ELSEVIER

Biophysical Chemistry 100 (2003) 437–452

Biophysical Chemistry

www.elsevier.com/locate/bpc

Isolation and characterization of the prokaryotic proteasome homolog HslVU (ClpQY) from *Thermotoga maritima* and the crystal structure of HslV☆

Hyun Kyu Song[a,1], Matthias Bochtler[b], M. Kamran Azim[a], Claudia Hartmann[a], Robert Huber[a], Ravishankar Ramachandran[a,*]

[a]*Abteilung Strukturforschung, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Planegg-Martinsried, Germany*
[b]*Max-Planck-PAN Joint Junior Research Program, International Institute of Molecular Biology and Cell Biology, ul. Ks. J. Trojdena 4, 02-109 Warsaw, Poland*

Received 15 March 2002; accepted 15 May 2002

Abstract

Heat-shock locus VU (HslVU) is an ATP-dependent proteolytic system and a prokaryotic homolog of the proteasome. It consists of HslV, the protease, and HslU, the ATPase and chaperone. We have cloned, sequenced and expressed both protein components from the hyperthermophile *Thermotoga maritima*. *T. maritima* HslU hydrolyzes a variety of nucleotides in a temperature-dependent manner, with the optimum lying between 75 and 80 °C. It is also nucleotide-unspecific for activation of HslV against amidolytic and caseinolytic activity. The *Escherichia coli* and *T. maritima* HslU proteins mutually stimulate HslV proteins from both sources, suggesting a conserved activation mechanism. The crystal structure of *T. maritima* HslV was determined and refined to 2.1-Å resolution. The structure of the dodecameric enzyme is well conserved compared to those from *E. coli* and *Haemophilus influenzae*. A comparison of known HslV structures confirms the presence of a cation-binding site, although its exact role in the proteolytic mechanism of HslV remains unclear. Amongst factors responsible for the thermostability of *T. maritima* HslV, extensive ionic interactions/salt-bridge networks, which occur specifically in the *T. maritima* enzyme in comparison to its mesophilic counterparts, seem to play an important role.

Keywords: HslV; HslU; Proteolytic activity; Crystal structure; Thermostability; Proteasomal homolog

Abbreviations: AAA, ATPases associated with a variety of cellular activities; AOR, aldehyde ferredoxin oxidoreductase from *Pyrococcus furiosus*; EcV, *E. coli* HslV; EcU, *E. coli* HslU; EcVU, *E. coli* HslVU; HiV, *H. influenzae* HslV; HiU, *H. influenzae* HslU; HiVU, *H. influenzae* HslVU; MBP-SulA, maltose-binding protein and SulA fusion protein; r.m.s., root-mean square; *T. maritima, Thermotoga maritima*; TmV, *T. maritima* HslV; TmU, *T. maritima* HslU; TmVU, *T. maritima* HslVU; Z-GGL-AMC, carbobenzoxy–Gly–Gly–Leu–7-amido-4-methylcoumarin.

☆ This paper is dedicated to the memory of Professor John T. Edsall.

*Corresponding author. Tel.: +49-89-8578-2678; Fax: +49-89-8578-3516.
E-mail address: rshankar@biochem.mpg.de (R. Ramachandran).
[1] Present address: Department of Cancer Biology, Dana-Farber Cancer Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA.

0301-4622/03/$ - see front matter
PII: S0301-4622(02)00297-1

1. Introduction

HslVU (ClpQY) is a prokaryotic ATP-dependent protease that has recently attracted a lot of interest as a simple model system for the eukaryotic 26S proteasome [1–4]. Its proteolytic component, HslV, is a dimer of hexamers that forms a central proteolytic chamber with active sites on the interior walls of the cavity [1]. As individual subunits share significant sequence [5] and structural similarity [1] with proteasomal β-subunits, HslV may be regarded as a 20S proteasomal core particle lacking the anti-chambers and displaying different symmetry. HslV and proteasomal β-subunits are members of the Ntn-family of hydrolases and have a threonine residue at the N-terminus after the processing of a methionine residue or a propeptide [1,6–8]. The proteolytic activity of HslV is stimulated by HslU, which possesses ATPase and 'unfoldase' activity and belongs to the Hsp100 family of ATPases [9]. The structure determination of HslU showed that it exists as a hexamer, in which each subunit consists of three domains termed N-terminal (N), intermediate (I), and C-terminal (C), respectively [10]. It also confirmed it as a member of the family of ATPases associated with a variety of cellular activities (AAA-ATPase). HslU is therefore a useful model system for the base part of the proteasomal 19S regulatory cap, which contains six components that belong to the AAA-ATPase family [11].

In spite of abundant structural and biochemical information about HslVU, many aspects of the enzyme mechanism are still a matter of debate [12]. Not only have different docking modes been observed in HslVU crystal structures from different sources, but different quaternary structures, singly and doubly capped species, are also known from the crystal structure of the *E. coli* enzyme [10,13,14]. Extensive mutagenesis studies on *E. coli* HslVU [15] could not resolve the ambiguity in quaternary association. Furthermore, a sodium or potassium ion was implicated in modulating the proteolytic activity of HslV by virtue of its proximity to the active site in the crystal structure of *H. influenzae* HslV (to be called HiV) [16]. Recently CodW, an HslV homolog from *B. subtilis*, was shown to be an N-terminal serine protease [17]. This is in stark contrast to HslVs from *E. coli* and *H. influenzae*, which have all been shown to exhibit an N-terminal threonine-dependent catalytic mechanism. This is surprising, as CodW possesses the catalytic threonine residue present in HslV from all organisms of known sequence.

In the current studies, we report the cloning and expression of HslV and HslU from *T. maritima* (to be called TmV and TmU, respectively). Their biochemical properties were characterized and the crystal structure of TmV was solved at high resolution. Comparisons of the available crystal structures of HslV, amongst other things, throw light on the probable factors for thermostability of TmV.

2. Materials and methods

2.1. Cloning and expression

T. maritima HslV and HslU were cloned using standard polymerase chain reaction (PCR) techniques. PCR primers were designed based on the reported *T. maritima* DNA sequences [34], viz. TM0521 and TM0522 for *hslV* and *hslU* respectively, such that *hslV* was flanked by NdeI and BamHI restriction sites while *hslU* was flanked by NdeI and NcoI sites. The resultant fragments were ligated into pET-12b or pET-20b expression vectors (Novagen) which had been digested with the appropriate restriction enzymes (New England Biolabs). The integrity of the resultant plasmids was verified by DNA sequencing. The plasmids were then transformed into BL21(DE3) or B834(DE3). Cells were grown at 37 °C. Expression of TmU was induced by the addition of 1 mM IPTG at OD (600 nm) = 0.7. Maximal expression of TmV was observed when 1 mM IPTG was added after the cell culture reached saturation. The cells were centrifuged and kept frozen at −20 °C until further use.

2.2. Purification

The same procedure was used to purify both TmV and TmU. The cell pellet was resuspended in ice-cold lysis buffer A (20 mM Tris–HCl, pH 7.5, 100 mM NaCl) and subsequently disrupted

by ultrasonication. A tablet of protease inhibitor cocktail (Roche Molecular Biochemicals), 1 mg of lysozyme and 0.5 mg of Dnase A per liter of culture were added before ultrasonification. The crude cell extract was centrifuged and the resultant supernatant was immersed in a water bath at 80 °C for 30 min and then placed in an ice bath for 5 min. Denatured *E. coli* proteins were removed by centrifugation and the supernatant was loaded onto a Hi-Trap Q-Sepharose anion exchange column (5 ml, Amersham Pharmacia) which was previously equilibrated with buffer A containing 1 mM EDTA. TmV and TmU eluted at approximately 250 and 350 mM NaCl, respectively. The final chromatographic step was performed using a Superose 6 preparation-grade column (Amersham-Pharmacia) previously equilibrated with buffer A containing 1 mM EDTA and 200 mM sodium chloride. The protein was concentrated to approximately 10 mg/ml and stored at 4 °C. N-terminal sequencing and electron spray ionization mass spectroscopic (ESI-MS) analyses were carried out to check the integrity of the protein samples.

