



Title	Studies on structure and molecular dynamics in proteins from hyperthermophilic archaeon
Author(s)	山本, 喬彦
Citation	大阪大学, 2006, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/46847
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

**Studies on structure and molecular dynamics in
proteins from hyperthermophilic archaeon**

(超好熱古細菌に由来する蛋白質の構造及び
分子ダイナミクスに関する研究)

Takahiko Yamamoto

Osaka University

2006

Preface

The studies presented here have been carried out under the direction of Professor Yasushi Kai at Department of Materials Chemistry, Graduate School of Engineering, Osaka University, Japan, from 2000 to 2006. The interest of this thesis is focused upon the elucidation of the relationship between structures and their function of the proteins derived from two kinds of hyperthermophilic archaeon, based on structural biology and molecular dynamics.

Takahiko Yamamoto

Department of Materials Chemistry,
Graduate School of Engineering,
Osaka University.

March, 2006

Contents

Preface

General Introduction	1
-----------------------------	----------

Chapter I:

Structural studies on TBP-interacting protein (*Tk*-TIP26) from the hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1

I.1	Introduction	3
I.2	Materials and methods	6
I.2.1	Overexpression, purification, and crystallization of native <i>Tk</i> -TIP26	6
I.2.2	Preparaion of selenomethionine-substituted (SeMet) crystals	7
I.2.3	Data collection	8
I.2.4	Structure determination	8
I.2.5	Overexpression and purification of histidine-tagged <i>Tk</i> -TBP	9
I.2.6	Preparation of <i>Tk</i> -TIP26/ <i>Tk</i> -TBP complex and SDS-PAGE analysis	10
I.2.7	Preparation of Figures	10
I.3	Results and discussion	12
I.3.1	Crystals of native and SeMet <i>Tk</i> -TIP26	12
I.3.2	Diffraction data	12
I.3.3	Phasing, density modification, and structure refinement	13
I.3.4	Overall structure	18
I.3.5	N-terminal domain	20
I.3.6	C-terminal domain	22
I.3.7	Dimer structure of <i>Tk</i> -TIP26	23
I.3.8	Formation of <i>Tk</i> -TIP26/ <i>Tk</i> -TBP complex	25
I.3.9	Inhibition mechanism of the interaction between TBP and TATA-DNA by <i>Tk</i> -TIP26	27
I.4	Summary of Chapter I	29

I.5	References	30
-----	------------	----

Chapter II:

Studies on structure and molecular dynamics of Thioredoxin Peroxidase (ApTPx) from aerobic hyperthermophilic archaeon *Aeropyrum pernix* strain K1

II.1	Introduction	35
II.2	Materials and methods	36
II.2.1	Preparation of SeC207S crystals	37
II.2.2	Preparation of ApTPx (C207S) crystals	37
II.2.3	Data collection	38
II.2.4	Structure determination and refinement	38
II.2.5	Optimization of the partial structure	39
II.2.6	Preparation of figures	39
II.3	Results and discussion	40
II.3.1	Crystals of ApTPx	40
II.3.2	Diffraction data	41
II.3.3	Structure determination and refinement of SeC207S	42
II.3.4	Structure determination and refinement for Form I and Form II of ApTPx	45
II.3.5	Structure of reduced form (SeC207S) of ApTPx	47
II.3.6	Comparison of monomer structures of ApTPx	52
II.3.7	Active site structure in the oxidized form (Form II)	54
II.3.8	Movement of Arg128 and His42	56
II.3.9	Electrostatic surfaces of the reduced and intermediate form of ApTPx	56
II.3.10	H ₂ O ₂ reduction mechanism by ApTPx	59
II.4	Summary of Chapter II	61
II.5	References	62
	Conclusion	66

List of publications	67
List of supplementary publication	68
Acknowledgement	69

General Introduction

"Archaea" is a third primary kingdom of life, which is distinct from Bacteria and Eukarya. The archaeal cells are evolutionarily more closely related to the eukaryotic cells than to the bacterial cells, especially in their transcription mechanism. In the eucaryotic cells, transcriptions of nuclear genes are initiated by RNA polymerase II and six general transcription factors, while archaeal cells require only three factors, which are homologous to eukaryal correspondent factors. This suggests that the increase of our understanding of archaeal transcription mechanism can be a prototype for understanding eucaryal transcription regulation mechanism.

Archaea live in various extreme environments and can be classified as thermophiles, halophiles and methanogens. Among them, hyperthermophilic archaea live in high temperature condition, whose optimum growth temperatures are 95°C or higher, and most proteins derived from these archaea show thermal stability. These characteristics suggest that proteins or enzymes from hyperthermophilic archaeon can be utilized to industrial application. This study focused upon proteins from two kinds of hypethermophlic archaeon found in Kodakara Island, Japan. One is *Thermococccs kodakaraensis* strain KOD1, and the other is *Aeropyrum pernix* strain K1. In this study, the relationship between structures and their function of the proteins from these archaea is discussed based on structural biology and on molecular dynamics.

"Chapter I" discusses TBP-interacting protein (*Tk-TIP26*) from the hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. TBP-interacting proteins (TIPs) are proteins that specifically interact with

TATA-binding protein (TBP), which is indispensable factor for eucaryal and archaeal transcription. TIPs exist in eucaryal and archaeal cells, some of them have been already characterized, and they involve transcription regulation. However, no three-dimensional structures of TIPs have been determined yet. In this study, I describe the three-dimensional structure of *Tk*-TIP26 and discuss its function based on structural biology in order to understand the transcription regulation mechanism by *Tk*-TIP26.

“Chapter II” discusses archaeal peroxiredoxin, Thioredoxin peroxidase (ApTPx) from the hyperthermophilic archaeon *Aeropyrum pernix* strain K1. Peroxiredoxin (Prx) is a thiol dependent peroxidase, which reduce H_2O_2 or alkyl peroxide to H_2O or alcohol, respectively. Prxs have been identified in bacterial and eucaryal cells and the three-dimensional structures of several kinds of Prxs have been already determined. Recently, the gene encoding Prx was also found in aerobic hyperthermophilic archaeon *Aeropyrum pernix* strain K1. This Prx is referred to as “Thioredoxin peroxidase (ApTPx)”, because it functions in Thioredoxin system. ApTPx contains three cysteine residues in a single polypeptide; C50, C207, and C213, and it reduces H_2O_2 to H_2O . Mutational analyses suggest that amino acid sequence of ApTPx resembles 1-Cys Prx, while it utilize two cysteine residue; C50 and C213 for its catalytic cycles, likewise 2-Cys Prx. In this chapter, I describe crystal structures of ApTPx prepared under three different H_2O_2 concentrations; no H_2O_2 , 0.1 mM H_2O_2 , and 1.0 mM H_2O_2 . Three different structures of ApTPx reveal the detailed reaction mechanism, in which H_2O_2 is reduced to H_2O by ApTPx.

Chapter I

Structural studies on TBP-interacting protein (Tk-TIP26) from the hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1

I.1 Introduction

TATA-box binding protein (TBP) plays a crucial role in basal transcription in eukaryotic and archaeal cells (Roeder 1996; Qureshi et al. 1997). In eukaryotic cells, RNA polymerases require multiple, accessory factors to interact with their respective promoters. In the case of the RNA polymerase II transcription system, at least six general transcription factors (GTFs) - TFIIA, TFIIB, TFIID (which is a large multi-subunit complex consisting of TBP and TBP-associated factors; (TAFs)), TFIIE, TFIIF, and TFIIH - are required to form a pre-initiation complex (PIC) (Roeder 1996). In archaeal cells, archaeal RNA polymerase requires at least two additional GTFs: TBP and transcription factor B (TFB) which are homologues of eukaryotic TBP, and transcription factor IIB (TFIIB), respectively (Bell and Jackson 2001; Bell et al. 2001b). Moreover, recent studies revealed that archaea have an eucaryal TFIIEa homologue (TFE), and TFE could play a stimulatory role in transcription initiation (Bell et al. 2001a; Bell and Jackson 2001; Bell et al. 2001b; Hanzelka et al. 2001). These results suggested that archaeal transcription shares high degrees of similarity with that of eucarya, and that it is simpler than eucaryal transcription. This simplicity of archaeal transcription could be a prototype in order to increase our understanding of eukaryotic transcription

mechanisms.

To further our understanding of transcription mechanisms, it is very important to understand transcription factors other than GTFs. Recent studies suggested that several new transcription factors could interact with TBP. Some of them are referred to as "TBP-interacting proteins" (TIPs), because they were isolated using the affinity chromatography with TBP(Makino et al. 1996). TIPs have been identified in both eucaryal and archaeal cells. In eukarya, the most studied TIPs are TIP120(Yogosawa et al. 1996) and TIP49(Kanemaki et al. 1997). TIP120 was isolated from rat liver and sequence analyses showed that the N-terminal region of TIP120 exhibited 44% of sequence similarity to that of *Drosophila* TAF80(Yogosawa et al. 1996). Kinetic studies and electrophoretic mobility shift assays showed that TIP120 activated the basal levels of transcription from various kinds of promoters, and also stimulated the DNA-binding activity of RNA polymerase II(Makino et al. 1999). TIP49 was also isolated from rat liver nuclear extracts using the affinity chromatography with TBP. The amino acid sequence of TIP49 contains an ATP-/GTP-binding P-loop, and exhibits sequence similarity of bacterial DNA helicase RuvB(Yamada et al. 2001). In archaea, eucaryal TIP49 homologues were identified in the genomes of *A. fulgidus* and *P. horikoshii*(Soppa 1999).TIP26 was also isolated and characterized in several species of *Thermococcales* (Matsuda et al. 1999).

In 1999, Matsuda *et al.* firstly identified TIP26 in the *Thermococcus kodakaraensis* strain KOD1 (*Tk*-TIP26). KOD1 was isolated from a solfatara at a wharf on Kodakara Island, Kagoshima, Japan. The growth temperature of this strain ranges from 65 to 100°C, with an optimal growth temperature of 95°C

(Morikawa et al. 1994; Fujiwara et al. 1998). *Tk*-TIP26 was isolated from cell lysates of this strain by affinity chromatography with TBP-agarose (Matsuda et al. 1999). Cloning of the gene encoding this protein showed that *Tk*-TIP26 is composed of 224 amino acid residues (molecular weight of 25,558) and exists in a dimeric form. Database analyses indicate that the orthologous genes are only present in the genomes of two genera, *Pyrococcus* and *Thermococcus*. The recombinant *Tk*-TIP26 and *Tk*-TBP interact with each other with an equilibrium dissociation constant, K_D of 1.24-1.46 μ M (Matsuda et al. 1999). Further studies on this protein revealed that in the absence of *Tk*-TFB, *Tk*-TIP26 prevents *Tk*-TBP from binding to TATA-DNA (Matsuda et al. 1999). On the other hand, in the presence of *Tk*-TFB, *Tk*-TIP26 does not inhibit the formation of the TFB/TBP/TATA-DNA ternary complex, but interacts with this complex to form the TIP26/TFB/TBP/TATA-DNA quaternary complex (Matsuda et al. 2001). These results suggest that TIP26 plays a role as new possible transcription regulation factor in *Thermococcales*.

Biological or biochemical studies have identified and characterized TIPs, but no 3rd-dimensional structures of TIPs have been determined, while this information is required for further understanding of the interaction between TIPs and TBP.

We report here the first crystal structure of TBP-interacting protein (*Tk*-TIP26) from the hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. In addition, we report the isolation of the *Tk*-TIP26/*Tk*-TBP complex. Characterization of this complex, and structural information about *Tk*-TIP26 and archaeal TBP gave us the insights on molecular mechanism of the interaction

between TIP26 and TBP.

I.2 Materials and methods

I.2.1 Overexpression, purification, and crystallization of native *Tk*-TIP26

The gene encoding *Tk*-TIP26 was cloned into pET-28a (Novagen). The resultant plasmid was used to transform *E. coli* HMS174(DE3)pLysS (Novagen). Cells were grown in NZCYM medium containing 100 µg/ml kanamycin at 37°C. At an OD₆₆₀ of 0.6, 1 mM IPTG was added to induce gene expression and the cultivation was continued for additional 8h. Cells were then harvested by centrifugation at 15,000 x g for 20 min at 4°C. The following procedures were carried out at 4°C.

Cells were suspended in 50 mM NaH₂PO₄ (pH 8.0) containing 500 mM NaCl, 50 mM imidazole, and 1 mM 2-mercaptoethanol (2-Me), disrupted by sonication, and centrifuged at 15,000 x g for 30 min. The supernatant was applied to a HiTrap™ Chelating HP column (Amersham Pharmacia Biotech) equilibrated with the same buffer as that for sonication lysis. The sample was eluted with a 100-250 mM imidazole gradient. The active fractions were combined and dialyzed stepwisely against the following buffers with 1 h-interval. 1) 50 mM NaH₂PO₄ (pH 8.0) containing 500 mM NaCl, 100 mM imidazole, 1 mM 2-Me, and 1 mM PMSF, 2) 50 mM NaH₂PO₄ (pH 8.0) containing 500 mM NaCl, 50 mM imidazole, 1 mM 2-Me, and 1 mM PMSF (twice), 3) 20 mM NaH₂PO₄ (pH 7.2) containing 500 mM NaCl, 50 mM Imidazole, and 1 mM DTT, and 4) 20 mM NaH₂PO₄ (pH 7.0) containing 500 mM NaCl and 1 mM DTT (twice). After dialysis, the sample was concentrated with

Centriprep YM-10 (Millipore) and applied to a HiLoad™ 26/60 Superdex 75™ prep grade column (Amersham Pharmacia Biotech) equilibrated with 20 mM NaH₂PO₄ (pH 7.0) containing 500 mM NaCl and 1 mM DTT. The active fractions were pooled, concentrated with Centriprep YM-10 (Millipore), and used for crystallization without the cleavage with a thrombin.

Crystals of *Tk*-TIP26 were obtained by the hanging-drop vapor-diffusion method. Initial condition was found using Crystal Screen 2™ (Hampton Research). After some steps of improvement, optimized crystallization was performed in the following condition: drop solution (4 µl) contained 5 mg/ml *Tk*-TIP26, 10 mM NaH₂PO₄ (pH 7.0), 250 mM NaCl, 0.5 mM DTT, 50 mM MES (pH 5.8), 6% (w/v) PEG 8000, 5% (v/v) MPD, and 5% (v/v) glycerol. Well solution (500 µl) contained 100 mM MES (pH 5.8), 7.5% (w/v) PEG 8000, 10% (v/v) MPD, and 10% (v/v) glycerol.

I.2.2 Preparation of selenomethionine-substituted (SeMet) crystals

The plasmid for overexpression of *Tk*-TIP26 mentioned above was used to transform *E. coli* B834(DE3)pLysS cells (Novagen). Transformed cells were initially precultured in LB medium containing 100 µg/ml kanamycin until OD₆₆₀ reached to 0.6. Cells were grown at 37°C in a modified minimal medium comprising 0.5% (wt/wt) glucose, 1 mM MgSO₄, 40 µg/ml each of all common amino acids except for methionine, 0.5 mg/ml each of all nucleic acids, 25 µg/ml of selenomethionine, and 100µg/ml kanamycin. When OD₆₆₀ reached to 0.6, 1mM IPTG was added to induce gene expression and the cultivation was continued at 25°C for additional 12 h. Cells were then harvested and the SeMet

protein was purified as described for the native protein.

Crystals of SeMet *Tk*-TIP26 were obtained by the hanging-drop vapor-diffusion method. Crystallization was performed in the following condition: drop solution (4 μ l) contained 7.5 mg/ml SeMet *Tk*-TIP26, 10 mM NaH₂PO₄ (pH 7.0), 250 mM NaCl, 0.5 mM DTT, 50 mM MES (pH 5.8), and 3% (w/v) PEG 8000. Well solution (500 μ l) contained 100 mM MES (pH 5.8) and 5% (w/v) PEG 8000.

I.2.3 Data collection

The native crystals were mounted directly from drops into the nitrogen stream at 100 K and the diffraction data were collected up to 2.3 Å on SPring-8 BL41XU beamline with a Mar CCD.

A MAD experiment of SeMet *Tk*-TIP26 was performed on BL40B2 beamline at SPring-8. The crystals of SeMet *Tk*-TIP26 were soaked into 100 mM MES (pH 5.8) containing 7% (w/v) PEG 8000 and 20% (v/v) glycerol for a few seconds and flash-frozen in the nitrogen-gas stream at 100 K. Three different wavelengths were chosen from inspection of the fluorescence spectrum of the crystal, corresponding to the inflection point, peak and low-energy remote. The diffraction data were recorded on ADSC Quantum 4 CCD detector.

All diffraction data were processed using MOSFLM (Leslie, 1997) and the CCP4 program suite (Collaborative Computational Project, Number 4, 1994), respectively.

I.2.4 Structure determination

The initial phases were estimated by MAD method using the SeMet data. Two of

the three selenium sites in an asymmetric unit were found using Bijvoet difference Patterson methods. These heavy atom parameters were refined, and phase calculation and density modification were processed using the program CNS(Brünger et al. 1998). The quality of the corresponding electron density map was excellent. An initial model was built using the program O (Jones et al. 1994) and TURBO-FRODO(Jones 1985). Secondary structures were assigned with DSSP(Kabsch and Sander 1983).

