Structure of the 20S proteasome from bovine liver at 2.75 Å resolution

A Doctoral Thesis
by
Masaki Unno

Submitted to the Faculty
of Science, Osaka University
Japan

February, 2002

ACKNOWLEDGEMENT

The present work has been carried out under the direction of Professor Tomitake Tsukihara of Institute for Protein Research, Osaka University. I would like to thank him for his incessant guidance and encouragement throughout this work. I am deeply indebted to Dr. Keiji Tanaka of The Tokyo Metropolitan Institute of Medical Science (Rinshoken) for his invaluable advice and discussion. I am also grateful to Dr. Tsunehiro Mizushima, Rinshoken, for his technical advice and helpful suggestion. I am deeply indebted to Professor Noritake Yasuoka and Associate Professor Yukio Morimoto, Himeji Institute of Technology, for his incessant encouragement and helpful advice to X-ray crystallography.

I am also grateful to Associate Professor Atsushi Nakagawa of Institute for Protein Research, Osaka University for his valuable suggestion. I also wish sincere thanks to Associate Professor Yoshiki Matsuura and Associate Professor Masami Kusunoki of Institute for Protein Research, Osaka University.

I am deeply grateful to Dr. Nobuo Kamiya and Dr. Masahide Kawamoto at SPring-8 (BL41XU), Japan Synchrotron Radiation Research Institute (JASRI) for their help in data collection.

I am deeply grateful to Professor Hiroki Nakagawa of Chiba University for his useful advice for purification of the proteasome. I also express my thanks to Mr. Yoshikazu Tomisugi of Kumamoto University and Mr. Miyuki Miyawaki of Chiba University for their valuable help for this work.

I am deeply grateful to Dr. Hiroaki Sakai of The Institute of Physical and Chemical Research (RIKEN). I also express thanks to Dr. Eiki Yamashita and Dr. Genji Kurisu of Institute for Protein Research, Osaka University for their useful help and discussion in X-ray structure study.

I express my thanks to all members in Tsukihara's laboratory for kind assistances and encouragement. I also express my thanks to all of my friends for their warm encouragement.

Finally, I thank my parents, my sister, my brother and my dead grandmother for their incessant understanding and encouragement.

海野昌喜

Masaki Unno

February, 2002

Table of contents

Chapter I. General introduction

I-1. Proteolysis and multisubunits proteases

Misfolded and malfunctioning proteins, prone to aggregation, must be degraded and removed to keep cellular structure. Moreover, protein degradation has a key regulatory role in many cellular pathways and therefore must be subject to spatial and temporal control.

A major way to control proteolysis is compartmentalization. For example, the lysosome is a specialized membrane-bounded compartment where control is exercised by vesicle sorting. Once internalized, substrate is degraded by relatively non-selective bulk process.

On the other way, several multisubunit proteases such as 20S proteasome have a common barrel-shaped architecture by self-compartmentalization (Lupas A. *et al.* 1997.). The architecture delimits an inner compartment, which is several nanometers in diameter and harbors the active sites. These proteases must work with other proteins that can recognize, bind and unfold target proteins that should be degraded, because only unfolded proteins can access to the active site of these multisubunit proteases. Therefore, accessory components or regulatory complexes need to bind catalytic components. These regulatory complexes contain ATPase subunits that render protein degradation energy dependent (Larsen CN. *et al.* 1997.). Eukaryotic 26S proteasome has PA700 (or 19S cap) as a regulatory complex. The PA700 contains ATPase of "AAA" (ATPase associated with a variety

of cellular activities) family.

I-2. The ubiquitin-proteasome system

The selective degradation of many short-lived proteins in eukaryotic cells is carried out by the ubiquitin-proteasome system (Fig.I-1). In this system, proteins are targeted for degradation by covalent ligation to ubiquitin, a 76-amino-acid-residue protein. Ubiquitin is joined reversibly to proteins by covalent linkage between the carboxyl-terminus of ubiquitin and lysine ε-amino groups of the acceptor proteins. Ubiquitinated proteins are in a dynamic state, subject to either further rounds of ubiquitin addition, ubiquitin removal by the 26S proteasome, which break down targeted substrates to short peptides. The degradation of polyubiquitinated proteins requires ATP hydrolysis. The core and proteolytic chamber of the 26S proteasome is formed by the 20S proteasome. Short peptides formed by the above processes can be further degraded to free amino acids by cytosolic peptidase (Hershko A. *et al.* 1998.).

The life span of many cellular proteins seems to be controlled by this proteolytic system. The proteolytic system is involved in many biological processes, including the removal of abnormal, misfolded or improperly assembled proteins, stress response (by processing or degrading transcriptional regulators), cell-cycle control (by destruction of cyclins and CDK inhibitors), metabolic adaptation (by destruction of transcription factors or metabolic enzymes), and

the cellular immune response (by generating antigenic peptides presented by major histcompatibility complex (MHC) class I molecules).

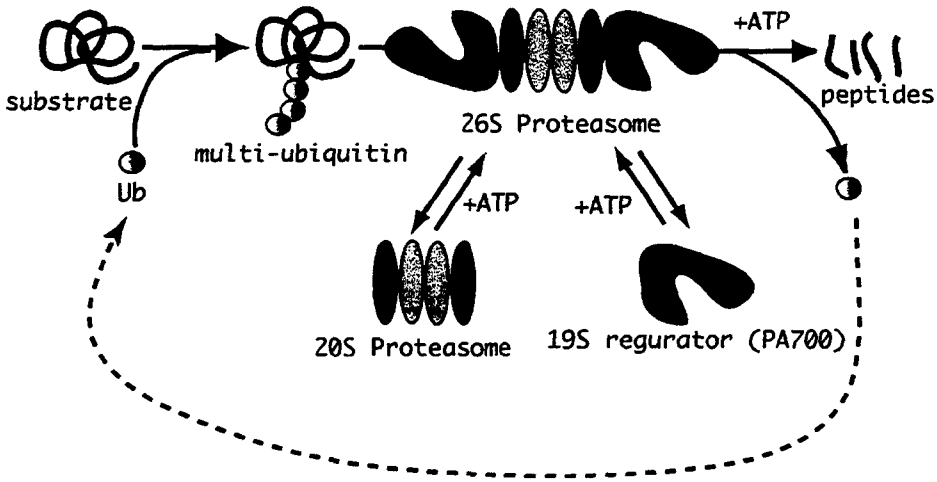


Fig. I-1 Ubiquitin-proteasome system.

I-3. The proteasome

The proteasomes are unusually large multisubunit proteolytic complexes, and that are widely recognized as the central enzymes of nonlysosomal protein degradation in both the cytsol and nucleus. They are involved in the degradation of misfolded proteins (Peters JM. 1994.) as well as in the degradation and processing of short-lived regulatory proteins (Palombella VJ. *et al.* 1994.).

I-3-1. The 20S proteasome

Four seven-subunits-rings stack into a cylinder to form the 20S proteasome. Archaeon proteasomes consists of 14 copies each of 2 distinct

but related polypeptides, α and β. On the other hand, eukaryotic proteasomes are built of 2 copies of each of 7 distinct α- and β-type subunits. In mammalian cells, additional three active β-type subunits are induced by γ-interferon. These additional subunits replace specific and closely related active subunits which are expressed constitutionally. The 20S proteasome of which three subunits are replaced is called "immunoproteasome" (Fig. I-2).

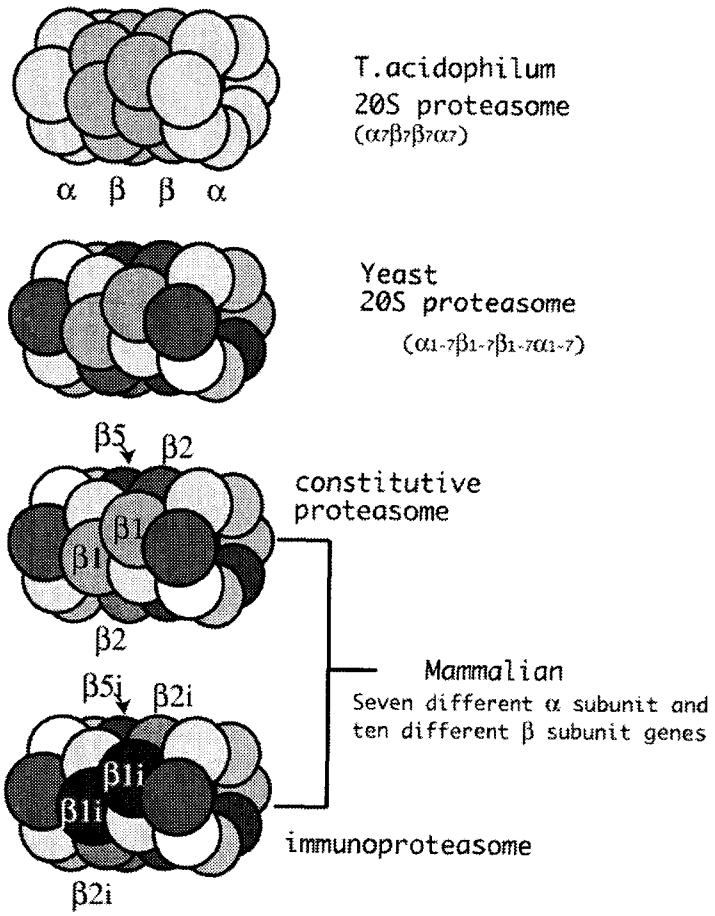


Fig.I-2 Caricatures of the 20S proteasomes

I-3-2. The immunoproteasome

Despite the diversity in numbers of different subunits, the molecular shape of 20S proteasome is strongly conserved from *T.acidophilum* to human enzymes (Püler G. *et al.*1994.). Additional three β subunits, β1i/LMP2, β2i/MECL1, and β5i/LMP7 are induced by the cytokine γ-interferon (Hisamatsu H. *et al.* 1996, Nandi D. *et al.* 1996, Groettrup M. *et al.* 1996, Früh K. *et al.* 1992, Akiyama K. *et al.* 1994, van Kaer LV. *et al.* 1994, Fehling HJ. *et al.* 1994, Cerundolo V. *et al.* 1995.) in mammalian proteasomes. All three inducible subunits of β1i, β2i and β5i are catalytically active and they replace constitutive active subunits β1, β2 and β5, respectively. β1i and β5i are encoded by genes in the MHC class II gene region adjacent to the transporters associated with antigen presentation (TAP) genes, which are involved transport of antigenic peptides from the cytosol to the endoplasmic reticulum (Ortiz-Navarrete V. *et al.* 1991, Glynne R. *et al.* 1991, Martinez CK. *et al.* 1991, Kelly A. *et al.* 1991.). The 20S proteasome that contains inducible three subunits, β1i, β2i and β5i, is mentioned as immunoproteasome and enhances proteasomal generation of MHC class I-binding peptides (Fehling HJ. *et al.* 1994, Pamer E. *et al.* 1998, Tanaka K. *et al.* 1998.).

I-3-3 The 26S proteasome

The 26S proteasome is a eukaryotic ATP-dependent protease that was found as a large multisubunit complex with a molar mass of approximately 2 MDa. It consists of a central

cylindrical 20S proteasome as a catalytic machine, and two large V-shaped regulatory complexes named PA700 which contain some ATPase-subunits. Thus, the shape of the 26S proteasome looks like dumbbell (Baumeister W. *et al.* 1997, Fig.I-3 (a)). PA700 is assembled from approximately 17 different polypeptides, ranging in molecular weight from 15 to 112 kDa, and possibly varying in composition depending on cell type and cell state (Ferrell K. *et al.* 1997, Rubin DM. *et al.* 1996.).

I-3-4 The football-like proteasome

In 1992, another activator protein of 20S proteasome was found and named PA28 (or 11S regulator). Degradation of small peptides by the proteasome is greatly stimulated by the purified PA28 protein in ATP-independent manner (Ma CP. *et al.* 1992, Dubiel W. *et al.* 1992.). Electron microscopic investigations indicate that PA28 is a ring-shaped particle, and that it associates with 20S proteasome ATP-independently by forming a football-like structure (Gray CW. *et al.* 1994, Fig.I-3 (b)). PA28 is a cytoplasmic complex formed by equal amounts of two different, but related 28 kDa subunits, which were named PA28α and PA28β (Ahn JY. *et al.* 1995.). PA28α and PA28β are suggested to form hetero-hexisamer or –heptamer complex with alternating α and β subunits.

In vitro studies showed that PA28 can enhance generation of antigenic peptides by inducing dual substrates cleavages by the 20S proteasome (Dick TP. *et al.* 1996.). A cytokine, γ-interferon, induces expression of both PA28 subunits. Genes for

PA28α and PA28β are found only in organisms which adaptive immune systems (Kohda K. *et al.* 1998.).

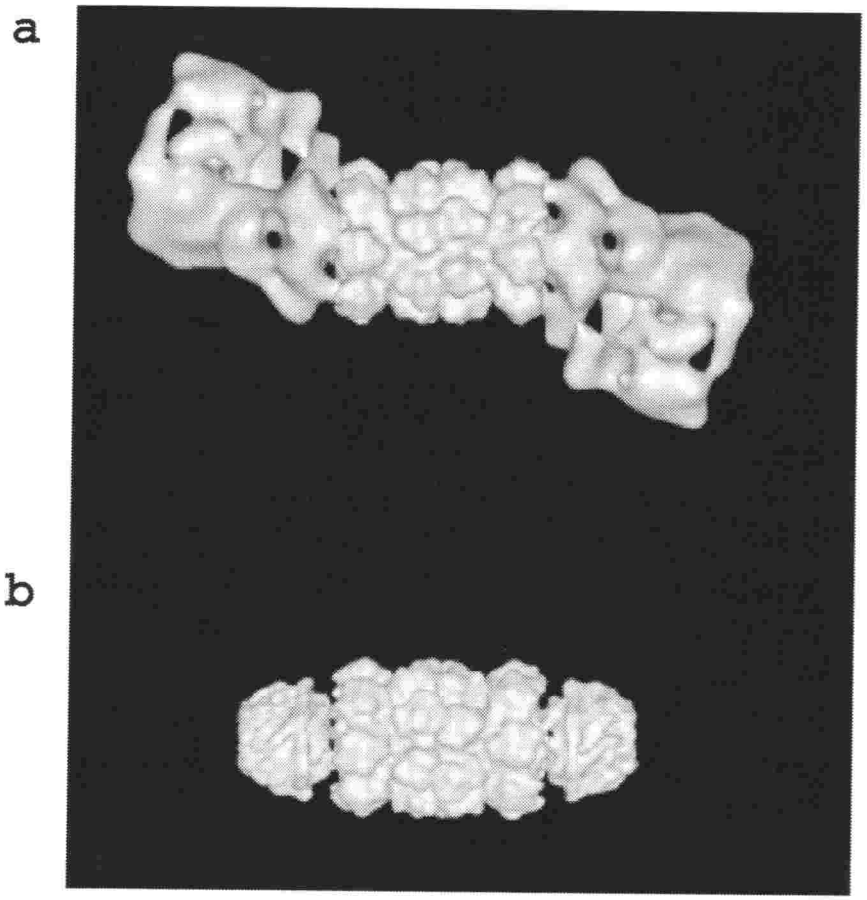


Fig.I-3 Composite models of the 26S proteasome and the football-like proteasome based on electron microscopy.
a. The 26S proteasome. b. The football-like proteasome.

I-3-5. The hybrid proteasome

It was found that PA28 and PA700 bind 20S

proteasome simultaneously (Hendil K. *et al.* 1998.). The complex was named "hybrid proteasome" (Tanahashi N. *et al.* 2000.). The hybrid proteasome seems to act more efficient proteolysis of some substrates; perhaps intact the cavity of the 20S proteasome, whose cleavage ability is modified by the PA28 complex. The hybrid proteasome is induced by γ-interferon, which induces the expression of PA28α and PA28β in a variety of cells (Ahn JY. *et al.* 1995, Tanahashi N. *et al.* 1997, Früh K. *et al.* 1999.). Therefore, the immunoproteasome and hybrid proteasome, both of which are induced by γ-interferon, may cooperate for the efficient production of MHC class I peptides in the soluble fraction of the cells in the response to the cell-mediated adaptive immunity.

I-4. The immune system

Our immune system saves us from certain death by infection. Any vertebrate born with a severely defective immune system will soon die unless extraordinary measures are taken to isolate it from a host of infectious agents - bacterial, viral, fungal, and parasitic. As part of their immune system, vertebrate cells display intracellular antigens at the cell surface to distinguish between infected and uninfected cells. This task falls to the MHC class I, which bond antigenic peptides of defined length (usually 8 – 10 amino acid residue long) and sequence. The peptide-MHC pair can then be recognized by specific T-cell receptors on cytotoxic T cells. The cytotoxic T cells can kill

other cells. They are important in host defense against most viral pathogens. One of proteasome's many jobs is to generate antigenic peptides.

Proteasomal protein substrate processing results in the generation of peptides of 8 – 11 residues in length, which are transported into the endoplasmic reticulum (ER) by the transporter associated with antigen presentation (TAP) complex. Within the ER, peptides bind to and stabilize the MHC class I heterodimers. Only fully assembled, peptide-loaded MHC molecules are translocated, through the Golgi apparatus, to the cell surface. By binding specifically to a given MHC allomorph – an MHC molecule encoded by a specific haplotype – loaded with a unique peptide, cytotoxic T cells with complementary T-cell receptors (TCRs) are stimulated to proliferate and destroy the infected target cell (Fig. I-4).

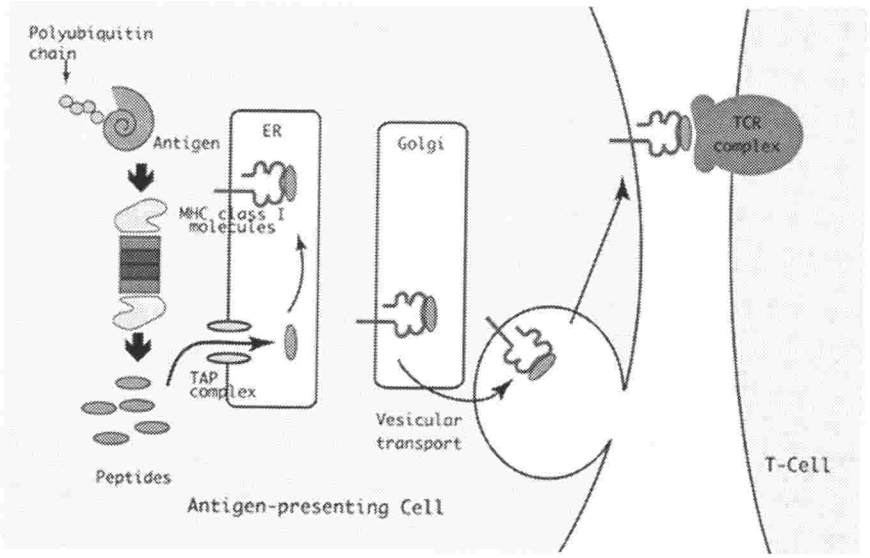


Fig. I-4 Antigen processing and presentation

I-5. Purpose of the present research

As mentioned above, the mammalian 20S proteasome is a catalytic core of the 26S proteasome, the football-like proteasome and the hybrid proteasome. The football-like proteasome and the hybrid proteasome are found only in mammalian cells. These proteasomes are key enzymes involved in many biological processes and are exceptionally huge proteases. The crystal structures of the 20S proteasome from archaebacterium *Thermoplasma acidophilum* and yeast *Saccharomyces cerevisiae* have been solved in 1995 and 1997, respectively (Löwe J. *et al.* 1995, Groll M. *et al.* 1997.). In higher eukaryotes, proteasomes are known to act as antigen-processing enzymes responsible for the generation of peptide ligands presented by MHC class I molecules (Tanaka K. *et al.* 1998.). Moreover, additional three subunits not found in the yeast are induced by the cytokine γ-interferon. Thus, we can know that mammalian proteasome is very important in immune response. In addition, in mammalian proteasomes, at least five different proteolytic activities have been detected (Orlowski M. *et al.* 1993.). However, there is no information on the tertiary structure of the mammalian 20S proteasome. To clarify their immunological role and the mechanism underlying the interaction with PA28 that does not present in the yeast and to understand catalytic systems different from the yeast, we have worked on analyzing the crystal structure of the mammalian 20S proteasome.

The 20S proteasome has a characteristic quaternary structure. The quaternary structure of

the *T.acidophilum* proteasome is similar to that of viruses. Many viruses are often made up with units which may have either icosahedral or helical symmetry. Forming these symmetrical structures are simple because of small kinds of subunits. Eukaryotic proteasomes are much more complex, having an (α1-α7, β1-β7, β1-β7, α1-α7) complex in four stacked rings and occupying unique locations. We are interested in this unique structure and tried revealing the reason why the 20S proteasomes in eukaryotes are built up correctly.

Chapter II. Purification and Crystallization of the 20S proteasome from bovine liver and brain

II-1. Introduction

At first, we purified and crystallized the 20S proteasome from bovine liver. In order to obtain a better yield of proteasome and produce X-ray grade crystals, protein purification was carried out using several columns. However, we found two types of proteasomes – constitutive forms and immunoproteasome – in the solution purified from bovine liver (data not shown). Thus, we tried changing the organs from which we purified the 20S proteasome from the liver to brain, in which the immunoproteasome does not exist. However, the crystals that were purified from bovine brain reflected X-ray, poorly (~ 4 Å). They had large unit cells and it was difficult to solve the structures (Chapter IX). Finally the crystal from bovine liver diffracting up to 2.75 Å was used for the structural analysis.

II-2. Purification of the 20S proteasome

The 20S proteasome is isolated from the bovine liver and brain in almost the same procedure. The method is described previously for their isolation from rat liver (Tanaka K. *et al.* 1988.). The protocol is summarized in Fig.II-1.

For assay of hydrolysis of the fluorogenic substrate Suc-Leu-Leu-Val-Tyr-MAC, it is incubated with the enzyme for 10 min at 37 °C in the presence of 0.05 % SDS in 100 mM Tris-HCl (pH 8.0). The

reaction was stopped by adding 100 µl, 10 % SDS and 2ml, 0.1 M Tris-HCl (pH 9.0), then the fluorescence of the reaction products was measured.

One kg of bovine liver was homogenized with the buffer solution containing 25 mM Tris-HCl , pH 7.5, 250 mM sucrose and 2-mercaptan ethanol at 4 °C in a cold room using a commercially available juicer-blender. The solution was centrifuged for 30 minutes at 6,800 × g and the resulting supernatant was centrifuged for 1 hour at 700,000 × g. The supernatant was purified by sequential chromatography on Fast Flow Q-Sepharose (Pharmacia, 50 × 500 mm) , Bio-Gel A-1.5m (Bio-Rad, 50 ×900 mm), hydroxyapatite HTP (Bio-Rad, 22 × 250 mm) and heparin-Sepharose CL-6B (Pharmacia, 22 × 250 mm). The 20S proteasomes exist in both the absorbed materials and non-absorbed materials for the Fast Flow Q-Sepharose. Although the first absorbed band contains much 20S proteasome, it contains much contaminant. Thus we removed the first absorbed materials. The non-absorbed materials were absorbed into Q-Sepharose chromatography again. The absorbed materials in the second Q-Sepharose chromatography were used to sequential steps. The standard buffer containing 25 mM Tris-HCl, pH 7.5, 20 % (v/v) glycerol and 10mM 2-mercaptan ethanol (buffer A) was used for the Q-Sepharose ion exchanger, the gel filtration and the affinity chromatography. Buffer B used for the hydroxyapatite chromatography contained 10 mM potassium-phosphate buffer, pH 7.0, 20 % (v/v) glycerol and 10 mM 2-mercaptan ethanol. The final enzyme solution from the chromatography steps was

dialyzed against buffer solution containing 25 mM Tris-HCl, pH 7.5, 1 mM dithiothreitol (DTT) (buffer C) and concentrated by ultracentrifugation for 15 hours at 165,000 × g. The resulting precipitated materials were dissolved in a small amount of buffer solution C and concentrated to 35 mg/ml by ultrafiltration (Morimoto Y. *et al.* 1995.) with CENTRICON. From 1kg of the bovine liver (brain), about 8 mg (4 mg) of the 20S proteasomes were ultimately isolated for crystallization in this procedure.