2.3. Assays

Proteins were quantified by their absorbance at 280 nm or by the method of Bradford [18] using bovine serum albumin as a standard.

Nucleotide hydrolytic activity was measured as described earlier [19]. In this method, the amount of inorganic phosphate formed is spectroscopically determined at 660 nm as a complex with malachite green and ammonium molybdate.

Peptide hydrolysis was assayed using carbobenzoxy–Gly–Gly–Leu–7-amido-4-methylcoumarin (Z-GGL-AMC; Bachem) as a substrate [3]. Enzyme samples (1 μg of HslV; 2 μg of HslU) were incubated for 20 min in buffer U (50 mM Tris, pH 7.5, 5 mM $MgCl_2$), 1 mM adenosine 5′-*O*-(3-thiotriphosphate) (ATP-γS) and 2 mM Z-GGL-AMC. The reaction was terminated by the addition of 1% SDS to a final volume of 1000 μl and the fluorescence of the resultant products was measured to quantitate them.

Resorufin-labeled casein (Roche Molecular Biochemicals) [20] or fluorescein isothiocyanate (FITC)-labeled casein [21] were used as model protein substrates. In the case of the former substrate, 10 μg of HslV and 25 μg of HslU were incubated for various time periods and temperatures in the presence of 0.4% resorufin-labeled casein in buffer U. Undigested substrate was precipitated using 5% trichloroacetic acid and removed by centrifugation. The absorbance of released peptides in the supernatant fractions was spectroscopically measured at 574 nm. The degradation of FITC-casein was measured using HPLC. Enzyme samples (15 μg of HslV; 37.5 μg of HslU) in buffer U were incubated for 45 min with 2 mM ATP-γS and 1 mM FITC-casein. The reaction was stopped by the addition of calpain inhibitor I (acetyl–Leu–Leu–norleucinal) to a final concentration of 1 mM.

Earlier reports have shown that SulA has a high tendency to aggregate in vivo, as well as in vitro, and therefore MBP-SulA (maltose binding protein and SulA fusion protein) was used as the substrate [22]. The pMal-p2-SulA plasmid and *E. coli* strain CSH26 were kindly donated by Dr A. Higashitani (National Institute of Genetics, Japan). MBP-SulA was overexpressed and purified as previously described [22]. The reaction mixture (60 μl) contained 4 μg of MBP-SulA, 1 μg of HslV, 2.5 μg of HslU, 0.02% Triton X-100, 1 mM DTT and 3 mM ATP in buffer U. After 5–7 h of incubation at various temperatures, the reaction was stopped by adding 35 μl of 50 mM Tris–HCl, pH 6.8, 0.1 M DTT, 2% SDS, 0.1% bromophenol blue and 10% glycerol, and its products were analyzed on 12% (w/v) slab gels containing SDS [23].

2.4. Crystallography

Crystals of TmV were grown at room temperature by the hanging-drop vapor-diffusion method. A hanging drop was prepared by mixing equal volumes of a 9 mg/ml protein solution and the reservoir solution [27.5% (w/v) MPD, 200 mM magnesium acetate and 100 mM sodium cacodylate, pH 6.5].

Diffraction data were collected using a CCD detector at the BW6 beamline, Deutsche Elektronen Synchrotron Center, Hamburg, Germany. The data were processed and scaled using DENZO and SCALEPACK [24]. Crystals belong to an orthorhom-

Table 1
Data collection and refinement statistics

Data collection statistics	
Space group	*I*222
Cell parameters (Å)	$a=77.53$, $b=93.75$, $c=146.17$
Resolution range[a] (Å)	17.0–2.1 (2.14–2.1)
Unique/total reflections	31 188/334 480
Redundancy	10.7
Completeness[a] (%)	99.3 (99.7)
R_{merge}[a,b] (%)	5.2 (24.9)
Refinement and model statistics	
R-factor/R_{free}[c] (%)	19.1/22.9
Resolution range (Å)	17.0–2.1
Modeled residues	171×3
Number of water molecules/Na^+ ions	198/3
R.M.S.D. bond lengths (Å)	0.005
R.M.S.D. bond angles (°)	1.24
Average *B* value (Å^2)	
Main-chain/side-chain/water/Na^+	26.8/31.5/35.9/22.7
Ramachandran outlier (%)	0

R.M.S.D., root-mean square deviation.

[a] Values in the parenthesis are for the reflections in the highest-resolution shell (2.14–2.1 Å).

[b] $R_{merge}=\sum_h\sum_i|I(h,i)-\langle I(h)\rangle|/\sum_h\sum_i I(h,i)$, where $I(h,i)$ is the intensity of the *i*th measurement of reflection *h* and $\langle I(h)\rangle$ is the average value over multiple measurements.

[c] $R=\sum\big||F_o|-|F_c|\big|/\sum|F_o|$, where R_{free} is calculated for a 10% test set of reflections.

bic space group, *I*222 and have cell parameters of $a=77.53$ Å, $b=93.75$ Å and $c=146.17$ Å. Molecular replacement was used to obtain initial phases using AMoRe [25] and the *E. coli* HslV (PDB code 1NED) model was used for calculations. The asymmetric unit contains three subunits of TmV. The SIGMAA-weighted ($2mF_o-DF_c$) and three-fold averaged map calculated after one round of rigid-body refinement showed unambiguous density for nearly all side chains. Models were rebuilt using O [26] and refined using CNS [27]. Engh and Huber parameters [28] were used throughout. Non-crystallographic symmetry restraints were imposed during the refinement cycles, but were released in the final round. Solvent molecules were added using 'F_o-F_c' maps using CNS. PROCHECK [29] was used to assess the quality of the model. Programs in the CCP4 suite [30], CHAIN [31] and GRASP [32] were used to analyze the structure. The data collection and the refinement statistics are summarized in Table 1. The coordinates and structure factors have been deposited in the Protein Data Bank under the accession code 1M4Y.

3. Results and discussion

3.1. Cloning and purification of TmV and TmU

TmV and TmU share 63.3 and 54.1% sequence identity to HslV and HslU from *E. coli* (to be called EcV and EcU), respectively. A sequence comparison of TmV with HslVs from several sources and the β-subunit of the 20S proteasome from *Thermoplasma acidophilum* is shown in Fig. 1a. TmV has a short extension of five amino acids preceding two threonine residues and a characteristic Gly residue at the −1 position. We generated constructs with/without this extension (Table 2). This putative propeptide was auto-processed, albeit inefficiently, in the *E. coli* expression system in the former case, resulting in active enzyme. While the unprocessed form exhibits mixed oligomeric states, the processed form exhibits a single oligomeric state (dodecamer). This situation is analogous to the *T. acidophilum* 20S proteasome when the α-subunits are not processed [33]. TmU was purified according to the same procedure as for