1.2.5 Overexpression and purification of histidine-tagged *Tk*-TBP

The gene encoding *Tk*-TBP was cloned into pET28a (Novagen). The resultant plasmid was used to transform *Escherichia coli* BL21(DE3)pLysS (Novagen). Cells were grown in NZCYM medium at 303 K comprising 100 µg/ml kanamycin. When an OD₆₆₀ reached 0.6, 1 mM IPTG was added to induce gene expression and cultivation was continued for an additional 12 h at 293 K. Cells were harvested by centrifugation at 15,000 g for 20 min. Cells were suspended in 50 mM NaH₂PO₄ (pH 8.0) containing 300 mM NaCl, 10 mM imidazole, and 1 mM 2-mercaptoethanol (2-Me), disrupted by sonication, and centrifuged at 15,000 g for 30 min. The supernatant was applied to a HiTrap™ Chelating HP column (GE Healthcare, Piscataway, NJ), equilibrated with the same buffer as that for sonication lysis. The sample was eluted along a 50-300 mM imidazole gradient. The active fractions were combined and concentrated with Centriprep YM-10 (Millipore), and applied onto a HiLoad 26/60 Superdex 75 prep-grade column (GE Healthcare, Piscataway, NJ), equilibrated with 50 mM Tris-HCl (pH 6.8), 150 mM NaCl and 1 mM DTT. The fractions containing *Tk*-TBP were pooled and

concentrated with Centriprep YM-10 (Millipore) to 8.0 mg/ml. The purified *Tk*-TBP was frozen in liquid nitrogen and stored at 193 K.

I.2.6 Preparation of *Tk*-TIP26/ *Tk*-TBP complex and SDS-PAGE analysis

In order to prepare the *Tk*-TIP26/*Tk*-TBP complex, equal volumes of purified *Tk*-TIP26 (0.6 mM) and *Tk*-TBP (0.3 mM) were mixed and incubated at 303 K for 30 minutes. The mixed solution was applied to the gel filtration column Superdex 200 10/30 HR column (GE Healthcare, Piscataway, NJ), equilibrated with 50 mM Tris-HCl (pH 6.8), 150 mM NaCl and 1 mM DTT. The active fractions (shown in Figure 4 (a)) were pooled, and used in the SDS-PAGE analysis. The stoichiometry between *Tk*-TIP26 and *Tk*-TBP was estimated by SDS-PAGE analysis under the following conditions. Several concentrations of purified *Tk*-TIP26 (8.81, 3.26, 2.29, and 2.21 μ M), *Tk*-TBP (7.70, 3.01, 2.57, and 1.92 μ M), and purified *Tk*-TIP26/*Tk*-TBP complex were applied to 12.5% polyacrylamide gel, PAGEL (ATTO), and proceeded to SDS-PAGE analysis at 20 mA constant current. The resultant gel was stained with coomassie brilliant blue (CBB) and intensities of the bands in the gel were estimated by the program Lane & Spot Analyzer (ATTO). Using these intensities, the calibration curves of *Tk*-TIP26 and *Tk*-TBP were prepared and these curves were used to estimate the concentrations of *Tk*-TIP26 and *Tk*-TBP in the complex.

I.2.7 Preparation of Figures

Figures of Figure I-4 (a), Figure I-5 (a), Figure I-5 (b), Figure I-6, Figure I-7 (a), Figure I-7 (b), Figure I-9 (a) and Figure I-9 (b) were prepared with the programs

MOLSCRIPT(Kraulis 1991), RASTER3D(Merrit and Murphy 1994), and GRASP(Nicholls et al. 1993). Figure I-4 (b) was prepared with ALSRIPT(Barton 1993), Figure 2 (c) with TopDraw(Collaborative Computational Project, Number 4, 1994; Bond 2003).

I.3 Results and discussion

I.3.1 Crystals of native and SeMet *7k-TIP26*

Crystals of native and SeMet *7k-TIP26* are shown in **Figure I-1 (a) and (b)**. Native crystals were grown up to maximum size (0.15 x 0.15 x 0.2 mm) in a week. SeMet crystals were grown up to maximum size (0.2 x 0.2 x 0.3 mm) in twelve hours.

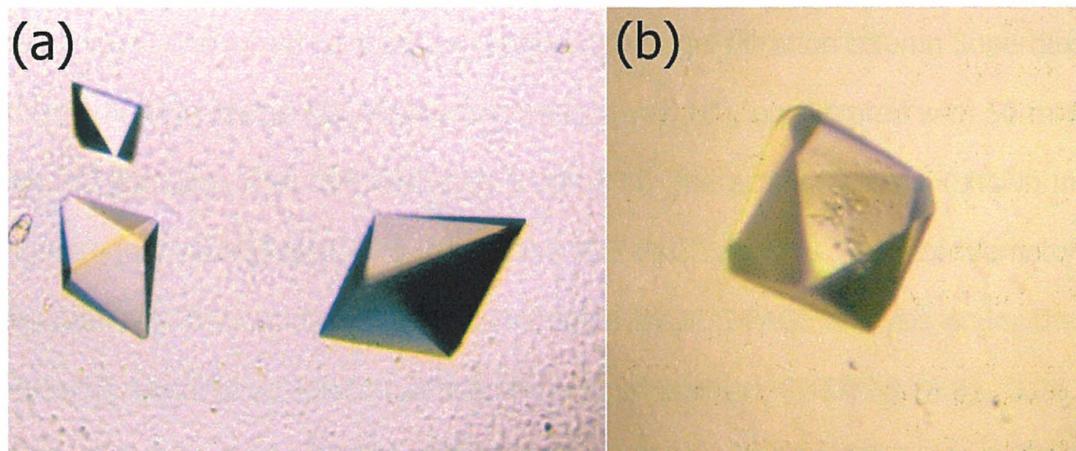


Figure I-1 Picture of crystals of *7k-TIP26* (a) Native crystals, and (b) SeMet crystals.

I.3.2 Diffraction data

Native crystals diffracted up to 2.3 Å and belong to tetragonal space group $P4_32_12$ with unit-cell parameters $a = 73.83$, and $c = 86.41$ Å. The value of the Matthews coefficient is 2.17 Å³/Da for a monomer in the asymmetric unit, corresponding to a solvent content of 43.0%, a typical value for protein crystals (Matthews, 1968).

SeMet-substituted crystals also belong to tetragonal space group $P4_12_12$ or $P4_32_12$ and its unit-cell parameters were $a = 74.34$, and $c = 86.38$ Å. The value of the Matthews coefficient is 2.14 Å³/Da for a monomer in the asymmetric unit,

corresponding to a solvent content of 42.0% (Matthews, 1968). The crystals diffracted up to 3.0 Å (peak) and 2.8 Å (inflection and remote), respectively. Details of all of the data processing and statistics were shown in **Table I-1**.

Table I-1 Diffraction data statistics of SeMet and Native 7k-TIP26

Diffraction data	SeMet			Native
Temperature (K)	100			100
X-ray source	SPring-8 BL40B2			SPring-8 BL41XU
Crystal system	Tetragonal			Tetragonal
Space group	$P4_32_12$			$P4_32_12$
Cell dimensions (Å)	$a = 74.34, c = 86.38$			$a = 73.83, c = 86.41$
V_M (Å ³ /Da)	2.14			2.17
Solvent content (%)	42			43
	Peak	Inflection	Remote	
Wavelength(Å)	0.9793	0.9795	0.9873	0.9686
Resolution(Å)	3.0 (3.16-3.00)	2.8 (2.95-2.80)	2.8 (2.95-2.80)	2.3 (2.32-2.20)
Completeness(%)	100 (100)	100 (100)	100 (100)	93.0 (94.1)
Unique reflections	5,243	6,407	6,246	11,625
Redundancy	31.6	20.2	20.1	4.6
$R_{sym}^{\#}$ (%)	7.6 (16.7)	7.1 (19.6)	7.3 (20.2)	5.6 (20.9)
$\langle I/\cdot(I) \rangle$	8.4 (3.0)	9.0 (3.5)	9.0 (3.5)	10.2 (2.4)

$^{\#} R_{sym} = \sum |I - \langle I \rangle| / \sum \langle I \rangle$ where I is the intensity of observation I and $\langle I \rangle$ is the mean intensity of the reflection.

1.3.3 Phasing, density modification, and structure refinement

The positions of selenium sites were determined by Bijvoet difference Patterson method. Bijvoet difference Patterson map using the diffraction data of peak wavelength were shown in **Figure I-2**. These maps strongly show the

existence of two selenium sites in an asymmetric unit. Using these heavy atom sites, initial phase was calculated by multiwavelength anomalous dispersion (MAD) method using the diffraction data whose resolution range was from 40 to 3.0 Å. Phasing statistics are shown in **Table I-2**. For the refinement, diffraction data of SeMet remote wavelength between 40 to 2.8 Å resolution were included in simulated annealing refinement with bulk solvent correction using CNS (Brünger et al. 1998). Ten percent of the reflections were kept separate to monitor R_{free} , and were not used in the refinement. When an R factor reached around 28%, the protein model was used to refine the native data up to 2.3 Å. After several rounds of manual rebuilding and refinement, the individual B factors were refined, along with the addition of solvent molecules. A zinc ion was placed on the site with approximately 5 sigma peak in an omit map. Zinc is the most probable metal, because it shows a tetrahedral coordination geometry, and the temperature factor of the zinc ion modeled at the site had been refined to a value similar to those of the ligand atoms. The final refinement statistics are shown in **Table I-3**. The current model contains regions of residues (-8 ~ 218) for a molecule in the asymmetric unit. No electron densities for the first 14 residues of the histidine-tag or C-terminal 6 residues were defined. These residues were eliminated from the refinement. Experimental electron density map using MAD phase and $2|F_o|-|F_c|$ map in the final refined model are shown in **Figure I-3 (a)** and **(b)**, respectively.

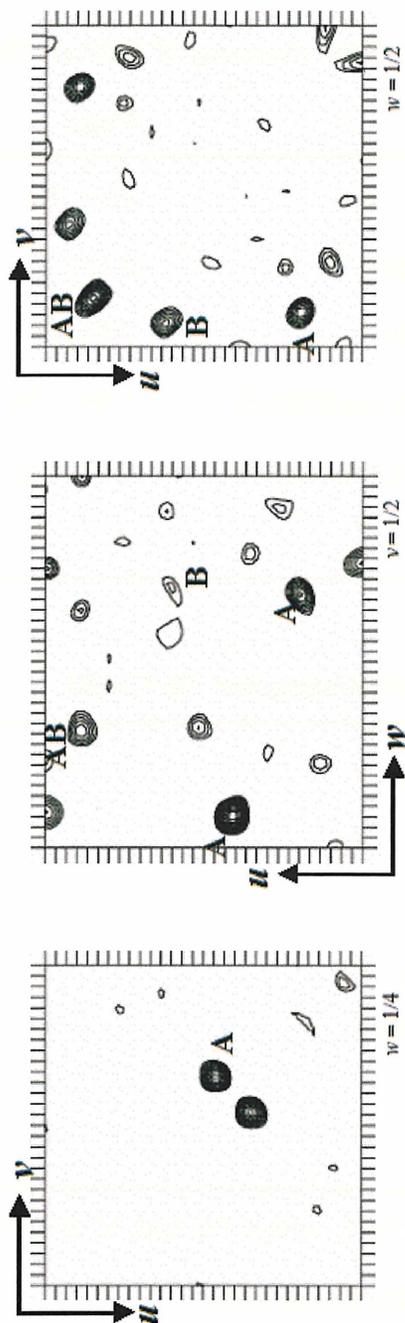


Figure I-2 Bijvoet difference Patterson map calculated from peak data at 4.0 Å resolution. The asymmetric part ($0 < u < 0.5$, $0 < v < 0.5$, $0 < w < 0.5$) of three Harker sections ($w = 1/4$, $v = 1/2$, and $w = 1/2$) are shown with contour levels starting 2σ and increasing by step of 0.5σ . The peaks A and B are Se-Se self vectors corresponding to two independent Se atoms in the asymmetric unit. The peaks AB are cross-vectors near the Harker sections.

Table I-2 Statistics for MAD phasing of SeMet *Tk*-TIP26

	Peak	Edge
Resolution (Å)	40 – 3.0	40 – 3.0
No. of sites (found/total)	2 (3)	2 (3)
R_{culis} (centric/acentric)	0.61/0.71	0.57/0.76
Phasing power (centric/acentric)	1.55/1.39	1.51/1.33
Figure of merit (centric/acentric)	0.64/0.32	0.62/0.28
Mean figure of merit	0.84/0.59	

Table I-3 Refinement statistics for *Tk*-TIP26

Resolution range (Å)	40 - 2.30
R_{cryst} (%)^a / R_{free} (%)^b	19.1/24.1
Water molecules	110
Metal ions	1
Glycerol molecules	2
Rmsd Bond length (Å)	0.005
Rmsd Bond angles (deg.)	1.2
Ramachandran (%)	
Favored	90.1
Allowed	9.9

^a $R_{\text{cryst}} = \Sigma ||F_o| - |F_c|| / \Sigma |F_c|$, where F_o and F_c are the observe and calculated structure factor amplitudes, respectively

^b R_{free} was calculated using a randomly selected 10% of the data set that was omitted through all stages of refinement

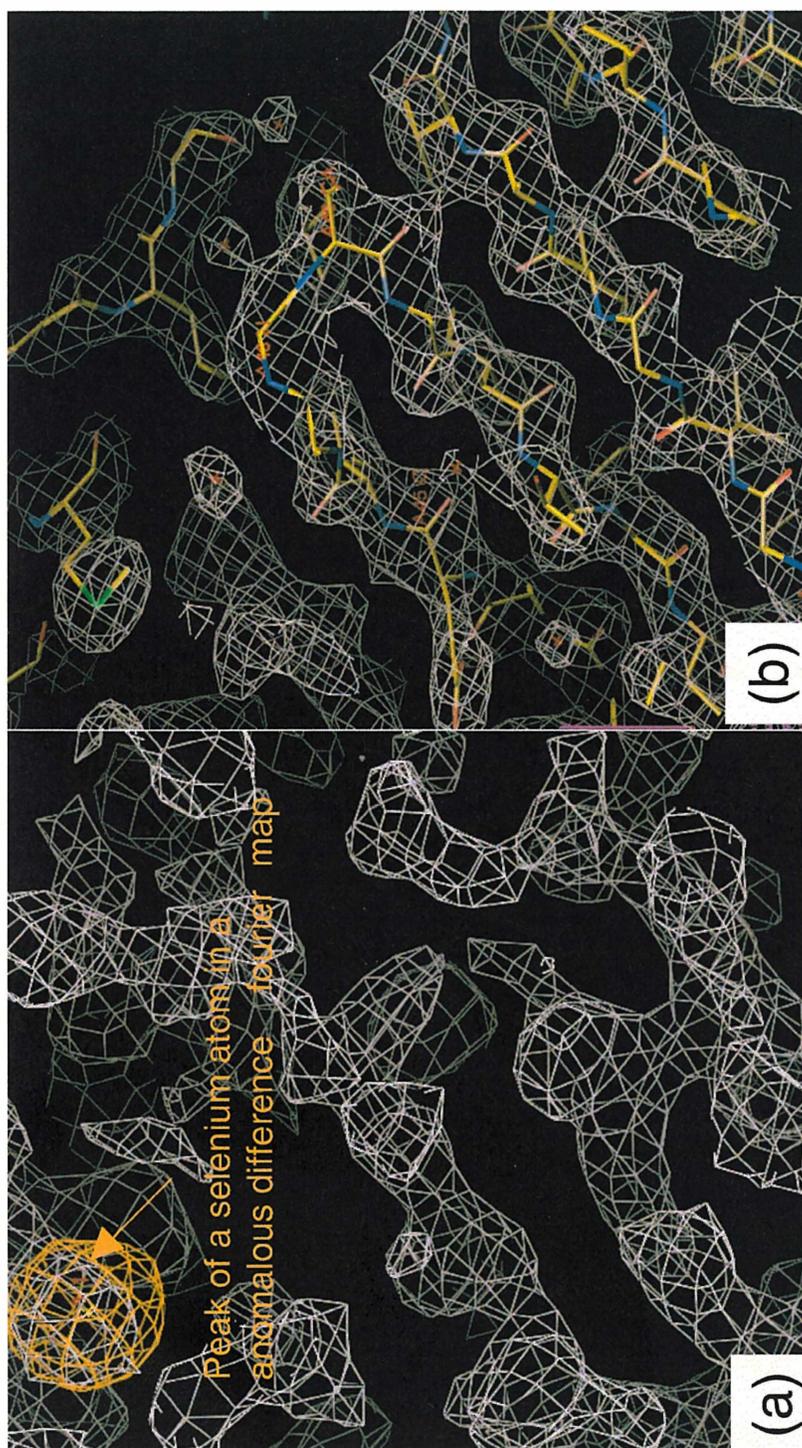


Figure I-3 Electron density maps of Tk-TIP26. (a) Experimental electron density map from MAD phase contoured at 1σ . (b) $2|F_o|-|F_c|$ electron map contoured at 1σ superimposed on the 2.3 \AA resolution coordinates of *Tk*-TIP26

I.3.4 Overall structure

We obtained $P4_32_12$ crystals of *Tk*-TIP26, and determined its structure at 2.3 Å resolution. The asymmetric unit of the crystal lattice contains one molecule of *Tk*-TIP26. The over-all structure of *Tk*-TIP26 comprises of seven α -helices (a1 - a7), twelve β -strands (b1 - b12), and two 3_{10} -helices (A1 and A2) with overall dimensions of $64 \times 49 \times 28$ Å (**Figure I-4 (a)**). Additional electron densities derived from a part of a histidine tag (his-tag) are located before Met1. The last five residues of the his-tag (Pro-5, Arg-4, Gly-3, Ser-2, and His-1) and the first three residues of *Tk*-TIP26 (Met1, Tyr2, and Ala3) form α -helix "a1" (**Figure I-4 (a) and (b)**). Side residues of Arg-4 and Ser-2 form hydrogen bonds with those of Glu109 and Arg188 in the counterparts of the symmetrically equivalent molecules, respectively. Previously, we obtained the crystals of *Tk*-TIP26 without the his-tag. However, they diffracted to only 3.2 Å or worse and they were not isomorphous (data not shown). The inter-molecular interactions derived from the his-tag could contribute to the packing stabilization of molecules in the crystal and the improvement of resolution limit.