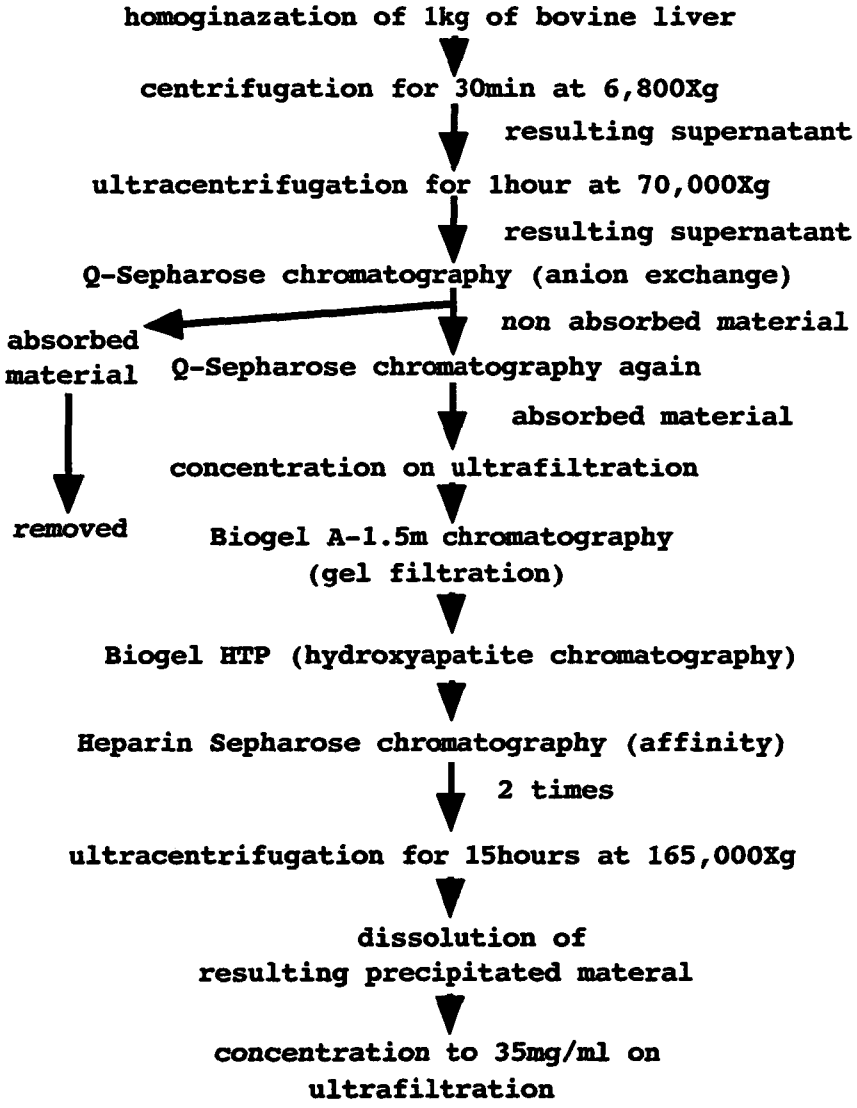


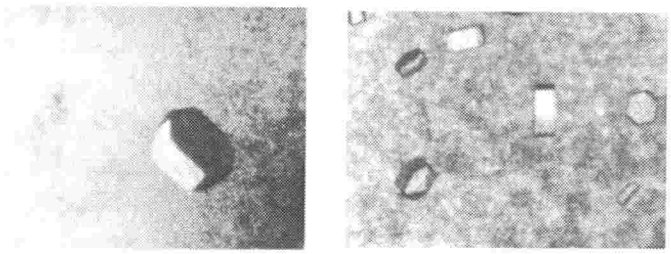
Fig.II-1 A scheme of purification procedure of the 20S proteasome.

II-3. Crystallization

Crystallization was performed by the hanging-drop vapor diffusion technique at 298 K (Morimoto Y. *et al.* 1995, Tomisugi Y. *et al.* 2000.) by mixing 3 μl of enzyme solution (35 mg/ml in 25 mM Tris-HCl, pH 7.5, 1mM DTT) and 3 μl of reservoir solutions. Crystals appeared within 2 days and grew to several sizes (Fig.II-2).

The best condition of the reservoir solutions in which the good crystals for x-ray experiments were obtained was different in the L.O.T of preparations. The pH of Na Cacodylate buffer was 6.5 ~ 7.0, the concentration of Na Cacodylate was 0.1 M ~ 0.15 M and the precipitant was 33 % ~ 37 % of 2-Methyl-2,4-pentanediol (MPD). The salt for crystallization was 0.15 M Mg acetate. The good crystals were obtained under the condition from the different combination.

Most of the crystals of the 20S proteasome disappeared or collapsed for 5 ~ 10 days. Moreover, drops in which the crystals of the 20S proteasome appeared always contained precipitate (Fig.II-2).



0.5 mm

Fig.II-2 Crystals of the bovine liver 20S proteasome

Chapter III. X-ray experiments

III-1. Introduction

To analyze protein structures at atomic resolution, we need good crystals that diffract x-rays at high resolution. However, most crystals obtained do not always diffracted x-rays up to an atomic resolution. Especially such a large macromolecular complex as the 20S proteasome, we have known that most of crystals of diffract X-rays poorly even if they are large and have beautiful shapes.

To reduce radiation damage to crystals during X-ray diffraction experiments, we used the cryo-technique in collecting intensity data of the crystals of the 20S proteasomes. This technique is not only to prevent radiation damages but also to allow easier treatment of crystals than use of the glass or quartz capillary tubes for room-temperature data-collection. Moreover, we are usually able to collect a whole data set from single flash-cooled crystals owing to the prolonged lifetime. Thus, systematic errors can be minimized. It is a great advantage they can be frozen and stored in liquid nitrogen until when we should use them even if the crystals were exposed once.

III-2. Freezing the crystal of the 20S proteasome

Protein crystals are usually soaked in a solutions containing cryo-protectant to prevent the crystals from having cracks upon freezing. The reservoir solution that crystallized the 20S proteasomes contains high MPD content (33-37 %).

MPD can be an antifreeze agent and the concentration is sufficient as a cryo-solution. Thus, we did not have to prepare other cryo-solutions.

The crystals of the 20S proteasomes were picked up by nylon loops and were immediately plunged into liquid nitrogen. The viscosity of reservoir solution was very high because of not only high MPD content but also precipitate around the crystals (Fig.II-2). To make matters worse, most of crystals were very small and number of crystals was large. Moreover, the crystals were easy to be deteriorated when cover glasses of the hanging-drop vapor diffusion method were opened. These factors made us difficult to pick up a crystal from a drop for X-ray experiments. We tried plunging the crystals into both liquid nitrogen and gaseous nitrogen. However, in our case, it was difficult to plunge the crystals into gaseous nitrogen. As mentioned, the crystals were so weak against the physical and chemical shocks and the lifetime of the crystals of the 20S proteasome was very short. When we tried freezing the crystals in gaseous nitrogen, we could pick up a few crystals from a drop. It required the place and the special equipment to freeze in gaseous nitrogen. During freezing a crystal, another crystals in the same drop were deteriorated. On the other hand, when we plunged the crystals into liquid nitrogen, we could pick up many crystals from a drop. Freezing in liquid nitrogen is simple and we can close a cover glass of hanging drop vapor diffusion technique immediately after the crystals were plunged into the liquid nitrogen filled in a mug-cup on the desk.

The resolution limits of X-ray diffracted by the crystals frozen in liquid nitrogen and gaseous nitrogen had no differences. Thus, most of the crystals were directly plunged into liquid nitrogen.

III-3. Intensity data collection

We had serious problems in the X-ray experiments. They were the oscillation ranges, X-ray exposure time and the camera distance. The restricted area of the detector we used limited camera length. The large unit cell in a crystal limited oscillation angles. Our crystals were usually deteriorated by X-ray radiation. Thus, we tried a number of experiments and only a few data could be used as intensity data (Fig.III-1 (a), III-1 (b)). Although we thought that we would collect some partial data and combine them to make the whole one at first, the isomorphism between these crystals was very poor. The X-ray experimental condition for a data used structure determination is listed in Table III-1.

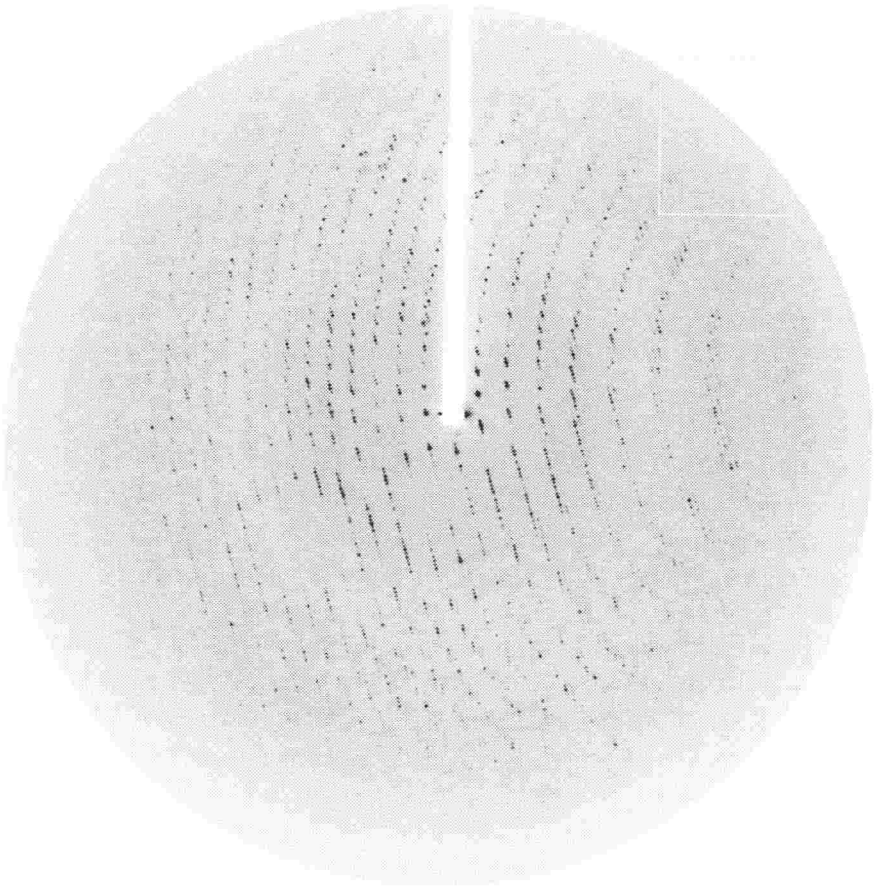


Fig.III-1 (a) X-ray diffraction pattern of the crystal of the 20S proteasome from bovine liver. An oscillation photograph taken at BL41XU of SPring-8, as described in the text.

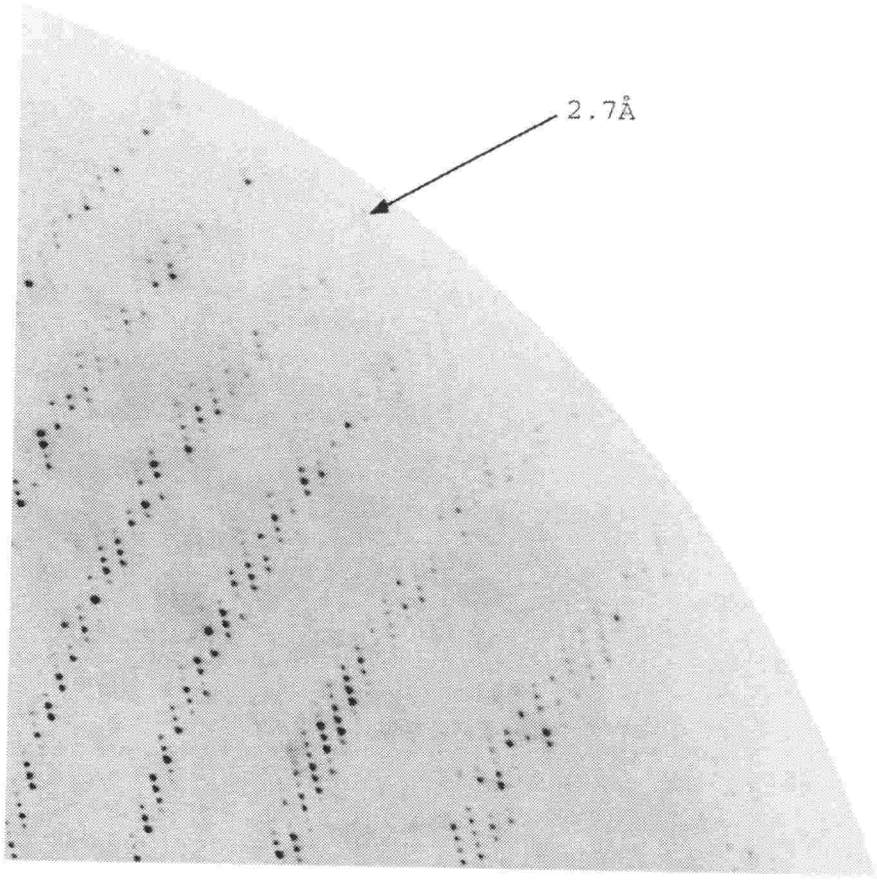


Fig.III-1 (b) Magnified x-ray diffraction pattern from the image in Fig.III-1 (a).

Table III-1 X-ray experimental condition

12th October 2000 at SPring-8 (Harima, JAPAN)
Beamline BL41XU
Detector Mar CCD 165
Wavelength 1.00 Å
Collimator 0.1 mm
Camera length 220 mm
Oscillation range 0.5 °
Exposure time 15 seconds
Temperature 100 K

III-4. Data processing and scaling

X-ray intensities were evaluated by using program MOSFLM (Leslie AGW. 1992.) and were scaled with the program SCALA of CCP4 (Collaborative Computational Project, Number 4, 1994.). Diffraction spots were found around 2.7 Å resolution but we used diffractions up to 2.75 Å resolution because of high R-value and low completeness. The statistics are listed in Table III-2.

Table III-2 Data collection statistics

Space group	$P2_12_12_1$
Cell constants	
a = 315.7 Å, b = 205.9 Å, c = 116.0 Å	
$\alpha = \beta = \gamma = 90°$	
Limiting resolution	2.75 Å (2.82-2.75 Å)
Measured reflections	733,250 (15,510)
Unique reflections	189,429 (10,087)
Completeness	96.3 %, (70.0 %)
$< I/\sigma >$	5.6 (1.8)
#R_{merge}	9.5 % (40.6 %)

#R_{merge} = $\sum|I - \langle I \rangle| / \sum \langle I \rangle$, where I is the observed intensity and <I> is the average intensity over symmetry equivalent measurements

Chapter IV. Structure determination

IV-I. Introduction

An easier way to obtain a preliminary model can be followed if the structure of a protein with a homologous amino acid sequence has already been established. The only successful technique has been termed by Rossmann the "molecular replacement method" (Rossmann MG. *et al.* 1972.). In essence it involves generating a preliminary model of the target crystal structure by orienting and positioning the search molecules within the unit cell of the target crystal so as to best account for the observed diffraction pattern. Phases can then be calculated from this model and combined with the observed amplitudes to give an approximate Fourier synthesis of the structure of interest. This in turn can be subject to any of the several forms of refinement to provide a final accurate structure of the target crystal.

If there is one molecule in the crystallographic asymmetric unit, then six parameters - three translations and three angles - fully describe how the search molecule is the unit cell. In principle one could simply search on these six variables to determine the model that gives best agreement between obtained structure factor (F_o) and calculated structure factor (F_c). In general such a search is beyond the capacity of even the fastest computers, although in special cases in which symmetry reduces the number of parameters such searches have been successful (Baldwin JM. 1980.). Rossmann and Blow realized that this

six-dimensional search could be reduced to a sequence of two three-dimensional searches (Rossmann MG. *et al.* 1962.) in which first the orientation and then the position of the search molecule are determined.

Actually, the eukaryotic 20S proteasome has a molecular 2-fold axis and a quasi 7-fold axis and they are perpendicular each other (Fig.IV-1). The sequences of the human and yeast 20S proteasome have high homologies. The identity ratios between corresponding subunits of the human and yeast 20S proteasomes are in the ranges from 46.2% to 57.1% for the constitutive α-subunits and from 42.5% to 63.7% for the constitutive β-subunits (e.g. Fig.IV-2). Mammals and the yeast constitute a multi-gene family and may have originated from a common ancestral gene. These subunits are listed in Table IV-1.

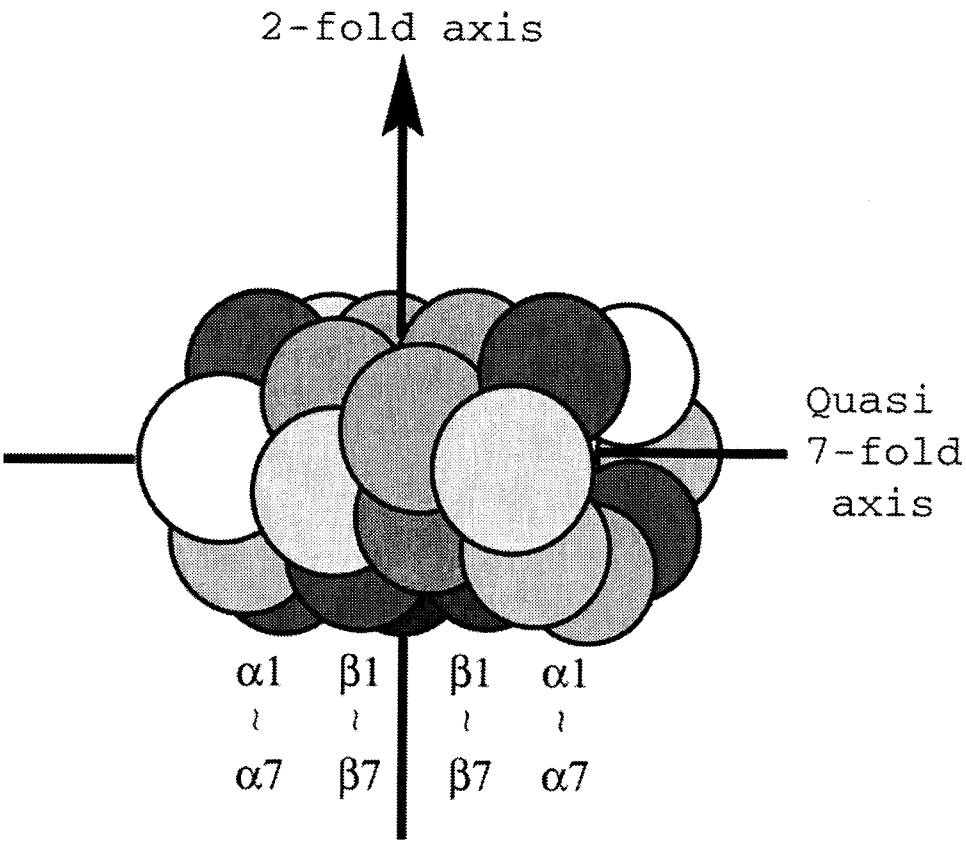


Fig.IV-1 Over all structure of eukaryotic 20S proteasome. Eukaryotic 20S proteasome has a 2-fold axis and a quasi 7-fold axis.

```
PSA6_HUMAN      MS---RGSSAGFDRHITIFSPEGRLYQVEYAFKAINQGGLTSVAVRGKDCAVIVTQKKVP
PSA6_YEAST      MSGAAAASAAGYDRHITIFSPEGRLYQVEYAFKATNQTNINSLAVRGKDCTVVISQKKVP
                **    .*:**:********************** ** .::*:*******:*:::*****
PSA6_HUMAN      DKLLDSSTVTHLFKITENIGCVMTGMTADSRSQVQRARYEAANWKYKYGYEIPVDMLCKR
PSA6_YEAST      DKLLDPTTVSYIFCISRTIGMVVNGPIPDARNAALRAKAEAAEFRYKYGYDMPCDVLAKR
                *****.:**:::* *:..** *:.*  .*:*. . **: ***:::*****::* *:*.**
PSA6_HUMAN      IADISQVYTQNAEMRPLGCCMILIGIDEEQGPQVYKCDPAGYYCGFKATAAGVKQTESTS
PSA6_YEAST      MANLSQIYTQRAYMRPLGVILTFVSVDEELGPSIYKTDPAGYYVGYKATATGPKQQEITT
                :*::**:***.* *****  : ::.:*** **.:** ****** *:****:* ** * *:
PSA6_HUMAN      FLEKKVKKK-----FDWTFEQTVETAITCLSTVLSIDFKPSEIEVGVVTVENPKFRILTE
PSA6_YEAST      NLENHFKKSKIDHINEESWEKVVEFAITHMIDALGTEFSKNDLEVGVATKD--KFFTLSA
                 **::.**.      : ::*:.** *** :  .*. :*. .::****.* :  **  *:
PSA6_HUMAN      AEIDAHLVALAERD
PSA6_YEAST      ENIEERLVAIAEQD
                 :*: :***:**:*
```

Fig.IV-2 A sequence alignment of α1 subunits of human (iota) and yeast (C2). We show this result as an example of high homologies among eukaryotic 20S proteasome subunits.

Table IV-1 Subunits of the 20S proteasome

Gene Name (Human)	M.W. (Da)	Others names		
			Mammal	Yeast
α-type subunits				
PSMA1	29555	α6	C2*	Pre5
	30239		Pros30*	
PSMA2	25898	α2	C3	Prs4/Y7
PSMA3	28433	α7	C8	Prs1/C1
PSMA4	29483	α3	C9	Prs5/Y13
PSMA5	26425	α5	zeta	Pup2/Doa5
PSMA6	27374	α1	iota/Pros27	Pus2/C7
PSMA7	27900	α4	C6-S/XPAC-7*	Pre6
			C6-L*	
β-type subunits				
PSMB1	26489	β6	C5	Prs3
PSMB2	22836	β4	C7	Pre1/C11
PSMB3	22931	β3	C10	Pup3
PSMB4	29192	β7	N3/beta	Pre4
PSMB5	22897	β5	X/MB1/epsilon/Lmp1 7	Pre2/Doa3
PSMB6	25315	β1	Y/delta/Lmp9	Pre3
PSMB7	29965	β2	Z/Mc14/Lmp19	Pup1
PSMB8	29769	β5i	Lmp-7-E1 /Ring10*	
	30354		Lmp7-E2*	
PSMB9	23245	β1i	Lmp2/Ring12	
PSMB10	28936	β2i	MECL-1 /Lmp10	

*The members of this pair of subunits are almost identical. Therefore, their mRNA may have arisen by alternative splicing or use of different transcription initiation sites.
i: immunoproteasome

IV-2. Search model

At first, the yeast 20S proteasome model (Groll M. *et al.* 1997, PDB code: 1RYP) was truncated to polyalanine except glycines. Next, we translated

the model molecule by the 2-fold local axis parallel to the crystallographic y-axis and the quasi 7-fold axis parallel to x-axis.

IV-3. Rotation function

The self-rotation function was calculated under several conditions, but no significant peaks were found.

The cross rotation function was calculated with X-PLOR (Brünger AT. 1992.). The calculation was performed with data having Patterson vector length between 60 and 30 Å. The reflections between 20 and 5 Å resolution were used. The model Patterson map in P1 cell was computed by placing the search model into 450 × 450 × 450 Å orthrhombic box and evaluating the structure factors and the fast Fourier transform (FFT) of the squared amplitudes. The highest peak was at $\theta_1 = 42.5$, $\theta_2 = 15.0$, $\theta_3 = 137.5$ in Eulerian angles. X-PLOR defines that θ_1 is the rotation around the z-axis, θ_2 around the new x-axis, and θ_3 around the new z-axis (Fig.IV-3).

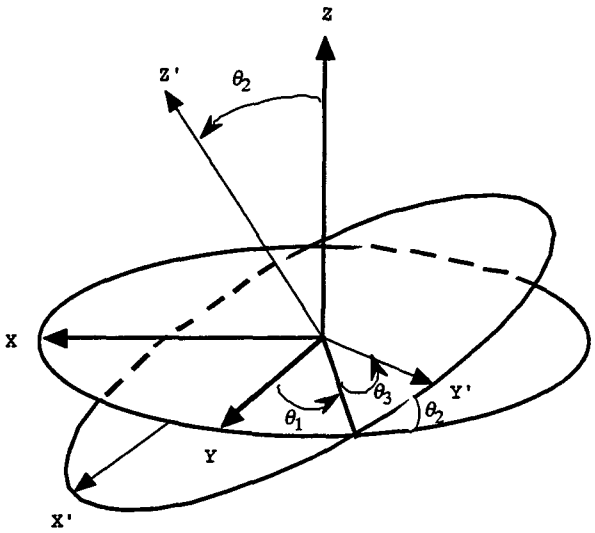


Fig.IV-3 Eulerian angles, θ_1, θ_2, θ_3 defined by X-PLOR.