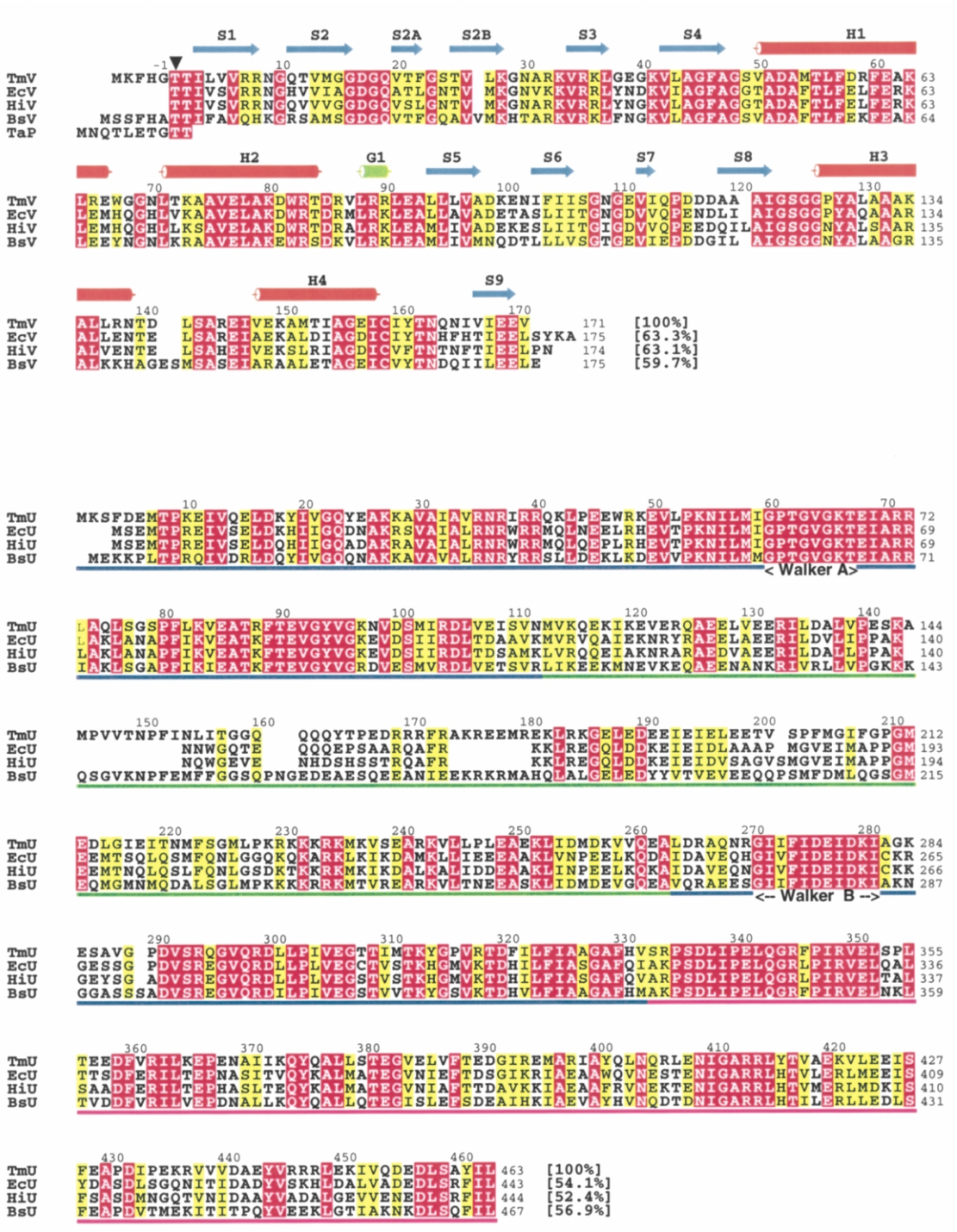

Fig. 1. (a) Sequence alignment of *T. maritima* HslV (TmV), *E. coli* HslV (EcV), *H. influenzae* (HiV), *B. subtilis* CodW (BsV) and the propeptide of the β-subunit of the proteasome from *T. acidophilum* (TaP). The catalytically important N-terminal threonine residue is marked by a black triangle, and the secondary structural elements, 11 β-strands (*cyan arrows*), four α-helices (*orange cylinders*), and a 3_{10}-helix (*green cylinder*), are indicated above the sequence for TmV. (b) Sequence alignment of *T. maritima* HslU (TmU), *E. coli* HslU (EcU), *H. influenzae* (HiU), and *B. subtilis* CodX (BsU). N-, I-, and C-domains are marked as blue, green, and purple bars, respectively. The Walker A- and B-motifs are labeled. In the sequences, identical residues are colored in red and homologous residues are colored in yellow. This figure was produced with ALSCRIPT [49].

Table 2
Molecular characteristics of *T. maritima* HslVU expressed in *E. coli*

Constructs	Expression vector	Host cell	Expression level	Oligomer[a]	Molecular mass		Comments
					Calculated	MS	
TmU	pET-20b	B834(DE3)	High	Hexamer	53052.4	53059.8	Minor degradation bands $_{156}$ITGGQQ, $_{165}$TPEDR, $_{224}$SGML
TmV							
Intact	pET-20b	B834(DE3)	Medium	Mixture (many states)	18932.8	ND	Mixture: intact (~80%) and processed (~20%) form[b]
Processed	pET-12b	BL21(DE3)pLysS	Low	Dodecamer	18332.0	18332.0	N-terminal sequencing: ~90% TTIL, ~10% MTTIL

ND, not determined.
[a] By gel filtration profile.
[b] By SDS-PAGE band intensity.

TmV outlined in Section 2. It exists as a hexamer, as verified by chromatographic methods (Table 2). Both proteins have moderately high expression levels in *E. coli*.

3.2. Hydrolysis of nucleotides by TmU

TmU hydrolyses ATP and other nucleotides at comparable rates, as summarized in Table 3 (also Fig. 2), in a temperature dependent manner. At physiological temperature, the ATPase activity is observed to be approximately 10% of the maximum, which lies between 75 and 80 °C and compares well with the optimum growth temperature of *T. maritima*, viz. 80 °C [34]. GTP and UTP were found to be hydrolyzed faster than ATP, whereas CTP was hydrolyzed more slowly compared to ATP. In addition, deoxynucleotides were also hydrolyzed by TmU. Moreover, TmU was also able to stimulate the amidolytic and caseinolytic activity of TmV in the presence of these other nucleotides (Table 3).

In contrast to bacterial activators such as EcU and CodX, which were biochemically characterized as ATPases, proteasome-activating nucleotidase (PAN) from an archaeon, *Methanococcus jannaschii*, hydrolyses both ATP and CTP [35]. In addition, an N-terminal deletion mutant, PAN(Δ1-73), possesses general nucleotide-hydrolysis activity [36], as observed in the case of TmU (Table

Table 3
Nucleotide dependence of *T. maritima* HslVU

	Pi generation[a]	Z-GGL-AMC	Casein
TmV + ATPγS	NA	–	–
TmVU + ATP	++++	++++	++++
TmVU (−ATP)	NA	+++	+++
TmVU + ADP	–	+++	++
TmVU + AMP-PNP	–	+++	++++
TmVU + ATPγS	–	++++	++++
TmVU + CTP	++	++++	++++
TmVU + GTP	+++++	++++	++++
TmVU + UTP	++++++	++++	+++++
TmVU + dATP	+	+++	+++
TmVU + dCTP	+++++	++++	++++
TmVU + dGTP	++++	++++	+++
TmVU + dTTP	++++++	++++	++++

The values are relative to the hydrolytic activity of TmVU at 70 °C as 100% (++++++, 150–200%; +++++, 110–150%; ++++, 80–110%; +++, 60–80%; ++40–60%; +, 20–40%; –, <20%); NA, not applicable.
[a] Nucleotide hydrolysis activity was measured in absence of HslV.