Tk-TIP26 is composed of two domains: N-terminal domain (-8 ~ 98) and C-terminal domain (99~ 224). Proteins whose folding is similar to that of *Tk*-TIP26 were searched with DALI (Holm and Sander 1995). As a result, the N-terminal domain of *Tk*-TIP26 shared a folding similarity to proteins with the topology of type II restriction endonucleases (see below). On the other hand, the C-terminal domain does not share folding similarity with any proteins whose structures have been determined to date.

I.3.5 N-terminal domain

The N-terminal domain of *Tk*-TIP26 comprises four α -helices (a1 – a4) and six β -strands (b1 – b6), forming an α/β structure, in which the central β -sheet, formed by six β -strands, is surrounded by three α -helices, a2, a3, and a4 (**Figure I-5 (a)**). As mentioned above, the folding of N-terminal domain of *Tk*-TIP26 is similar to that of proteins with the topology of type II restriction endonucleases. Among these, archaeal holliday junction resolvase Hjc from *Pyrococcus furiosus* (*Pfu*-Hjc, PDBID: 1GEF)(Nishino et al. 2001a; Nishino et al. 2001b) is the most topologically similar (**Figure I-5 (a)** and **(c)**). Interestingly, the superimposition between the N-terminal domain of *Tk*-TIP26 and monomeric *Pfu*-Hjc shows an r.m.s.d. value for C $_{\alpha}$ carbon atoms of 3.5 Å, even though their sequence similarity is below 10%. However, the electrostatic surface of the N-terminal domain of *Tk*-TIP26 is remarkably different from that of *Pfu*-Hjc. The N-terminal domain of *Tk*-TIP26 is mainly negatively charged, while *Pfu*-Hjc is positively charged (**Figure I-5 (b)**). In addition, the arrangements of *Tk*-TIP26 in the dimer are notably different from that of *Pfu*-Hjc(Nishino et al. 2001a; Nishino et al. 2001b). *Tk*-TIP26 forms a dimer via the interaction between the N-terminal domain and C-terminal domain in the counterpart of the dimer (**Figure 4 (a)**), while *Pfu*-Hjc is dimerized with a single domain. The contact of the dimeric *Tk*-TIP26 involves hydrogen bonds (see below), while that of *Pfu*-Hjc involves hydrophobic interaction(Nishino et al. 2001a; Nishino et al. 2001b). These results suggest that *Tk*-TIP26 does not have the same function as in *Pfu*-Hjc, even though these two proteins share similar topology.

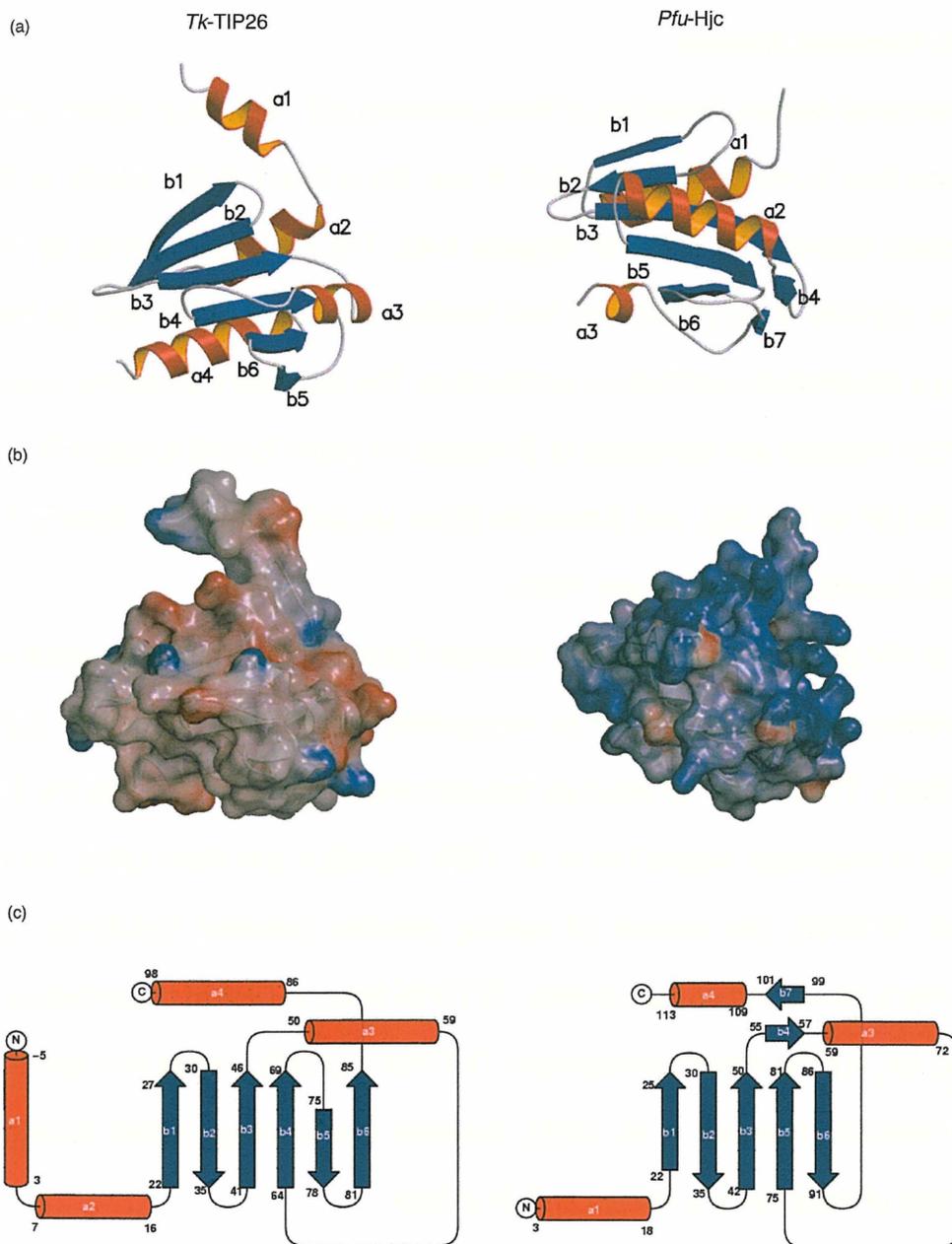


Figure I-5 Structural comparison between the N-terminal domain of *Tk-TIP26* (left image), and *Pfu-Hjc* (right image). (a) Ribbon diagrams. (b) Electrostatic surfaces generated by GRASP. Coloring is according to the local electrostatic potential range from -10kT^{-1} (red) to $+10\text{kT}^{-1}$ (blue). (c) Topology diagrams.

I.3.6 C-terminal domain

The C-terminal domain comprises of three α -helices (a5 - a7), six β -strands (b7 – b12), and two 3_{10} -helices (A1 and A2). A zinc ion is found in the vicinity of the loop region between b9 and b10 (**Figure I-6**). The ion is coordinated by two cysteine (Cys149 and Cys152) and two histidine residues (His164 and His168), forming a tetrahedral coordination architecture. The secondary structures around these four residues are composed of β -strands b9 (from Tyr140 to Lys147), b10 (from Tyr154 to Asp156), and α -helix a6 (from Leu160 to Arg165), forming the classical Cys₂His₂-type zinc finger motif.

Most of the proteins containing the classical Cys₂His₂-type zinc finger motif are transcription factors that function by recognizing specific DNA sequences (Laity et al. 2001). The number of residues between the second cysteine and the first histidine is invariantly twelve (Lee et al. 1989; Pavletich and Pabo 1991). In the case of *Tk*-TIP26, the number of spacing residues between Cys-X₂-Cys and His-X₃-His is eleven rather than twelve. *Tk*-TIP26 binds to DNA in the presence of the zinc ion, but the interaction between *Tk*-TIP26 and DNA is weak and seems to be non-specific (Matsuda et al. 1999). However, the topology of this domain is composed of “exactly” a β -hairpin and an α -helix.

The zinc finger motif in *Tk*-TIP26 is an unusual Cys₂His₂-type zinc finger and is not functional as a DNA-binding module, probably because *Tk*-TIP26 possesses only a single zinc finger motif. Zinc finger proteins usually contain multiple zinc finger motifs, which are required for specific interaction with DNA (Pavletich and Pabo 1991). In addition, this motif should not have the critical role for the stability of overall structure, because its mutant protein C149A/C152A could be

overexpressed and purified as in wild-type protein (Matsuda et al. 1999).

This zinc finger motif is the first structure of Cys₂His₂-type zinc finger motif among archaeal proteins. However, The role of the motif in *Tk*-TIP26 is not unknown and should be elucidated.

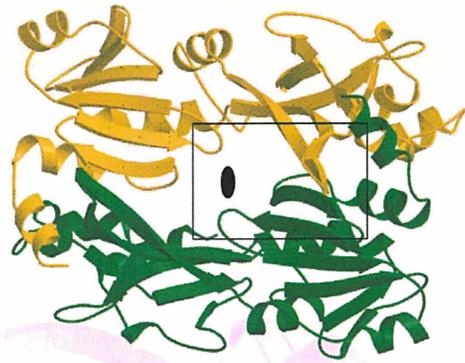


Figure I-6 Ribbon diagram of Zn-finger motif. Electron density map covered zinc ion is Simulated Annealed Omit Map contoured at 5.0 σ .

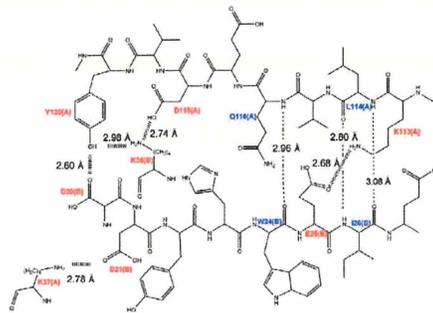
I.3.7 Dimer structure of *Tk*-TIP26

Previously, it was reported that *Tk*-TIP26 exists in a dimeric form (Matsuda et al. 1999). The structure of dimeric *Tk*-TIP26 is shown in **Figure I-7 (a)**. Two molecules of *Tk*-TIP26 are related by the crystallographic two-fold axis and the contact surface area is calculated to be 2,260 Å². The dimer interface is mainly composed of the β -strands b1 and b8 in the counterpart of the dimer (**Figure I-7 (b) and (c)**). The contact involves hydrogen bonds between side residues of Lys37, Tyr120, Asp118 and Lys113 in one molecule, and Asp21, Asp20, Lys36 and Glu25 in the other, respectively. Hydrogen bonds are formed between main chains of Gln116 and Leu114 in one molecule, and Trp24 and Ile26 in the other, respectively (**Figure I-7 (b) and (c)**).

(a)



(c)



(b)

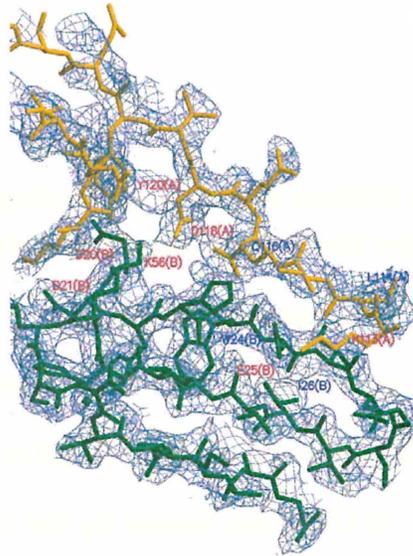
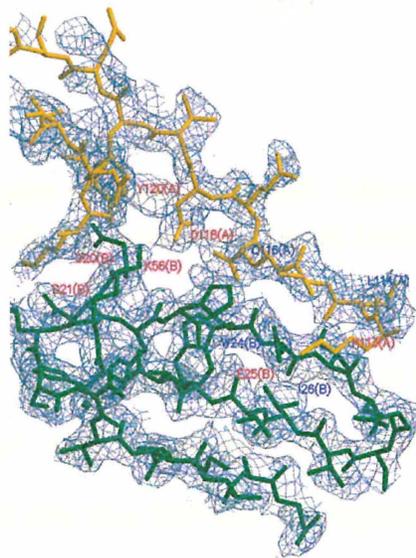


Figure I-7 (a) Dimer structure of *7k*-TIP26. The black ellipse in the middle of the dimer represents a crystallographic 2-fold axis. The area in the black box is magnified in (b). (b) Closed-up view of the dimer interface in stereo projection. The $2|F_o| - |F_c|$ electron density map was contoured at 1.0σ . (c) The details of the dimer interactions. The residues involving the interaction between side residues are colored in red, and residues involving the interaction between main chains are colored in blue.

I.3.8 Formation of *Tk*-TIP26/ *Tk*-TBP complex

Purified *Tk*-TIP26 and *Tk*-TBP were mixed and incubated at 303 K for 30 minutes. After incubation, this mixture was applied to gel filtration column chromatography to purify. The chromatograph shows a single peak with molecular weight of approximately 158 kDa, showing that *Tk*-TIP26/*Tk*-TBP complex is obtained (**Figure I-8 (a)**). This value is in agreement with an estimate of molecular weight of this complex using dynamic light scattering (DLS) (data not shown). The peak fraction was applied to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The resultant gel shows just two bands (**Figure I-8 (b)**): one derived from *Tk*-TIP26 (approximate molecular mass of 27 kDa), and the other, from *Tk*-TBP (approximate molecular mass of 23 kDa). For the estimation of the stoichiometry of the interaction between *Tk*-TIP26 and *Tk*-TBP, several concentrations of *Tk*-TIP26, *Tk*-TBP, and the *Tk*-TIP26/*Tk*-TBP complex were loaded to SDS-PAGE on the same gel. The resultant gel was analyzed for band intensities in order to produce calibration curves of *Tk*-TIP26 and *Tk*-TBP (data not shown). The calibration curves indicate that in the lane of *Tk*-TIP26/*Tk*-TBP complex, 0.126 pmol of *Tk*-TIP26 and 0.0675 pmol of *Tk*-TBP were present, i.e. the molar ratio of the interaction between *Tk*-TIP26 and *Tk*-TBP is approximately 2:1. Together with the estimation of molecular weight of *Tk*-TIP26/*Tk*-TBP complex, *Tk*-TIP26 and *Tk*-TBP interact with stoichiometry of 4:2.

As mentioned above, *Tk*-TIP26 usually exists in dimeric form (Matsuda et al. 1999). On the other hand, biochemical and crystallographic studies reported that TBP usually exists in a dimeric form in solution and crystal structures (Nikolov and

Burley 1994; Coleman et al. 1995; DeDecker et al. 1996). These results suggest that 2 dimer molecules of *Tk*-TIP26 interact with 1 dimer molecule of *Tk*-TBP.

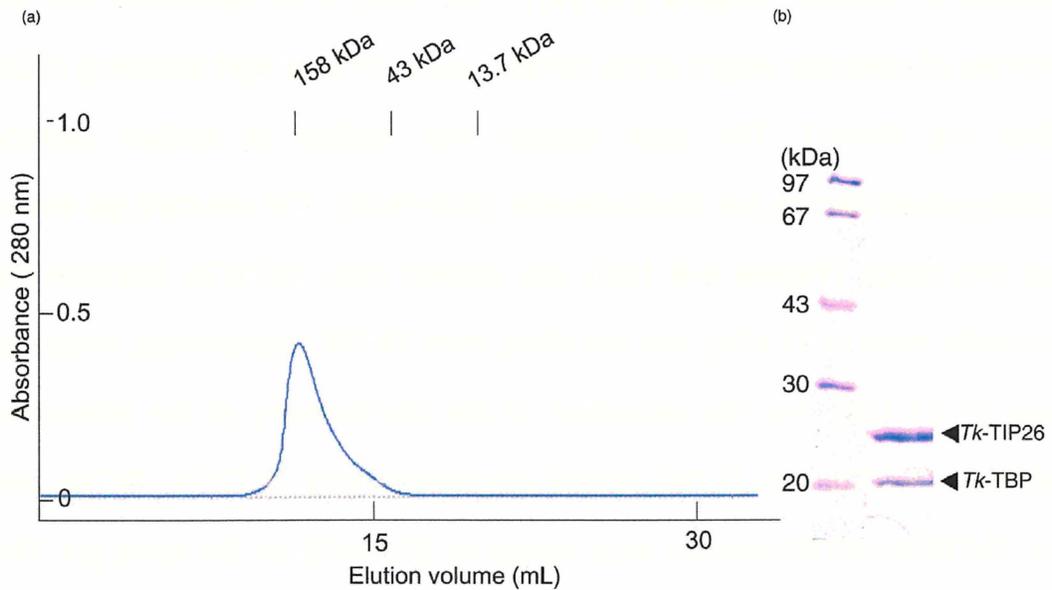


Figure I-8 Interaction between *Tk*-TIP26 and *Tk*-TBP. (a) The elution pattern of *Tk*-TIP26/*Tk*-TBP complex in gel filtration column chromatography. (b) SDS-PAGE of peak fractions of the chromatograph in (a).