IV-4. Translation function

After calculating the cross rotation function to fix the orientation of a test molecule, the translational position of the molecule must be determined. The translation function was calculated with X-PLOR using the reflections between 15 and 4 Å resolutions. The highest peak was at (0.082, 0.129, 0.425) in the fractional coordinates (Fig.IV-4). This showed a unique solution with a correlation value of 0.3097 and *R*-factor of 46.8 %.

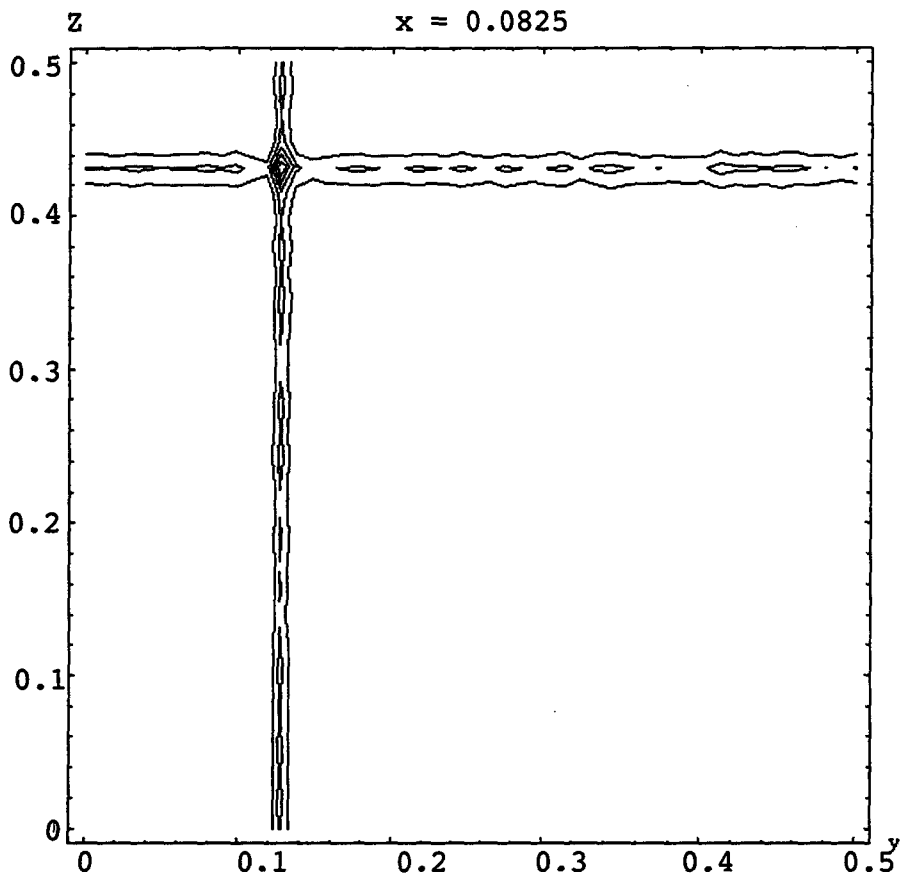


Fig.IV-4 Translation function in the plane x = 0.0825.

IV-5. Rigid body refinement

The rotational and translational freedoms are reduced to six by treating the molecule as a rigid body in the rigid body method (Sussman JL. 1985, Head-Gordon T. *et al.* 1991.)

The procedure was as follows:

1. The whole molecule was regarded as a rigid body and rigid body refinement was performed.

2. The monomers were regarded as rigid bodies, which were related by non-crystallographic symmetry (NCS) and rigid body refinement was performed.
3. Each of 28 subunits was regarded as a rigid body and rigid body refinement was performed.

These procedures (2 ~ 3) were done under NCS restraints. After rigid body minimization, *R*-factor was reduced to 45.29 %.

Chapter V. Phase refinement/extension

V-1. Introduction

After a first set of phases is obtained and an electron density map is calculated, the electron density map is the interpreted of the map in terms of the polypeptide chain. If this is successful and the major part of the chain can indeed be followed in the electron density map, refinement of the structure can be begun. However, insufficient quality of the electron density map might hamper a complete and unambiguous tracing of the polypeptide chain, increasing the risk of introducing errors in the model, which can not be easily removed during refinement. In such a case refinement should be preceded by a process to improve the quality of the map through improvement of the protein phase angles. During phase improvement, all available information on the structure should be used. For example, in this work, non-crystallographic symmetry (NCS) within a molecule of the 20S proteasome is known and the electron density of each part related by NCS can be averaged.

V-2. Used data

We used the observed amplitudes from 70 Å to 2.75 Å and the calculated amplitudes and phases from 70 Å to 6.0 Å. To suppress the model bias, $[(2m|\mathbf{F}_N| - D|\mathbf{F}^C_P|)\exp(i\alpha^C_P)]$ was used as Fourier coefficient (Read RJ. 1986.). Definitions of terms and notation are listed in Table V-1.

Table V-1. Definitions of terms and notation

$$\mathbf{F}_N = \sum_{j=1}^{P} f_j \exp(2\pi i \mathbf{s}.\mathbf{r}_j) + \sum_{j=P+1}^{N} f_j \exp(2\pi i \mathbf{s}.\mathbf{r}_j)$$

where $\mathbf{s}$ is the reciprocal-lattice vector ($|\mathbf{s}| = 2\sin\theta/\lambda$) and the $\mathbf{r}_j$ are the atomic coordinates (in Å)

$= \mathbf{F}_P + \mathbf{F}_Q$, where the P atoms constitute the partial structure and the Q atoms the missing structure

$= |\mathbf{F}_N| \exp(i\alpha_N)$

$$\mathbf{F}^c{}_P = \sum_{j=1}^{P} f_j \exp(2\pi i \mathbf{s}.(\mathbf{r}_j + \Delta\mathbf{r}_j))$$, where $\Delta\mathbf{r}_j$ are positional errors

= structure factor of partial structure with errors

$D = \langle \cos 2\pi \mathbf{s}.\Delta\mathbf{r} \rangle$ (Luzzati V. 1952)

ε = correction factor of partial structure with errors

$$\Sigma_N = \sum_{j=1}^{N} f^2{}_j = \langle |\mathbf{F}_N|^2 / \varepsilon \rangle, \quad \Sigma_Q = \langle |\mathbf{F}_Q|^2 / \varepsilon \rangle$$

$$\mathbf{E}_N = \mathbf{F}_N / (\varepsilon \Sigma_N)^{1/2}$$

$$\sigma_A = D(\Sigma_P / \Sigma_N)^{1/2}$$

$m = \langle \cos(\alpha_N - \alpha^c{}_P) \rangle$

$= I_1(X)/I_0(X)$ for non-centric reflections, where I_0 and I_1 are the zero and first-order modified Bessel functions, respectively, or

$= \tanh(X/2)$ for centric reflections, where

$X = 2|\mathbf{F}_N||\mathbf{F}_P|/\varepsilon\Sigma_Q$ for a partial structure with no errors (Woolfson MM. 1956, Sim GA. 1959, 1960)

$= 2\sigma_A|\mathbf{E}_N||\mathbf{E}_P|/(1-\sigma_A^2)$ for a partial structure with no errors (Srinivasan R. 1966)

V-3. NCS Averaging

In this crystal, only one 20S proteasome molecule is contained in the asymmetric unit. However, the 20S proteasome contains two equal parts. We can divide the molecule into two groups; α_1-α_7, β_1-β_7 and β_1-β_7, α_1-α_7. It was easy to know the orientation and position of these molecules.

At first, we defined the envelope of the molecules. We made two envelopes, α- and β-rings, from the polyalanine model after the molecular replacement. The masks were generated from a sphere around all atoms. We set the sphere radius at 7 Å for building a mask. We removed the overlap regions with its crystallographic and the NCS related partners.

The electron density in the molecules, related by this NCS, was essentially equal, although the difference in the contact with neighbors might cause some deviation from exact equality.

V-4. Solvent Flattening

Because the solvent molecules except ones internally in the protein molecules have dynamic nature, its time-averaged electron density has a low constant. Thus we can remove the noise peaks in the solvent region simply by setting the electron density to a low constant value.

The molecular boundary was updated by each calculating cycle. In this work, a new map was constructed such that the density at each grid point is proportional to the weighted sum of the

diffraction signals (density above a certain background level within a sphere of radius R from that grid point in the original map,

$$\rho'_j = K \sum_{(I-R)} w_i \rho_i$$
$$w_i = 1 - r_{ij}/R, \quad r_{ij} < R \text{ and } \rho_i > 0$$
$$w_i = 0, \quad r_{ij} > R \text{ or } \rho_i < 0$$

where r_{ij} is the distance between grid points i and j, K is an arbitrary constant, and ρ constants no ρ^0 term. The negative densities are omitted giving them zero weights (Wang BC. 1985.).

V-5. Histogram matching

The method of histogram matching uses the frequency distribution of electron density values. It appears that these distributions as a function of ρ are fairly independent of the protein. A poor image has a different distribution and this is then scaled to the standard distribution and this process results in an improved image (Lunin VY. *et al.* 1988, 1990, 1991, Zhang *et al.* 1990.). The frequency distribution of electron density levels calculated at grid points is plotted as a function of ρ for the poor electron density map and compared with the standard plot. Next, the plots are divided into bins containing an equal number of grid points. The bin boundaries for the poor map are ρ_i, ρ_{i+1}, with i = 1 to n and n ≈ 100, and for the standard map are ρ'_i, ρ'_{i+1}. By scaling with a factor $a = (\rho'_{i+1} - \rho'_i)/(\rho_{i+1} - \rho_i)$ the bin width from ρ_i to ρ_{i+1} is given the correct value. The bin is then moved over a distance b such that ρ_i moves ρ'_i and ρ_{i+1} to ρ'_{i+1}.

From $\rho'_i = a \times \rho_i + b$ it can easily be derived that $b = (\rho_{i+1}\rho'_i - \rho_i\rho'_{i+1})/(\rho_{i+1} - \rho_i)$.

The electron densities in each bin are corrected with the appropriate *a* and *b*. This results in an improved electron density map from which new phase angles can be calculated.

V-6. Result

In this work, we extended and improved phase angles with NCS averaging, solvent flattening and histogram matching from 6.0 Å to 2.75 Å resolution by using program DM (Cowtan K. 1994.) of CCP4 program package. The final density modification free-*R* value (in no way comparable to the refinement free-R) was 28.4 % and NCS averaging correlations between α-rings and between β-rings were 88.2 % and 90.1 %, respectively.

Chapter VI. Modeling and model refinement

VI-1. Introduction

To know the three-dimensional structure of the mammalian 20S proteasome at atomic resolution, we have to construct an atomic model which is thought to provide a reasonable "fit" to the electron density.

VI-2. Amino acid sequence

Since the amino acid sequences of bovine 20S proteasome were not known, the amino acid sequences of the human proteasome were substituted for those of the bovine enzyme. The amino-acid sequences of the human and yeast proteasomes were taken from SWISSPROT (Bairoch A. *et al.* 2000). The N-terminal residue of any subunit is assigned as 1st residue for the bovine 20S proteasome. Thus the numbering system is different from that of the yeast β-type subunits (Groll M. *et al.* 1997.). Primary structure alignments among the subunits were performed with program CLUSTALW (Thompson JD. *et al.* 1994.)

VI-3. Constructing the model

Model building was done on a SGI workstation using the program TURBO-FRODO (Jones TA. 1978).

Although the model used in molecular replacement was the polyalanine model, we could recognize the electron densities of side chains after the density modifications. We judged that we succeeded in phase determination and thus we tried constructing a model. The 20S proteasome has about 7,000 amino acids, but our duties are about 3,500

amino acids because the independent region is half of the molecule. It was hard to fit all 3,500 amino acids to the electron densities. The crystal consisted of constitutive proteasomes without any contamination of immunoproteasome.

Some α-helices and β-strands of the yeast proteasome were found at different places from the electron densities regarded as α-helices and β-strands, respectively. Thus, at first, we cut a few parts of the model molecule and then, we tried using simulated annealing (Kirkpatrick S. *et al.* 1983.) with program package of CNS (Brünger AT. *et.al.*, 1998). After simulated annealing, it was a little easier to construct the model.

Most parts of electron density map were so clear that side chains were assigned unambiguously, but small parts were too poor to build side chain structures. Alanine residues were left at these parts with poor electron density. They are listed in Table VI-1. Since some N-termini and C-termini lost electron density, their structural models were not built.

Table VI-1 Residues whose main chain structures were determined and those of which side chains remained as Ala because of this poor electron density

	Residues remained as Ala in model
α1 2-245/246	K45, D46, K55, D62, E146, K181, K186, F187, D188, E196, R245
α2 1-233/233	R3, K17, V35, Q51, K52, S53, I54, R59, W138, N139, E140, R142, M162, K175, R176, Y177, L183, L192, L194, K195, F198, E199, Q201, V225, K226
α3 2-255/261	S13, I36, N51, I52, K54, Q177, Y179, M184, I194, L197, M201, K205, K210, K222, K229, Q230, K231
α4 2-245/248	V26, I40, V42, L43, E46, K47, K48, S49, V50, L53, Q54, E56, K61, N68, I83, F136, D137, F138, D139, T141, Q146, K157, N159, R163, K166, R169, E170, K174, I181, E182, T183, D185, T187, I188, K189, I192, K193, L195, L196, E197, V198, V199, K204, E207, L208, M211, R212, R213, S216, K218, E223, E224, I225, E226, K227, Y228, V229, I232, E233, K234, E235, K236, E237, E238, N239, E240, K241, K242, K243, Q245
α5 8-241/241	K86, K197, K192, E208, K209, K231
α6 4-240/263	K208, K216, D217, E237, E238, P240
α7 1-245/254	N143, E203, K205, K207, R223, E245
β1 1-202/205	T201, L202
β2 1-220/234	K9, K180, K205, K215
β3 2-205/205	R48
β4 1-199/201	K41, R95, E109, K185
β5 1-201/204	K106, E150, I185
β6 1-213/213	V6, K45, V161, L166, K200
β7 1-217/219	R44, K156, E170, K195, S216

VI-4. Structure refinement

Once model building was finished, the

structure was refined under the NCS restraints. The sigma-A weighted electron density maps (2Fo-Fc and Fo-Fc) (Read RJ. 1986.) or composite omit maps (Read RJ. 1986, Bhat TN. 1988, Hodel A. *et al.* 1992, Brünger AT. *et al.* 1997) were calculated to reduce biased electron density caused by the initial structure. These calculations were performed with program CNS (Brünger AT. *et al.* 1998.). The model revised in the maps was refined with simulated annealing and temperature factor refinement. The maps were calculated again. We repeated these procedures again and again. After several cycles, we were aware that these maps showed some parts being different between NCS-related regions. Thus, NCS restraint was imposed on the whole molecule except for the residues 40-43 and 49-54 of subunit α2 and the residues 236-241 of subunit α6 during the refinement. Temperature factor refinements were done without NCS restraints on the way of these refinement steps.

In the middle of refinement steps, water molecules were picked up automatically by using program CNS and checked these water molecules with graphics by looking at the difference Fourier map. The distances between waters and any Os and Ns were from 2.6 Å to 3.6 Å. The number of water molecules was 166. The hydrogen bonding distance we set was a little long, but it was reasonable because of high temperature factors of the model. Magnesium ions were picked up by using program PEAKMAX and WATPEAK of CCP4 program package (Collaborative Computational Project, Number 4. 1994.). The distances of Mg from the proteasome and waters

picked up were 1.5 Å - 2.5 Å and checked the positions of the ions graphically. The number of Mg ions was 30. These were restrained with NCS. The used map in picking water molecules and Mg ions was Fo-Fc map.

The final structural model with R = 25.0 % and R_{free} = 29.4 % was obtained by iterative modeling and CNS refinement at 2.75 Å resolution. The statistics of phasing and refinement are summarized in Table VI-2.

Table VI-2 Model refinement statistics

†R_{cryst}	25.0 %/ ‡R_{free} 29.4 %
R.m.s. bond length	0.008 Å
R.m.s. bond angle	1.4°
Average B	
alpha ring1	84.1 Å^2
beta ring1	64.0 Å^2
beta ring2	59.2 Å^2
alpha ring2	68.1 Å^2
Ramachandran plot (%)	
Most favored	83.9
Additional allowed	14.7
Generously allowed	1.2
Disallowed	0.1

†R_{cryst} = $\sum||Fobs| - |Fcalc||/\sum|Fobs|$. ‡$R_{free}$ is the same as R_{cryst} but for a 5 % subset of all reflections that were never used in crystallographic refinement.

Chapter VII. Structure of the 20S Proteasome from bovine liver

VII-1. Introduction

The crystal structural analysis of the 20S proteasome from bovine liver was performed to elucidate structure-function relationships and structural organization of the mammalian proteasome.

In VII-2, we discuss the general structure of the bovine 20S proteasome and the differences between the structures of the bovine proteasomal subunits and the corresponding yeast proteasomal subunits.

Proteasomes are known as members of N-terminal nucleophile superfamily. In VII-3, the structures of three active sites are discussed.

The mammalian and the yeast proteasome have N-terminal threonine in the β7 subunits. In VII-4, discovery of the novel active site is discussed.

In VII-5, we discuss the structure of the predicted immunoproteasome.

Eukaryotic proteasomes are in both the cytosol and nucleus. They have nuclear localization signals (NLSs). Some NLSs are found in the characteristic motives which contain many lysines and glutamic acids in small region (KEKE motif). In VII-6, we discuss KEKE motif and NLS.

Outer rings of 20S proteasomes are consisted of α(-type) subunits. Substrates are carried through the central pore which is formed with α(-type) subunits to the central chamber which is formed with β(-type) subunits. In VII-7, we discuss

the central pore of α-ring.

In VII-8, we discuss the quaternary structure of the eukaryotic proteasomes, which are consisted of 14 different subunits.

VII-2. General architecture and subunit structures in comparison with the yeast 20S proteasome

The overall shape of the bovine 20S proteasome was an elongated cylinder having large central cavities and narrow constrictions (Fig. VII-1). The overall dimensions of bovine 20S proteasome are about 150 Å in length and about 115 Å in diameter. The α-type subunits are located at the ends, whereas the β-type subunits form the two equatorial rings. The general architecture was completely conserved among the *T.acidophilum* (Löwe J. *et al.* 1995), the yeast (Groll M. *et al.* 1997) and the mammalian 20S proteasomes. An identical arrangement of subunits in the yeast 20S proteasome was confirmed in the bovine 20S proteasome (Fig. VII-2).

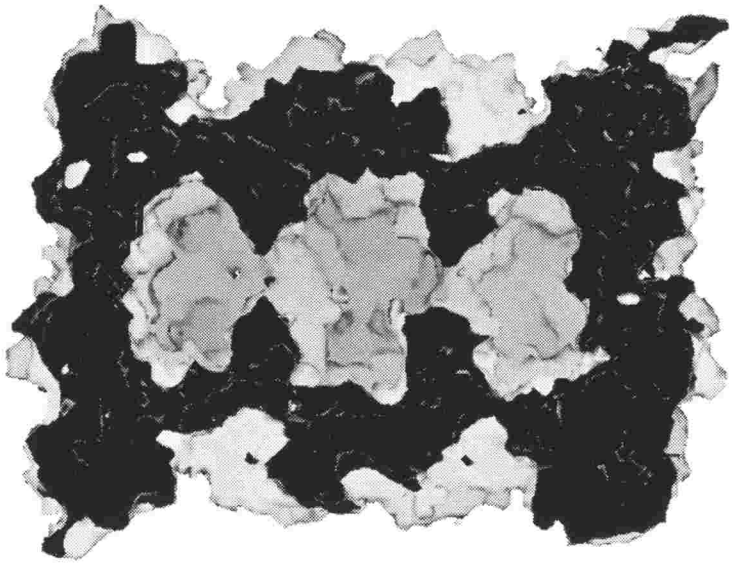


Fig. VII-1 View of the 20S proteasome from bovine liver cut open along the quasi-sevenfold axis; the three compartments are clearly visible. The gate is shut.

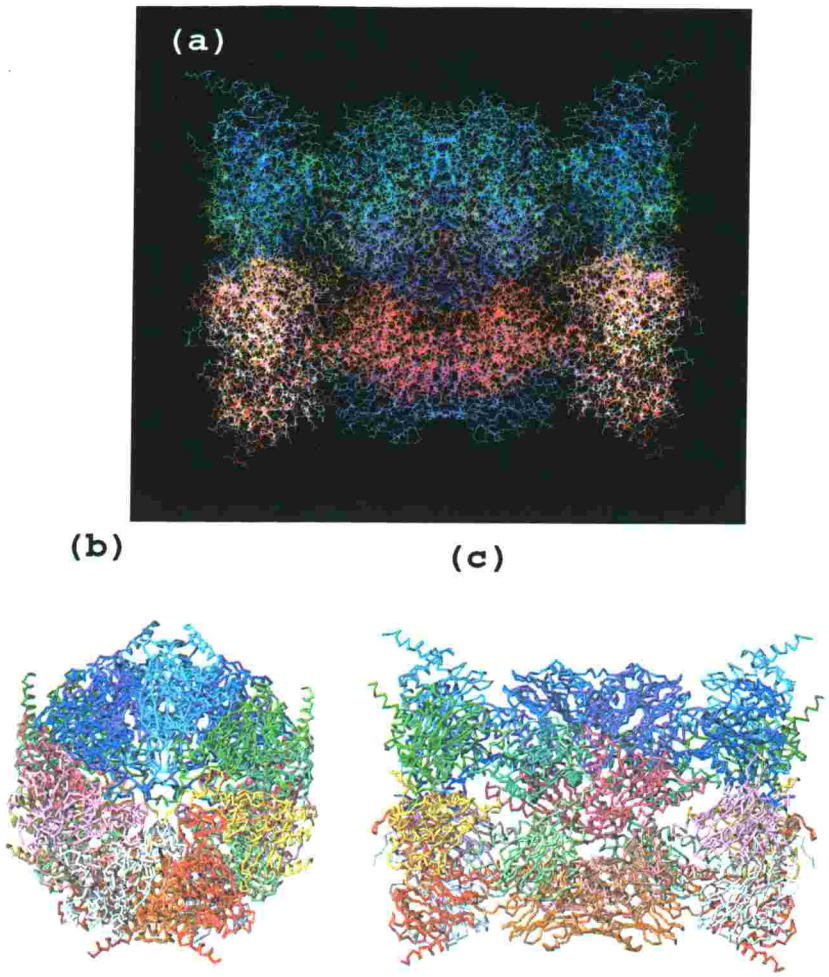


Fig. VII-2 Side view (a) of all atoms and top view (b) and side view (c) of Cα drawings of the bovine 20S proteasome consisting of 28 subunits, $[(\alpha 1\text{-}\alpha 7)(\beta 1\text{-}\beta 7)]_2$. Fourteen different subunits are in different colors; α1, red; α2, yellow; α3, green; α4, sky blue; α5, blue; α6, pink; α7, gray; β1, orange; β2, dark sea green; β3, medium sea green; β4, dark sky blue; β5, purple; β6, magenta; and β7, light pink.

All the α subunits have a β-sandwich structure formed by two five-antiparallel β-sheets consisting of strands S1 to S10. The sandwich structure is surrounded by helical layers of helices H1 and H2 at one side and H3, H4 and H5 at the other side as are those of the *T.acidophilum* and the yeast proteasomes. Any α subunits have another helical structure of H0 at the N-terminus side of S1. Each α subunit of the bovine enzyme exhibits a structural difference from the corresponding subunit of the yeast proteasome in the secondary structure segments of H5 except for the α7 subunits (Fig. VII-4 (a)).