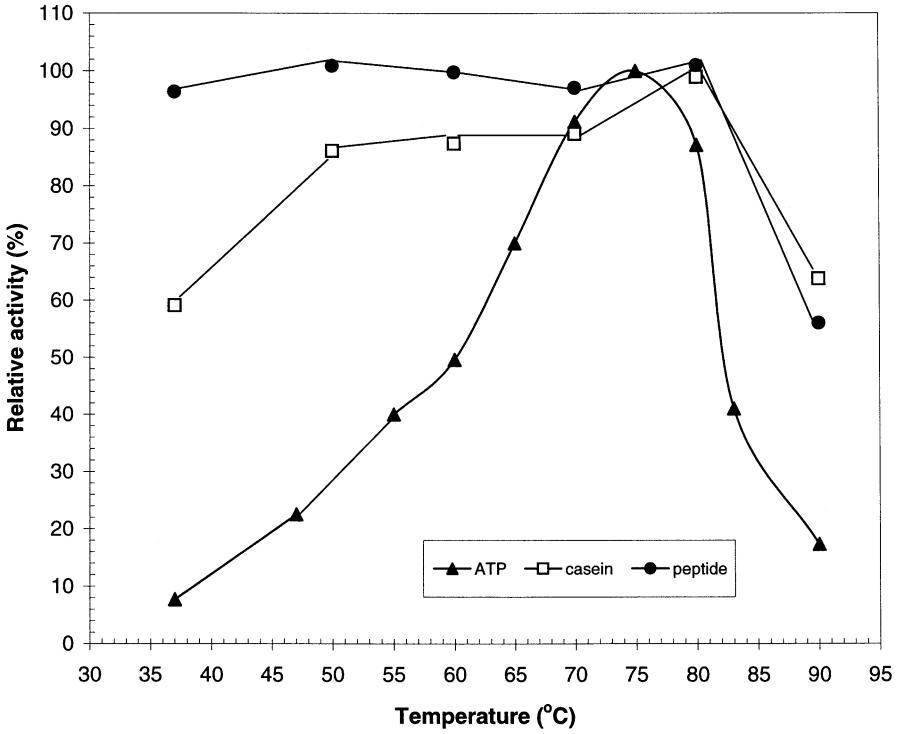

Fig. 2. Temperature dependence of *T. maritima* HslVU activity; ATPase, caseinolytic and amidolytic activity is indicated by triangles, squares and circles, respectively. All values are relative to the corresponding activity at 80 °C, which was set to 100% (see Section 2 for assay conditions).

3). It is known that the backbone amide and carbonyl groups of Ile18 are important determinants for the adenine ring in EcU–nucleotide complexes, and the region around Ile18 is quite well conserved among the several HslUs (Fig. 1b). The NTPase activity exhibited by TmU is consequently intriguing. Ongoing structural studies of TmU and its complexes with various nucleotide analogs are expected to clarify some of these less understood aspects.

3.3. Stimulation of the proteolytic activity of TmV by TmU

To determine whether TmV and TmU form an active proteolytic system in vitro, their various activity values were examined. Towards this end, we made use of three different substrates, viz. a small chromogenic peptide (Z–Gly–Gly–Leu–AMC), casein as a model flexible substrate and MBP-SulA. In *E. coli*, SulA is known to be a natural substrate of HslVU [37]. Currently, a homolog for SulA in *T. maritima* has not yet been identified. Nevertheless, it was of interest to determine whether TmVU could degrade *E. coli* SulA, as this would mean that TmU could also specifically recognize this substrate and unfold it.

The peptide and casein degradation assays were performed in the presence of the non-hydrolyzable ATP analog, ATP-γS, as reported earlier [15], while the MBP-SulA assay requires ATPase activity. TmVU was found to hydrolyze the small chromogenic peptide without any temperature dependence. The degradation of casein, on the other hand, was found to be temperature-dependent (Fig. 2). We initially pre-incubated TmVU and *E. coli* MBP-SulA for 30 min at 37 °C and then increased the assay temperature to 70 °C as the ATPase activity of TmU is substantial at higher temperatures, although it is still approximately

Table 4
Cross-activity of HslVU between *T. maritima* and *E. coli*

	Complex formation[a]	Z-GGL-AMC			Casein			MBP-SulA	
		37 °C	50 °C	70 °C	37 °C	50 °C	70 °C	30 °C	70 °C
TmU: TmV	No	++++	++++	++++	+++	+++++	+++++	–	–
TmU: EcV	No	++++	++++	+++	++	+	+	–	–
EcU: TmV	No	++++	++++	++++	+++	+++++	+++++	–	–
EcU: EcV	No	++++	++++	++++	++++	++	++	++++	–

The values are relative to TmVU activity at 70 °C as 100% (+++++, 110–150%; ++++, 80–110%; +++, 60–80%; ++, 40–60%; +, 20–40%; –, <20%).
[a] By gel filtration.

10% of the maximum at physiological temperatures. It was observed, however, that TmVU could not degrade *E. coli* MBP-SulA (Table 4). We suggest that *T. maritima* HslU does not recognize *E. coli* SulA; but as a cautionary note it must be pointed out that MBP-SulA has a tendency to aggregate under high-temperature assay conditions, although the assay at physiological temperatures is fairly well established.

We had reported earlier [12] that peptides derived from the carboxy terminus of HslU stimulate the amidolytic and caseinolytic activity of HslV to levels observed with wild-type HslU itself. Those experiments support a mechanism wherein the carboxy termini of HslU bind to pockets of HslV in order to activate it for proteolytic activity [13]. In order to check for the same feature in TmVU, we synthesized an octapeptide corresponding to the carboxy terminus of TmU. We found that this peptide could support amidolytic activity like full-length TmU, thereby suggesting that the activation mechanism is conserved across different species. This conclusion is in line with the results of sequence comparisons of the carboxy termini of HslU from different sources [12]. We also found that an octapeptide corresponding to the carboxy terminus of *E. coli* HslU stimulates the amidolytic activity of TmV.

3.4. Cross-reactivity between the E. coli and T. maritima complexes

It is known from studies on *B. subtilis* HslVU (CodWX) that the CodW-EcU hybrid hydrolyzed *E. coli* SulA is nearly as actively as the EcVU complex at 37 °C, although the CodX-EcV hybrid could not degrade the substrate [17]. It was also reported that CodX-EcV and CodW-EcU hybrids could actively hydrolyze Z-GGL-AMC and casein. It was suggested, therefore, that CodX could not recognize *E. coli* SulA.

We therefore correspondingly examined the activity of the TmV-EcU and EcV-TmU hybrid complexes against different substrates. We found that for peptide and casein substrates both hybrids were as active as native TmVU and EcVU complexes themselves (Table 4). We then examined the activity of the TmV-EcU and EcV-TmU hybrids against *E. coli* MBP-SulA. It is expected a priori that since EcU recognizes SulA, at least the TmV-EcU hybrid should exhibit activity against this substrate. Surprisingly, the TmV-EcU hybrid (as also EcV-TmU) did not exhibit any activity against MBP-SulA at both 37 and 70 °C.

Although the results of the amidolytic and caseinolytic activity values agree well with studies on the *B. subtilis* complex, the MBP-SulA activity of the TmV-EcU hybrid does not. Lack of activity of the TmV-EcU hybrid complex against SulA at 70 °C might be due to reduced activity of EcU and/or aggregation of MBP-SulA at high temperature. On the other hand, the low activity of TmV at 37 °C could probably be responsible for the lack of degradation of the MBP-SulA substrate at physiological temperature (Table 4).

3.5. Crystal structure of TmV

The crystal structure of TmV was solved using EcV as a model in molecular replacement calcu-

lations. The model accounts for all 171 residues of TmV, three sodium ions and 198 water molecules. It has been refined to a free *R*-value of 22.9% for data in the range 17.0–2.1 Å and has good stereochemistry. The main chain is unambiguously defined and the electron density for the side chains of the following residues is either poor or absent: Phe22, Arg35, Arg86, Arg89 and Arg90. Among 447 non-glycine and non-proline residues, the numbers of residues lying in the most favored, additionally allowed, generously allowed and disallowed regions in the Ramachandran plot are 410 (91.7%), 37 (8.3%), 0 and 0, respectively. The refinement statistics, as well as model quality parameters are summarized in Table 1. A representative omit map is shown in Fig. 3.