I.3.9 Inhibition of the interaction between TBP and TATA-DNA by *Tk*-TIP26

In the absence of *Tk*-TFB, *Tk*-TIP26 prevents *Tk*-TBP from binding to TATA-DNA (Matsuda et al. 1999; Matsuda et al. 2001). We compared the structures and the electrostatic surfaces of *Tk*-TIP26 and TBP from *Pyrococcus woesei* (DeDecker et al. 1996) (*Pw*TBP, PDBID:1PCZ), which shares 83% of amino acid similarity with *Tk*-TBP. In the structure of dimeric *Tk*-TIP26, negatively charged cleft is formed after formation of the dimer. (**Figure I-9 (a) left image**). On the other hand, dimeric *Pw*TBP possesses specifically positively charged surfaces composed of Arg78, Lys82, Lys86, Lys88, Lys93, and Lys95 on both ends of the surface of the dimeric *Pw*TBP (**Figure I-9 (a) right image**). Subsequently, a model of *Tk*-TIP26/TBP complex was predicted (**Figure I-9 (b)**). In this model, dimeric TBP is sandwiched by two molecules of dimeric *Tk*-TIP26 from both ends of the surface, forming a 4:2 complex. Recent studies suggested that the equilibrium between TIP26/TBP complex and TBP/TATA-DNA complex inclines toward the fraction of the former complex in the absence of TFB(Matsuda et al. 2001). In addition, the 4:2 complex cannot access to TATA-DNA in the absence of TFB(Matsuda et al. 1999), while dimeric TBP can interact with TATA-DNA by the dissociation to monomers(Coleman et al. 1995). These results suggest that the formation of the 4:2 complex would stabilize TBP in dimeric form, although it has not been elucidated how the interaction between the monomers of dimeric TBP would be affected by the formation of the 4:2 complex yet. This is the first structural basis of the inhibition mechanism where *Tk*-TIP26 prevents TBP from binding to TATA-DNA in the absence of TFB (**Figure I-9 (c)**).

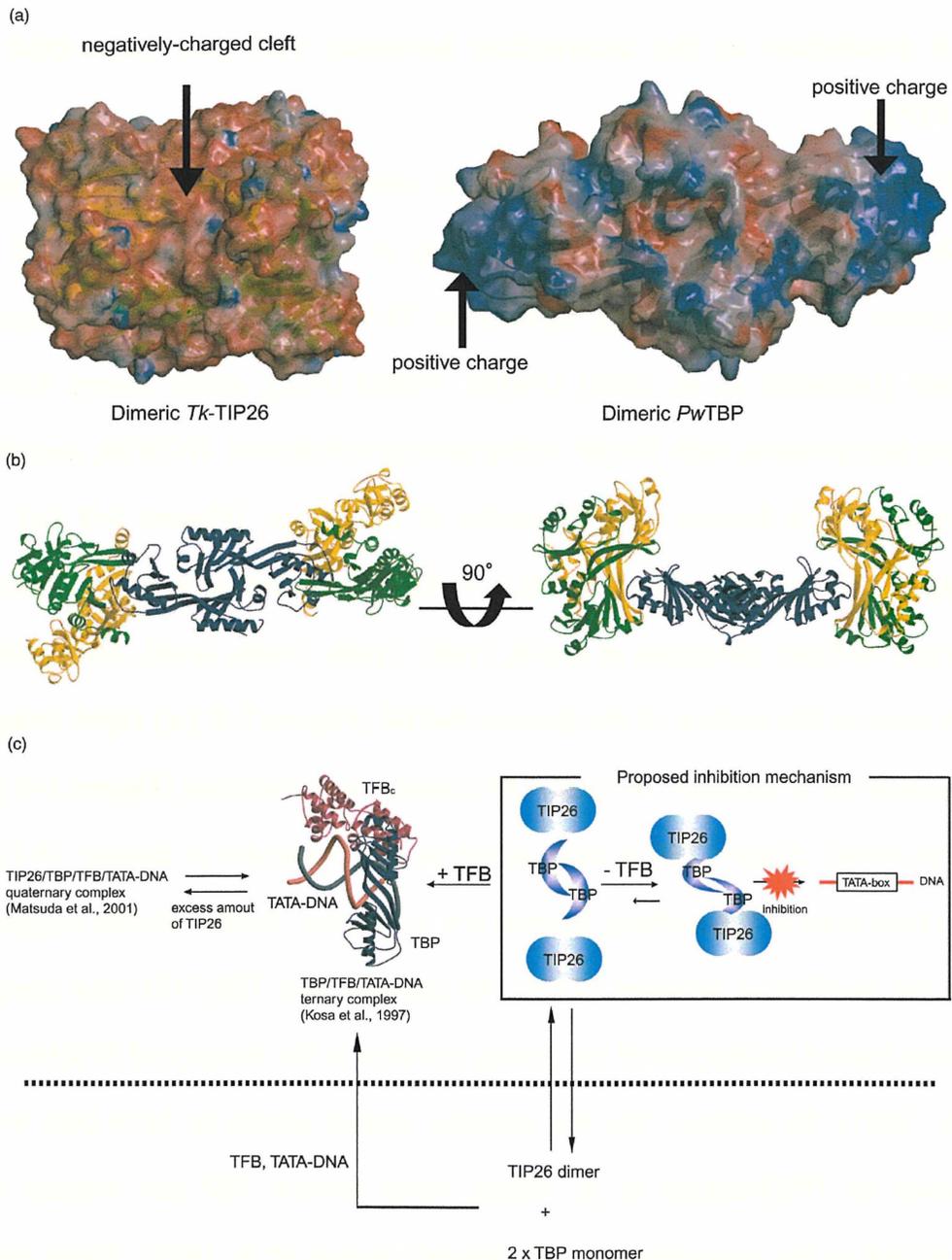


Figure 6 Proposed inhibition mechanism of the interaction between TBP and TATA-DNA by *Tk*-TIP26. (a) Electrostatic surfaces of dimeric *Tk*-TIP26 and dimeric *Pw*TBP generated by GRASP(Nicholls et al. 1993). Coloring is according to the local electrostatic potential range from -10 kT^{-1} (red) to $+10 \text{ kT}^{-1}$ (blue). (b) Model of *Tk*-TIP26/TBP complex. Two dimer molecules of *Tk*-TIP26 were colored as in Figure 4 (a). Dimeric *Pw*TBP was colored with navy. The view of *Pw*TBP in the left side was same as the one in Figure 6 (a). (c) Schematic draw of the inhibition mechanism in which *Tk*-TIP26 prevents TBP from binding to TATA-DNA.

I.4 Summary of Chapter I

The structural studies on *Tk*-TIP26 allowed us to build a putative model of the *Tk*-TIP26/TBP complex, providing insights in the molecular mechanisms of the interaction between TIP26 and TBP. These insights will give us a clue for the understanding of the interaction between TBP and other TBP-interacting proteins. Crystallographic studies on the *Tk*-TIP26/*Tk*-TBP complex and the *Tk*-TIP26/*Tk*-TBP/*Tk*-TFB/TATA-DNA quaternary complex are in progress to increase our understanding of the transcription regulation mechanism by TIP26.

I.5 References

- Barton, G.J. 1993. Alscript: a tool to format multiple sequence alignments. *Protein Eng.* **6**: 37-40.
- Bell, S.D., Brinkman, A.B., van der Oost, J., and Jackson, S.P. 2001a. The archaeal TFIIIE α homologue facilitates transcription initiation by enhancing TATA-box recognition. *EMBO Rep* **2**: 133-138.
- Bell, S.D., and Jackson, S.P. 2001. Mechanism and regulation of transcription in archaea. *Curr Opin Microbiol* **4**: 208-213.
- Bell, S.D., Magill, C.P., and Jackson, S.P. 2001b. Basal and regulated transcription in Archaea. *Biochem Soc Trans* **29**: 392-395.
- Bond, C.S. 2003. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**: 311-312.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N.S., et al. 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 904-925.
- Collaborative Computing Project Number 4 (CCP4), 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 760-763.
- Coleman, R.A., Taggart, A.K., Benjamin, L.R., and Pugh, B.F. 1995. Dimerization of the TATA binding protein. *J Biol Chem* **270**: 13842-13849.
- DeDecker, B.S., O'Brien, R., Fleming, P.J., Geiger, J.H., Jackson, S.P., and Sigler, P.B. 1996. The crystal structure of a hyperthermophilic archaeal TATA-box binding protein. *J Mol Biol* **264**: 1072-1084.
- Fujiwara, S., Takagi, M., and Imanaka, T. 1998. Archaeon *Pyrococcus kodakaraensis* KOD1: application and evolution. *Biotechnol Annu Rev* **4**: 259-284.
- Hanzelka, B.L., Darcy, T.J., and Reeve, J.N. 2001. TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eucaryal transcription factor TFIIIE α . *J Bacteriol* **183**: 1813-1818.
- Holm, L., and Sander, C. 1995. Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **20**: 478-480.
- Jones, T.A. 1985. Diffraction methods for biological macromolecules. Interactive computer graphics: FRODO. *Methods Enzymol.* **115**: 157-171.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1994. Improved methods for building protein models in electron density maps and the location of

- errors in these models. *Acta Crystallogr. sect. A* **47**: 110-119.
- Kabsch, W., and Sander, C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**: 2577-2637.
- Kanemaki, M., Makino, Y., Yoshida, T., Kishimoto, T., Koga, A., Yamamoto, K., Yamamoto, M., Moncollin, V., Egly, J.M., Muramatsu, M., et al. 1997. Molecular cloning of a rat 49-kDa TBP-interacting protein (TIP49) that is highly homologous to the bacterial RuvB. *Biochem Biophys Res Commun* **235**: 64-68.
- Kosa, P.F., Ghosh, G., DeDecker, B.S., and Sigler, P.B. 1997. The 2.1-Å crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. *Proc Natl Acad Sci U S A* **94**: 6042-6047.
- Kraulis, P.J. 1991. Molscript: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**: 946-950.
- Laity, J.H., Lee, B.M., and Wright, P.E. 2001. Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol* **11**: 39-46.
- Lee, M.S., Gippert, G.P., Soman, K.V., Case, D.A., and Wright, P.E. 1989. Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* **245**: 635-637.
- Leslie, A. G. W. 1997. MOFSLM Users Guide: MOSFLM Version 5.50. MRC Laboratory of Molecular Biology, Cambridge, England.
- Makino, Y., Yogosawa, S., Kayukawa, K., Coin, F., Egly, J.M., Wang, Z., Roeder, R.G., Yamamoto, K., Muramatsu, M., and Tamura, T. 1999. TATA-Binding protein-interacting protein 120, TIP120, stimulates three classes of eukaryotic transcription via a unique mechanism. *Mol Cell Biol* **19**: 7951-7960.
- Makino, Y., Yoshida, T., Yogosawa, S., and Tamura, T. 1996. [Detection of TBP-interacting proteins (TIPs) and demonstration of a novel complex containing TBP and ATPases]. *Tanpakushitsu Kakusan Koso* **41**: 1170-1177.
- Matsuda, T., Fujikawa, M., Haruki, M., Tang, X.F., Ezaki, S., Imanaka, T., Morikawa, M., and Kanaya, S. 2001. Interaction of TIP26 from a hyperthermophilic archaeon with TFB/TBP/DNA ternary complex. *Extremophiles* **5**: 177-182.
- Matsuda, T., Morikawa, M., Haruki, M., Higashibata, H., Imanaka, T., and Kanaya,

- S. 1999. Isolation of TBP-interacting protein (TIP) from a hyperthermophilic archaeon that inhibits the binding of TBP to TATA-DNA. *FEBS Lett* **457**: 38-42.
- Merrit, E.A., and Murphy, M.E. 1994. Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 869-873.
- Morikawa, M., Izawa, Y., Rashid, N., Hoaki, T., and Imanaka, T. 1994. Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic *Pyrococcus* sp. *Appl Environ Microbiol* **60**: 4559-4566.
- Nicholls, A., Bharadwaj, R., and Honig, B. 1993. GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64**: A166.
- Nikolov, D.B., and Burley, S.K. 1994. 2.1 Å resolution refined structure of a TATA box-binding protein (TBP). *Nat Struct Biol* **1**: 621-637.
- Nishino, T., Komori, K., Ishino, Y., and Morikawa, K. 2001a. Dissection of the regional roles of the archaeal Holliday junction resolvase Hjc by structural and mutational analyses. *J Biol Chem* **276**: 35735-35740.
- Nishino, T., Komori, K., Tsuchiya, D., Ishino, Y., and Morikawa, K. 2001b. Crystal structure of the archaeal holliday junction resolvase Hjc and implications for DNA recognition. *Structure (Camb)* **9**: 197-204.
- Pavletich, N.P., and Pabo, C.O. 1991. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**: 809-817.
- Qureshi, S.A., Bell, S.D., and Jackson, S.P. 1997. Factor requirements for transcription in the Archaeon *Sulfolobus shibatae*. *Embo J* **16**: 2927-2936.
- Roeder, R.G. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21**: 327-335.
- Soppa, J. 1999. Transcription initiation in Archaea: facts, factors and future aspects. *Mol Microbiol* **31**: 1295-1305.
- Yamada, K., Kunishima, N., Mayanagi, K., Ohnishi, T., Nishino, T., Iwasaki, H., Shinagawa, H., and Morikawa, K. 2001. Crystal structure of the Holliday junction migration motor protein RuvB from *Thermus thermophilus* HB8. *Proc Natl Acad Sci U S A* **98**: 1442-1447.
- Yamamoto, T., Matsuda, T., Sakamoto, N., Matsumura, H., Inoue, T., Morikawa, M., Kanaya, S., and Kai, Y. 2003. Crystallization and preliminary X-ray analysis of TBP-interacting protein from the hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. *Acta Crystallogr D Biol*

Crystallogr **59**: 372-374.

Yogosawa, S., Makino, Y., Yoshida, T., Kishimoto, T., Muramatsu, M., and Tamura, T. 1996. Molecular cloning of a novel 120-kDa TBP-interacting protein.

Biochem Biophys Res Commun **229**: 612-617.

Chapter II

Studies on structure and molecular dynamics of Thioredoxin Peroxidase (ApTPx) from aerobic hyperthermophilic archaeon *Aeropyrum pernix* strain K1

II.1 Introduction

The peroxiredoxin (Prx) constitutes a family of antioxidant proteins that act as peroxidases that reduce hydrogen peroxide and alkyl peroxide to water and the corresponding alcohol, respectively (Schroder and Ponting 1998; Rhee et al. 2001; Wood et al. 2003b). Prxs participate in the antioxidative mechanism called the "thioredoxin system" (Holmgren 1985; 1989). The typical 2-Cys Prxs have two conserved redox-active cysteines, the peroxidatic cysteine and the resolving cysteine. The peroxidase reaction results in an intersubunit disulfide bond between these redox-active cysteines concomitant with the reduction of peroxide substrate. The tertiary structures for Prxs from several sources have been reported, distinct quaternary structures, i.e., monomer (Declercq et al. 2001), dimer (Choi et al. 1998; Hirotsu et al. 1999), and toroid-shaped decamer (Alphey et al. 2000; Schroder et al. 2000; Wood et al. 2002; Wood et al. 2003a; Kitano et al. 2005; Pineyro et al. 2005), being observed depending on the subtype and redox state. However, the tertiary structure of a hyperthermophilic Prx has not been reported to date.

We previously identified the genes for thioredoxin peroxidase, thioredoxin, and

thioredoxin reductase in the genome of *A. pernix* K1 (Jeon and Ishikawa 2002; 2003), which lives at the highest temperature among the aerobic organisms whose genome sequences are available (Sako et al. 1996; Kawarabayasi et al. 1999). We presented the first evidence of a set of proteins involved in the thioredoxin system in archaea by confirming the activity (Jeon and Ishikawa 2002; 2003). Mutational analyses of thioredoxin peroxidase from *A. pernix* K1 (ApTPx) showed that peroxidatic Cys50 and resolving Cys213 form an intersubunit disulfide bond upon oxidation and that another cysteine residue (Cys207) is not essential for the peroxidase function (Jeon and Ishikawa 2003). Although we had proposed a ring structure consisting of eight subunits (Jeon and Ishikawa 2003), we found a fivefold axis and five twofold axes in a crystal of ApTPx (C207S mutant) in a later study (Nakamura et al. 2005). This suggested that the overall structure of ApTPx consists of decameric ring, which is homologous to Prxs from human (Schroder et al. 2000), *Salmonella typhimurium* (Wood et al. 2002; Wood et al. 2003a), and *Crithidia fasciculata* (Alphey et al. 2000).