While the N-terminal residues 2-16 before the reverse turn and H0 of bovine α1 were visible in the electron density map, the seven N-terminal residues of the corresponding segment of the yeast subunit are disordered. In the bovine α1 subunit, helices H3 and H4 were connected by a short linkage with a five-residue deletion in comparison with the long loop structure of the yeast α1 subunit. The bovine α2 subunit had a turn between S9 and S10, while the yeast α2 subunit has a long insertion forming a loop. The bovine α3 subunit had a longer H5 helix by three turns at the C-terminus than the yeast α3 subunit. The loop structure between S9 and S10 of the bovine α3 subunit was shorter by two residues than that of the yeast subunit. The α4 subunits of the bovine and the yeast proteasome had different loop structures between S2 and S3 with the same number of residues. Both α5s of the bovine and the yeast proteasomes had long loops between H2 and S5 each with a unique conformation at the

central part of the α-ring. The α6 subunit of the mammalian proteasome had a longer polypeptide sequence by 28 residues at the C-terminus than that of the yeast α6 subunit. The H5 helix of the bovine α6 subunit was longer than that of the yeast. However, C-terminal residues 241-263 of the bovine enzyme were not visible because of autolysis or a disordered conformation. The N-termini of bovine α7 was longer than that of the yeast subunit by four residues and had a different orientation from the yeast subunit. The helix H3 of bovine α7 is shorter than that of yeast α7 by one turn.

The structures of all the β subunits of the bovine had a β-sandwich structure similar to the α subunits consisting of the strands of S1 to S10 like those of the yeast and the *T.acidophilum* proteasomes. However, the bovine β-subunits exhibited structural variations at the N-terminal residues before Thr1 in the numbering system of the *T.acidophilum* proteasome and at the C-terminal residues after S10.

Each structure of the β subunit of the bovine enzyme had a high similarity to that of the corresponding subunit of the yeast enzyme in the secondary structure segments of H1 to H4 and S1 to S10 (Fig. VII-4 (b)). The post-translational processing of eukaryotic β subunits found in the yeast proteasome was confirmed in the structure of the bovine 20S proteasome. Three active subunits of β1, β2 and β5 had the N-terminal residues of Thr as did the *T.acidophilum* β subunits and the yeast active β subunits. The bovine subunits of β6 and β7, respectively had ten and eight extra N-terminal

residues in comparison with the active subunits in the electron density map. These were partly processed during the structural organization. The bovine β4 subunits preserved a Met residue at the N-terminus. The Met residue of the bovine β3 subunit at the N-terminus was not located in the clear electron density map probably because of its disordered conformation. The C-terminus following S10 of the bovine β1 subunit was inserted between trans β1 (β1′) and trans β7 (β7′) with a long extended conformation instead of the two-turn helix of the yeast β1 subunit. The β2 subunit of both enzymes had a long C-terminal arm consisting of about 25 residues which wound around the β3 subunit. Glu91 of the human β2 was replaced by an Arg residue in the bovine β2 subunit based on the electron density map, though Glu91 was well conserved among eukaryotes. Any notable conformational difference was not detected between the β3 subunits of the bovine and the yeast except for a residue at the amino terminus. The bovine β4 subunit was longer than that of yeast by two residues at the C-terminus with a long extended structure instead of the hooked structure of the yeast β4 subunit. The bovine β5 terminated in an H5 α-helix shorter by one residue than the H5 of the yeast β5 subunit which is followed by a C-terminal loop with eight residues. Between H3 and H4 of both β6 subunits, there were long loops. However, the loop of the bovine enzyme was shorter than that of the yeast by eight residues. The bovine β7 subunit of the bovine 20S proteasome terminated at a helical structure, while the yeast β7 subunit has an extra eight-residue loop at the C-terminal

side of the helix. The N-terminus of the bovine β7 subunit was in a different orientation from that of the yeast β7 subunit. The loop between H1 and H2 of the bovine β7 subunit was shorter than that of the yeast by four residues.

The overlay of the alpha-carbon (Cα) tracings shows some insertions and the considerable variability in the core structures, which is less in the α-type and more in the β-type subunits (Fig.VII-4 (a), (b)).

(a)

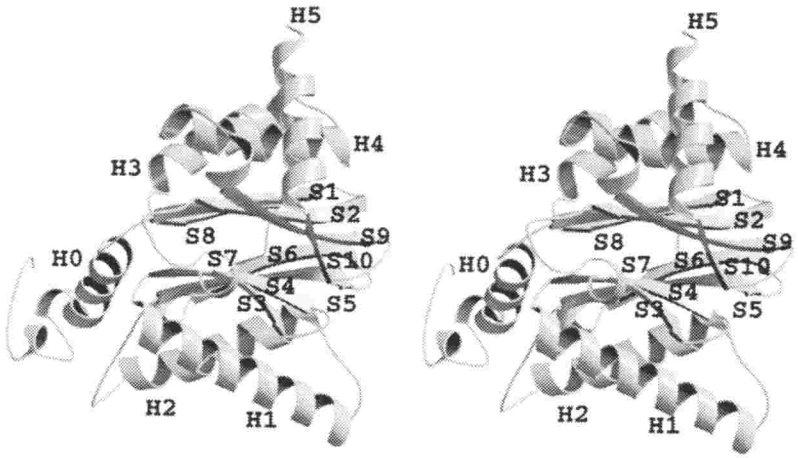


Fig. VII-3 (a) Stereo view of the ribbon diagram of the α1 subunit as the typical α-type subunits of the bovine 20S proteasome.

(b)

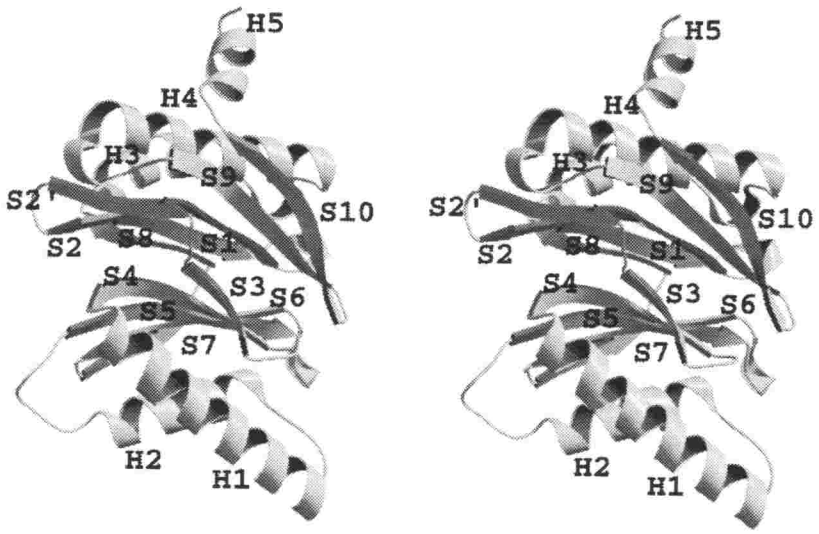


Fig. VII-3 (b) Stereo view of the ribbon diagram of the α5 subunit as the typical β-type subunits of the bovine 20S proteasome.

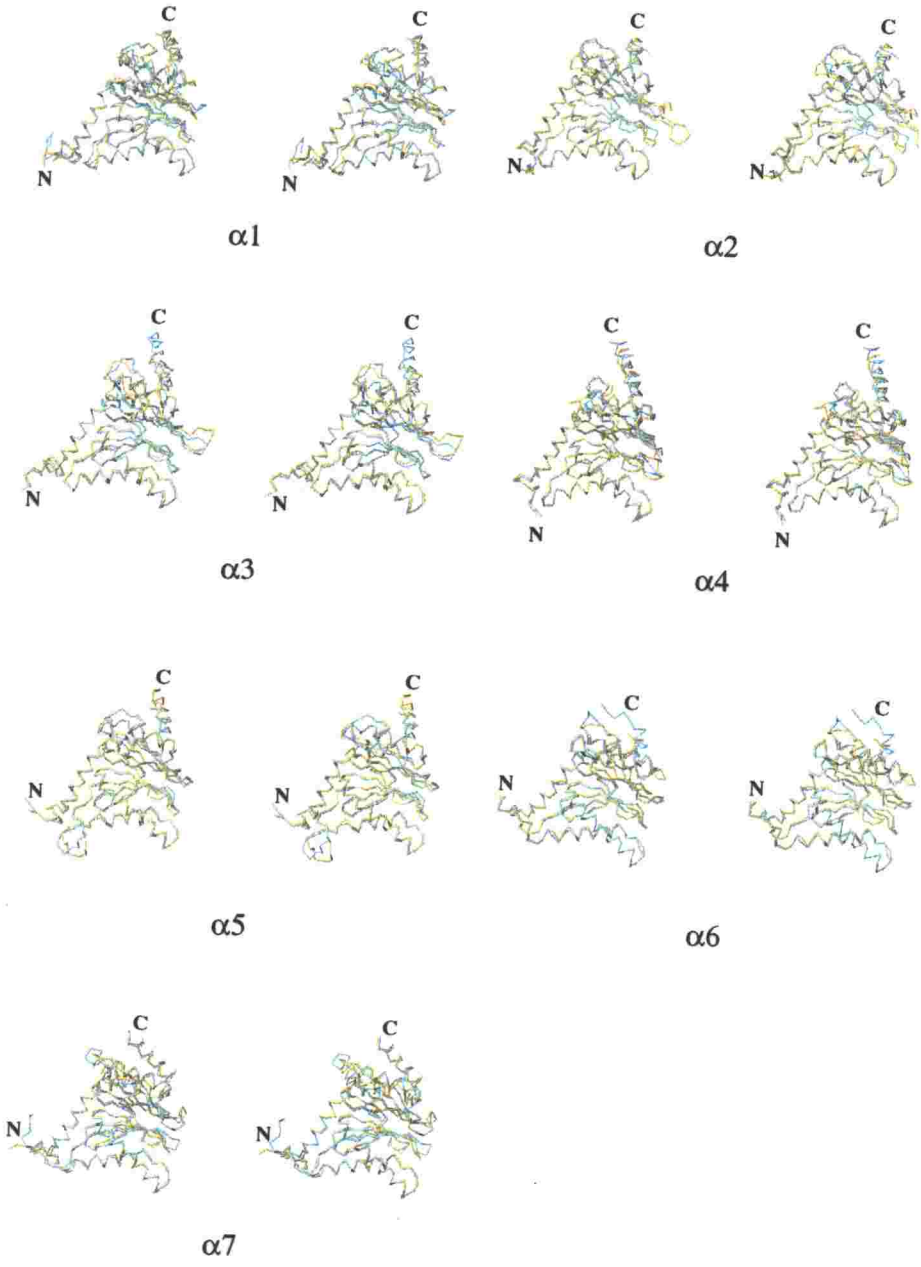


Fig. VII-4 (a) Stereo view of superposition of corresponding α-type subunits of the bovine and the yeast proteasomes. Sky blue represents the bovine subunits and yellow shows the yeast subunits. All the subunits have almost the same orientations.

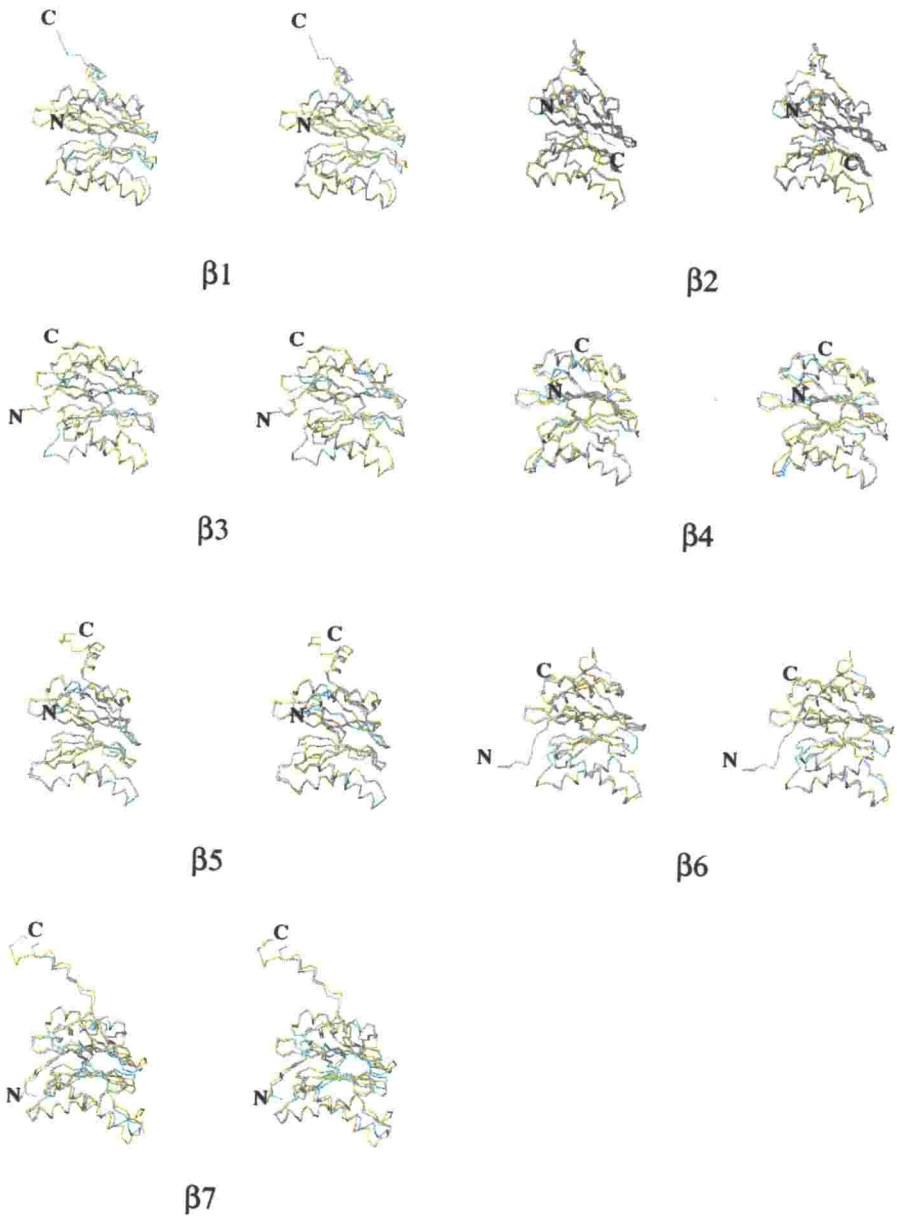


Fig. VII-4 (b) Stereo view of superposition of corresponding β-type subunits of the bovine and the yeast proteasomes. Sky blue represents the bovine subunits and yellow shows the yeast subunits. All the subunits have almost the same orientations.

VII-3. Structural conservation of Ntn-hydrolase active sites

The structures of the Ntn-hydrolase active sites of the bovine and the yeast proteasome were compared by superposing functional amino acid residues of Thr1, Glu17, Arg19, Lys33, Ser129, Asp166, Ser169 and Gly170, where the numbering system followed that of the β subunit of the *T.acidophilum* proteasome (Löwe J. *et al.* 1995)

Superposition of the bovine β1, β2 and β5 around the active sites are depicted by a stereoscopic drawing in Fig. VII-5 (a). R.m.s. deviation for the all atoms of the functional residues of the bovine β1-β2, β1-β5 and β2-β5 were 0.353 Å, 0.329 Å and 0.360 Å, respectively. The locations of these functional residues in each subunit were well conserved among the active β subunits.

The β-ring of the bovine proteasome was superposed on the β-ring of the yeast proteasome by the least-square fitting of main chain atoms. The displacements between equivalent atoms are shown by the colors of the main chain atoms of the bovine β-ring depicted by a stereoscopic pair (Fig. VII-5 (b)). R.m.s. deviation for the all atoms of the functional amino acids between the corresponding subunits of the human and the yeast 20S proteasome were 0.259 Å, 0.250 Å and 0.244 Å for the β1, β2 and β5 subunits, respectively. The catalytic sites including the substrate binding regions of β1, β2 and β5 were well conserved between the corresponding bovine and the yeast β subunits

in the quaternary structure. Not only the main chain atom positions but also the side chain atom positions were well conserved.

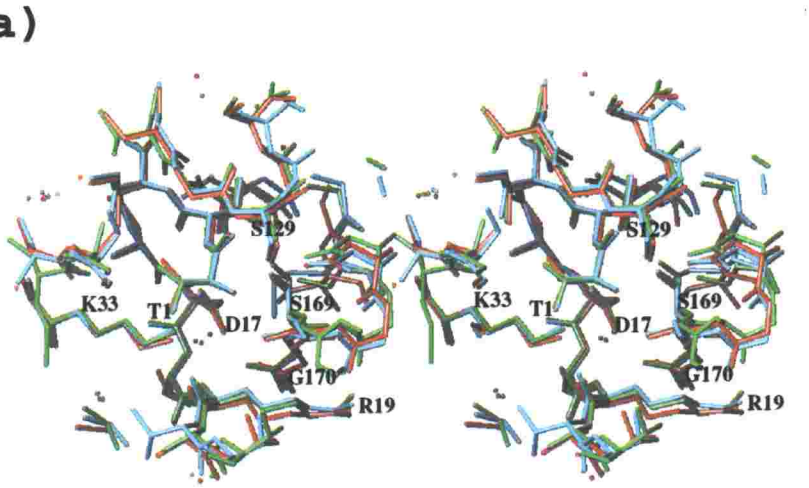


Fig. VII-5 Conserved structure of the Ntn-hydrolase active sites.
(a) Stereo view of the superposition of three active sites of the bovine β1 (sky blue), β2 (red) and β5 (green). The functional amino acids of Thr1 Glu17, Lys33, Ser129, Asp166, Ser169 and Gly170 were applied for the least squares fitting. They are shown by the single letter notation with the residue number equivalent to that of *T. acidophilum* β subunit. The functional groups of the Ntn-hydrolase active sites were conserved well not only in the primary structure but also in the tertiary structure.

(b)

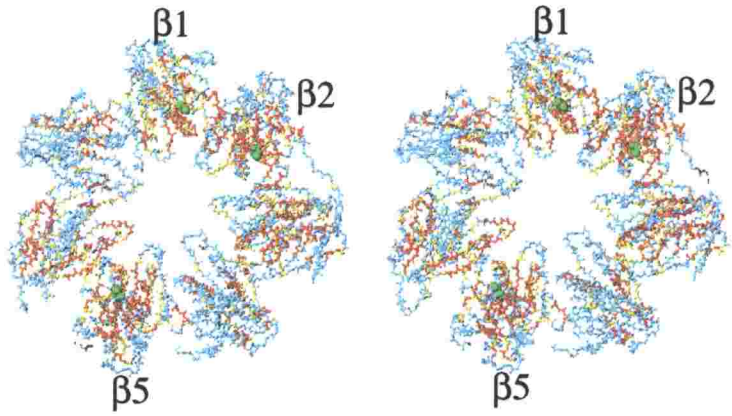


Fig. VII-5 (continued). (b) Structural difference between the bovine and the yeast β-rings is shown by a stereoscopic drawing of the main chain atoms of the bovine β subunits viewed from the trans-β ring. Two structures were superposed by the least-squares fitting of the equivalent main chain atoms. Displacements between equivalent atoms are shown by colors. Red, smaller than 0.2 Å; Yellow, 0.2 Å ~ o.4 Å; Sky blue, larger than 0.4

We found water molecules in all catalytic subunits close to Thr1 Oγ and N. Thr1 N is a proton acceptor when Thr1 Oγ attacks to an electrophilic center. Thr1 N is positioned to serve as a proton shuttle from Thr1 Oγ to substrates.

The distances between two Thr1 Cα atoms of the active subunits, β1, β2, and β5 (Table VII-1). The shortest distance between two active sites – Cα of Thr1 – is about 26 Å, which may be spanned by a hepta- or octapeptide in expanded conformation. Longer peptides may be generated by two active sites at a wider distance, shorter ones could be due to premature deacylation, rebinding, and processing.

Other β-type subunits – the β3, β4, β6 and β7 subunits – are similar to active subunits in respect to the conformations around the active sites, though the β3 and β4 subunits are not processed and the β6 and β7 subunits are intermediately processed. The β3, β4 and β6 subunits have Asp and Lys at the positions corresponding to Asp17 and Lys33 of active subunits, respectively. Only the β7 subunit does not have Lys at the positions of Lys33 of active subunits. Instead of Lys, the b7 subunit has an Arg in this position. The cleavage sites of these subunits except the β7 subunit are not Thr. The β7 subunit has Thr at the cleavage site, but far from the position corresponding to Thr1 of active subunits.

Table VII-1. Distance between Cα of Thr1 of catalytic subunits

β1	β2	28.8 Å
β1	β5	64.3 Å
β1	β1′	25.5 Å
β1	β2′	48.1 Å
β1	β5′	59.7 Å
β2	β5	65.1 Å
β2	β2′	64.7 Å
β2	β5′	41.6 Å
β5	β5′	47.9 Å

VII-4. Novel Ntn-hydrolase active site of the β7 subunit

The structure around the N-terminal Thr1 (corresponding to Thr-8 of the yeast β7 subunit) of the bovine β7 subunit is depicted by a stereoscopic pair in Fig. VII-6 (a). The location of the following functional groups satisfied the requirement for the Ntn-hydrolase active site. The terminal amino group of Thr1 formed a hydrogen bond with Asn104 Oδ. Thr1 Oγ-H had a hydrogen bond with Asp59 Oδ. Arg91 of β1 formed a salt bridge with Asp56 of β7. The oxyanion hole was formed by the Tyr88 OH of β1 or Arg99 Nη. A water molecule was found near Thr1 and must be replaced by a substrate upon formation of the enzyme-substrate complex (Fig. VII-6 (b)). Although the location of functional groups was different from those of β1, β 2 and β5, a reasonable reaction mechanism (Oinonen C. *et al.* 2000) for substrate hydrolysis was developed by considering the location of these functional groups as shown in Fig. VII-7. These structural features suggested that the Thr1 of bovine β7 was a novel

Ntn-hydrolase active site.

All the functional groups of Thr1, Asp56, Arg99 and Asn104 of the β7 subunits, and Tyr88 of the β1 subunits were conserved in amino acid sequence between the bovine and the yeast enzymes. The three-dimensional location of these functional groups of both enzymes had a similarity except for Thr1. Thr-8 of the yeast β7 subunit (Thr1 of the bovine β7) exhibiting a different conformation from the Thr1 of the bovine β7 subunit (Fig. VII-5 (b)) was able to be converted to the same conformation as the Thr1 of the bovine enzyme. Thus, the yeast proteasome must share the novel Ntn-hydrolase active site with the bovine proteasome.

There is a hollow around this putative active center, which is much smaller than the S1 pockets of β1, β2 or β5 (Fig. VII-6 (c)). The hollow could not accommodate bulky residues. This active site must have a small neutral amino acid-preferring (SNAAP) activity.

Thr1 of the bovine β7 subunit at the cleavage site was far from the positions of Thr1 of the active subunits when they were superimposed by main chain atoms. The direction of the N-terminal main chain of β7 indicated that the new active site was not in the chamber formed by the two β-rings but was near the gate formed by the α- and β-rings (Fig. VII-6 (d)).

