884.
[21] Zwickl, P., Lottspeich, F. and Baumeister, W. (1992) FEBS Lett. 312, 157–160.
[22] Dick, L.R., Moomaw, C.R., DeMartino, G.N. and Slaughter, C.A. (1991) Biochemistry 30, 2725–2734.
[23] Wenzel, T. and Baumeister, W. (1993) FEBS Lett. 326, 215–218.
[24] Dahlmann, B., Kühn, L. and Reinauer, H. (1983) FEBS Lett. 160, 243–247.
[25] Covey, T.R., Bronner, R.F., Shushan, B.I. and Henion, J. (1988) Rapid Commun. Mass Spectrom. 2, 249–256.
[26] Mann, M., Meng, C.K. and Fenn, J.B. (1989) Anal. Chem. 61, 1702–1708.
[27] Orlowski, M., Cardozo, C. and Michaud, C. (1993) Biochemistry 32, 1563–1572.
[28] Driscoll, J., Brown, M.G., Finley, D. and Monaco, J.J. (1993) Nature 365, 262–264.
[29] Gaczynska, M., Rock, K.L. and Goldberg, A.L. (1993) Nature 365, 264–267.
[30] Sumper, M., Oesterhelt, D., Riepertinger, C. and Lynen, F. (1969) Eur. J. Biochem. 10, 377–387.
[31] Guo, H.C., Jardetzky, T.S., Garrett, T.P.J., Lane, W.S., Strominger, J.L. and Wiley, D. (1992) Nature 360, 364–366.