In this chapter, we describe the tertiary structure of a selenomethionine derivative of the C207S mutant of ApTPx (SeC207S) in the reduced form determined at 2.0 Å resolution. We also determined the structures of C207S mutant ApTPx treated with two different concentration of H₂O₂, which is a substrate of ApTPx. We discuss the characteristic features of this hyperthermostable Prx and discuss the reaction mechanism where H₂O₂ is reduced to H₂O.

II.2 Materials and methods

II.2.1 Preparation of SeC207S crystals

Expression and purification of SeC207S were performed as described (Nakamura et al. 2005) except that the *E. coli* cells were cultivated in a modified M9 medium (Doubie 1997). Crystals of SeC207S were obtained by the hanging-drop vapor-diffusion method. A 4 μ l of drop comprising 10 mg/ml SeC207S, 5 mM DTT, 50 mM imidazole-HCl (pH 6.5), and 500 mM sodium acetate was equilibrated against 450 μ l of a reservoir solution comprising 60 mM imidazole-HCl (pH 6.5) and 600 mM sodium acetate at 20°C. Crystals were soaked in a solution of x1.25 concentration of the reservoir solution containing 20% (v/v) ethylene glycol, and then flash-frozen in the nitrogen-gas stream at 100 K.

II.2.2 Preparation of ApTPx (C207S) crystals

Crystals of ApTPx were obtained by the hanging-drop vapor-diffusion method in reducing condition. 2 μ l of protein solution containing 10 mg/ml of ApTPx and 20 mM Tris-HCl (pH 7.5) was mixed with 400 μ l of reservoir solution comprising 100 mM sodium acetate (pH 4.8), 200 mM calcium chloride, 10% (v/v) isopropanol, 1 mM DTT at 20°C. Crystals of ApTPx were soaked to two distinct conditions of solution for cryogenic experiment. The one comprised of 0.1 mM H₂O₂, 25% (v/v) glycerol, and reservoir solution used in crystallization (solution I), the other, 1 mM H₂O₂, 25% (v/v) glycerol, and reservoir solution used in crystallization (solution II). After soaking into solution I and II, crystals of ApTPx were flash-frozen in the nitrogen-gas stream at 100 K.

II. 2.3 Data collection

A MAD experiment of SeC207S was performed at BL38B1 beamline at SPring-8 (Harima, Japan). Four different wavelengths were chosen from inspection of the fluorescence spectrum of the crystals, corresponding to the edge, peak, low-energy remote, and high-energy remote. The diffraction data were recorded on Jupiter210 CCD detector (Rigaku/MSK). Diffraction data were indexed, integrated, and reduced with HKL2000 program suite (Otwinowski and Minor 1997).

Diffraction data of ApTPx soaked in solution I and solution II were collected at BL41XU and BL38B1 at SPring-8 (Harima, Japan), using ADSC quantum315 CCD detector and Jupiter210 CCD detector (Rigaku/MSK), respectively. Diffraction data were indexed, integrated, and reduced with HKL2000 program suite (Otwinowski and Minor 1997).

II.2.4 Structure determination and refinement

Structure of reduced form of ApTPx was solved by multiwavelength anomalous dispersion (MAD) method using SeC207S crystals. Selenium sites were searched with SOLVE (Terwilliger and Berendzen 1999). The positions of selenium sites were refined and initial phase was calculated using SHARP (de La Fortelle and Bricogne 1997). The phase was improved with SOLOMON (Abrahams and Leslie 1996). Initial model was built using program O (Jones et al. 1991) and refined with CNS (Brünger et al. 1998).

Structures of H₂O₂-soaked ApTPx were solved by molecular replacement using MOLREP (Collaborative Computational Project 1994) with the reduced structure of

ApTPx as a search model. These structure were refined using CNS(Brünger et al. 1998) and Refmac5(Collaborative Computational Project 1994). Stereochemical analysis, secondary structure assignment, structural comparisons, and calculation of the accessible surface area were carried out using *PROCHECK*(Laskowski et al. 1993), *DSSP*(Kabsch and Sander 1983), *CHIMERA*(Pettersen et al. 2004), and *GRASP*(Nicholls et al. 1991), respectively.

II.2.5 Optimization of the partial structure

Partial structure of catalytic domain of oxidized ApTPx was optimized by density functional theory using the program Gaussian(P. M. W. Gill 1993).

II.2.6 Preparation of figures

Figure II-3, Figure II-4 (b), Figure II-5, Figure II-6, Figure II-7 were prepared using Molscrip(Meritt and Murphy 1994) and Raster3D(Meritt and Murphy 1994).

Figure II-8 was prepared with Pymol(<http://www.pymol.org>).

II.3 Results and discussion

II.3.1 Crystals of ApTPx

The pictures of crystals of SeC207S and ApTPx are shown in **Figure II-1 (a) and (b)**. SeC207S crystals were grown up to its maximum size (0.2 x 0.3 x 0.15 mm) in twelve hours and crystals of ApTPx were grown up to its maximum size (0.2 x 0.15 x 0.05 mm) in a day.

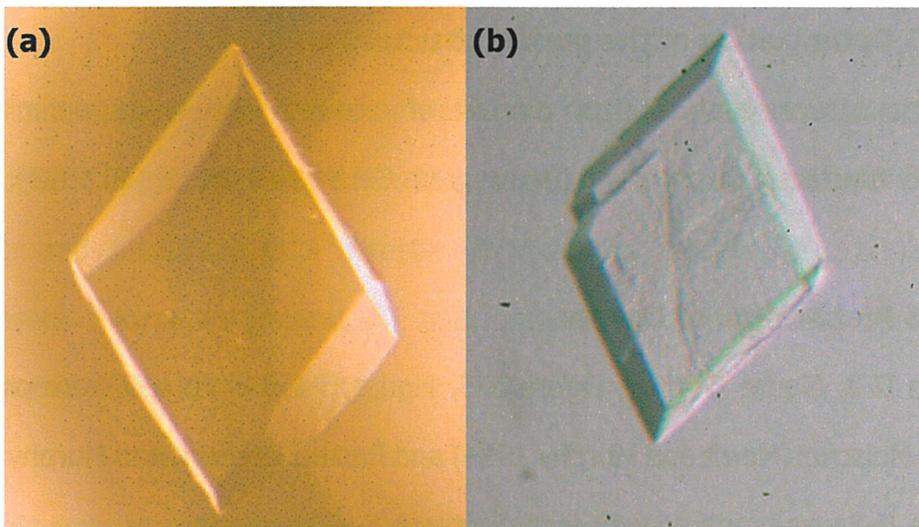


Figure II-1 Crystals of ApTPx (a) SeC207S and (b) C207S

II.3.2 Diffraction data

Crystal of SeC207S diffracted to 2.44 Å (peak, edge, and high-energy remote wavelenghts) and 2.00 Å (low-energy remote), respectively. The crystal belonged to the space group $P1$ with cell dimensions of $a = 75.6$ Å, $b = 101.8$ Å, $c = 102.9$ Å, $\alpha = 105.4^\circ$, $\beta = 105.3^\circ$, and $\gamma = 93.3^\circ$. Diffraction data statistics for SeC207S are shown in **Table II-1**.

Table II-1 Data collection statistics for SeC207S

	Peak	Edge	High remote	Low remote
X-ray source	BL38B1, SPring-8	BL38B1, SPring-8	BL38B1, SPring-8	BL38B1, SPring-8
Wavelength (Å)	0.97915	0.97948	0.976	0.984
Space group	$P1$			
Cell dimensions (Å, °)	$a = 75.6, b = 101.8, c = 102.9, \alpha = 105.4, \beta = 105.3, \text{ and } \gamma = 93.3$			
Resolution range (Å) ^a	50 - 2.44 (2.53 - 2.44)	50 - 2.44 (2.53 - 2.44)	50 - 2.43 (2.52 - 2.43)	30 - 2.00 (2.07 - 2.00)
R_{merge} (%) ^{a, b}	6.0 (9.8)	5.0 (9.0)	5.4 (9.7)	8.4 (18.8)
$I/\sigma(I)$	16.6	16.5	15.9	8.7
Total reflections	400,490	399,388	403,717	569,297
Unique reflections	106,539	106,743	108,010	191,063
Completeness (%) ^a	98.3 (95.9)	98.3 (95.0)	98.3 (94.3)	95.7 (86.1)

^aValues in parentheses are for the highest resolution shell.

^b $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$, where I is the intensity of observation I and $\langle I \rangle$ is the mean intensity of the reflection.

Crystals of ApTPx soaked with solution I (Form I) and solution II (Form II) diffracted to 2.4 Å and 2.0 Å, respectively. Both Form I and Form II belonged to the space group $P1$ with cell dimensions of $a = 76.2$ Å, $b = 103.4$ Å, $c = 104.6$ Å,

$\alpha = 105.8^\circ$, $\beta = 105.2^\circ$, $\gamma = 92.7^\circ$, and $a = 76.2 \text{ \AA}$, $b = 103.3 \text{ \AA}$, $c = 105.0 \text{ \AA}$, $\alpha = 106.0^\circ$, $\beta = 105.1^\circ$, $\gamma = 92.7^\circ$, respectively and both of the crystals contained ten monomers in the asymmetric unit. Data statistics for Form I and Form II are shown in **Table II-2**.

Table II-2 Diffraction data statistics for Form I and Form II crystals of ApTPx

	Form I (0.1 mM H ₂ O ₂)	Form II (1.0 mM H ₂ O ₂)
X-ray source	SPring-8 BL41XU	SPring-8 BL38B1
Space group	<i>P</i> 1	<i>P</i> 1
Cell parameters (\AA, $^\circ$)	$a = 76.2, b = 103.4, c = 104.6$ $\alpha = 105.8, \beta = 105.2, \gamma = 92.7$	$a = 76.2, b = 103.3, c = 105.0$ $\alpha = 106.0, \beta = 105.1, \gamma = 92.7$
Resolution range^a (\AA)	50 – 2.4 (2.49 – 2.40)	50 – 2.0 (2.07 – 2.00)
No. of total reflections	242,709	339,192
No. of unique reflections	115,754	198,017
$\langle I/\sigma(I) \rangle^a$	8.5 (2.74)	8.4 (1.96)
Completeness^a (%)	87.2 (71.4)	92.1 (80.0)
$R_{\text{merge}}^{a,b}$ (%)	8.8 (22.6)	6.3 (24.3)

^aValues in parentheses are for the highest resolution shell.

^b $R_{\text{merge}} = \sum |I - \langle I \rangle| / \sum I$, where I is the intensity of observation I and $\langle I \rangle$ is the mean intensity of the reflection.

II.3.3 Structure determination and refinement of SeC207S

Initial phase of SeC207S was estimated by multiwavelength anomalous dispersion (MAD) method. 27 of 40 possible selenium sites were found using program SOLVE (Terwilliger and Berendzen 1999). These sites were refined and

initial phase was calculated. Resulting electron density map had excellent quality and initial model was built on this map. After several rounds of refinement and manual fitting, R and R_{free} values were converged to 16.4% and 17.1%, respectively against 2.0 Å data. Phasing statistics, experimental and refined electron density map, and refinement statistics are shown in **Table II-3**, **Figure II-2**, and **Table II-4**, respectively.

Table II-3 Phasing statistics for SeC207S

	Peak	Edge	High-remote
Resolution range (Å)	50 – 2.44	50 – 2.44	50 – 2.44
No. of site (found/ total)	29/40	29/40	29/40
R_{cullis}^a (iso./ano.)	0.816/0.634	0.765/0.839	- /0.738
Phasing Power^b (iso./ano.)	0.883/2.098	1.054/0.839	- /1.553
Figure of merit^c	0.487		

^a $R_{\text{cullis}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$ for centric reflections, where $|F_{\text{obs}}|$ = observed heavy atom structure factor amplitude, and $|F_{\text{calc}}|$ = calculated heavy atom structure factor amplitude.

^bPhasing power = root-mean-square ($|F_{\text{h}}|/E$), where $|F_{\text{h}}|$ = heavy atom structure factor amplitude, and E = residual lack of closure error.

^cFigure of merit = $|F(\text{hkl})_{\text{best}}|/F(\text{hkl})$

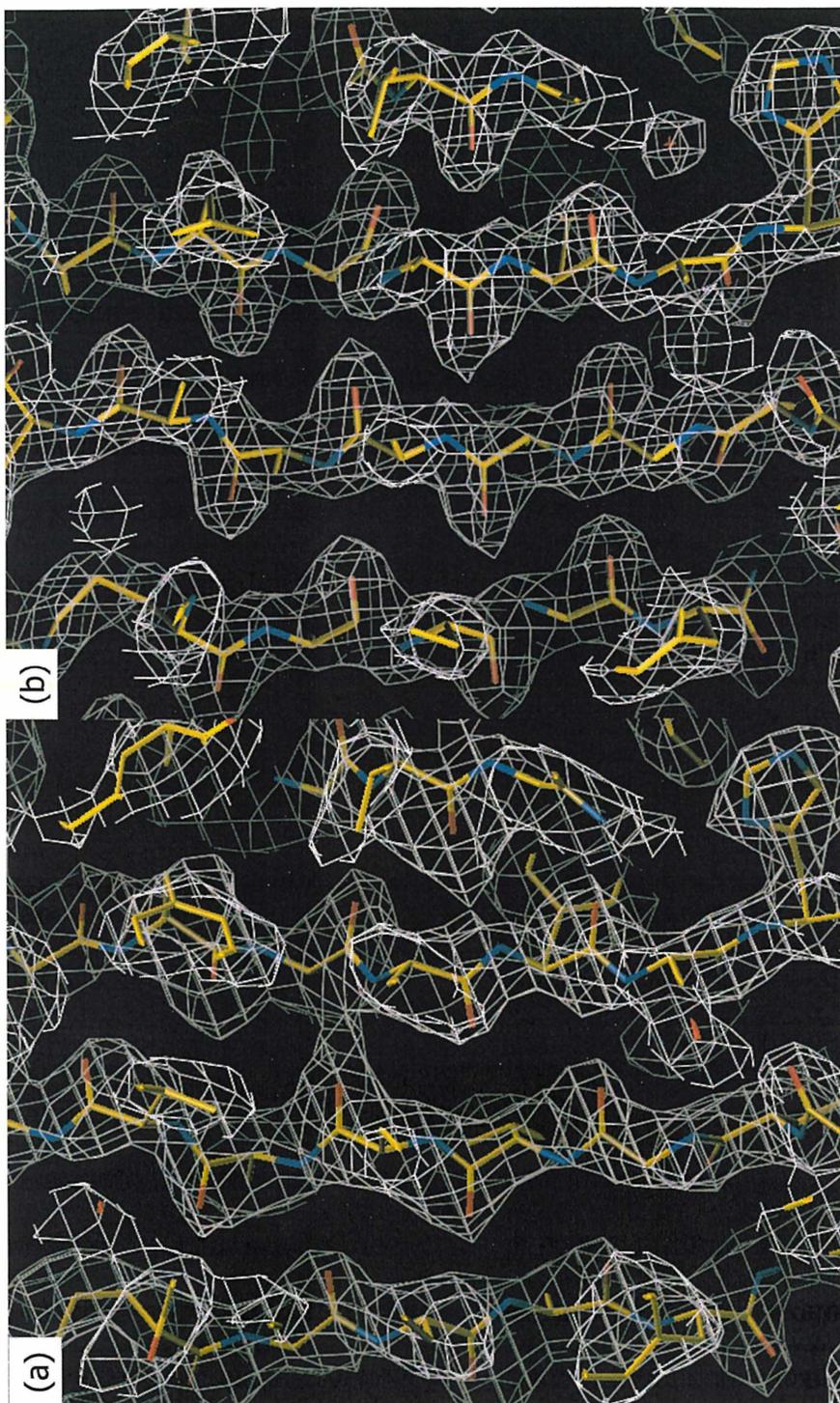


Figure II-2 Electron density maps of SeC207S. (a) Experimental electron density map from MAD phase contoured at 1σ . (b) $2|F_0| - |F_c|$ electron map contoured at 1σ superimposed on the 2.0 Å resolution coordinates of SeC207S.

II.3.4 Structure determination and refinement for Form I and Form II of ApTPx

Structures of Form I and Form II of ApTPx were determined by molecular replacement using a dimer structure of SeC207S as a starting model. For the refinement, diffraction data of Form I and Form II between 40 to 2.4 Å resolution, and between 50 to 2.0 Å resolution were included in simulated annealing refinement with bulk solvent correction using CNS (Brünger et al. 1998), respectively. Five percent of the reflections were kept separate to monitor R_{free} , and were not used in the refinement. After some rounds of refinement process, the structure of Form I of ApTPx was refined using Refmac5(Collaborative Computational Project 1994). During the refinement process, extra positive peaks connected to the S^γ atom of Cys50 was observed in the $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ maps in the all of the ten molecules in the asymmetric unit. These positive suggested that Cys50 was oxidized to cysteine sulfenic acid. All of the Cys50 residues in the asymmetric unit were replaced with cysteine sulfenic acid, and refined using the program Refmac5(Collaborative Computational Project 1994). Final refinement statistics for Form I and Form II of ApTPx are shown in **Table II-4**.