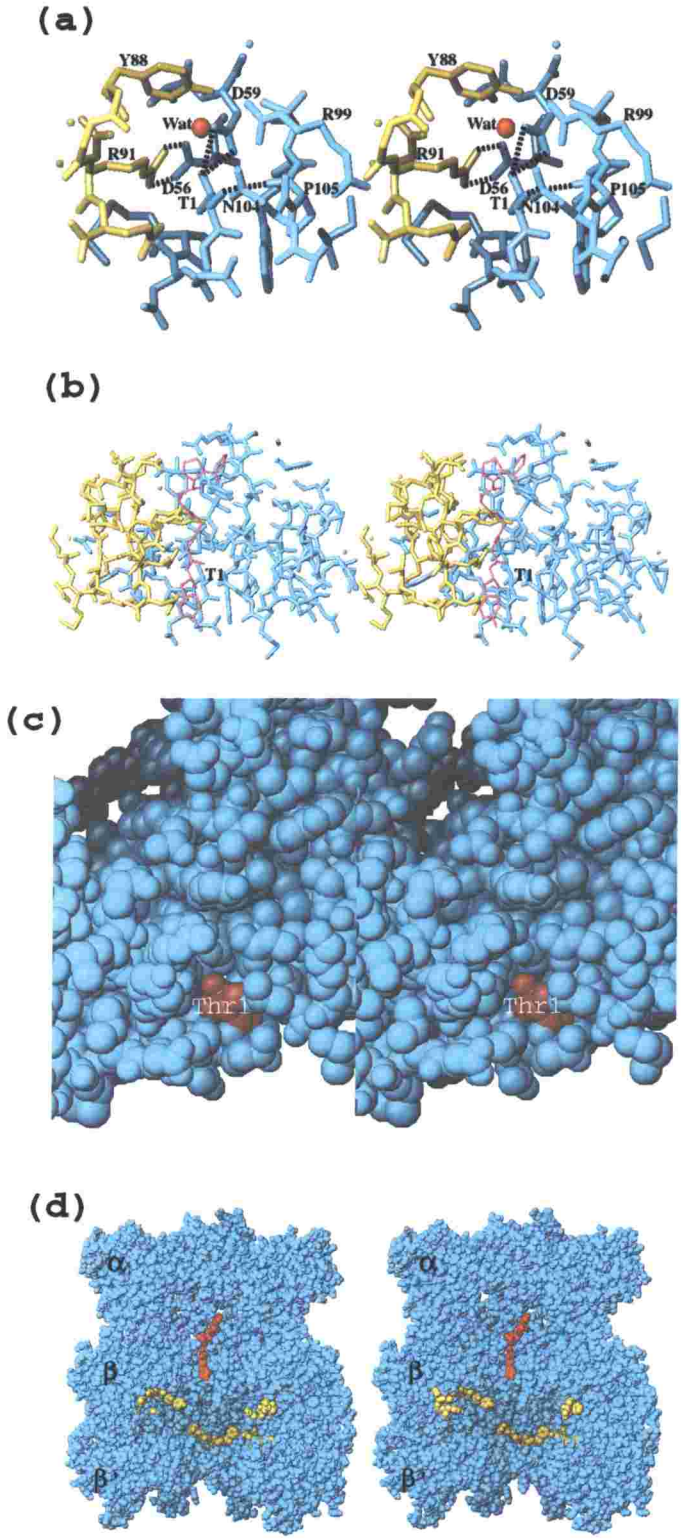

(a)
Y88
D59
R99
Wat
R91
D56
T1
N104
P105
(b)
T1
(c)
Thr1
(d)
α
β
β'

Fig. VII-6 (previous page) Stereo views of the new Ntn-hydrolase active site projected along almost the same directions.

(a) Residues included in the catalysis are shown by the single letter notation with a residue number of the bovine subunit. β1 and β7 subunits are colored yellow and sky blue, respectively. The red sphere is a water molecule, which is located near Thr1 of β7.

(b) Predicted structure of a complex of the bovine enzyme and a model substrate Phe-Gly-Pro-Ala-Gly-Gly-Tyr fitted by X-PLOR is shown by a stereoscopic pair. β1, β7 subunits and the model substrate are depicted in yellow, sky blue and red, respectively. The orientation is almost the same as in (a).

(c) The active site was drawn by a space-filling model to show its hollow surface. Thr1, colored red, is located at the bottom of this hollow.

(d) The predicted substrate structures in the 20S proteasome. The red model is the substrate of the novel β7 active site; the three yellow models are those of β1, β2 and β5 active sites. The α-ring, β-ring and β'-ring are depicted in part for convenience. The novel Ntn- catalytic site is in a different position and a different orientation (red) from the other three active sites (yellow).

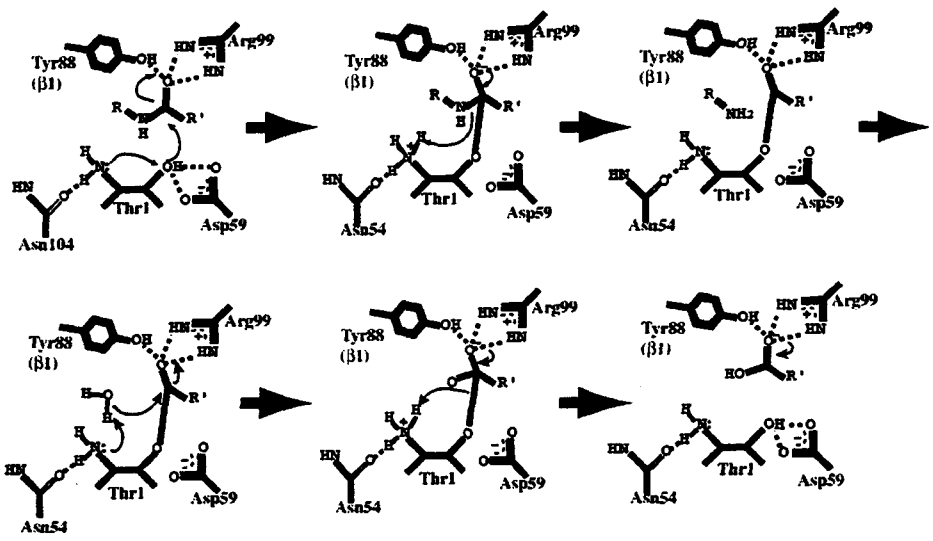


Fig. VII-7 Putative catalytic mechanism of the novel active site in β7.
The reaction begins when the nucleophilic oxygen of Thr1 donates its proton to its own α-amino group and attacks the carbonyl carbon of the substrate. The negatively charged tetrahedral intermediate is stabilized by hydrogen bonding. The acylation step is complete when the α-amino group of Thr donates a proton to the nitrogen of the scissile peptide bond. The covalent bond between a part of the substrate and the enzyme is formed, and the part of the substrate is released. The deacylation step begins when the hydroxyl group of water attacks the carbonyl carbon of the acyl-enzyme product and the basic α-amino group of the nucleophile accepts the proton from the water molecule. The negatively charged intermediate is stabilized, as in the acylation step. The reaction is completed when the α-amino group donates a proton to the nucleophile.

VII-5. Predicted structure of immunoproteasome and substrate specificities

The sequence identities of the human β1i, β2i and β5i subunit compared to the human β1, β2 and β5 are 59.2 %, 57.7 % and 68.6 %, respectively, while those compared to the yeast β1, β2 and β5 are 46.0 %, 46.0 % and 57.6 %, respectively. Any γ-interferon-induced subunit of the human 20S proteasome has significantly higher sequence similarity with the corresponding subunit of the human constitutive proteasome than with that of the yeast. Therefore, the X-ray structure of the bovine 20S proteasome was used for prediction of the three-dimensional structure of the immunoproteasome.

An initial model of the immunoproteasome was made by exchanging the residues of constitutive subunits with those of the inducible subunits with their standard side chain conformations taken from the dictionary of TURBO-FRODO (Jones TA. 1978). The initial model was refined by positional refinement using X-PLOR (Brünger AT. 1992) without observed **F** data. In this step, the remaining constitutive subunits were fixed and water molecules were not included.

The structure of the immunoproteasome predicted from the bovine enzyme was principally consistent with that from the yeast enzyme (Groll M. *et al.* 1997, Kloetzel PM. *et al.* 2001). Each S1 pocket of the bovine 20S proteasome was formed by two subunits as was that of the yeast enzyme (Groll M. *et al.* 1997). The S1 pocket of the active site

around Thr1 of the bovine β1 subunit was mainly formed by residues Thr20, Thr31, Thr35, Arg45, Ala49 and Gln53 of the β1 subunit and Tyr114, His116, Ser118, and Asp120 of the β2 subunit. That of the β2 active site was formed by Ala20, Cys31, His35, Gly45, Ala49 and Asp53 of the β2 subunit and Asp125, Met131 and Cys129 of the β3 subunit, and that of the β5 active site consisted of Ala20, Val31, Ile35, Met45, Ala49 and Ser53 of the β5 subunit and Ser129, Asp125, Gln131 and Asp133 of the β6 subunit (Fig VII-9). Val20, Phe31, Ser35 and Leu49 of the β1i subunit and His114 and Ser120 of the β2i subunit were located in the S1 pocket of the β1i active site instead of Thr20, Thr31, Thr35 and Arg45 of the β1 subunit, and Tyr114 and Asp120 of the β2 subunit. The electrostatic surface potential was calculated and displayed with GRASP (Nicholls A. *et al.* 1991) The character of the S1 pocket of β1i was more apolar than that of β1 as shown in Fig. VII-8 (a) and (b). Thus, PGPH activity would be reduced and chymotryptic activity would be enhanced in the γ-interferon-induced 20S proteasome.

In the S1 pockets of β2, there was not such a large structural change on introduction of the γ-interferon-induced subunits (β2i) as was the case when β1i. Ser32 and Asp53 of β2 were replaced by Glus in the β2i subunit. Because the basic residue of the substrate is preferably located in the pocket as shown in Fig. VII-8 (c), the trypsin-like activity of the β2i subunit might be enhanced.

Ser53 in the S1 pocket of β5 subunit was replaced by Gln53 in the β5i S1 pocket. It was not clear how β5i subunit effect on the ability in immune

response.

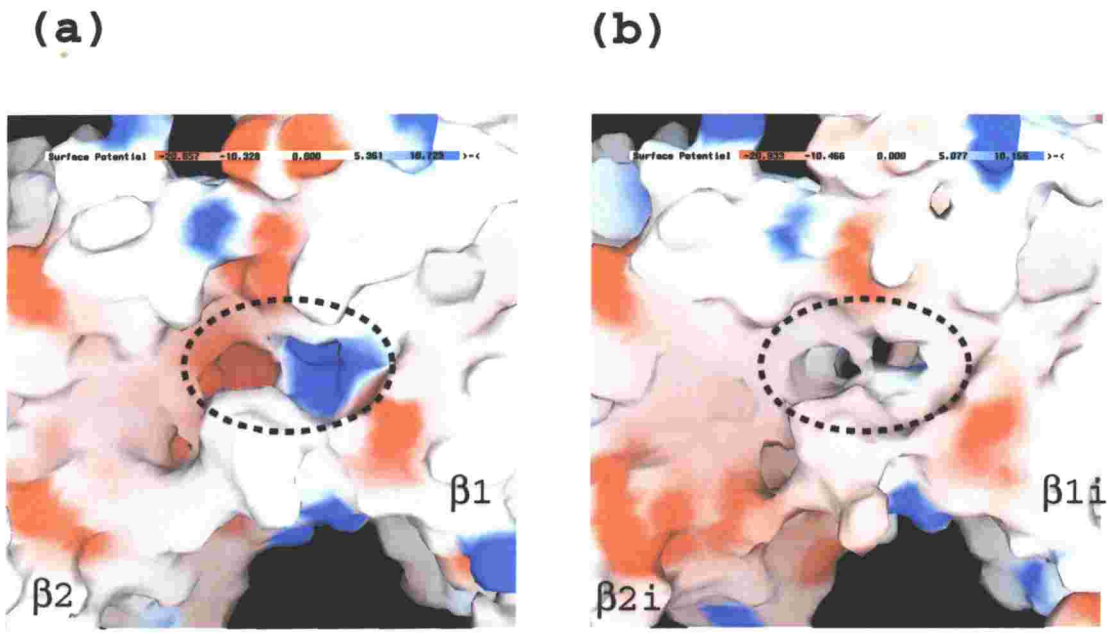


Fig. VII-8 Structure comparison of the active centers of the constitutive proteasome with the active centers of the immunoproteasome
(a), (b) Electrostatic surface potential of S1 pockets for the β1 and β1i active sites. Red and blue indicate negative and positive potentials, respectively. The S1 pocket is surrounded by a dotted circle. The S1 pocket of the β2 active center is formed by both β1 and β2 subunits.

(a) The S1 pocket of the β1 active center is charged positively at the β1 subunit side and negatively at the β2 subunit side.

(b) The S1 pocket of the β1i active center of the immunoproteasome is almost apolar.

(c)

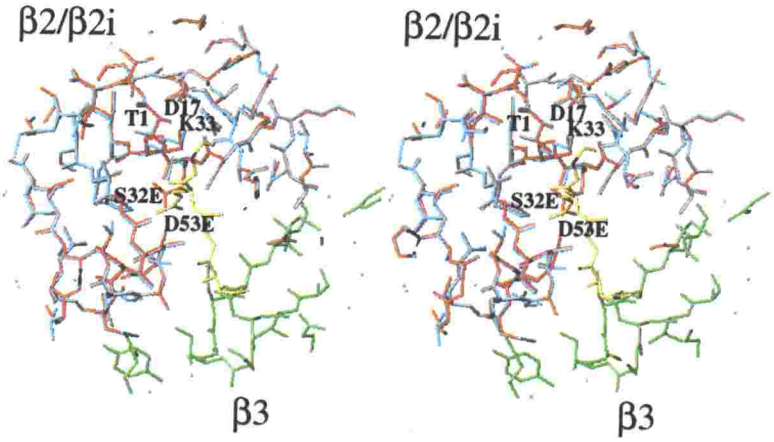


Fig. VII-8 (continued) (c) Structure comparison of the β2 active center with the β2i active is shown by a stereo pair. Amino acid residues are represented by the single letter notations with residue numbers. Both Ser32 and Asp53 in the constitutive subunit β2 (sky blue) are replaced by Glu in inducible subunit β2i (red). The green peptide is a part of the β3 subunit and the yellow one is a model substrate.

VII-6. KEKE motives and nuclear localization signals

The C-terminal helices of the α3 and α4 subunits extended out from the core with high flexibility. These regions contained lysine (K)- and glutamic acid (E)-rich KEKE motifs (Voges D. *et al.* 1999, Realini C. *et al.* 1994) in the α helices. Interestingly, the C-terminus of the α4 subunit in the 20S core particle in *Arabidopsis* can bind an Snf1-related protein kinase (SnRK) in association with the subunit of an Skp1-cullin-F-box (SCF) complex (Farrás R. *et al.* 2001). The C-terminus of the human α4 subunit interacts with the hepatitis B virus X protein (HBX) to regulate transcriptional activation of the hepatitis B virus (HBV) (Huang J. *et al.* 1996). The structures of the C-terminal segments of the α3 and α4 subunits were partially visible. Assuming helical conformations for these segments, their whole conformations were predicted (Fig. VII-9). The exposed helical rods with a diameter of about 11-16 Å exhibited a specific charge distribution on the molecular surfaces of the rods. The helical rods must interact with the SCF complex and the HBX by ionic interactions.

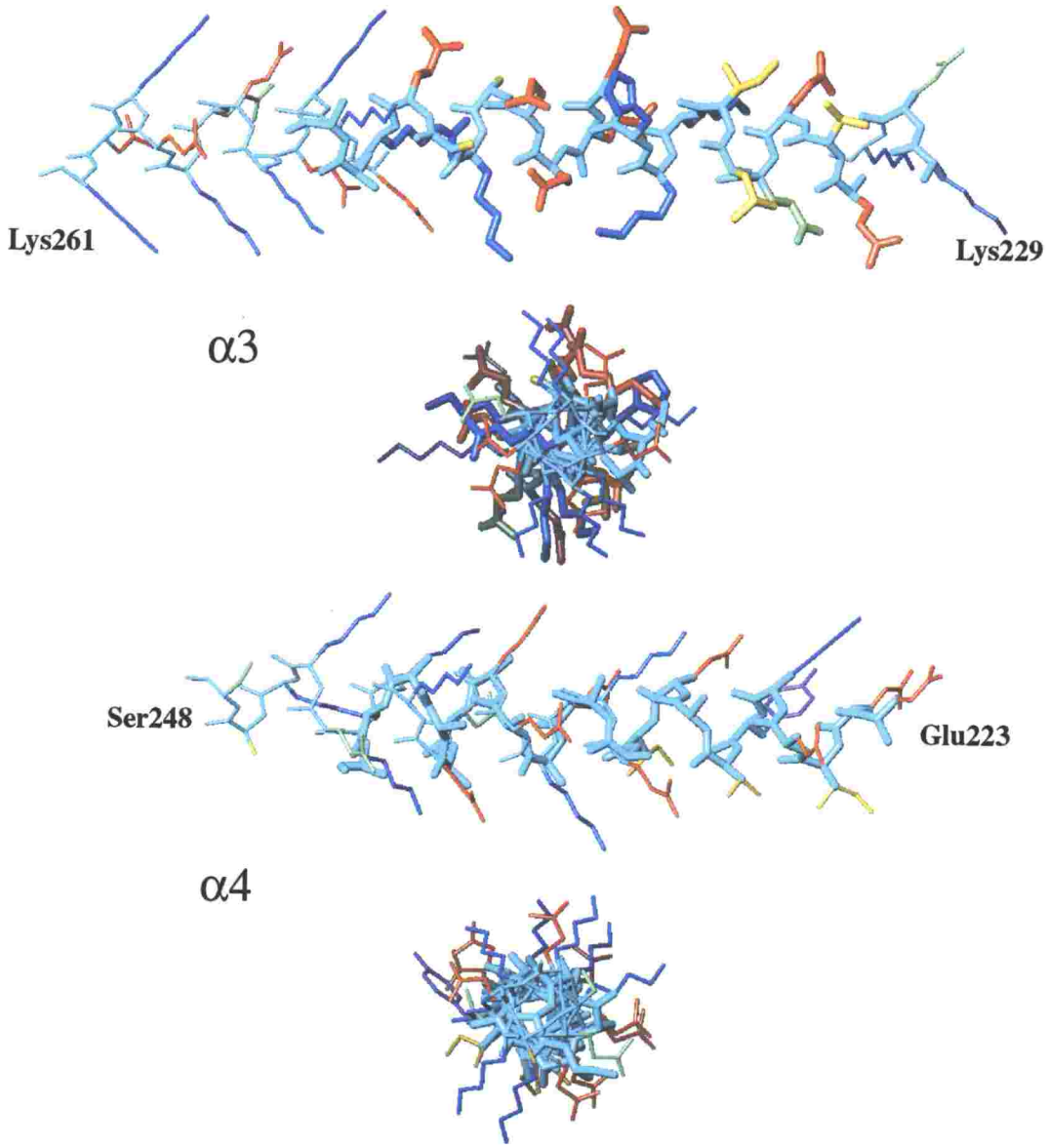


Fig. VII-9 Side views and top views of the C-terminal helices of α3 and α4 subunits.
These structures determined by X-ray crystallography are drawn by bold lines and the predicted structure are depicted by thin lines. The bones for main chains , basic residues, acidic residues, aliphatic residues, aromatic residues and the other are colored in sky blue, blue, red, yellow, magenta and light green, respectively.

The consensus sequence X-X-**K**-**K(R)**-X-**K(R)**, where X is any residue, has been proposed as a nuclear localization signal (NLS) (Roberts B. *et al.* 1989). Both of the C-termini of the α3 and α4 subunits contained this motif, too. The motif was also contained in the α1 and α2 subunits. The NLS, the residues 179-184 of α1, is in an α helix, which is not a C-terminal helix. On the other hand, the NLS, the residues 47-52 in α2, is not in a helix. This NLS segment was removed from NCS restraints in structure refinement and exhibited high flexibility judging from the temperature factors. The electron density map of the part containing the NLS was unclear. These NLSs were at the surface of the molecule (Fig. VII-10). It has been proved that all of the NLS sequences from human proteasomal subunits function as nuclear localization signals (Tanaka K. et al. 1990, Nederlof PM. *et al.* 1995). These NLS sequences are the same type as the basic amino acid cluster motif seen in simian virus 40 (SV40) T antigen. The crystal structure of the complex of importin-α and the NLS peptide of the SV40 T antigen has been reported (Fontes MR. *et al.* 2000, Conti E. *et al.* 2000), and the NLS peptide does not form a helix. Moreover, the crystal structure of importin-β bound to the importin-β binding (IBB) domain of importin-α has been analyzed (Cingolani G. *et al.* 1999). In this crystal structure, the IBB domain contains an α helix which binds to importin-β. Furthermore, the crystal structure of the autoinhibition of importin-α, in which an internal NLS, which is a part of the IBB

domain bound to the NLS binding site, has been studied (Kobe B. 1999). The NLS region contained in the IBB domain is not a helix. From these reports, the proteasome has the possibility of being imported to the nucleus by the importin-α. In this case, the NLS of α2 is caught by the importin-α with a small conformational change around the NLS to expose the signal to the molecular surface or the NLS helices of α1, α3 or/and α4 come loose from the helices to become extended forms fitting the conformation of importin-α.

One more possibility is that the 20S proteasome is directly caught by importin-β and imported to the nucleus. Because the C-terminal helices of the α3 and α4 subunits of the bovine 20S proteasome are positively charged, the importin-β can wrap these helices as the crystal structure of the importin-β-IBB complex (Cingolani G. *et al.* 1999). We show the models in which the proteasome is caught by the importin-β (Fig VII-11).

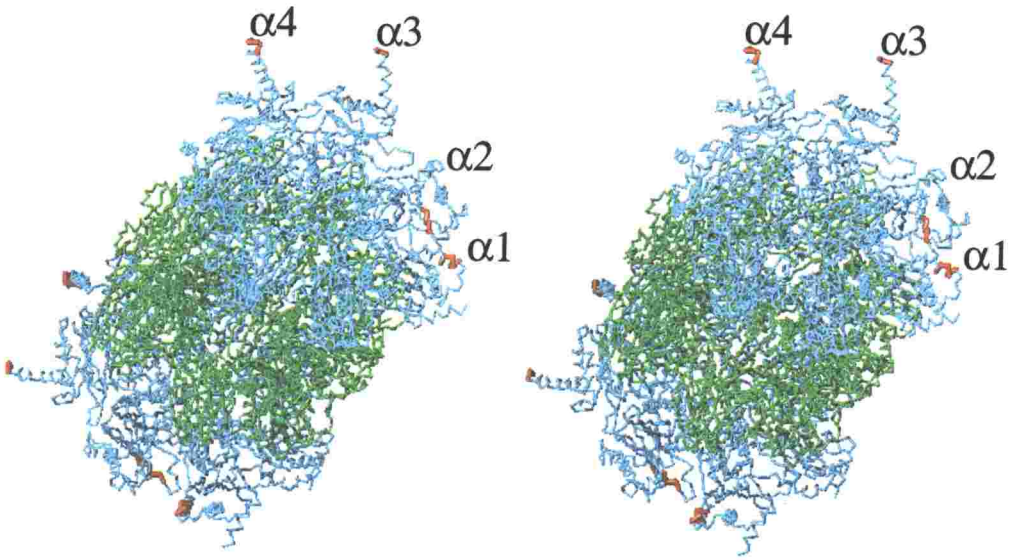


Fig. VII-10 Locations of NLSs are depicted by a stereo view of the bovine 20S proteasome. α subunits, β subunits and NLSs are colored sky blue, green and red, respectively. The NLSs of the α1, α3 and α4 subunits are helices, and that of the α2 subunit is in an extended conformation. NLSs are at the molecular surface.

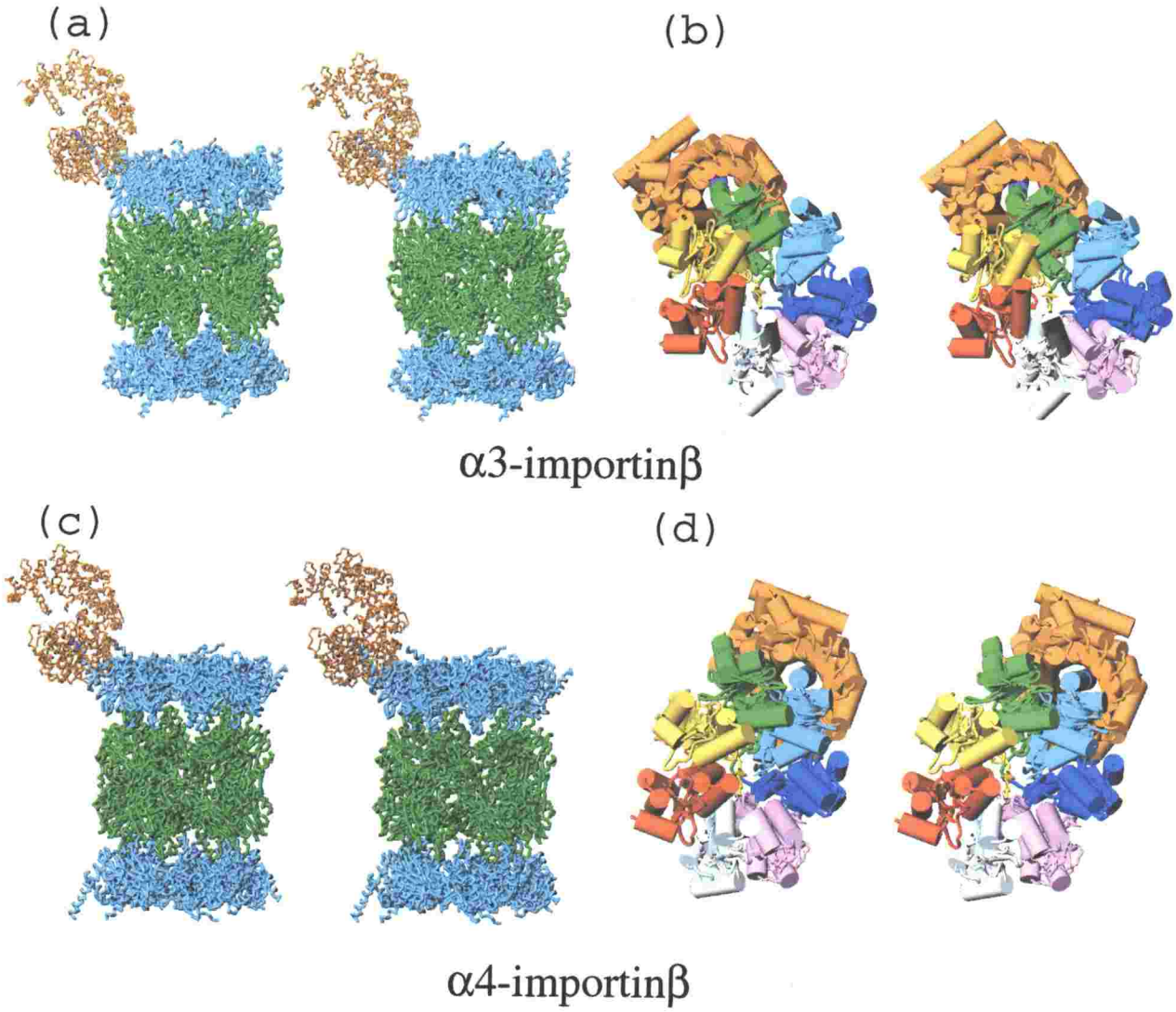


Fig. VII-11 Stereo view of putative complex model of the 20S proteasome and the importin-β.