The protomer consists of the typical fold exhibited by β-subunits of the proteasome [6] and it forms a double-doughnut-shaped dodecamer [1]. The entrance to the proteolytic chamber of TmV is formed by a hydrogen-bonded turn (followed by a long helix, H2) and a short 3_{10}-helix, G1. Its average diameter as measured between C^{α} atoms is 20.5 Å, a value that is slightly greater than that of the α-ring (17 Å) and significantly smaller than the β-ring (27 Å) of the archaeal proteasome [6]. In summary, the tertiary and quaternary structures of TmV are similar to EcV and HiV (Figs. 4 and 5) [1,13]; however, several differences in details do exist.

3.6. Comparison of free HslV structures

A superposition of the C^{α} atoms of free TmV, EcV and HiV, refined at 2.1, 3.8 and 1.9 Å, respectively, is shown in Fig. 4. The root-mean square (R.M.S.) deviation between free TmV and EcV is 0.94 Å for 171 C^{α} atom pairs, with residues 10, 38, 39, 83–90, 93–95, 100, 101, 116–121, 138–141, 163 and the C-terminus of TmV deviating by >1.0 Å at the C^{α} atom. Between free TmV and HiV, the R.M.S. deviation is 0.93 Å for 171 C^{α} atom pairs, with residues 27, 38, 39, 85–90, 100, 101, 103, 109–111 and 115–118 of TmV exhibiting deviations greater than 1 Å at the C^{α} atoms. These residues are located mostly on loops (Fig. 1a, Fig. 4).

Recently [16], it has been reported that two alternate sets of inter-subunit interactions in the HiV quaternary structure result in 'quasi-equivalent' packing within the assembly. We examined the TmV structure and observed that inter-subunit interactions of TmV subunits are similar to only one of the two alternate sets of interaction observed in HiV; specifically, where Asp52 makes a salt bridge with Arg83 of an adjacent subunit in the hexameric ring. Consequently, the entrance of TmV exhibits a circular shape with the distance across three pairs of C^{α} atoms among the six subunits being 19.5, 20.0 and 21.9 Å, respectively (Fig. 5). Although the average diameter as measured between C^{α} atoms is nearly the same for EcV, HiV and TmV, at 19.9, 19.3 and 20.5 Å, respectively, the HiV structure is very elliptical, with values of 13.1, 19.1 and 25.7 Å [14] (17.9, 19.7 and 22.0 Å in the case of EcV) due to its asymmetry. Hence, the TmV structure, in comparison with those of EcV and HiV, supports the suggestion that there is inherent flexibility in the intersubunit interactions of HslV.

K^+/Na^+ ion-binding sites were identified in the crystal structures of HiV, and it was further proposed that these monovalent ions are likely interchangeable with the divalent Mg^{2+} ion [14]. TmV crystallizes in the presence of 200 mM magnesium acetate and 100 mM sodium cacodylate buffer, and the buffer in which the protein was stored contained 200 mM sodium chloride. We therefore carefully looked for a cation-binding site analogous to that in HiV, and observed that a solvent peak in each TmV subunit exhibited a geometry inconsistent with water and occurred at the same position as the proposed K^+/Na^+ binding site (near residues 157, 160 and 163) in HiV, and consistent with that expected for an Na^+ ion [38]. This site in TmV also exhibited distorted octahedral coordination formed by three main-chain carbonyl oxygen atoms and three water molecules, as also reported in the HiV structure (Fig. 3b). It is not easy to distinguish between Na^+ and Mg^{2+} ions in the current structure, but geometric analysis supports that it is likely a Na^+ ion. The mean distance between oxygen and Na^+ atoms is 2.43 Å in the Cambridge Structural Database and 2.57 Å in the Protein Data Bank

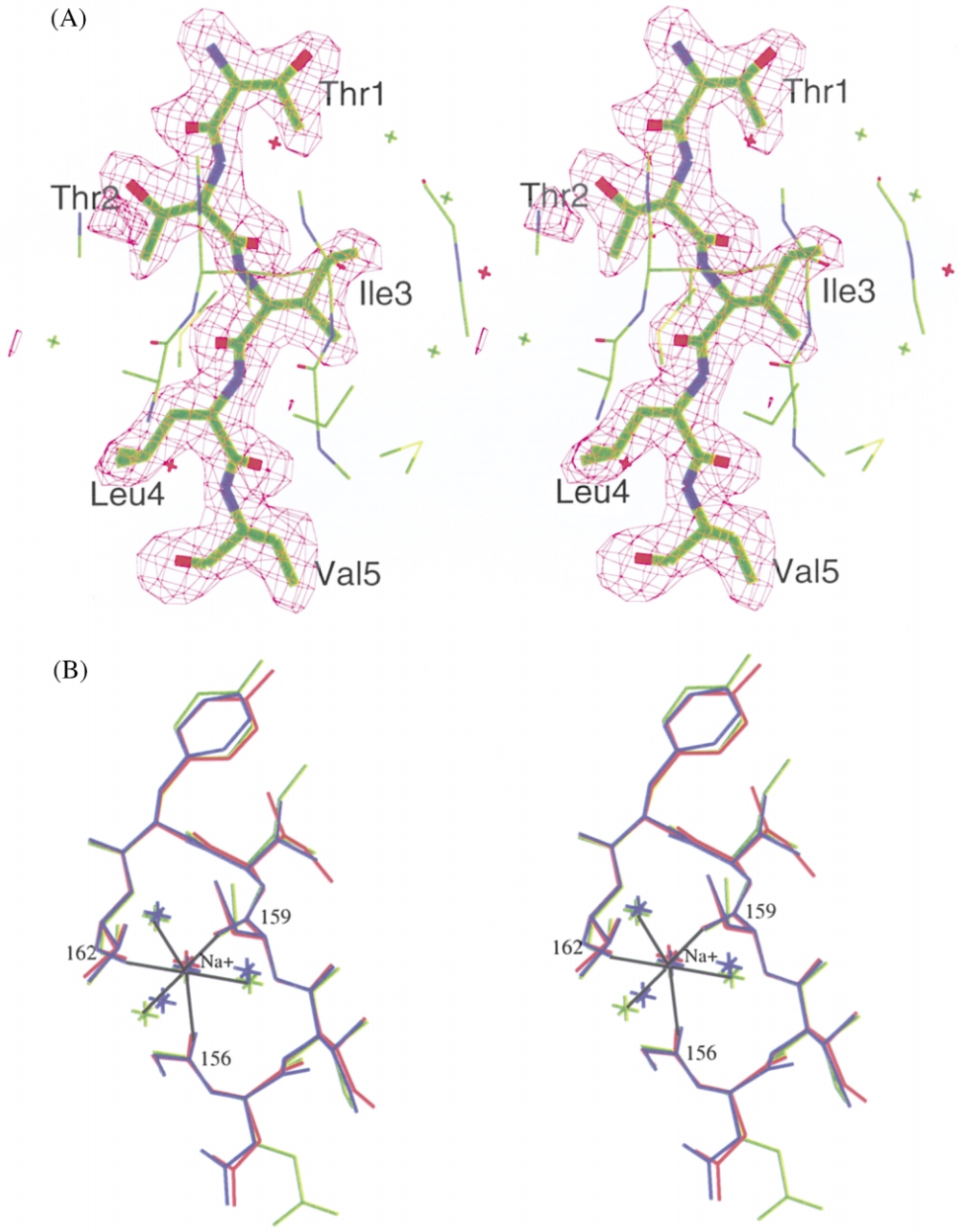

Fig. 3. (a) Stereo diagram showing the simulated annealed omit map calculated using data in the range 17–2.1 Å and contoured at 3σ. Omitted residues (five N-terminal residues) are drawn in thick lines and labeled. Carbon, nitrogen and oxygen atoms are colored green, blue and red, respectively. (b) Stereo diagram of the cation-binding site in TmV (*green*), EcV (*red*) and HiV (*blue*) taken from the first HslV chain in the respective coordinates. Na^+ and the water molecules are indicated by crosses. Metal–protein and metal–water interactions in TmV are marked. Some residues of TmV are labeled for clarity.