Table II-4 Refinement statistics for SeC207S, Form I, and Form II of ApTPx

	Reduced (SeC207S)	Form I (0.1 mM H₂O₂)	Form II (1.0 mM H₂O₂)
Resolution range (Å)	40 – 2.0	40 – 2.4	40 – 2.0
<i>R</i>_{cryst} (%) / <i>R</i>_{free} (%)	16.4 / 17.3	15.3 / 22.9	16.3 / 22.0
Water molecules	1,189	542	1,129
RMSD bond length (Å)	0.046	0.024	0.03
RMSD bond angles (deg.)	3.4	2.1	2.4
Ramachandran (%)			
Favored	90.5	89.2	91.6
Allowed	9.5	10.4	8.1
Disallowed	0	0.4	0.2

^a $R_{\text{cryst}} = \frac{\sum ||F_o| - |F_c||}{\sum |F_c|}$, where F_o and F_c are the observe and calculated structure factor amplitudes, respectively

^b R_{free} was calculated using a randomly selected 10% of the data set that was omitted through all stages of refinement

II.3.5 Structure of reduced form (SeC207S) of ApTPx

The monomer structure of SeC207S is presented in **Figure II-3 (a)**. The monomer structure can be divided into a main domain (Pro2 – Leu167 and Arg220 – Ala245) and an arm domain (Lys168 – Ser219). The arm domain is composed of $\alpha 6$, an antiparallel β -sheet comprising three β -strands ($\beta 9$, $\beta 10$, and $\beta 11$, see **Figure II-3 (a)**), and a neighboring unstructured region. The arm domain and subsequent C-terminal region are unique to ApTPx (discussed later).

ApTPx forms homodimers (**Figure II-3 (b)**). The dimer interface is created between the main domains. The arm domains stick out from the main body of the dimer (circles in **Figure II-3 (c)**). In addition to a 14-stranded twisted intersubunit β -sheet (**Figure II-3 (c)**, bottom), the N-terminal $\beta 1$ in both subunits form an antiparallel β -sheet (arrow in **Figure II-3 (b)**) in the homodimer structure. This N-terminal β -sheet is unique to ApTPx and thus is not found in mesophilic Prxs.

The overall structure of ApTPx comprises a decameric ring consisting of five homodimers (**Figure II-3 (d)**). There are two types of dimer-dimer interactions: one (the MM interaction) involves main domains and the other (the AM interaction) involves arm and main domains (**Figure II-3 (e)**). Because the arm domain is characteristic of ApTPx, the AM interaction is unique to ApTPx. In particular, the carbonyl oxygen of Thr191, the side chain oxygen of Glu193, and the side chain oxygens of Asp209 are involved in hydrogen bond networks with the next subunit (**Figure II-3 (f)**). These hydrogen bond networks stabilize the AM interaction.

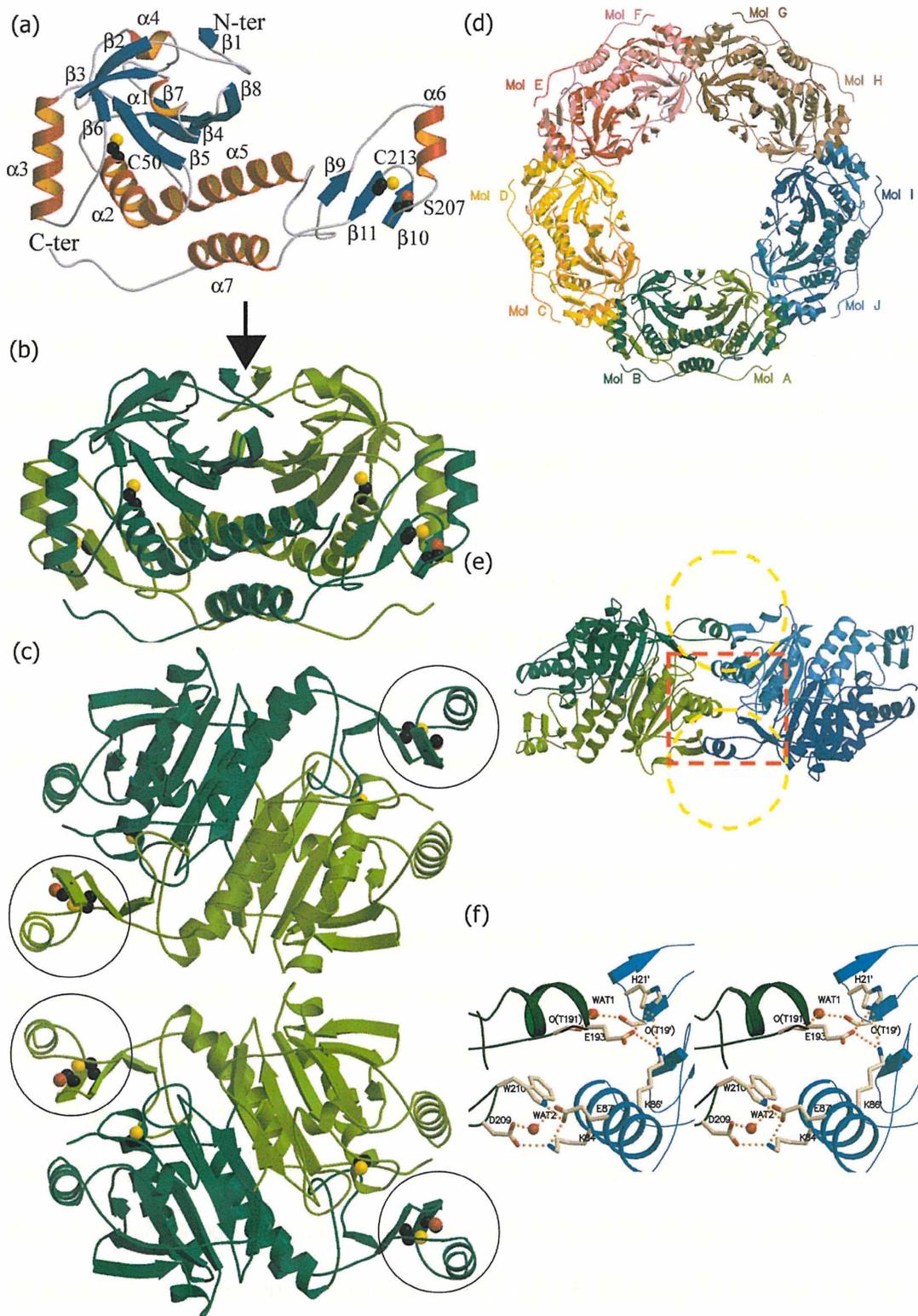


Fig. II-3. Representation of the tertiary structure. (a) The monomer structure of SeC207S is drawn. The side chains of Cys50, Ser207, and Cys213 are represented as balls and sticks. (b) The dimer structure.

(Figure II-3 continued) The dark green subunit is in the same orientation as in panel (a). The arrow indicates the antiparallel β -sheet formed by the β 1 strands of the subunits. (c) The dimer structure from the perpendicular direction as to in panel (b). The top and bottom views are from opposite directions to each other. The circles indicate the arm domains, which stick out of the main body of the dimer (see text). The side chains of Cys50, Ser207, and Cys213 are represented as balls and sticks in panels (a), (b), and (c). (d) The decamer structure from the same direction as in panel a. The names of the subunits are indicated. The outer and inner diameters of the ring are indicated by double-headed arrows. (e) Two homodimers (AB and IJ) are presented from the direction indicated by an open arrow in panel (d). The positions of the AM interactions and the MM interaction are indicated by red circles and a yellow square, respectively. (f) Close-up of the interface of the AM interaction indicated by the upper circle in panel (e). The hydrogen bond networks involving Thr191, Glu193, Asp209 are depicted.

Assembly of the subunits of ApTPx is characterized by the large buried accessible surface area. The formation of a decameric ring of ApTPx results in a buried surface of 5,391 Å² per monomer (43.7% of the monomer surface). This proportion is significantly higher than those in other decameric Prxs (35.4, 37.0, and 29.7% for PrxII, AhpC, and TryP, respectively). These findings suggest that subunit-subunit interactions are more dominant in ApTPx than in the mesophilic Prxs. Thermostable proteins often gain their stability through the formation of large complexes (Vieille and Zeikus 2001; Maeda et al. 2002). In this sense, the greater buried surface area would contribute to the hyperthermostability of ApTPx.

The polypeptide of ApTPx (250 amino acids) is significantly larger than those of

mesophilic Prxs (~ 200 residues). Sequence alignment allowed assignment of an additional region of ApTPx as the C-terminal region (**Figure II-4 (a)**). The main chain conformations in the N-terminal segment of the main domain are well conserved in these proteins, as a consequence of high homology between them (see legend to **Figure II-4 (a)**). However, the arm domain and the subsequent C-terminal segment are characteristic of ApTPx (**Figure II-4 (b)**). The characteristic C-terminal $\alpha 7$ of ApTPx is rich in polar residues and participates in intramolecular helix-helix interactions. The arm domains contribute to the interactions between dimers (the AM interactions). The AM interactions are not found in mesophilic decameric Prxs and thus are unique to ApTPx. These findings suggest that the characteristic arm domain strengthens the decamer formation of ApTPx.

Ser207 (Cys in the wild-type) and Cys213 are located on $\beta 10$ and $\beta 11$ in the arm domain, respectively, the distance between the α carbons being 4.2 Å. The closeness may enable the intramolecular disulfide bond formation between Cys207 and Cys213 in the wild-type at a particular stage in the reaction cycle. The behaviors of these SH groups of ApTPx during the reaction cycle remain unclear at present. In order to elucidate the behaviors of the SH groups, we are currently trying to determine the crystal structure of ApTPx in the oxidized form.

(a)

```

ApTPx  MPGSIPLIGERFPEMEYTTDHGYIKLPDHYYS--QGKWFVLFSPADFTPVYCTTEFVSFA  - 58
PrxII  MASGNARIGKPPAPDFKATAVVDG-AFKEVKLSDYKGGKYVYVLFYPLDFTFVCPTEIIAFS
TryP   MSCGAARLNHPAPEFDDMALMPNGTFKKYSLSSYKGGKYVYVLFYPMDFTFVCPTEIIQFS
AhpC   -----SLINTKIKPFKNQAFKNG-EFIEVTEKDTTEGRWSVFFFPADFTFVCPTELDGVA

ApTPx  RRYEDFQRLGVDLIQLSYDSVFSHIKWKEWIERHIGVRIP-FPIIADPQGTVARRLGLLH  -117
PrxII  NRAEDFRKLGCEYLGYSDSQFTHLAWIINTPRKEGGLPLNIPLLADYTRRLSEDYGYLK
TryP   DDAKRFAEINTEVISCSCDSEYSHLQWTSVDRKKGGLGPMaipMLADKTKAIARAYGYLD
AhpC   DHYEELQKLGVDYYSYSTDTHFTHKAWHSSSE---TIAKIKYAMIGDPTGALTRNFONMR

ApTPx  AESATHYRQVFIYDARGVIRTMLYPMELGRLYDEILRIYKALKLGDLSLKRAYPADWPN  -177
PrxII  TDEG-IAYRGLFIIDGKGYLRQITVNDLPVGRSYDEALRLVQAFQYTDEH-----
TryP   EDSG-YAYRQVFIIDPNGKLRQIIINDMPIGRNWEEVIRLYEALQFVEEH-----
AhpC   EDEG-LADRATFVYVDPQGIQATEVTAEGIGRDASDLLRKIKAAQYVAAHP-----

ApTPx  NEIIGEGLIYPPPTTEDQARARMESGQYRCLDWWFCWDTPASRDDVEEARRYLRRAAEK P  -237
PrxII  -----GEYCPAGWKKPGSDTIK P
TryP   -----GEYCPANWKKGDAKKE
AhpC   -----GEYCPAKWKEGATLAP

ApTPx  AKLLYEEARTH LH  -250
PrxII  NVDDSKEYFSKH
TryP   GH-----
AhpC   SLDLVGKI-----

```

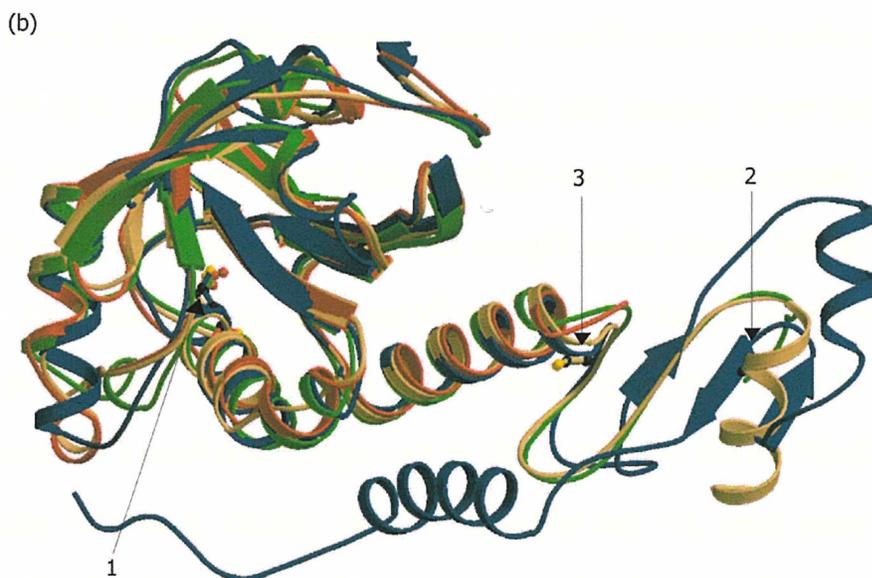


Figure II-4 Comparison with mesophilic Prxs. (a) Multiple alignments of the protein sequences were performed with *CLUSTALW* (Higgins et al. 1996) Red and blue characters indicate α -helices and β -strands, respectively. The numbers are given as for the sequence of ApTPx. Cys207, which is replaced by Ser in the C207S mutant, is underlined. Sequence homology between ApTPx and PrxII, AhpC, and TryP in the N-terminal segment of the main stromal/metabolic region N-terminal to a triangle) is 62, 56, and 54%, respectively.

(Figure II-4 continued) ApTPx, thioredoxin peroxidase from *A. pernix*; PrxII, peroxiredoxin II from human; AhpC, AhpC from *Salmonella typhimurium*; TryP, tryparedoxin peroxidase from *Crithidia fasciculata*. (b) Superimposition of the four Prxs aligned in panel a. The view of ApTPx structure is the same as in Fig. 1a. blue, ApTPx; yellow, Prx II; red, TryP; green, AhpC.

II.3.6 Comparison of monomer structures of ApTPx

Both of H₂O₂-soaked structures of ApTPx formed (α_2)₅ decamer as in the reduced form of ApTPx. RMSD values of both Form I and Form II of ApTPx compared with reduced form for Ca atoms were 0.77 Å and 1.14 Å, respectively. RMSD value when Ca atoms of Form I were compared with those of Form II was 0.27 Å. These results suggest that Both of H₂O₂-soaked structures of ApTPx are structurally distinct from the reduced form, and that Form I and Form II are structurally similar to each other. **Figure II-5** shows monomer structures of the reduced form, Form I, and Form II of ApTPx and their corresponding $|F_o| - |F_c|$ omit maps of Cys50 contoured at 3 σ . **Figure II-5 (a)** and **Figure II-5 (b)** show that Cys50 are not oxidized in both of the reduced and Form I structures of ApTPx, because there are not additional electron density around S^γ atom of Cys50. However, **Figure II-5 (c)** suggests that side residue of Cys50 in Form II structure is oxidized from cysteine (Cys50-SH) to cysteine sulfenic acid (Cys50-SOH), because there is an additional positive peaks in the $|F_o| - |F_c|$ omit map around S^γ atom of Cys50. That is, Form II structure of ApTPx was changed to the oxidized form by the reduction of H₂O₂.