(a) , (c) Sky blue and green indicate α subunits and β subunits of the bovine proteasome, respectively. Importin-β is colored in orange. In the case of the α3 helix is caught (a) and the α4 helix is caught (C) by importin-β.

(b), (d) In these figures, only α-rings and importin-β were depicted. Seven different subunits are in different colors; α1, red; α2, yellow; α3, green; α4, sky blue; α5, blue; α6, pink; and α7, gray.

VII-7. The central pore of the α-ring

The N-terminal polypeptide with 12 residues of *T.acidophilum* α subunits is disordered in the central pore (Löwe J. *et al.* 1995), while the corresponding N-termini of the bovine and the yeast α subunits have tertiary structures. The bovine 20S proteasome had no obvious way for the substrate to access the active center chamber because the extreme N-termini of the α2, α3, and α4 subunits and a loop structure of α5 subunit filled the central pore as in the yeast 20S proteasome (Fig.s VII-12 (b), VII-13). The ordered structure of the *T.acidophilum* α subunit around the central pore was shared by all the α subunits of the bovine and the yeast 20S proteasome. However, N-terminal polypeptides of the bovine α subunits corresponding to the disordered N-terminal regions of the *T.acidophilum* α subunit had individual tertiary structures (Fig.s VII-13, VII-14) with higher temperature factors than their β-sandwich core structures as did those of the yeast α subunit. These structural features suggested that the N-terminal residues corresponding to the disordered N-terminal residues of the *T.acidophilum* α subunit could change their conformations in binding regulatory complexes and upon introduction of a substrate as proposed by Groll *et al.* in 1997 and 2000 and Whittby *et al.* in 2000.

Both α5s of the bovine and the yeast proteasomes had long loops consisting of five and six insertion residues, respectively, between H2 and S5. The loops of the bovine and the yeast α5

subunits (Groll *et al.* 1997) exhibited individual conformations (Fig.s VII-12 (b), VII-15) with high temperature factors at the central part of the α-rings. The extra loops inserted between H2 and S5 exhibited different conformations in the bovine and yeast α5 subunits. The loop must change conformation as the N-terminal residues of the α subunits when a substrate accesses the chamber.

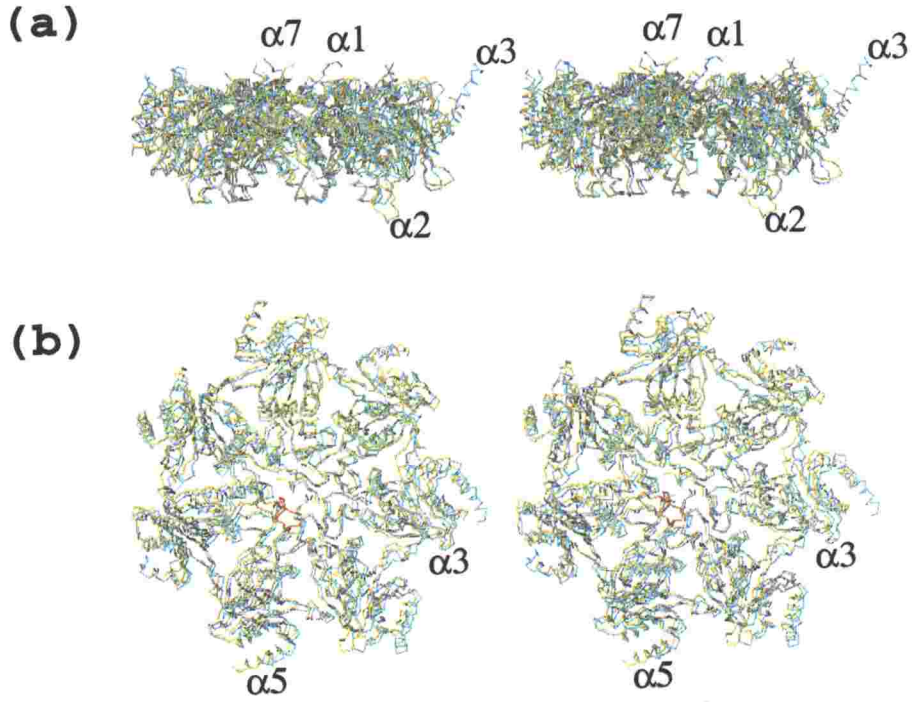


Fig. VII-12 Structural comparisons of the bovine α- and β-rings with the yeast α and β-rings are shown by stereoscopic drawings. The bovine subunits are sky blue and the yeast subunits, yellow. Several subunit names are given in each drawing.

(a) A side view of the α-rings clearly shows structural differences for the α1, α3, α5 and α7 subunits. Each α-ring interacted with a β-ring at the bottom site.

(b) A top view of the α-rings looking down to emphasize the difference near the pore, residues 124-134 of the bovine α5 are drawn in red and the corresponding residues of the yeast α5 subunit are in brown, respectively.

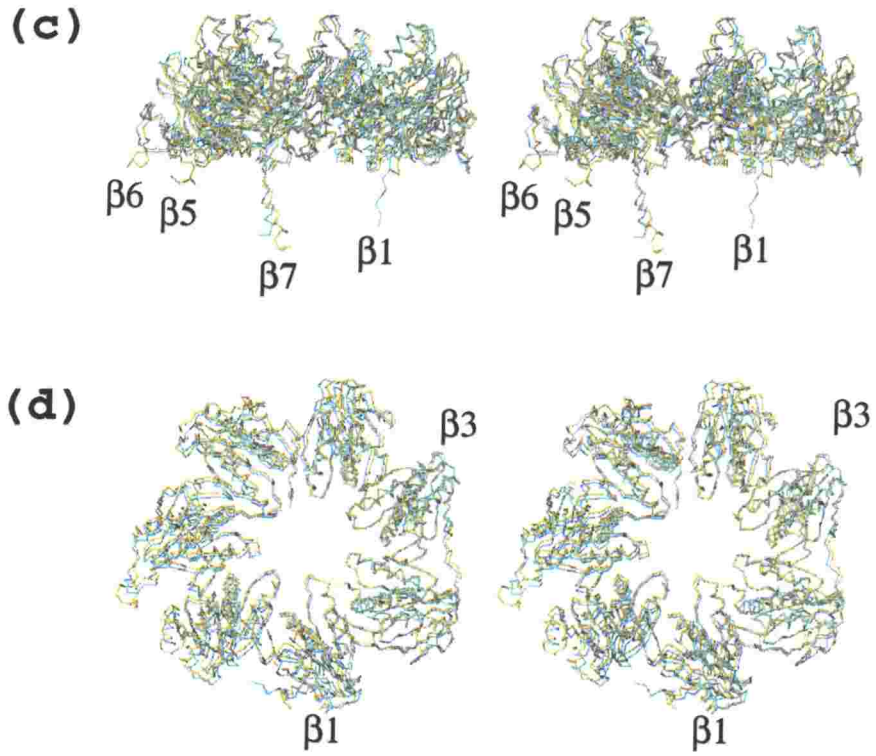


Fig. VII-12 (continued)

(c) A side view of the β-rings. The bottom site interacted with the trans β-ring related by two-fold symmetry.

(d) A top view of the β-rings looking down from the trans β-ring.

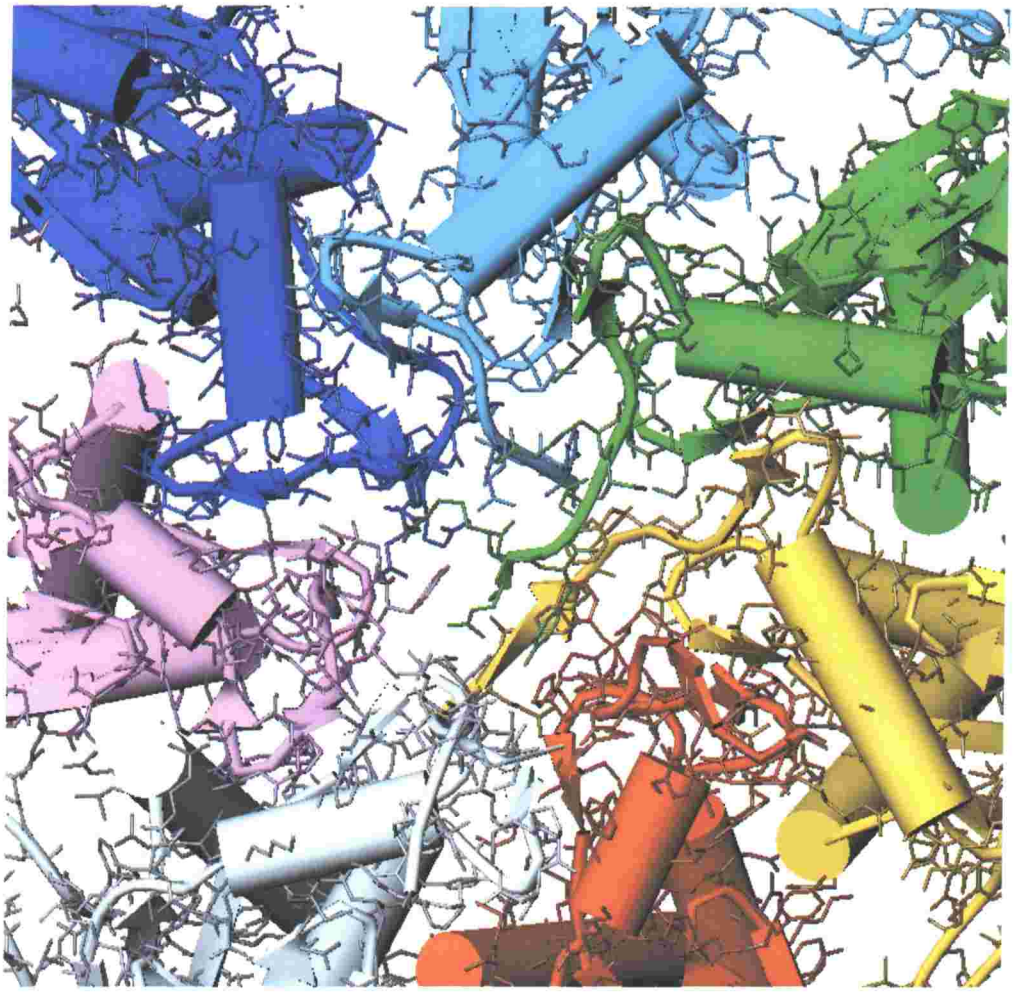


Fig. VII-13 Axial surface of the bovine proteasome emphasizing how the pore is occluded by specific α subunit tails. Seven different subunits are in different colors; α1, red; α2, yellow; α3, green; α4, sky blue; α5, blue; α6, pink; and α7, gray.

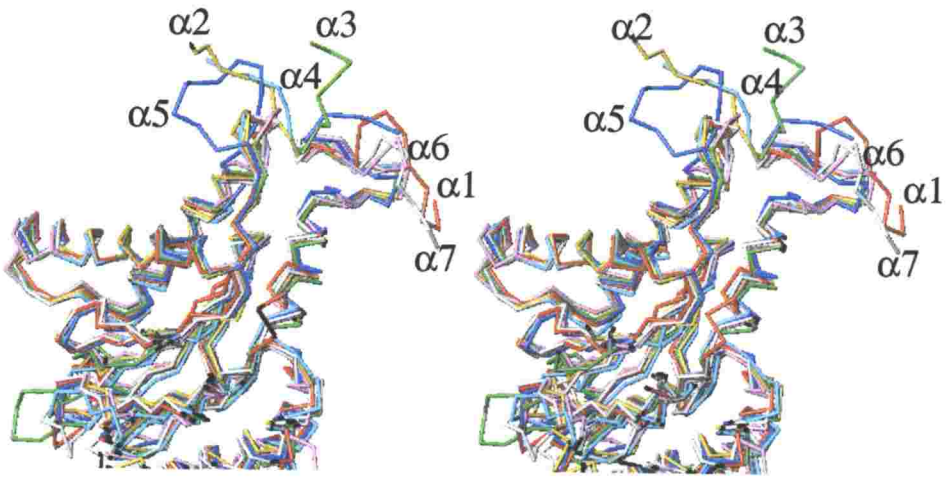


Fig. VII-14 Structural differences among the bovine subunits around the central pore are shown by a stereo view of the superposition of the bovine α subunits (coloring as in Fig. VII-13).

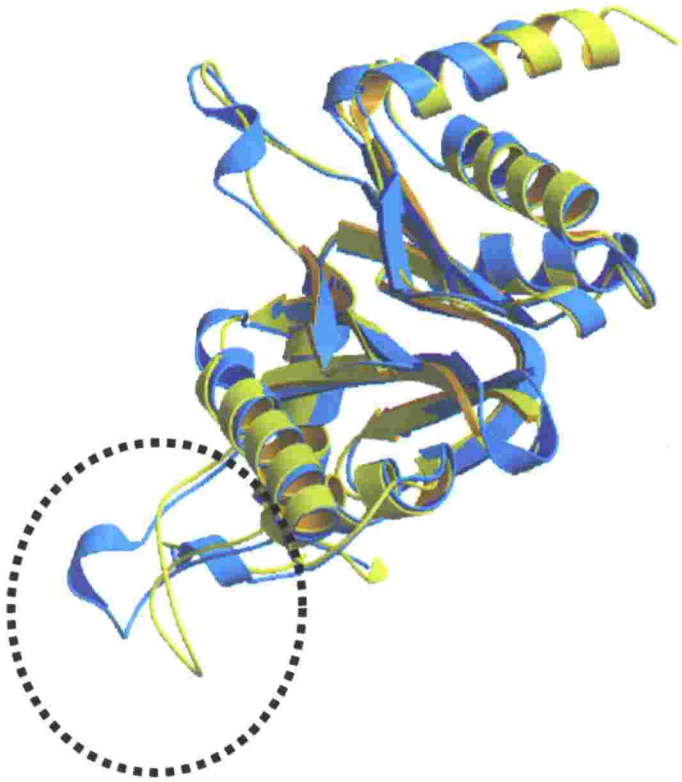


Fig. VII-15 The superposition of the α5 subunits of the bovine and the yeast proteasomes. Sky blue represents the bovine subunits and yellow shows the yeast subunits.

VII-8. Interaction across the rings and interchange of the constitutive and inducible subunits

The bovine α2 subunit lacked a long loop between the two strands of S9 and S10 of the yeast α2 subunit which interacts with residues 183-186 of Asp-Ala-Glu-Tyr and Lys40 of the yeast β2 subunit (Fig.s VII-4 (b), VII-12 (c), (d), VII-16 (a), VII-17, VII-18). These amino acid residues were equivalent to the amino acid residues 182-185 of Lys-Leu-Asp-Phe and Asn40 of the constitutive β2 subunit and those of 182-185 of Gly-Ala-Lys-Leu and Lys40 of the induced β2i subunit. These sequences for the three β2 subunits were completely different. If the bovine α2 subunit had a loop structure equivalent to the loop between S9 and S10 of the yeast α2 subunit, the bovine enzyme would be able to include neither the subunits of β2 nor β2i. The deletion of the loop in the bovine α2 subunit was requested for the bovine 20S proteasome to accommodate both subunits of β2 and β2i at the same position of the β-ring. The mammalian β5 and β5i subunits did not have seven extra residues at the C-terminal of the yeast β5 subunit, which form close contacts with the β3′ and β4′subunits (Fig.s VII-4 (b), VII-12 (c), (d), VII-16 (b), VII-18). These specific interactions in the yeast 20S proteasome were not seen in the bovine 20S proteasome structure. A loop between helix H4 and H5 of the bovine β6 subunit is shorter than that of the yeast β6 (Fig.s VII-4 (b), VII-12 (c), (d), VII-16 (c), VII-18). The long loops contact the long C-terminal extensions of β2′. The contact is tighter in the yeast rather than in the bovine.

The C-terminal region of β1 of the bovine 20S proteasome was longer than that of the yeast. The C-terminal region of the bovine β1 subunit inserted between the β1′ and β7′ subunits exhibited a flexible extended conformation with poor inter-subunit interactions such as a few hydrogen bonds (Fig.s VII-4 (b), VII-12 (c), (d), VII-16 (d), VII-18). Because the C-terminal extension of β1i was expected to have a flexible structure, it could be accommodated between β1i′ and β7′.

The long C-terminal regions of the β7 subunits of the bovine and the yeast subunits were located between both β1′ and β2′. However, the contacts between β7 and the two trans subunits of β1′ and β2′ were so loose that few hydrogen bonds existed among them. The bovine β7 subunit did not reach the opposite α-ring, while the yeast β7 C-terminal segment is so long that it contacts the trans α1 (α1′) subunit beyond the trans β-ring. Thus, the C-terminal region would not prevent the inducible subunits from being included in the 20S particle (Fig.s VII-4 (b), VII-12 (c), (d), VII-16 (e), VII-18).

The bovine 20S proteasome was less abundant in inter-subunit interactions concerning β1, β2 and β5 than that of the yeast. The residues included inter-subunit interactions with β1, β2 and β5 of the mammalian 20S proteasome were highly conserved between each pair of the γ—interferon-inducible subunits and the corresponding constitutive subunits. Exchange of the constitutive and the inducible subunits of the mammalian 20S proteasome was performed by preserving the intrinsic

inter-subunit interactions concerning β1, β2 and β5 or their corresponding inducible subunits and by removing extra inter-subunit interactions which are recognized only in the yeast 20S proteasome.

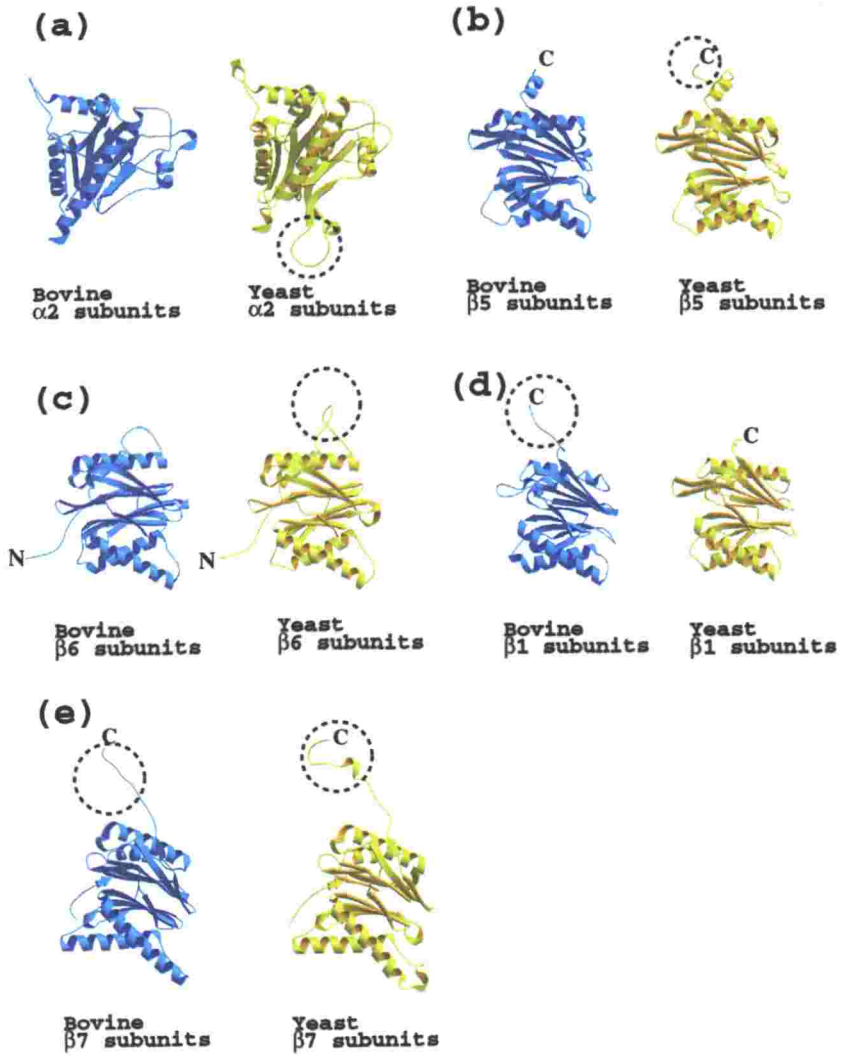


Fig. VII-16. Structural differences between the bovine and the yeast proteasomal subunits, those participate in forming quaternary structures. (a) α2, (b) β5, (c) β6, (d) β1 and (e) β7 subunits of the bovine 20S proteasome are colored in sky blue and of the yeast 20S proteasome are colored in yellow.

(a) (b)

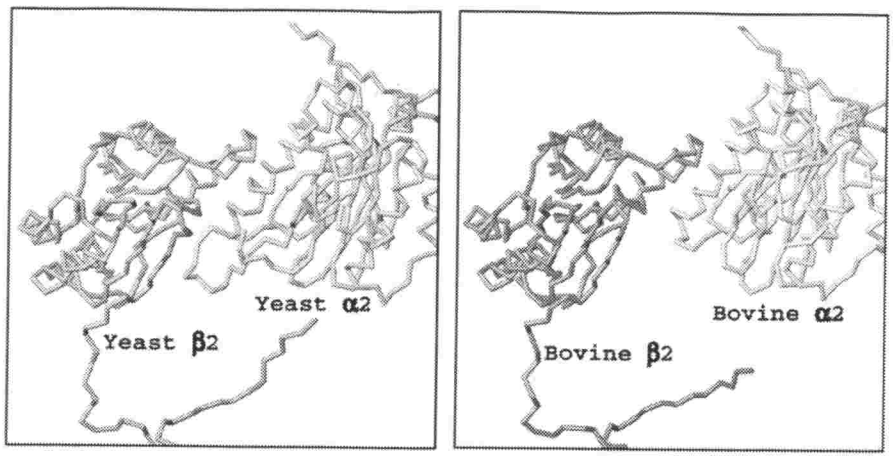


Fig. VII-17 Interaction between the α2 and β2 subunits of both of the yeast (a) and the bovine (b) 20S proteasomes.
A long loop between the two strands of S9 and S10 of the yeast α2 subunit interacts with residues 183-186 of D-A-E-Y and K40 of the β2 subunit. These amino acid residues were equivalent to the amino acid residues 182-185 of K-L-D-F and N40 of the constitutive β2 subunit and those of 182-185 of G-A-K-L and K40 of the induced β2i subunit.