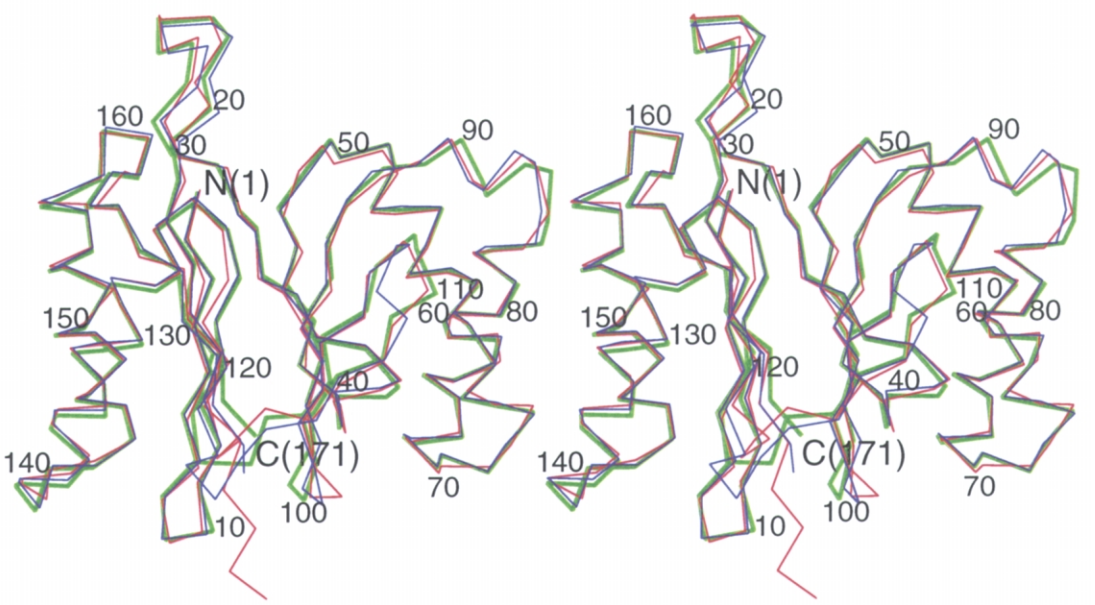

Fig. 4. Stereo diagram showing the C^{α} atom superposition of free HslV structures. Green, red and blue lines represent TmV, EcV and HiV, respectively. Every 20th residue of TmV is labeled.

[50]. The mean metal-donor distance between Mg^{2+} and the main-chain carbonyl oxygen is 2.07 Å [39]. The mean distance between the ion and oxygen atoms in our structure is 2.40 Å (ion–carbonyl oxygen, 2.25 Å; ion–water oxygen, 2.54 Å). This value is closer to that expected for a Na^{+} ion than Mg^{2+} in the database, and moreover is nearly the same as that observed in the Na^{+}-containing HiV structure [16].

In view of the HiV and TmV structures, we re-examined the free EcV and its co-crystal structures, which were solved to resolutions of 3.8 and 2.8 Å, respectively [1,10]. Even in the latter structure, it is difficult to unambiguously distinguish between a solvent site and a K^{+}/Na^{+} ion-binding site. Nevertheless, a site assigned earlier to a water molecule occurs at the same position as that of the cation in the HiV and TmV structures (Fig. 3b). The exact role of the cation in the proteolytic mechanism of HslV is currently unclear [16].

3.7. Structural basis for the thermostability of TmV

With the availability of crystal structures of HslV from three different sources, it is instructive to compare them to understand the causes for the high thermostability of TmV. The thermal denaturation temperatures (T_m) of TmV and EcV were determined by circular dichroism studies to be $>$ 110 and 70 °C, respectively (unpublished results, data not shown). We examined several factors that are thought to be important determinants of thermostability in proteins (for a recent review see [40]), and these are summarized in Table 5 and discussed in greater detail below.

An examination of the primary structures of the three enzymes did not reveal any striking difference in the amino acid composition. The number of residues in TmV is smallest at 171 residues, compared to 175 and 174 residues for EcV and HiV, respectively (Fig. 1a, Table 5). It has been suggested that shorter loops may contribute to the resistance of proteins towards thermal unfolding, as observed in the case of citrate synthase from *Thermoplasma acidophilum* [41]. The number of Arg residues or the ratio Arg/(Arg+Lys) has been reported to be higher in some thermostable proteins [42]. The number of Arg residues in TmV, EcV and HiV is 12, 10 and 11, while the Arg/(Arg+Lys) ratio is 0.55, 0.50 and 0.55, respec-

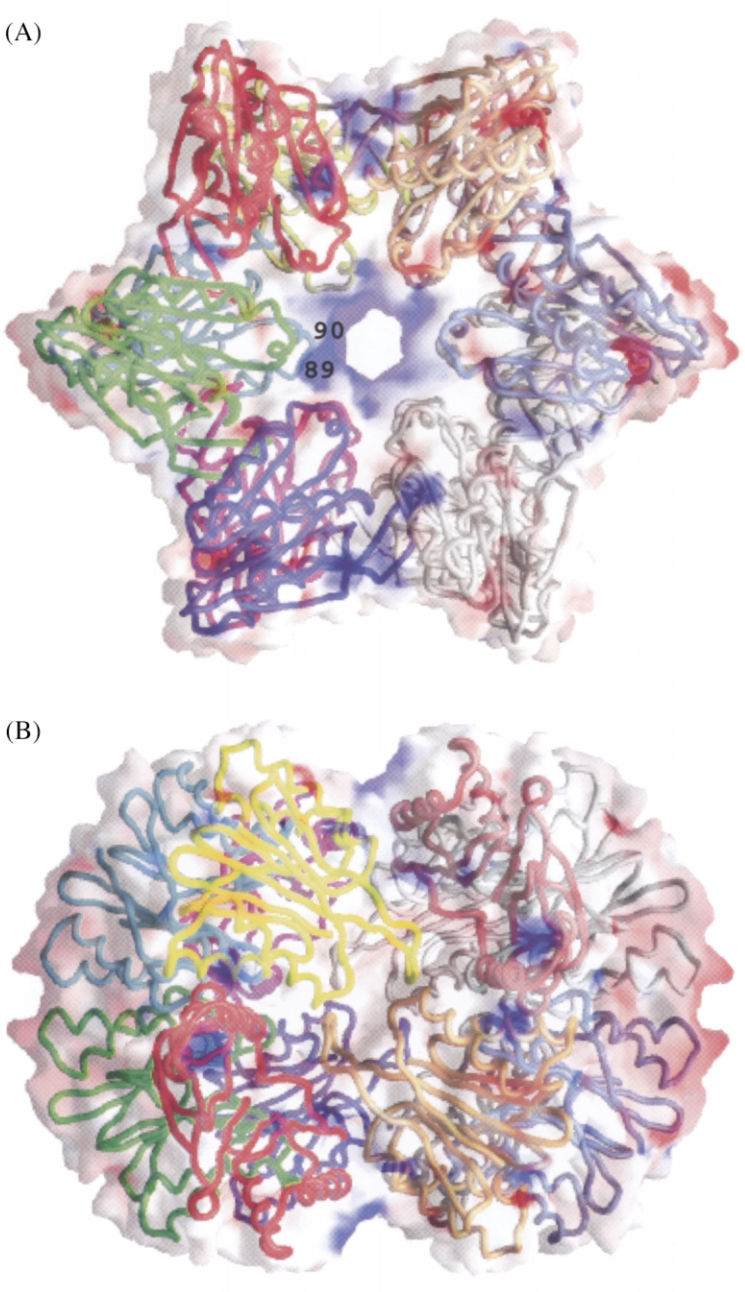

Fig. 5. (a) View along the six-fold axis of the TmV dodecamer. (b) Side view of the dodecamer. Arg89 and Arg90, which form part of a positively charged cluster of residues lining the entrance pore of TmV, are indicated. Positively and negatively charged transparent surfaces are shown in blue and red, respectively. Each polypeptide chain is colored differently for clarity. The figure was drawn using GRASP [32].