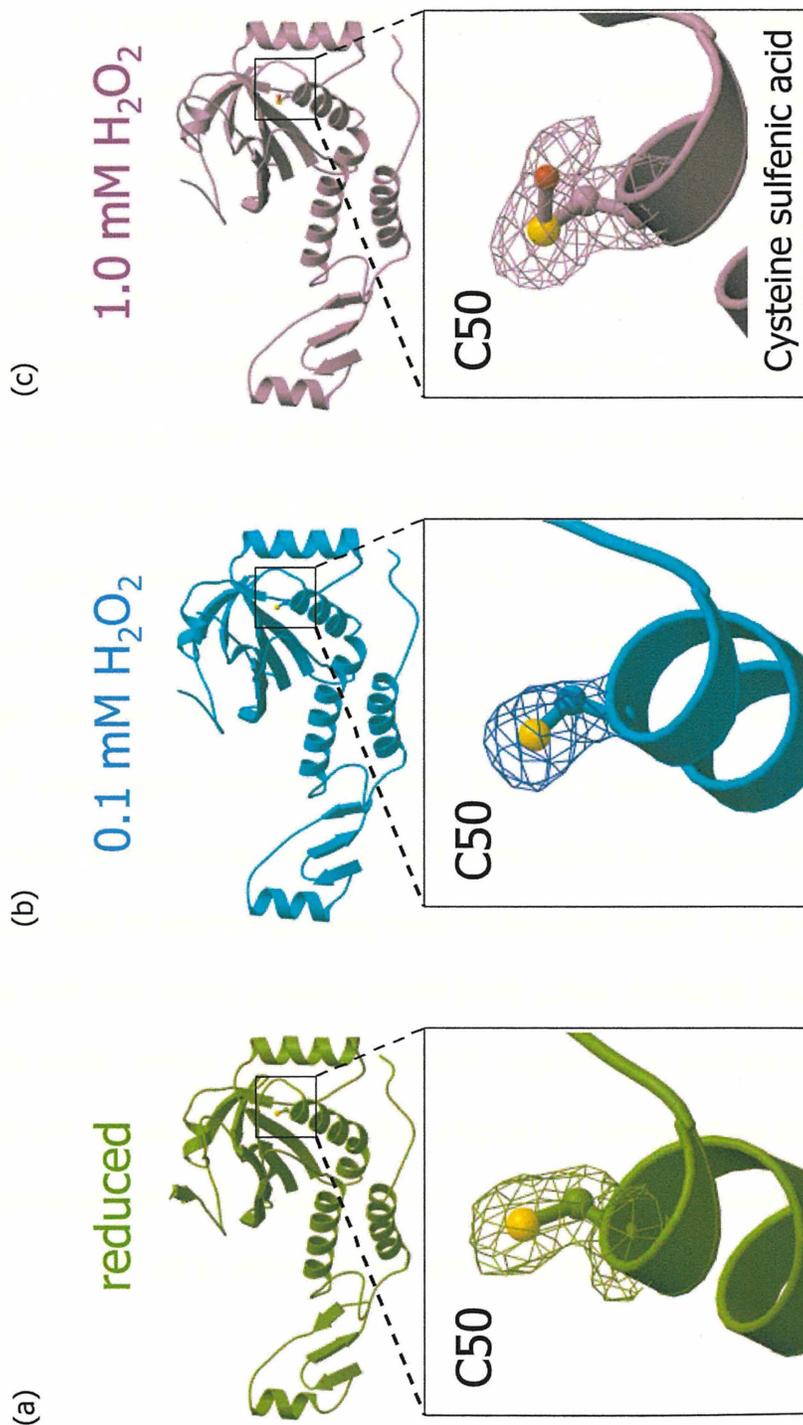


Figure II-5 Monomer structures and their corresponding magnified views around active site Cys50 imposed onto $|F_o| - |F_c|$ omit maps of Cys50 contoured at 3σ . (a) reduced form. (b) Form I (0.1 mM H_2O_2 -soaked), and (c) Form II (1.0 mM H_2O_2 -soaked).

II.3.7 Active site structure in the oxidized form (Form II)

Figure II-6 (a) shows the active site structure of oxidized ApTPx. His42, which does not interact with S^Y atom of Cys50 in the reduced structure (discussed below), approaches to Cys50-SOH with 2.15 Å of the bond distance between N^{δ1} atom of His42 and S^Y atom of Cys50-SOH. These two atoms are very adjacent when it is thought as the hydrogen bonding. In addition, the bond angle formed by O¹ atom, S^Y atom, and C_β atom of Cys50-SOH is 90°. In order to understand the bonding state between S^Y atom of Cys50-SOH and its surroundings, the partial structure containing Cys50-SOH and imidazole group, which is homologous structure of Hi42 were extracted, hydrogen atoms were added to the partial structure, and it was optimized by density function method. **Figure II-6 (b) left** and **right** show X-ray and optimized structure respectively. **Table II-5** shows bond angles and distances in the x-ray and optimized structure shown in **Figure II-6 (b)**. These parameters have good agreement in both structures. In the optimized structure, sulfur atom forms covalent bonds with four atoms; one oxygen, one carbon, one nitrogen, and one hydrogen atoms. Taken into account for the bond angle formed by oxygen, sulfur, and carbon atoms, this sulfur atom forms hypervalent structure. These results suggest that the hypervalent sulfur atom is formed in the active site of the oxidized form of ApTPx.

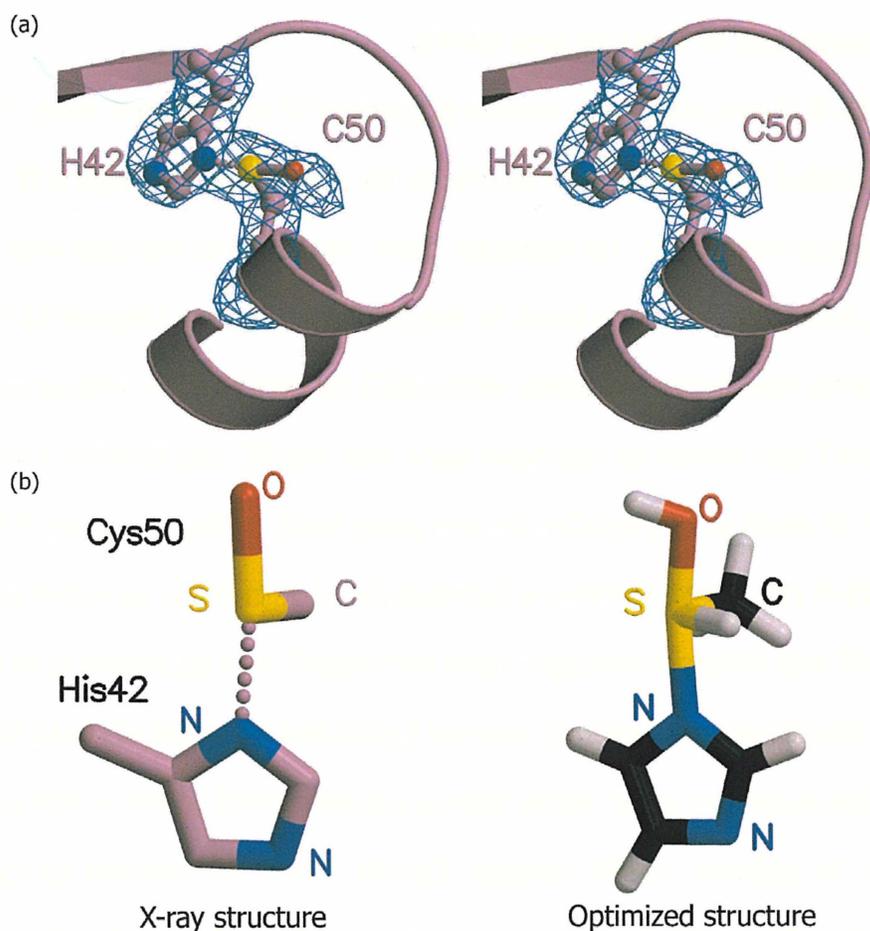


Figure II-6 Active site structure of oxidized form (Form II) of ApTPx. (a) Stereo diagram of active site structure of ApTPx superimposed onto a $|F_o| - |F_c|$ omit map contoured at 3σ . (b) Partial model of X-ray structure (left) and optimized structure (right).

Table II-5 Bond angles and distances in X-ray and optimized structure around His42 and Cys50-SOH

	X-ray structure	Optimized structure
Angle (O-S-C) ($^\circ$)	90.0	91.5
Angle (O-S-N) ($^\circ$)	173.3	170.3
Distance (O-S) (\AA)	1.87	1.80
Distance (S-N) (\AA)	2.15	1.97

II.3.8 Movement of Arg128 and His42

Figure II-7 shows structures around active site of the reduced form, Form I, and the oxidized form (Form II) of ApTPx. In the reduced form (**Figure II-7 (a)**), His42 interacts with Asp45 *via* hydrogen bond. Arg126 also interacts with Cys50 (Cys-SH), forming a hydrogen bond. In the oxidized structure (**Figure II-7 (c)**), Arg126 forms new hydrogen bonds with Glu53 by breaking the interaction with Cys50. His42 also breaks its hydrogen bond with Asp45, and forming a covalent bond with S^γ atom of Cys50-SOH. On the other hand, in the structure of Form I (**Figure II-7 (b)**), Arg126 forms a hydrogen bond with Glu53 as in the oxidized form. His42 also breaks a hydrogen bond with Asp42, and approaching to Cys50. However, the distance between S^γ atom of Cys50 and N^{δ1} atom of His42 is 2.9 Å, which is the hydrogen bond rather than a covalent bond found in the oxidized structure. In addition, Cys50 remains Cys-SH form, not oxidized to cysteine sulfenic acid. These results suggest that the structure of Form I is that of the intermediate state between the reduced and the oxidized form.

II.3.9 Electrostatic surfaces of the reduced and the intermediate form of ApTPx

Electrostatic surfaces of the reduced and the intermediate state were compared and shown in **Figure II-8**. In the reduced form (**Figure II-8 (a)**), Arg126 is exposed to the molecular surface of ApTPx and the electrostatic surface of the part is charged to mainly positive. On the other hand, in the intermediate structure (**Figure II-8 (b)**), Arg126 withdraws from the molecular surface of ApTPx, and resulting structural changes by

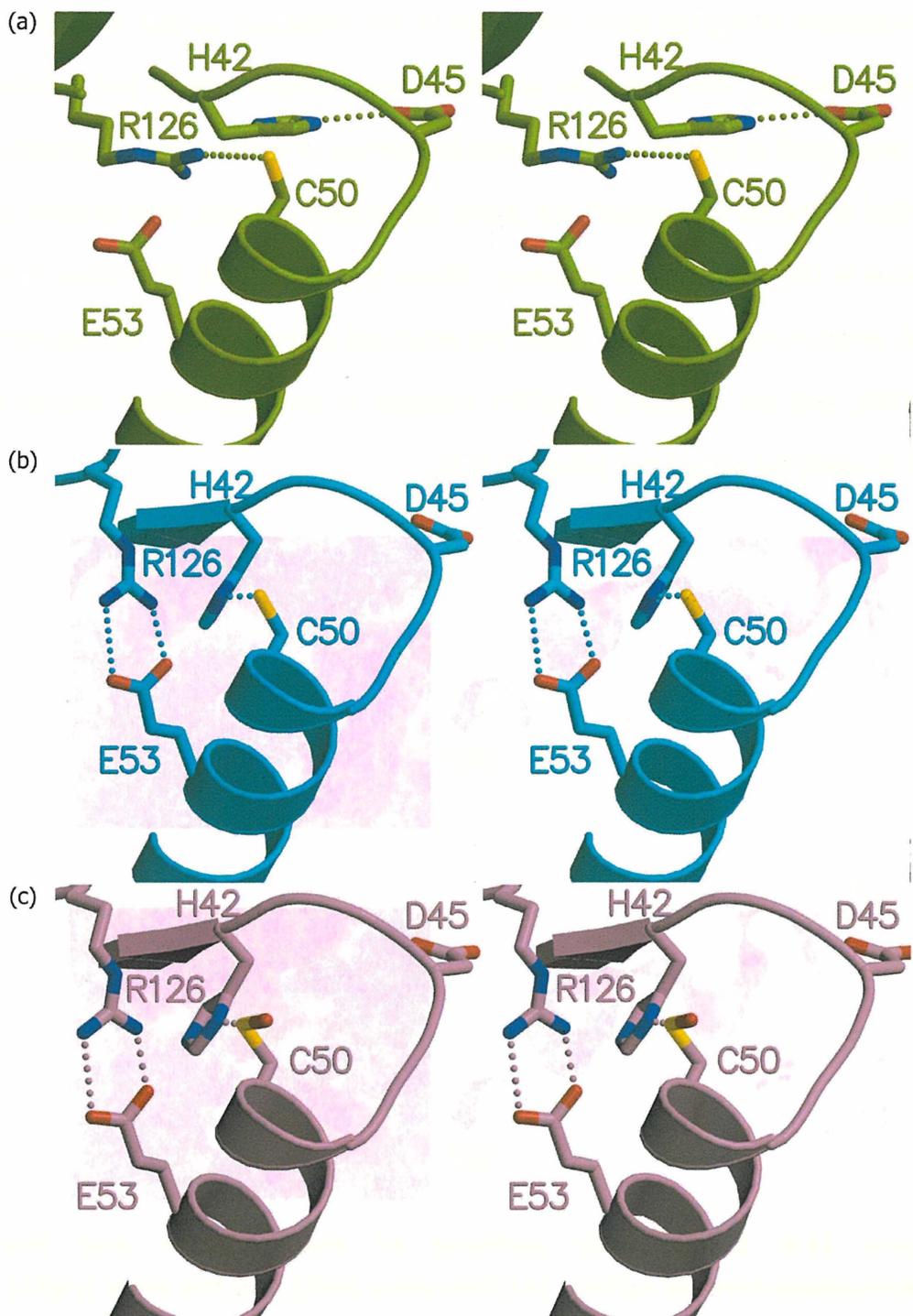


Figure II-7 Stereo views of the structures around active site. (a) reduced form, (b) intermediate form (Form I), and (c) oxidized form (Form II).

the movement of Arg126 alters the shape of the molecular surface around the active site pocket. The active site pocket becomes deeper than in the reduced form and Cys50 is exposed to the molecular surface of ApTPx. The electrostatic potential of the surface around the active site pocket changes from positive to negative in the intermediate structure. These results suggest that some H_2O_2 signal, which is not elucidated, cause the structural change on the reduced form of ApTPx, and the structure of ApTPx changes to the intermediate structure, which is ready for accept the substrate H_2O_2 .

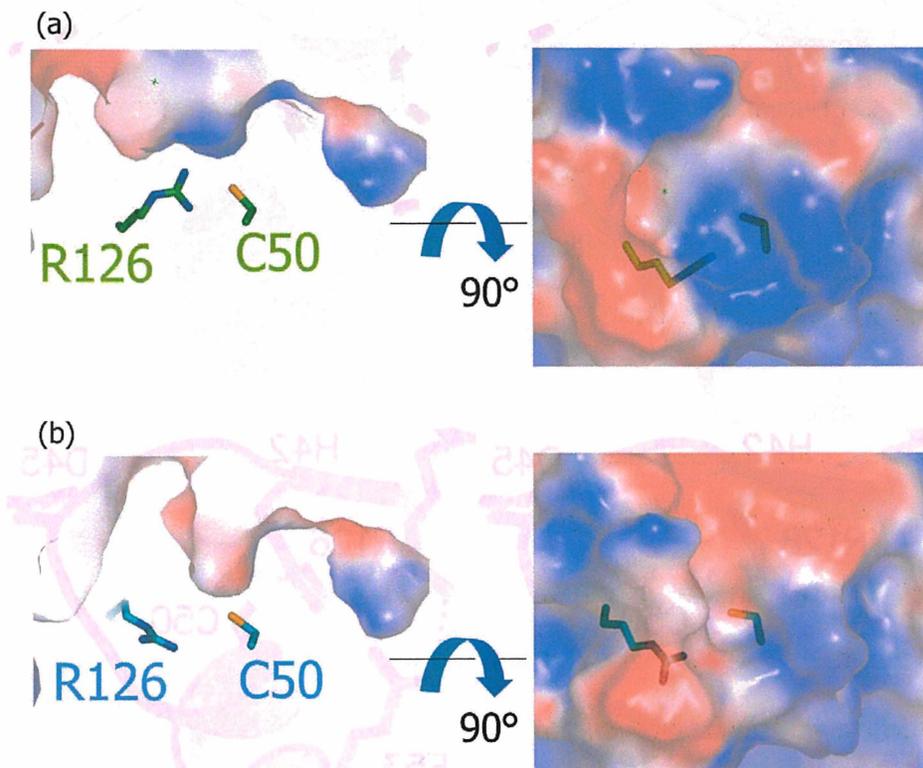


Figure II-8 Electrostatic surfaces of the reduced and the intermediate form of ApTPx. (a) Side view (left) and top view (right) of the reduced form. (b) Side view (left) and top view (right) of the intermediate form.

II.3.10 H₂O₂ reduction mechanism by ApTPx

Structural differences between the reduced, the intermediate, and the oxidized forms gave us the definite clues to increase our understanding of the reaction mechanism, in which H₂O₂ was reduced to H₂O by ApTPx. **Figure II-9** shows the proposed reaction mechanism based on three kinds of structures mentioned above. In the reduced state (**Figure II-9 (I)**), Cys50 and Asp45 interact with Arg126 and His42 *via* hydrogen bonds, respectively. Because of some H₂O₂ signal, which has not been elucidated, guanidium group of Arg126 extracts the hydrogen atom of the thiol group of Cys50, and Cys50 is deprotonated to become a thiolate anion (Cys50-SH⁻). At the same time, His42 breaks its interaction with Asp45 and forms a hydrogen bond with S^γ atom of Cys50, stabilizing the thiolate anion of Cys50 (**Figure II-9 (II)**). After the formation of the intermediate state shown in **Figure II-9 (II)**, the thiolate anion of Cys50 nucleophilically attacks one of the oxygen atoms of H₂O₂ **Figure II-9 (III)**, reducing H₂O₂ to H₂O. Cys50 is oxidized, and the hydrogen atom bound to the N^{δ1} atom of His42 shifts onto the S^γ atom of Cys50, forming the oxidized form, which consists of the hypervalent sulfur structure (**Figure II-9 (IV)**). After that, Cys213 in another molecule of the dimer approaches to Cys50 and they form inter-subunit disulfide bond (**Figure II-9 (V)**), producing H₂O. Finally, the inter-subunit disulfide bond is reduced by Thioredoxin, and ApTPx finishes one cycle of the reaction.