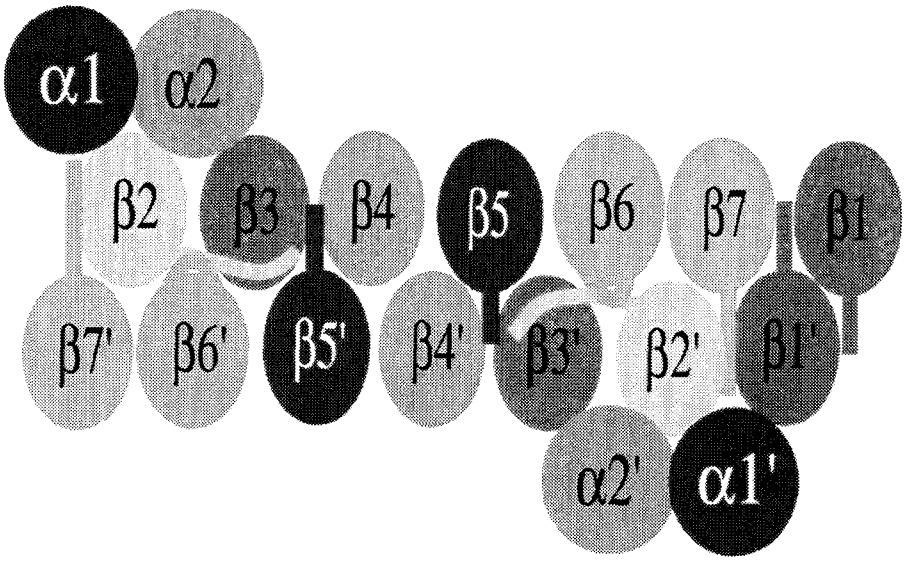


Fig. VII-18 Schematic representation of the eukaryotic 20S proteasome focused on β-type subunits.
Rolled-off cylinder envelope shows the subunit topology around β subunits and trans-β (β') subunits.

Chapter VIII. Structural organization of eukaryotic 20S proteasomes

VIII-1. Introduction

As mentioned above, the arrangement of subunits of the bovine and the yeast 20S proteasome was identical (Fig.VII-2). Eukaryotic 20S proteasomes are formed with the seven different α-type and seven different β-type subunits. Not only in the α-type subunits and in the β-type subunits, but also among the α and β subunits, their folds are well conserved. However, each subunit of eukaryotic proteasomes is located at its proper position in the assembly. It is another aim of the present research to elucidate how to built the correct quaternary structure of the enzyme complex.

Assembly of *Thermoplasma* protesomes was analyzed by expression of genes in *Escherichia coli* (Zwickl P. *et al.* 1992.). The *Thermoplasma* β subunit is synthesized as a precursor with an N-terminal propeptide of eight amino acid residues, which is autocatalytically cleaved off during proteasome assembly (Seemüller E. *et al.* and Baumeister W. 1996.). When expressed alone, *Thermoplasma* α subunits assemble into seven-member rings, whereas the precursor of β subunits and processed β subunits lacking propeptides do not form an ordered structure by themselves (Zwickl P. *et al.* 1994.). Co-expression of the α and β genes yields fully assembled and proteolytically active proteasomes, independent of the presence or absence of the β precursors. In vitro, *Thermoplasma* proteasomes can also reassemble after complete

dissociation, reemphasizing that the propeptide is not essential for assembly (Grziwa A. *et al.* 1994.).

Assembly in eukaryotic proteasomes is not fully understood, but requires a highly orchestrated process to ensure proper positioning of each of the 14 different subunits. It has been reported that the proteasome assembly depends on additional factors that probably act as chaperones and are only transiently associated with the nascent complex (Schmidtke G. *et al.* 1997, Ramos PC. *et al.* 1998). Assembly of some eukaryotic proteasome subunits was studied in vitro. When expressed alone, the human α7 subunit assembles into seven-member rings (Gerards WL. *et al.* 1997.). In contrast, the human α1 and α6 subunits do not form rings when expressed alone, but both can co-assemble with α7 into heterooligomeric rings (Gerards WLH. *et al.* 1998.). Thus, the α7 subunit is probably important early in assembly.

The first detected intermediate in eukaryotic proteasome assembly contains all the α-type subunits and a subset of the β-type subunits (β2, β3 and β4) in a 300-kDa or 13S complex (Schmidtke G. *et al.* 1997.). Subsequent incorporation of the remaining β-type subunits (β1, β5-β7) triggers the fast dimerization of these precursor complexes into the processing-component complexes. The final step in proteasome maturation is the processing event, which has been suggested to occur in the fully assembled cylinder (Seemüller E. *et al.* 1995, Chen P. *et al.* 1996, Schmidtke G. *et al.* 1996.).

We studied the proteasome assembly by the

structural analysis based on the X-ray structure of the bovine and the other 20S proteasomes. In this chapter, we discuss how the eukaryotic proteasomal subunits assemble correctly.

VIII-2. Contact areas

At first, the contact areas between two subunits of all crystal structures of proteasomes (Table VIII-1) were calculated with program AREAIMOL (Lee B. *et al.* 1971.) in CCP4 program package. There were differences in contact areas depend on subunit arrangement.

As listed in Table VIII-1, the α-α contacts were more intimate than the β-β contacts or the α-β contacts. Thus it is proper that α-rings are formed up at first and β -type subunits are attached to the α-rings (Frentzel S. *et al.* 1994, Patel S. *et al.* 1994, Schmidtke G. *et al.* 1997, Yang Y. *et al.* 1995, Nandi D. *et al.* 1997).

Hypothetical contact areas between each pair of two α-type subunits by replacing one of them with the other α-subunits were calculated per all of the seven α-type subunits. There were not noticeable differences for those contact areas (data not shown). From only these calculations, there were no contradictions even when a certain subunits were replaced by another subunit.

Table.VIII-1 Interaction area between subunits ($Å^2$)

	Bovine	Yeast	*T.acidophilum*
α1-α2	2779	2823	2695
α2-α3	2749	2768	
α3-α4	2776	3244	
α4-α5	2588	2721	
α5-α6	2480	2661	
α6-α7	2705	2649	
α7-α1	2560	2744	
β1-β2	1175	1270	1628
β1i-β2i	1071		
β2-β3	3099	3233	
β2i-β3	2974		
β3-β4	1506	1131	
β4-β5	1175	1348	
β4-β5i	1098		
β5-β6	1051	1495	
β5i-β6	1285		
β6-β7	1876	1924	
β7-β1	1479	1213	
β7-β1i	1485		
α1-β1	880	857	989
α1-β1i	938		
α1-β2	1179	1322	1061
α1-β2i	1151		
α2-β2	809	1414	
α2-β2i	878		
α2-β3	1454	1069	
α3-β3	964	958	
α3-β4	949	1074	
α4-β4	901	816	
α4-β5	747	649	
α4-β5i	747		
α5-β5	1141	826	

α5-β5i	1055		
α5-β6	919	1105	
α6-β6	768	717	
α6-β7	1102	1182	
α7-β7	652	827	
α7-β1	837	903	
α7-β1i	879		
β1-β1′	1611	1274	
β1i-β1i′	1558		
β1-β7′	2270	2720	1621
β1i-β7′	2253		
β2-β7′	1530	1826	1332
β2i-β7′	1498		
β2-β6′	2448	2879	
β2i-β6′	2412		
β3-β6′	1455	1555	
β3-β5′	1645	1950	
β3-β5i′	1670		
β4-β5′	1311	1598	
β4-β5i′	1235		
β4-β4′	2246	1396	
	Bovine	Yeast	*T.acidophilum*

VIII-3 Hydrogen bonds

Hypothetical α-rings (VIII-2) did not show any notable change in contact areas between two subunits comparing with the correct α-ring. This result does not tell us why the eukaryotic proteasomal subunits occupy their unique location in the 20S proteasomes. Since hydrogen bonds can fix two subunits together, numbers of hydrogen bonds between two subunits were evaluated.

Because our structure had rather high temperature factors, a criterion for hydrogen bond distance between non-hydrogen atoms was 2.0 Å - 3.5 Å.

At first, we picked up possible hydrogen bonds in the bovine and the yeast 20S proteasome with program CONTACT in CCP4 program package (Collaborative Computational Project, Number 4. 1994.). Next, these hydrogen bonds were inspected in computer graphics with Turbo-Frodo (Jones TA. 1978.). And we rejected or added the hydrogen bonds.

VIII-3-1. The α-rings

An α subunit was replaced by another α subunit with smallest r.m.s.d. against the concerned α subunit to built a hypothetical α-ring. Consequently seven hypothetical α-rings were constructed. The number of hydrogen bonds for each α subunit of the real and the hypothetical α-rings are listed in Table VIII-2.

Any α subunit except for the α2 subunit has more hydrogen bonds than the corresponding subunit of the hypothetical α-ring. The hypothetical model where the α2 subunit was replaced by the α4 subunit has more hydrogen bonds than the α-ring structure determined experimentally. However, the hypothetical α-ring – α1-α4-α3-α4–α5-α6-α7-α1 – of the bovine proteasome was not advantageous in making the α - β complex as in the next section. Thus, such an incorrect α-ring would be converted quickly to the intact α-ring.

Table VIII-2 Numbers of hydrogen bonds between subunits in the original α-rings and in the incorrect models in which subunits were replaced

	Bovine	Yeast
α7-α1-α2	26	30
α7-α1*-α2	14 α1* = α3	20 α1* = α4
α1-α2-α3	30	29
α1-α2*-α3	33 α2* = α4	19 α2* = α4
α2-α3-α4	35	34
α2-α3*-α4	25 α3* = α1	27 α3* = α4
α3-α4-α5	37	37
α3-α4*-α5	24 α4* = α2	27 α4* = α3
α4-α5-α6	30	28
α4-α5*-α6	23 α5* = α1	27 α5* = α2
α5-α6-α7	31	27
α5-α6*-α7	21 α6* = α2	17 α6* = α7
α6-α7-α1	31	31
α6-α7*-α1	23 α7* = α3	26 α7* = α1

VIII-3-2. The α-β and β-cis-β interactions

The hydrogen bonds between α subunits and β subunits and between β subunits and cis-β subunits which are in a half proteasome were inspected to elucidate inter-subunit interactions within β-ring and across the rings. Seven hypothetical half proteasome were constructed by replacing each β subunit with another β subunit as were the case of hypothetical α-rings. The numbers of hydrogen bonds around β subunits are listed in Table VIII-3.

In each hypothetical half proteasome, hydrogen bonds around the substituted subunit were decreased except for the yeast hypothetical β-ring where β5 subunit was replaced by the β3 subunit. The interactions between the β2 and β3 subunits are unique (Table VIII-3, Fig.VIII-1) among β-cis-β interactions in both of the bovine and the yeast 20S proteasome structures. The long C-termini of the β2 arms embrace the β3 subunits and form antiprallel β-sheets with the last strands S12 of the β3 subunits. Since these interactions are strong, the β3 subunits in eukaryotes must be caught by the β2 subunits at the initial stage of assembly and have no chance to bind with other subunits.

The main chain of the β3* subunit collided with the main chain of the β2 subunit in the bovine proteasome, the β1* subunit collided with the β2 subunit, and the β7* subunit collided with the α6 subunit in the yeast proteasome. Thus, these were unreasonable models.

Let us remember the contradiction that was mentioned in VIII-3-1. It was about the α2 subunit in the bovine proteasome. The α4 subunit could be replaced with the α2 subunit based on the number of hydrogen bonds. The tight β2-β3 interaction may solve this problem. The α2 contacted with both of the β2 and β3 subunits, and the total number of hydrogen bonds of the α2-β2 and α2-β3 was 12. When the α2 was replaced by the α4 – the most similar subunit -, the number of hydrogen bonds was 8. Thus the tight β2-β3 complex prefers the α2 subunit rather than the α4 subunit. Further, the binding surface potential in the β2-β3 complex is negative

and is a little positive in the α2 subunit, whereas is a little negative in the α4 subunit (Fig. VIII-2). The hypothetical α-ring – α1-α4-α3-α4-α5-α6-α7-α1 – is not favorable for a complex formation with the β2-β3 complex. Thus the lifetime of the hypothetical α-ring would be short.

Table VIII-3 Numbers of hydrogen bonds around original β subunits and incorrect β subunits.

Bovine

	α7	α1	β7	β2	Total
β1	4	2	11	11	28
β1*	2	4	5	0	11

	α1	α2	β1	β3	Total
β2	6	5	11	22	44
β2*	0	3	3	3	9

	α2	α3	β2	β4	Total
β3	7	3	22	9	41
β3*	3	5	22	5	35

	α3	α4	β3	β5	Total
β4	4	10	9	4	27
β4*	8	1	5	5	19

	α4	α5	β4	β6	Total
β5	10	11	4	11	36
β5*	5	1	4	6	16

	α5	α6	β5	β7	Total
β6	8	2	11	9	30
β6*	5	1	11	9	26

	α6	α7	β6	β1	Total
β7	12	2	9	11	34
β7*	6	0	9	2	17

Yeast

	α7	α1	β7	β2	Total
β1	4	3	5	11	23
β1*	3	1	4	10	18

	α1	α2	β1	β3	Total
β2	10	9	11	24	54
β2*	1	7	8	9	25

	α2	α3	β2	β4	Total
β3	4	3	24	8	39
β3*	6	3	12	2	23

	α3	α4	β3	β5	Total
β4	9	4	8	9	30
β4*	10	4	4	1	19

	α4	α5	β4	β6	Total
β5	4	6	9	13	32
β5*	2	6	10	15	33

	α5	α6	β5	β7	Total
β6	7	4	13	15	39
β6*	2	2	9	9	22

	α6	α7	β6	β1	Total
β7	3	7	16	5	31
β7*	8	3	2	6	19

β1*, β2*, β3*, β4*, β5*, β6* and β7* are β5, β5, β4, β3, β1, β7 and β4 in the bovine proteasome and β2, β1, β5, β5, β3, β7 and β6 in the yeast proteasome, respectively.

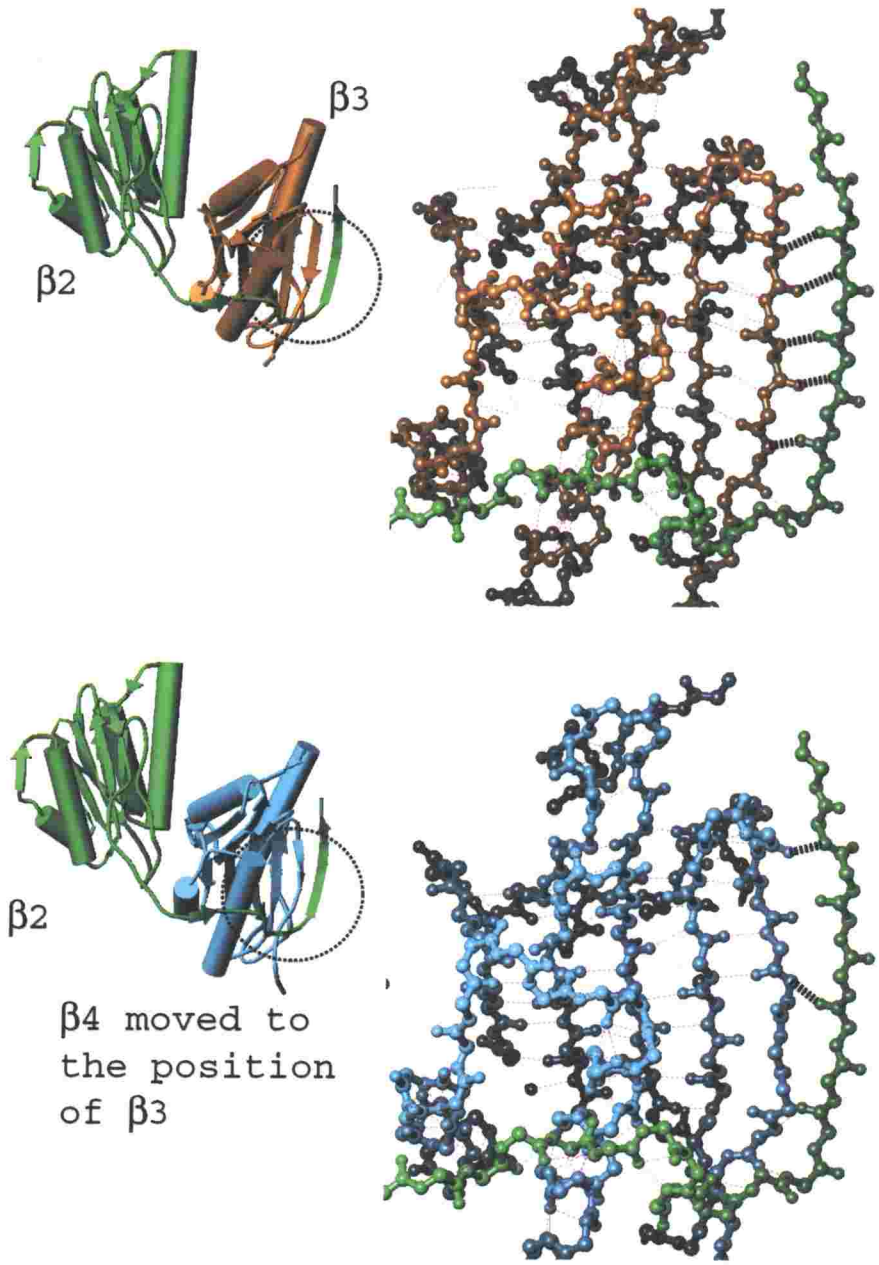


Fig.VIII-1 The tight β2-β3 complex in the bovine proteasome
Green, orange and sky blue indicate β2, β3 and β3*, respectively. - β3* is β4 moved to the β3 position. In original complex, the β2 subunit embraces the β3 subunit and make an antiprallel β sheet with the last strand of the β3 but does not make such a β sheet with the β3*.

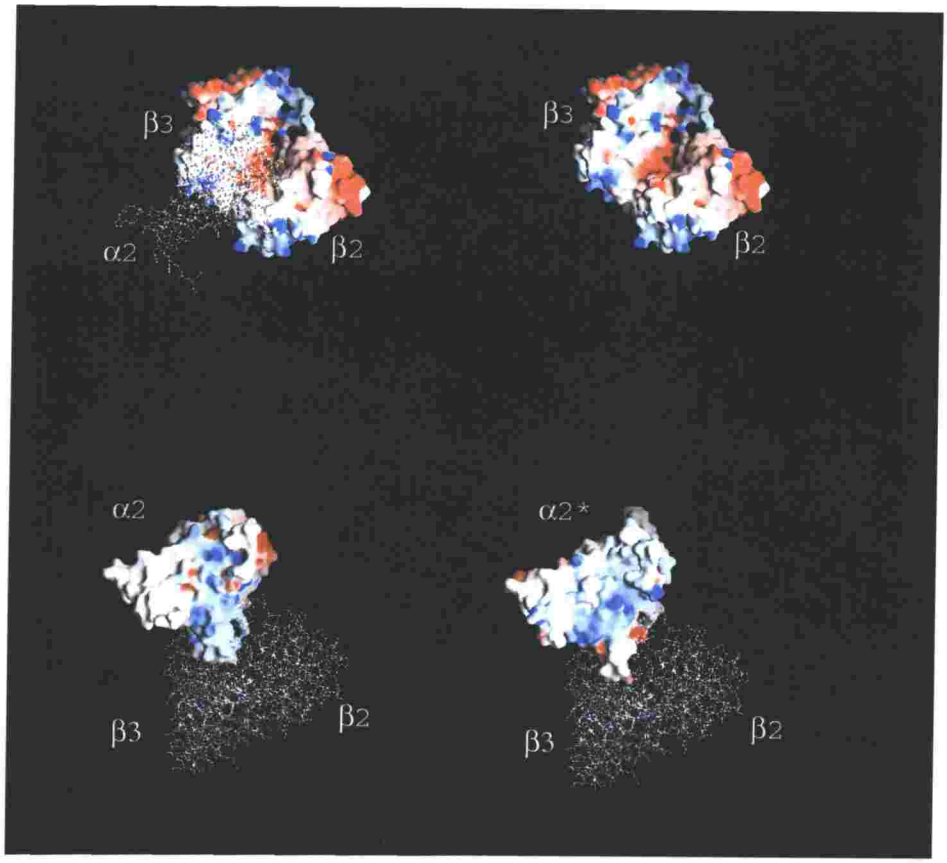


Fig. VIII-2 Electrostatic surface potential of the tight β2-β3 complex, α2 and α4.
Red and blue indicate negative and positive potentials, respectively.

VIII-3-3. The β-trans-β interactions

Now, we knew that eukaryotic 20S proteasome was established as the reasonable [α1-α7, β1-β7] complexes – half proteasomes -, correctly. Finally, we investigated about the interactions between two different β-rings (β-trans-β interactions) from the view of hydrogen bonds. The method for picking up hydrogen bonds was the same as above. In eukaryotic 20S proteasomes, the β1 subunit contacts with both of the trans-β1 (β1′) and trans-β7 (β7′) subunits. The β2 subunit contacts with both of the β7′ and trans-β6 (β6′) subunits. The β3 subunit contacts with both of the β6′ and trans-β5 (β5′) subunits. And the β4 subunit contacts with both of the β5′ and trans-β4 (β4′) subunits. In short, molecular 2-fold axes pass through from between the β1 and β1′ subunits to between the β4 and β4′ subunits. The dimeric structure was comprised with six hypothetical dimeric structures in terms of hydrogen bonds.

Six hypothetical dimeric structures were built by rotating one of the half proteasomes by $2\pi/7$, $2\times2\pi/7$, $3\times2\pi/7$, $4\times2\pi/7$, $5\times2\pi/7$ and $6\times2\pi/7$, respectively. And the hydrogen bonds were picked up by the same way (as VIII-3-1).

In Table VIII-4, the numbers of hydrogen bonds between the β-rings and the trans-β rings from both crystal structures were listed. The hypothetical dimer showed decrease in number of the hydrogen bonds and some collisions between main chains. In the $4\times2\pi/7$ rotated models of the bovine and yeast proteasomes, there were no collisions, but the other incorrect models had some collisions between

main chains (Table VIII-5).

From this result, only original models of the 20S proteasome and the $4\times2\pi/7$ rotated models were allowed to build. The original model has more hydrogen bonds to make dimeric structure than the $4\times2\pi/7$ rotated model (Table VIII-4). Thus, the dimeric structure observed in the crystal structure must be preferentially constructed.

Table VIII-4. Numbers of hydrogen bonds for the models of β-rings and trans-β rings of the bovine and yeast proteasomes.

	0	1 × 2π/7	2 × 2π/7	3 × 2π/7	4 × 2π/7	5 × 2π/7	6 × 2π/7
Bovine	180	136	108	90	97	109	88
Yeast	200	101	134	108	104	118	86

Table VIII-5. Collisions of the β-trans-β interactions of the model β-rings.

	0	1 × 2π/7	2 × 2π/7	3 × 2π/7	4 × 2π/7	5 × 2π/7	6 × 2π/7
Bovine	no	β2 and β7′	β3 and β1′ β5 and β6′	β4 and β7′	no	β6 and β1′	β5 and β2′ β7 and β1′
Yeast	no	β2 and β7′	β2 and β1′ β5 and β6′	β4 and β7′ β5 and β7′	no	β4 and β2′ β6 and β1′ β7 and β7′	β4 and β3′ β5 and β2′

VIII-4. Hierarchy of organization of quaternary structure of the eukaryotic 20S proteasome

From these results of calculating contact areas and counting hydrogen bonds between subunits, we speculated that eukaryotic 20S proteasome is built by following steps (Fig. VIII-3).

At first, the correct α-rings and the tight β2-β3 complexes are separately formed. Second, α-rings and the tight β2-β3 complexes bind to form intermediate complexes. In this step, α2 subunits (in the α-rings) are caught by the tight β2-β3 complexes. Third, the residual β-type subunits bind to the intermediate complexes in the correct positions to build the half proteasomes. Finally, the half proteasomes recognize themselves and bind together to form dimmers. The β and trans-β subunits recognize each other by their specific feature of the C-termini of the β1, β7 and β2 subunits (Fig.s VII-12 (c), (d)).

In our structural investigations, we ignored prosequences of the β-type subunits, some of which are known to work as chaperons (Chen P. *et al.* 1996, Zuhl F. *et al.* 1997). Water molecules and Mg ions were ignored upon counting hydrogen bonds. However, our calculations showed the hypothetical structure had some disadvantages in inter-subunit interactions. Our speculation is consistent with some biochemical studies (Nandi D. *et al.* 1997, Schmidt M. *et al.* 1997).