Table 5
Comparison of thermostability factors

	TmV	EcV	HiV
Total number of residues per chain[a]	171	175	174
Number of Arg residues	12	10	11
[Arg/(Arg+Lys)] ratio	0.55	0.50	0.55
Number of deamination sites per chain[a]	3	3	3
Total number of H-bonds (monomer)	5720 (449)	5500 (447)	5736 (460)
Volume of cavities[b] (Å^3)	0	211	155
Number of cavities	0	20	13
Relative surface area[c]	0.95	1.03	0.98
Optimized packing[d]	0.53	0.49	0.51
Number of salt bridges (intra/inter)[e]	140/52	88/16	80/16
Number of salt bridges per residue	0.094	0.050	0.046

[a] See Fig. 1a for details.
[b] Probe-accessible volume with probe radius 1.4 Å using program GRASP [32].
[c] Relative surface area $=A_o/A_c$ ($A_c=15.0\times N^{0.866}$, where N is the total number of protein atoms).
[d] Efficiency of packing is given by the fraction of atoms in a protein with zero accessible surface area.
[e] Ionic interactions between charged side-chains within 4 Å. Calculations are for the dodecamer unless stated otherwise.

tively (Table 5). The differences amongst them are not particularly significant. The deamidation of asparagine residues is known to be a determinant of heat inactivation of proteins [43]. The formation of the major product of deamidation, isoaspartate, frequently occurs at Asn–Gly, Asn–Ser or Asp–Gly sequences when they lie in regions of the polypeptide that are highly flexible [44]. TmV, EcV and HiV have three potential sites of deamidation, and these are located in strictly conserved regions of the sequence (Fig. 1a). Therefore, isoaspartate formation is not expected to be the major factor for heat inactivation of EcV and HiV.

An examination of the tertiary and quaternary structures shows that the total number of hydrogen bonds per protomer, as well as the dodecamer, does not follow expected trends (Table 5). Another important factor implicated for the protein stability is the size and number of cavities in the structure. We looked for the presence of cavities in the enzymes from the three sources using GRASP [32] and a probe radius of 1.4 Å. We found that there are no cavities with a significant probe-accessible volume in TmV, while the EcV and HiV structures have significant vacant volumes formed by a number of small cavities (Table 5). This suggests that TmV has a tightly packed core compared to EcV and HiV.

It is expected that minimization of the ratio of surface area of a protein to its enclosed volume should increase protein stability by simultaneously reducing unfavorable surface energy and increasing the packing interactions in the interior, although there are ambiguities about the latter as a general factor responsible for thermotolerance. This is supported by the fact that enzymes from hyperthermophiles have a relatively small solvent-exposed surface area [45]. The relative surface area is a good indicator of the extent of surface area minimization. The observed accessible surface area (A_o) of TmV is 60 721 Å^2 and the expected area (A_c) is calculated to be 63 657 Å^2 ($A_c=15.0\times N^{0.866}$, where N is the total number of protein atoms). The ratio A_o/A_c of 0.95 observed in TmV is relatively low when compared to many other protein molecules [45]. Another simple indicator for evaluating packing efficiency is the fraction of atoms in a protein with zero accessible surface area. For example, a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase (AOR) from *Pyrococcus furiosus* exhibits a value of ~0.55, which is significantly higher than average [45]. In TmV this value is 0.53,

approximately as high as in AOR. Hence, the relative surface area and packing efficiency in TmV follow expected trends for thermostable proteins (Table 5) and are expected to contribute towards the temperature stability of TmV.

Salt bridges are known to be important determinants of thermostability in proteins [45–47]. For example, in AOR the number of 'ion-pairs per residue' was found to be more than two-fold the average for other proteins (~0.085 vs. ~0.040) [45]. The factor 'ion-pairs per residue' was defined as the difference between the numbers of attractive and repulsive ionic interactions that occur between polar atoms of charged side chains within 4 Å, divided by the total number of residues [45]. Moreover, the distribution of residues involved in ionic interactions in hyperthermophilic enzymes exhibit differences compared to those from mesophilic sources. That is, hyperthermophilic enzymes were found to exhibit extensive salt-bridge networks or ion pair clusters [48]. There are a total of 192 salt bridges in the TmV dodecameric structure (Table 5). It exhibits extensive intra- and inter-subunit ionic interactions and salt bridge networks when compared to EcV and HiV, where the total number of salt bridges is 104 and 96, respectively. The number of 'ion-pairs per residue' in TmV is 0.094, a value which is significantly higher than the average value (~0.040) observed in other proteins and even higher than that observed in AOR (~0.085). Arg138 is noteworthy, as it is involved in both intra- and inter-subunit interactions (intra, Arg138···Asp117···Lys134; inter, Arg138···Glu157). EcV and HiV do not exhibit these interactions. Another noteworthy feature, amongst other ionic interactions, is that a salt bridge, Asp81···Lys33, occurring in EcV and HiV is extended into a network in TmV, consisting of interactions involving Asp85···Arg59···Asp81···Lys63···Glu66. Similarly, Arg36···Glu169 in EcV and HiV is extended into Lys36···Glu169···Arg32 in TmV.

In summary, the tighter packing, larger fraction of buried atoms, minimized surface area and increased ionic interactions, amongst other factors, seem to provide a good explanation for the thermostability of TmV as compared to its mesophilic counterparts. Among these factors, the large number of intra- and inter-subunit ionic interactions/salt-bridge networks observed in TmV appears to play a major role.

Acknowledgments

The authors would like to thank H. Bartunik and G. Bourenkov for help with data collection; S. Körner for mass spectroscopic analysis; P. Zwickl for discussion on the NTPase activity; J. Kaiser for gift of *T. maritima* genomic DNA; M. Harding, University of Edinburgh, for discussions on geometry of metal ion–protein interactions; and A. Higashitani (National Institute of Genetics, Japan) for the gift of MBP-SulA fusion protein clone. R.R. and K.A are recipients of fellowships from The Alexander von Humboldt Foundation.

References

[1] M. Bochtler, L. Ditzel, M. Groll, R. Huber, Crystal structure of heat shock locus V (HslV) from *Escherichia coli*, Proc. Natl. Acad. Sci. USA 94 (1997) 6070–6074.
[2] M. Bochtler, L. Ditzel, M. Groll, C. Hartmann, R. Huber, The proteasome, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 295–317.
[3] M. Rohrwild, G. Pfeifer, U. Santarius, et al., HslV-HslU: a novel ATP-dependent protease complex in *Escherichia coli* related to the eukaryotic proteasome, Proc. Natl. Acad. Sci. USA 93 (1996) 5808–5813.
[4] S.J. Yoo, J.H. Soel, D.H. Shin, et al., Purification and characterization of the heat shock proteins HslV and HslU that form a new ATP-dependent protease in *Escherichia coli*, J. Biol. Chem. 271 (1996) 14035–14040.
[5] S.E. Chuang, V. Burland, G. Plunkett, D.L. Daniels, F.R. Blattner, Sequence analysis of four new heat-shock genes constituting the hslTS/ibpAB and hslVU operons in *Escherichia coli*, Gene 134 (1993) 1–6.
[6] J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister, R. Huber, Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution, Science 268 (1995) 533–539.
[7] M. Groll, L. Ditzel, J. Lowe, et al., Structure of 20S proteasome from yeast at 2.4 Å resolution, Nature 386 (1997) 463–471.
[8] M. Groll, W. Heinemeyer, S. Jager, et al., The catalytic sites of 20S proteasomes and their role in subunit maturation: a mutational and crystallographic study, Proc. Natl. Acad. Sci USA 96 (1999) 10976–10983.
[9] E.C. Schirmer, J.R. Glover, M.A. Singer, S. Lindquist, HSP100/Clp proteins: a common mechanism explains