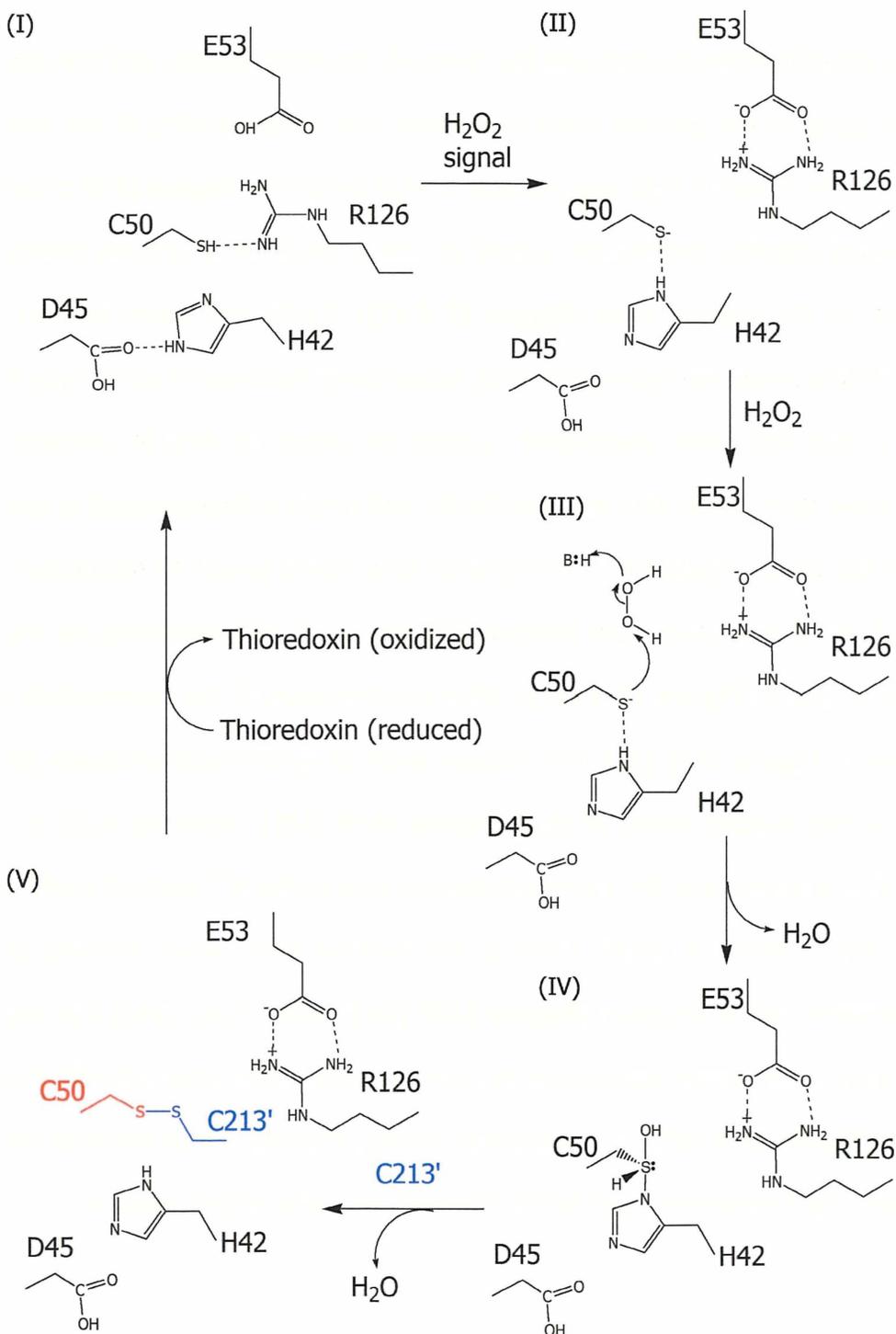


Figure II-9 H₂O₂ reduction mechanism by ApTPx

II.4 Summary of Chapter II

In this Chapter, we described the first crystal structure of archaeal peroxiredoxin (ApTPx) determined in its three forms: the reduced, the intermediate, and the oxidized form. In the reduced structure, we described the distinctive features when compared with mesophilic peroxiredoxines. In the intermediate form, the movement of Arg126, which should be caused by some H_2O_2 signal, and the accompanying structural changes drive the pocket connecting to the active site open for the substrate H_2O_2 to easily access to Cys50. In the oxidized form, we firstly found the structure containing a hypervalent sulfur in the protein structure. These three structures revealed the detailed reaction mechanism where H_2O_2 is reduced by ApTPx.

II. 5 References

- Abrahams, J.P., and Leslie, A.G. 1996. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr* **52**: 30-42.
- Alphey, M.S., Bond, C.S., Tetaud, E., Fairlamb, A.H., and Hunter, W.N. 2000. The structure of reduced tryparedoxin peroxidase reveals a decamer and insight into reactivity of 2Cys-peroxiredoxins. *J Mol Biol* **300**: 903-916.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N.S., et al. 1998b. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 904-925.
- Choi, H.J., Kang, S.W., Yang, C.H., Rhee, S.G., and Ryu, S.E. 1998. Crystal structure of a novel human peroxidase enzyme at 2.0 Å resolution. *Nat Struct Biol* **5**: 400-406.
- Collaborative Computational Project, N. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**: 760-763.
- de La Fortelle, E., and Bricogne, G. 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol* **276**: 472-494.
- Declercq, J.P., Evrard, C., Clippe, A., Stricht, D.V., Bernard, A., and Knoops, B. 2001. Crystal structure of human peroxiredoxin 5, a novel type of mammalian peroxiredoxin at 1.5 Å resolution. *J Mol Biol* **311**: 751-759.
- Doublet, S. 1997. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* **276**: 523-530.
- Echalier, A., Trivelli, X., Corbier, C., Rouhier, N., Walker, O., Tsan, P., Jacquot, J.P., Aubry, A., Krimm, I., and Lancelin, J.M. 2005. Crystal structure and solution NMR dynamics of a D (type II) peroxiredoxin glutaredoxin and thioredoxin dependent: a new insight into the peroxiredoxin oligomerism. *Biochemistry* **44**: 1755-1767.
- Higgins, D.G., Thompson, J.D., and Gibson, T.J. 1996. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* **266**: 383-402.
- Hirotsu, S., Abe, Y., Okada, K., Nagahara, N., Hori, H., Nishino, T., and Hakoshima, T. 1999. Crystal structure of a multifunctional 2-Cys peroxiredoxin heme-binding protein 23 kDa/proliferation-associated gene product. *Proc Natl Acad Sci U S A* **96**: 12333-12338.

- Holmgren, A. 1985. Thioredoxin. *Annu Rev Biochem* **54**: 237-271.
- Holmgren, A. 1989. Thioredoxin and glutaredoxin systems. *J Biol Chem* **264**: 13963-13966.
- Jeon, S.J., and Ishikawa, K. 2002. Identification and characterization of thioredoxin and thioredoxin reductase from *Aeropyrum pernix* K1. *Eur J Biochem* **269**: 5423-5430.
- Jeon, S.J., and Ishikawa, K. 2003. Characterization of novel hexadecameric thioredoxin peroxidase from *Aeropyrum pernix* K1. *J Biol Chem* **278**: 24174-24180.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47 (Pt 2)**: 110-119.
- Kabsch, W., and Sander, C. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**: 2577-2637.
- Kawarabayasi, Y., Hino, Y., Horikawa, H., Yamazaki, S., Haikawa, Y., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., et al. 1999. Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1. *DNA Res* **6**: 83-101, 145-152.
- Kitano, K., Kita, A., Hakoshima, T., Niimura, Y., and Miki, K. 2005. Crystal structure of decameric peroxiredoxin (AhpC) from *Amphibacillus xylanus*. *Proteins* **59**: 644-647.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK - a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**: 283-291.
- Li, S., Peterson, N.A., Kim, M.Y., Kim, C.Y., Hung, L.W., Yu, M., Legin, T., Segelke, B.W., Lott, J.S., and Baker, E.N. 2005. Crystal Structure of AhpE from *Mycobacterium tuberculosis*, a 1-Cys peroxiredoxin. *J Mol Biol* **346**: 1035-1046.
- Maeda, N., Kanai, T., Atomi, H., and Imanaka, T. 2002. The unique pentagonal structure of an archaeal Rubisco is essential for its high thermostability. *J Biol Chem* **277**: 31656-31662.
- Merrit, E.A., and Murphy, M.E. 1994. Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 869-873.
- Nakamura, T., Matsumura, H., Inoue, T., Kai, Y., Uegaki, K., Hagihara, Y., Ataka, M.,

- and Ishikawa, K. 2005. Crystallization and preliminary X-ray diffraction analysis of thioredoxin peroxidase from the aerobic hyperthermophilic archaeon *Aeropyrum pernix* K1. *Acta Crystallographica Section F* **61**: 323-325.
- Nicholls, A., Sharp, K.A., and Honig, B. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**: 281-296.
- Otwinowski, Z., and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**: 307-326.
- P. M. W. Gill, B.G.J., J. A. Pople, and M. J. Frisch. 1993. The Performance of the Becke-Lee-Yang-Parr (B-LYP) Density Functional Theory with Various Basis-Sets. *J. Chem. Physics* **98**: 5612 - 5626.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.
- Pineyro, M.D., Pizarro, J.C., Lema, F., Pritsch, O., Cayota, A., Bentley, G.A., and Robello, C. 2005. Crystal structure of the tryparedoxin peroxidase from the human parasite *Trypanosoma cruzi*. *J Struct Biol* **150**: 11-22.
- Rhee, S.G., Kang, S.W., Chang, T.S., Jeong, W., and Kim, K. 2001. Peroxiredoxin, a novel family of peroxidases. *IUBMB Life* **52**: 35-41.
- Sako, Y., Nomura, N., Uchida, A., Ishida, Y., Morii, H., Koga, Y., Hoaki, T., and Maruyama, T. 1996. *Aeropyrum pernix* gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100 degrees C. *Int J Syst Bacteriol* **46**: 1070-1077.
- Sarma, G.N., Nickel, C., Rahlfs, S., Fischer, M., Becker, K., and Karplus, P.A. 2005. Crystal structure of a novel *Plasmodium falciparum* 1-Cys peroxiredoxin. *J Mol Biol* **346**: 1021-1034.
- Schroder, E., Littlechild, J.A., Lebedev, A.A., Errington, N., Vagin, A.A., and Isupov, M.N. 2000. Crystal structure of decameric 2-Cys peroxiredoxin from human erythrocytes at 1.7 Å resolution. *Structure Fold Des* **8**: 605-615.
- Schroder, E., and Ponting, C.P. 1998. Evidence that peroxiredoxins are novel members of the thioredoxin fold superfamily. *Protein Sci* **7**: 2465-2468.
- Terwilliger, T.C., and Berendzen, J. 1999. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55 (Pt 4)**: 849-861.
- Vieille, C., and Zeikus, G.J. 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**:

1-43.

Wood, Z.A., Poole, L.B., Hantgan, R.R., and Karplus, P.A. 2002. Dimers to doughnuts: redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* **41**: 5493-5504.

Wood, Z.A., Poole, L.B., and Karplus, P.A. 2003a. Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* **300**: 650-653.

Wood, Z.A., Schroder, E., Robin Harris, J., and Poole, L.B. 2003b. Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem Sci* **28**: 32-40.

Conclusion

In this study, it was inquired that the relationship between structures and their function of the proteins from two kinds of archaea identified in Kodakara Island, Japan, based on structural biology and on molecular dynamics.

In Chapter I, the first crystal structure of TBP-interacting protein (*7k-TIP26*) was determined. It was also provided that the first molecular mechanism where *7k-TIP26* prevents TBP from binding to TATA-DNA in the absence of TFB.

In Chapter II, the first crystal structure of archaeal peroxiredoxin (ApTPx) was determined in its three forms: the reduced form, the intermediate form, and the oxidized form. These structural results, combined with the results of computational chemistry, revealed the detailed H₂O₂ reduction mechanism by ApTPx.

List of publication

- 1) Crystallization and preliminary X-ray analysis of TBP-interacting protein from the hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1.
Yamamoto, T., Matsuda, T., Sakamoto, N., Matsumura, H., Inoue, T., Morikawa, M., Kanaya, S., & Kai, Y.
Acta Crystallogr D Biol Crystallogr. 2003, 59: 372-374.

- 2) Crystal structure of TBP-interacting protein (*Tk-TIP26*) and implications for its inhibition mechanism of the interaction between TBP and TATA-DNA.
Yamamoto, T., Matsuda, T., Inoue, T., Matsumura, H., Morikawa, M., Kanaya, S., & Kai, Y.
Prot. Sci., in press.

- 3) Crystal Structure of Thioredoxin Peroxidase from Aerobic Hyperthermophilic Archaeon *Aeropyrum pernix* K1.
Nakamura, T., Yamamoto, T., Inoue, T., Matsumura, H., Kobayashi, A., Hagihara, Y., Uegaki, K., Ataka, M., Kai, Y., & Ishikawa, K.
Proteins, in press.

- 4) A novel reaction mechanism via hypervalent sulfur intermediate from the structure of oxidized-form of archaeal Thioredoxin peroxidase.
Yamamoto, T., Nakamura, T., Abe, M., Inoue, T., Matsumura, H., Hagihara, Y., Ishikawa, K., Ataka, M., & Kai, Y.
in preparation.

List of supplementary publication

1) Crystal structure of intein homing endonuclease II encoded in DNA polymerase gene from Hyperthermophilic Archaeon *Thermococcus kodakaraensis* strain KOD1.

Matsumura, H., Takahashi, H., Inoue, T., Yamamoto, T., Hashimoto, H., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., & Kai, Y.

Proteins, in contribution.

2) Crystal structure of thermostable lipase from *Geobacillus sp.* T1 and its activation mechanism by cation- π stacking interaction.

Takahiko Yamamoto, Raja Noor Zaliha Raja Abd. Rahman, Leow Thean Chor, Abu Bakar Salleh, Mahiran Basri, Kazufumi Takano, Shigenori Kanaya, Hiroyoshi Matsumura, Tsuyoshi Inoue, & Yasushi Kai.

in preparation.

Acknowledgement

The author wishes to express his special gratitude to Professor Yasushi Kai for his enthusiastic guidance, helpful suggestion, and encouragement throughout the studies. The author sincerely appreciates to Associate Professor Tsuyoshi Inoue for his continuous encouragement, stimulating discussion, and interesting suggestion for the accomplishment of this work. He acknowledges Lecturer Nobuko Kanehisa and Assistant Professor Hiroyoshi Matsumura for their kind encouragement and helpful advices.

In Chapter I, He is grateful to Professor Shigenori Kanaya (Department of Materials and Life Science, Graduate School of Engineering, Osaka University), Professor Masaaki Morikawa (Laboratory for Nanosystems Physiology, Research Institute for Electronic Science, Hokkaido University), and Dr. Tomoki Matsuda (Laboratory of Environmental Molecular Biology, Graduate School of Environmental Earth Science, Hokkaido University) for their stimulating discussion and profound advices. He is also grateful to Mr. Seiji Hashima (ATTO Co. Ltd.) for his kind help in the data analyses of SDS-PAGE.

In Chapter II, the author acknowledges Dr. Mitsuo Ataka, Dr. Kazuhiko Ishikawa, Dr. Kouichi Uegaki, Dr. Yoshihisa Hagihara, and Dr. Tsutomu Nakamura (Research Institute for Cell Engineering, National Institute of Advanced Industrial Science and Technology (AIST)) for stimulating discussion and continuous guidance. He also wishes to express special acknowledgement for Lecturer Manabu Abe (Department of Applied Chemistry, Graduate School of Engineering, Osaka University) for his kind help in calculation of optimized structure of

hypervalent sulfur compounds and his profound discussion of reaction mechanism in ApTPx.

The author is grateful to Dr. Keiko Miura, Dr. Masahide Kawamoto, Dr. Kazuya Hasegawa, Dr. Nobutaka Shimizu, and Dr. Hisahide Sakai (Japan Synchrotron Radiation Research Institute (JASRI)) for their kind help in the collection of the X-ray diffraction data at BL38B1, BL40B2 and BL41XU in SPring-8, and to Professor Tomitake Tsukihara, Professor Atsushi Nakagawa and Dr. Eiki Yamashita, and Dr. Masato Yoshimura (Institute for Protein Research, Osaka University) for their kind help in fundamental data collection at BL44XU in SPring-8.

Special thanks are given to the author's co-workers Mr. Junichi Nakao (from April 2001 to March 2004), Ms. Yumiko Iga (from April 2004 to March 2005), and Mr. Tadashi Goto (from April 2005 to March 2006) for their stimulating and fruitful cooperation. He is also grateful to Dr. Hiroshi Hashimoto, Dr. Eichi Mizohata, Dr. Yosuke Okano, and Dr. Xie Yong and all the other members of research group of Kai Laboratory for their kind assistance and interesting discussions. The author expresses his special thanks to the center of excellence (21COE) program "Creation of Integrated EcoChemistry of Osaka University".