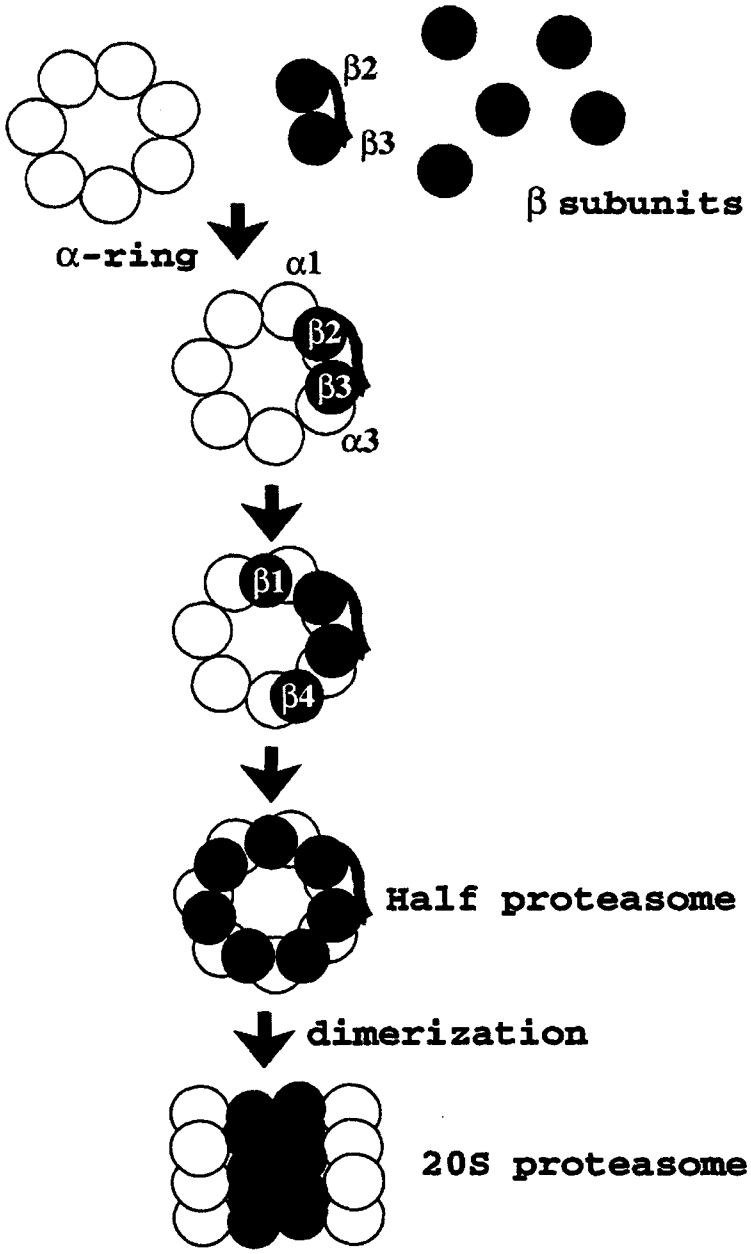


Fig. VIII-3 The maturation pathway of eukaryotic 20S proteasomes.

VIII-5. The quaternary structure of the immunoproteasome

The contact areas between the subunits of the predicted immunoproteasome (VII-5) listed in Table VIII-1 are as large as those of the constitutive proteasome. The numbers of hydrogen bonds around the model of inducible subunits are listed in Table VIII-6. Compared with hypothetical proteasomes (see Table VIII-3), they are much larger and enough to form the complex. Between the β-ring and trans-β-ring of the immunoproteasome, there were 170 hydrogen bonds. It is much larger than those of the hypothetical complexes and near that of the constitutive one (see Table VIII-4). These suggest that the inducible subunits can replace the constitutive subunits.

Table VIII-6. The numbers of hydrogen bonds around the inducible subunits of the β1i, β2i and β5i

	α7	α1	β7	β2i	Total
β1i	2	4	8	6	20

	α1	α2	β1i	β3	Total
β2i	6	6	6	24	42

	α4	α5	β4	β6	Total
β5i	10	7	7	7	31

Chapter IX. Other crystals

Besides the intensity data used in the structure determination, we have collected a number of intensity data that were not as good as the data could be used for structure determination. These crystals were poor isomorphism among them and exhibited poor resolutions. Moreover, since these crystals were quickly deteriorated by X-ray, it was difficult to collect the full data set. The statistics of these data collection are listed in Table IX-1.

Table IX-1. Other crystals

Crystal name	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	Space group	Completeness (%)	Resolution (Å)	R-factor (%)
Ps10	320.2	194.2	121.7	90.0	90.0	90.0	$P2_12_12_1$	70.2	4.0	16.4
5ps20	320.9	193.1	122.6	90.0	90.0	90.0	$P2_12_12_1$	68.4	3.3	8.0
5ps02	319.1	191.7	123.0	90.0	90.0	90.0	$P2_12_12_1$	33.3	3.0	8.5
5ps30	319.3	193.0	122.2	90.0	90.0	90.0	$P2_12_12_1$	21.0	3.0	8.6
8p126	316.7	205.7	115.7	90.0	90.0	90.0	$P2_12_12_1$	29.2	3.4	6.8
11p20	319.1	203.1	360.0	90.0	90.0	90.0	$P2_12_12_1$	96.1	4.3	11.8
8p108	317.9	207.1	116.3	90.0	90.0	90.0	$P2_12_12_1$	71.4	3.5	7.3
11p17	318.4	203.2	363.6	90.0	90.0	90.0	$P2_12_12_1$	89.3	3.7	9.9
11p13	319.7	193.2	121.4	90.0	90.0	90.0	$P2_12_12_1$	64.7	3.5	11.6
15p10	316.6	205.7	115.7	90.0	90.0	90.0	$P2_12_12_1$	96.0	3.45	6.7
16ip8	318.6	193.0	121.2	90.0	90.0	90.0	$P2_12_12_1$	81.6	4.0	5.3
15p08	205.5	231.6	632.9	90.0	89.9	90.0	C2	11.0	2.8	6.3
4np06	318.2	202.7	361.3	90.0	90.0	90.0	$P2_12_12_1$	99.9	3.8	14.0
5np07	320.6	201.7	358.6	90.0	90.0	90.0	$P2_12_12_1$	91.7	3.2	20.3
8p123	317.6	206.2	115.8	90.0	90.0	90.0	$P2_12_12_1$	98.3	4.0	10.5

Chapter X. Conclusion

Crystals of the 20S proteasome, a large multi catalytic protease, from bovine liver have been obtained from an enzyme precipitation using MPD as a precipitant. Many types of crystals – different cell dimensions or different space groups - were obtained from bovine liver and brain. The best crystals diffracted X-rays up to 2.7 Å resolution. The crystals of the enzyme belongs to cell dimensions a = 315.7 Å, b = 205.9 Å, and c = 116.0 Å, containing one molecule in an asymmetric unit.

X-ray diffraction data used for solving the structure was collected with the synchrotron radiation at the beamline of BL41XU in the SPring-8, Harima, Japan using charge coupled device (CCD). The structure was determined with the molecular replacement (MR) method. A polyalanine model truncated from the yeast 20S proteasome (Groll M. *et al.* 1997) was used as a search model in phase determination. MR was calculated at 5 Å resolution and the phase was extended from 6 Å to 2.75 Å in 400 small steps using solvent flattering, Histogram matching and non-crystallographic symmetry (NCS) averaging with program DM of CCP4 program package. A free R-factor for 5 % reflections in each shell was reduced from 0.559 to 0.284. The NCS relation between two monomers in an asymmetric unit was determined, and correlation coefficient was 0.89 in the final stage of the phase refinement. The electron density map obtained clearly showed side chains and most of the model was built.

In our crystal, only the constitutive subunits were contained, though purified solution

contained both of the constitutive and the "immuno" proteasomes (data not shown).

Cleavage sites of the proforms of the bovine β1, β2, β5, β6 and β7 subunits were equivalent to those of the yeast enzyme. The functional groups of the N-terminal hydrolase active sites were conserved well among the bovine, the yeast and the *T.acidophilum* 20S proteasomes (Groll M. *et al.* 1997, Löwe J. *et al.* 1995). The central pore of the α-ring of the bovine enzyme was closed by the ordered N-terminal residues of seven α subunits and the loop structure of the α5 subunit (Fig. VII-13) as in the case of the yeast (Groll M. *et al.* 1997).

A novel Ntn-hydrolase active site was assigned for the β7 subunit by inspecting the location of functional groups in the molecule. The active site cleft was much shallower than the S1 pockets of β1, β2 or β5 (Fig. VII-6 (c)). The active site may have SNAAP activity.

Several polypeptide segments with nuclear localization signals (NLS) were located at the molecular surface (Fig. VII-10). These polypeptide segments were expected to interact with importin-α upon transportation across the nuclear pore.

The structure of the immunoproteasome was predicted from the bovine proteasome consisting of constitutive subunits. Substrate specificities of the immunoproteasome predicted from the yeast 20S proteasome structure (Groll M. *et al.* 1997, Kloetzel PM. 2001) were confirmed by the present prediction based on the bovine 20S proteasome. In the immunoproteasome, the β1i active site reduced PGPH activity and enhanced chymotrypsin-like

activity and the β2i active site enhanced trypsin-like activity. Thus, the immunoproteasome was more advantageous for producing peptides for antigen presentation to MHC class I molecules.

The bovine proteasome gained the ability to interchange mutually the constitutive subunits and the inducible subunits by losing several inter-subunit interactions detected in the yeast proteasome.

The α-rings of the bovine proteasome had much higher temperature factors than those of the yeast proteasomes. Not only PA700 but also PA28 bind to mammalian 20S proteasomes as their activators, though counterpart genes of PA28s are not present in yeast. Thus, α-ring might have higher flexibility to bind both of those activators which have completely different structures.

The general architecture was completely conserved among the *T.acidophilum* (Löwe J. *et al.* 1995), the yeast (Groll M. *et al.* 1997) and the mammalian 20S proteasomes. Furthermore, an identical arrangement of subunits in the yeast 20S proteasome was confirmed in the bovine 20S proteasome. It is interesting that the eukaryotic 20S proteasomes do not mistake arranging their subunits. We found one of the factors that the specific hydrogen bonds between subunits decide the subunit arrangement.

Between a certain α-type subunit and the neighbor α-type subunit have much larger interactions than the α-β interactions, β-cis-β interactions (except β2-β3 interactions) or the β-trans-β interactions. Thus the α-rings must be

formed at first. The β2-β3 interaction is unique. The long C-terminal of the β2 subunits embraces the β3 subunits. Between these subunits, there are many hydrogen bonds. The C-terminal of the β2 subunits forms an antiparallel β sheet with a strand of the β3 subunits. These tight β2-β3 complexes are formed at early stage, too. The α-ring and the tight β2-β3 complex combine and the other subunits are stuck to the complexes. This is a half of the 20S proteasomes. The half proteasome binds together correctly and the 20S proteasome is built up. Numbers of hydrogen bonds told us that the incorrect arrangements are unfavorable to form complexes.

This is the first structure of the mammalian proteasome. We hope this structure can be applied in medical fields. However, to know the role of the 20S proteasome in the immune system, we must exactly reveal the structure of the immunoproteasome, not the predicted model, and analyze the differences between the constitutive and "immuno" proteasomes. Thus, we are planning to make crystals of the immunoproteasome as a next step.

Reference

Ahn JY, Tanahashi N, Akiyama K, Hisamatsu H, Noda C, Tsnaka K, Chung CH, Shimbara N, Willy PJ, Mott J. D, Slaughter CA, DeMartino GN. 1995. *FEBS Lett.* 366:37-42

Akiyama K, Yokota K, Kagawa S, Shimbara N, Tamura T, Akioka H, Nothwang HG, Noda C, Tanaka K and Ichihara A. 1994. *Science* 265:1231-1234

Bairoch A. and Apweiler R. 2000. *Nucleic Acids Res.* 28:45-48

Baldwin JM. 1980. *J. Mol. Biol.* 136:103

Baumeister W, Lupas A. 1997. *Curr. Opin. Strucut. Biol.*7:273-278

Baumeister W, Walz J, Zühl F, Seemüller E. 1998. *Cell* 92:367

Bhat TN. 1988. *J Appl Crystallogr* 21:279-281

Brünger A. 1992. *X-PLOR Version 3.1. A System for X-Ray Crystallography and NMR*

Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS. 1998. *Acta Crystallogr. D Biol. Crystallogr.* 54:905-921

Brünger AT, Adams PD and Rice LM. 1997. *Structure* 5:325-336

Cerundolo V, Kelly A, Elliot T, Trowsdale J and Townsend A. 1995. *Eur. J. Immunol.* 25:554-562

Chen P and Hochstrasser M. 1996. *Cell* 86:961-972

Cingolani G, Petosa C, Weis K and Müller CW. 1999. *Nature* 399:221-229.

Collaborative Computational Project, Number 4. 1994.*Acta Crystallogr. D.* 50:760-763.

Conti E and Kuriyan J. 2000. *Structure* 8:329-338.

Cowtan K.1994. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*. 31:34-38.

Dick TP, Ruppert T, Groettrup M, Kloetzel PM, Kuehn L, et al. 1996. *Cell* 86:253-62

Dubiel W, Pratt G, Ferrell K, Rachesteiner M. 1992. *J. Biol. Chem.* 267:22369-77

Farrás R, Ferrando A, Jásik J, Kleinow T, Ökrész L Tiburcio A, Salchert K, del Pozo C, Schell J and Koncz C. 2001. *EMBO J. 20*:2742-2756

Fehling HJ, Swat W, Laplace C, Kühn R, Rajewsky K, Müller U and von Boehmer H. 1994. *Science* 265:1234-1237

Ferrell K, Dubiel W. 1997. *Mol. Biol. Rep.* 24:83-88

Fontes MR, The T. and Kobe B. 2000.*J Mol Biol.* 297:1183-94.

Frentzel S, Pesold-Hurt B, Seeling A and Kloetzel PM. 1994. *J. Mol. Biol.* 236:975-981

Früh K, Yang Y, Arnold D, Chambers J, Wu L, Waters JB, Spies T and Patterson PA. 1992. *J. Biol. Chem.* 267:22131-22140

Früh K and Yang Y. 1999. *Curr. Opin. Immunol.* 11:76-81

Gerards WL, Enzlin J, Haner M, Hendriks IL, Aebi U et al. 1997. *J. Biol. Chem.* 272: 10080-86

Gerards WLH, Dejong WW. Bloemendal H, Bolelens W. 1998. *J. Mol. Biol.* 275:113-21

Glynne R, Powis SH, Beck A, Kelly A, Kerr lA and Trowsdale J. 1991. *Nature.* 353:357-360

Gray CW, Slaughter CA, DeMartino GN. 1994. *J. Mol. Biol.* 236:7-15

Grziwa A, Maack S, Pühler G, Wiegand G, Baumeister W, Jaenicke R. 1994 *Eur. J. Biochem.* 223:1061-67

Groettrup M, Kraft R, Kostka S, Standera S, Stochwasser R and Kloetzel PM. 1996. *Eur. J. Immunol.* 26:863-869

Groll M, Ditzel L, Löwe J, Stock D, Bochtler M, Bartunik H. D, Huber R. 1997. *Nature.* 386:463-471

Groll M, Bajorek M, Köhler A, Moroder L, Rubin DM, Huber R, Glickman MH and Finley D. 2000. *Nat. Struct. Biol.* 7:1062-1067

Head-Gordon T and Brooks CL. (1991) *Biopolymers* 31: 77-100

Hendil KB, Khan S, Tanaka K. 1998. *Biochem J.* 332:749-754

Hershko A, Ciechanover A. 1998. *Annu. Rev. Biochem.* 67:425-79

Hisamatsu H, Shimbara N, Saito Y, Kristensen P, Hendil KB, Fujiwara T, Takahashi E, Tanahashi N, Tamura T, Ichihara A and Tanaka K. 1996. *J. Exp. Med.* 183:1807-1816

Hodel A, Kim SH and Brunger AT. 1992. *Acta Cryst.* A48:851-859

Huang J, Kwong J, Sun EC-Y and Liang TJ. 1996. *J. Virol.* 70:5582-5591

Jones TA. 1978. *J. Appl. Cryst.* 11:268-272

Kelly A, Powis SH, Glynne R, Radley E, Beck S and Trowsdale J. 1991. Nature 353:667-668

Kirkpatrick S, Gelatt C. D. Jr., Vecchi M. P. 1983. *Science* 220:671-680

Kleinschmidt JA, Hügle B, Grund C, Franke WW. 1983. *Eur. J. Cell Biol.* 32:143-56

Kloetzel PM. 2001. *Nat. Rev.* 2, 179-187

Kobe B. 1999. *Nat. Struct. Biol.* 6:388-397

Kohda K, Ishibashi T, Shimbara N, Tanaka K, Matsuda

Y, Kasahara M. 1998. *J. Immunol.* 160:4923-35

Larsen CN, Finley D. 1997. *Cell* 91:431

Lee B, Richards FM. 1971. *J.Mol.Biol.* 55:379-400

Leslie AGW. 1992. *In Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography* 26.

Löwe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R. 1995. *Science.* 268:533-539

Lunin VY. 19988. *Acta Crystallogr.* A44:144-150

Lunin VY, Urzhumtsev AG and Skovoroda TP. 1990. *ActaCrystallogr.* A46:540-544

Lunin VY and Skovoroda TP. 1991. *Acta Crystallogr.* A47:45-52

Lunin VY and Verboslova EA. 1991. *Acta Crystallogr.* A47:45-52

Lupas A, Flanagan JM, Tamura T, Baumeister W. 1997. *Trends Biochem. Sci.* 22:399

Luzatti V. 1952. *Acta Cryst.* 5:802-810

Ma CP, Slaughter CA, DeMartino GN. 1992. *J. Biol. Chem.* 267:10515-10523

Martinez CK, Monaco JJ. 1991. *Nature.* 353:664-667

Martinez CK, Monaco JJ. 1993. *Mol. Immunol.* 30:1177-1183

Morimoto Y, Mizushima T, Yagi Y, Tanahashi N, Tanaka K, Ichihara A, Tsukihara T. 1995. *J. Biochem.* 117:471-474

Nandi D, Jiang H. and Monaco JJ. 1996. *J. Immunol.* 156:2361-2364

Nandi D, Woodward E, Ginsburg DB and Monaco JJ. 1997. *EMBO J.* 16:5363-5375

Nederlof PM, Wang HR and Baumeister W. 1995. *Proc. Natl. Acad. Sci. USA.* 92:12060-12064

Nicholls A, Sharp KA and Honing B. 1991. *Protein*

11:281-296
Oinonen C. and Rouvinen J. 2000. *Protein Science* 9:2329-2337
Orlowski M, Cardozo C, Michaud C. 1993. *Biochemistry.* 32:1563-72
Ortiz-Navarrete V, Seeling A, Gernold M, Frentzel S, Kloetzel PM and Hämmerling GJ. 1991. *Nature* 353:662-664
Palombella VJ, Rando OJ, Goldberg AL, Maniatis T. 1994. *Cell* 78:773-785
Patel SD, Monaco JJ. and McDevitt HO. 1994. *Proc. Natl. Acad. Sci. USA* 91:296-300
Pamer E, Cresswell P. 1998. *Annu. Rev. Immunol.* 16:323-58
Peters JM. 1994. *Trends Biochem. Sci.* 19:377-382
Püler G, Pitzer F, Zwickl P, Baumeister W. 1994. *Syst. Appl. Microbiol.* 16:734-741
Ramos PC, Hockendrff J, Johnson ES, Varshavsky A, Dohmen RJ. 1998. *Cell* 92:489-99
Read RJ. 1986. *Acta Cryst.* A42:140-149
Realini C, Roberts SW. and Rechsteiner M. 1994. *FEBS Lett.* 348:109-113.
Rechesteiner M. 1998. *Ubiquitin and the Biology of the Cell.* 147-189
Rice LM and Brunger AT. 1997. *Proteins: Structure, Function, and Genetics* 19:277-290
Rossmann MG. 1972. *Int. Sci. Rev. Ser.* 13
Rossmann MG, Blow DM. 1962. *Acta Crystallogr.* 15:26
Roberts B. 1989. *Biochim. Biophys. Acta* 1008:263-280
Rubin DM, van Nocker S, Glickman M, Coux O, Wafes I, et al. 1996. *Nature* 379:655-57
Schmidtke G, Kraft R, Kostka S, Hensklein P, Frömmel

C, Löwe J, Huber R, Kloetzel PM and Schmidt M. 1996. *EMBO J.* 15:6887-6898

Schmidtke G, Schmidt M and Kloetzel PM. 1997. *FASEB J.* 11:1235-43

Seemüller E, Lupas A, Stock D, Löwe J, Huber R and Baumeister W. 1995. *Science* 268:579-582

Seemüller E, Lupas A, Baumeister W. 1996. *Nature* 382:468-72

Sim GA. 1959. *Acta Cryst.* 12:813-815

Sim GA. 1960. *Acta Cryst.* 13:511-512

Srinvasan R. 1966. *Acta Cryst.* 20:143-144

Sussam JL. 1985. *Method in ENZYMOLOGY* 115:271-303

Tanahashi N, Muarakami Y, Minami Y, Shimbara N, Hendil K. B, Tanaka K. 2000. *J. Biol. Chem.* 275:14336-14345

Tanahashi N, Yokota K, Ahn J. Y, Chung C. H, Fujiwara T, Takahashi E, DeMartino G. N, Slaughter C. A, Toyonaga T, Yamamura K, Shimbara N, Tanaka K. 1997. *Genes Cells* 2:26410-26417

Tanaka K, Kasahara M. 1998. *Immunol. Rev.* 163:161-176

Tanaka K, Yoshimura T, Kumatori A, Ichihara A, Ikai A, Nishigai M, Kameyama K, Takagi T. 1988. *J. Biol. Chem.* 263:16209-16217

Tanaka K, Yoshimura T, Tamura T, Fujiwara T, Kumatori A and Ichihara A. 1990. *FEBS Lett.* 271:41-46.

Thompson JD, Higgins DG, Gibson TJ. 1994. *Nucleic Acids Res* 22:4673-4680

Tomisugi Y, Unno M, Mizushima T, Morimoto Y, Tanahashi N, Tanaka K, Tsukihara T, Yasuoka N. 2000. *J. Biochem.* 127:941-943

van Kaer LV, Ashton-Rickardt PG, Eichelberger M,

Gaczynska M, Nagashima K, Rock KL, Goldberg AL, Doherty AL and Tonegawa S. 1994. *Immunity* 1:533-541

Voges D, Zwickl P. and Baumeister W. 1999. *Annu. Rev. Biochem.* 68:1015-1068

Wang BC. 1985. *Method in Enzymology* 1985:90-112

Whitby FG, Masters EI, Kramer L, Knowlton JR, Yao Y, Wang CC and Hill CP. 2000. *Nature* 408:115-120.

Woolfson MN. 1956. *Acta Cryst.* 9:804-810

Yang Y, Früh K, Ahn K and Peterson PA. 1995. *J. Biol. Chem.* 270:27687-27694

Zuhl F, Seemuller E, Golbik R and Baumeister W. 1997. *FEBS Lett.* 418:189-194

Zwickl P, Lottspeich F and Baumeister W. 1992. *FEBS Lett.* 312:157-60

Zwickl P, Klinz J and Baumeister W. 1994. *Nat. Struct. Biol.* 1:765-70

List of publication

1. Unno, M., Mizushima, T., Morimoto, Y., Tomisugi, Y., Tanaka, K., Yasuoka, N. and Tsukihara, T. Structure determination of the constitutive 20S proteasome from bovine liver at 2.75 Å resolution (Dec 15, 2001. accepted). *J. Biochem.*

2. Tomisugi, M., Unno, M., Mizushima, T., Morimoto, Y., Tanahashi, N., Tanaka, K., Tsukihnara, T. and Yasuoka, N. (2000). New Crystal Forms and Low Resolution Structure Analysis of 20S Proteasome from Bovine Liver. *J. Biochem.* **127**, 941-943

3. Mizushima, T./ Unno, M. (2001). Progress in the three dimensional structural analysis of proteasome. *Experimental Medicine* **19**, 141-146

Not directly related

4. Matsuura, T., Unno, M., Sakai, H., Tsukihara, T. and Norioka, S. (2001). Purification and crystallization of Japanese pear S-RNAse associated with gametopythic self-incompatibility. *Acta. Cryst.* **D57**, 172-173

5. Matsuura, T., Sakai, H., Unno, M., Ida, K., Sato, M., Sakiyama, F. and Norioka, S. (2001). Crystal structure at 1.5 Å resolution of *Pyrus pyrifolia pistil* rebonuclease responsible for gametophytic self-incompatibility. *J. Biol. Chem.* **276**, 45261-45269