diverse functions, Trends Biochem. Sci. 21 (1996) 289–296.
[10] M. Bochtler, C. Hartmann, H.K. Song, G.P. Bourenkov, H.D. Bartunik, R. Huber, The structures of HsIU and the ATP-dependent protease HsIU-HsIV, Nature 403 (2000) 800–805.
[11] M.H. Glickman, D.M. Rubin, O. Coux, et al., A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3, Cell 94 (1998) 615–623.
[12] R. Ravishankar, C. Hartmann, H.K. Song, R. Huber, M. Bochtler, Functional interactions of HslV (ClpQ) with the ATPase HslU (ClpY), Proc. Natl. Acad. Sci. USA 99 (2002) 7396–7401.
[13] M.C. Sousa, C.B. Trame, H. Tsuruta, S.M. Wilbanks, V.S. Reddy, D.B. McKay, Crystal and solution structures of an HslUV protease–chaperone complex, Cell 103 (2000) 633–643.
[14] J. Wang, J.J. Song, M.C. Franklin, et al., Crystal structures of the HslVU peptidase-ATPase complex reveal an ATP-dependent proteolysis mechanism, Structure 9 (2001) 177–184.
[15] H.K. Song, C. Hartmann, R. Ramachandran, et al., Mutational studies on HslU and its docking mode with HslV, Proc. Natl. Acad. Sci. U S A 97 (2000) 14103–14108.
[16] M.C. Sousa, D.B. McKay, Structure of *Haemophilus influenzae* HslV protein at 1.9 Å resolution, revealing a cation-binding site near the catalytic site, Acta Crystallogr. D 57 (2001) 1950–1954.
[17] M.S. Kang, B.K. Lim, I.S. Seong, et al., The ATP-dependent CodWX (HslVU) protease in *Bacillus subtilis* is an N-terminal serine protease, EMBO J. 20 (2001) 734–742.
[18] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding, Anal. Biochem. 72 (1976) 248–254.
[19] P.A. Lanzetta, L.J. Alvarez, P.S. Reinach, O.A. Candia, An improved assay for nanomole amounts of inorganic phosphate, Anal. Biochem. 100 (1979) 95–97.
[20] Y.V. Matsuka, S. Pillai, S. Gubba, J.M. Musser, S.B. Olmsted, Fibrinogen cleavage by the *Streptococcus pyrogenes* extracellular cysteine protease and generation of antibodies that inhibit enzyme proteolytic activity, Infect. Immun. 67 (1999) 4326–4333.
[21] S.S. Twining, Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes, Anal. Biochem. 143 (1984) 30–34.
[22] A. Higashitani, Y. Ishii, Y. Kato, K. Horiuchi, Functional dissection of a cell-division inhibitor, SulA, of *Escherichia coli* and its negative regulation by Lon, Mol. Gen. Genet. 254 (1997) 351–357.
[23] U. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[24] W. Minor, D. Tomchick, Z. Otwinowski, Strategies for macromolecular synchrotron crystallography, Structure 8 (2000) R105–110.
[25] J. Navaza, An automated package for molecular replacement, Acta Crystallogr. A 50 (1994) 157–163.
[26] T.A. Jones, J.-Y. Zou, S.W. Cowan, M. Kjeldgaard, Improved methods for binding protein models in electron density maps and the location of errors in these models, Acta Crystallogr. A 47 (1991) 110–119.
[27] A.T. Brünger, P.D. Adams, G.M. Clore, et al., Crystallography & NMR system: a new software suite for macromolecular structure determination, Acta Crystallogr. D 54 (1998) 905–921.
[28] R. Engh, R. Huber, Accurate bond and angle parameters for X-ray protein structure refinement, Acta Crystallogr. A 47 (1991) 392–400.
[29] R. Laskowski, M. MacArthur, E. Hutchinson, J. Thorton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[30] Collaborative Computational Project No 4, The CCP4 suite: programs for protein crystallography, Acta Crystallogr. D 50 (1994) 760–763.
[31] J.S. Sack, CHAIN: a crystallographic modeling program, J. Mol. Graph. 6 (1988) 244–245.
[32] A. Nicholls, R. Bharadwaj, B. Honig, GRASP—graphical representation and analysis of surface properties, Biophys. J. 64 (1993) A166.
[33] P. Zwickl, J. Kleinz, W. Baumeister, Critical elements in proteasome assembly, Nat. Struct. Biol. 1 (1994) 765–770.
[34] K.E. Nelson, R.A. Clayton, S.R. Gill, et al., Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*, Nature 399 (1999) 323–329.
[35] P. Zwickl, D. Ng, K.M. Woo, H.P. Klenk, A.L. Goldberg, An archaebacterial ATPase, homologous to ATPases in the eukaryotic 26S proteasome, activates protein breakdown by 20S proteasomes, J. Biol. Chem. 274 (1999) 26008–26014.
[36] H.R. Wilson, M.S. Ou, H.C. Aldrich, J. Maupin-Furlow, Biochemical and physical properties of the *Methanococcus jannaschii* 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome, J. Bacteriol. 182 (2000) 1680–1692.
[37] I.S. Seong, J.Y. Oh, S.J. Yoo, J.H. Seol, C.H. Chung, ATP-dependent degradation of SulA, a cell division inhibitor, by the HslVU protease in *Escherichia coli*, FEBS Lett. 456 (1999) 211–214.
[38] J.P. Glusker, Structural aspects of metal liganding to functional groups in proteins, Adv. Protein Chem. 42 (1991) 1–76.
[39] M.M. Harding, Geometry of metal–ligand interactions in proteins, Acta Crystallogr. D 57 (2001) 401–411.
[40] A. Karshikoff, R. Ladenstein, Ion pairs and the thermotolerance of proteins from hyperthermophiles: a 'traf-

fic rule' for hot roads, Trends Biochem. Sci. 26 (2001) 550–556.
[41] R.J. Russell, D.W. Hough, M.J. Danson, G.L. Taylor, The crystal structure of citrate synthase from the thermophilic archaeon, *Thermoplasma acidophilum*, Structure 2 (1994) 1157–1167.
[42] N.T. Mrabet, A. Van den Broeck, I. Van den Brande, et al., Arginine residues as stabilizing elements in proteins, Biochemistry 31 (1992) 2239–2253.
[43] T.J. Ahern, A.M. Klibanov, The mechanisms of irreversible enzyme inactivation at 100 °C, Science 228 (1985) 1280–1284.
[44] D.W. Aswad, M.V. Paranandi, B.T. Schurter, Isoaspartate in peptides and proteins: formation, significance, and analysis, J. Pharm. Biomed. Anal. 21 (2000) 1129–1136.
[45] M.K. Chan, S. Mukund, A. Kletzin, M.W. Adams, D.C. Rees, Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase, Science 267 (1995) 1463–1469.
[46] K.Y. Hwang, H.K. Song, C. Chang, et al., Crystal structure of thermostable α-amylase from *Bacillus licheniformis* refined at 1.7 Å resolution, Mol. Cells 7 (1997) 251–258.
[47] M. Machius, G. Wiegand, R. Huber, Crystal structure of calcium-depleted *Bacillus licheniformis* α-amylase at 2.2 Å resolution, J. Mol. Biol. 246 (1995) 545–559.
[48] K.S. Yip, K.L. Britton, T.J. Stillman, et al., Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family, Eur. J. Biochem. 255 (1998) 336–346.
[49] G.J. Barton, ALSCRIPT: a tool to format multiple sequence alignments, Protein Eng. 6 (1993) 37–40.
[50] M.M. Harding, Metal-ligand geometry relevant to proteins: sodium and potassium, Acta Cryst. D Biol. Crystallogr. 58 (2002) 872–874.