T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Structural studies on RNA related enzymes : X-ray analyses of RNA 3'-terminal phosphate cyclase and two threonyl-tRNA synthetases with complementary functions, from crenarchaea
著者(和文)	清水了
Author(English)	Satoru Shimizu
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第7505号, 授与年月日:2009年3月26日, 学位の種別:課程博士, 審査員:関根 光雄
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第7505号, Conferred date:2009/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Powered by T2R2 (Science Tokyo Research Repository)

Structural studies on RNA related enzymes; X-ray analyses of RNA 3′-terminal phosphate cyclase and two threonyl-tRNA synthetases with complementary functions, from crenarchaea

A Doctoral Thesis

by

Satoru Shimizu

Submitted to the

Department of Life Science

Graduate School of Bioscience and Biotechnology

Tokyo Institute of Technology

2009

Abstract

The non-coding RNAs, such as rRNA, mRNA, tRNA etc, play the essential parts of the life system. Many kind of proteins are involved in producing such functional RNAs. Although several these proteins were revealed the tertiary structures, their detailed mechanisms are not fully understood. In this study, other enzymes related to RNA processing, RNA 3′-terminal phosphate cyclase and threonyl-tRNA synthetases from crenarchaea, have been subjected to reveal the tertiary structures.

1. X-ray analyses of RNA 3′-terminal phosphate cyclase from *Sulfolobus tokodaii* and its complexes

In pre-tRNA splicing, endonuclease cleaves RNA at first into two chains with the 3′-phosphomonoester end and the 5′-hydroxyl end, respectively, and then RNA ligase requires the two RNA chains to be modified at the ends. It is thought that RNA 3′-terminal phosphate cyclase (Rtc) catalyzes the ATP-dependent conversion of the 3′-phosphomonoester end to the 2′,3′-cyclic phosphodiester end in the presence of Mg^{2+} ion. This reaction may occur though the following two steps:

(i) Rtc + ATP → Rtc-AMP + PP*i*

(ii) RNA-3′p +Rtc-AMP → RNA-N>p + Rtc + AMP

In the first step, Rtc and ATP react with each other to form a covalently bound Rtc-AMP. In the second step, when a truncated RNA with a phosphate group at the 3-terminus is bound to Rtc-AMP, RNA-3′pp5′A intermediate is formed and then the cyclization reaction between the 2′-hydroxyl group and the 3′-phosphate group probably occurs non-enzymatically to generate RNA with the 2′,3′-cyclic end (RNA-N>p). Although the X-ray structure of Rtc from *E. coli* was reported, the substrate binding site and the catalytic reaction mechanism is not clear. In this study, X-ray analyses of Rtc from *Sulfolobus tokodaii* (*St*-Rtc) and its complexes with ATP in the following states have been performed, in order to reveal the detailed mechanism of Rtc catalysis. The selenomethionine derivative (*St*-Rtc*) has been at first solved by MAD method. The others are (1) *St*-Rtc* in complex with Mn^{2+} ion (*St*-Rtc*+Mn) to identify Mg^{2+} ion binding site, (2) *St*-Rtc* in complex with ATP and no Mg^{2+} ion (*St*-Rtc*+ATP) to get an insight of ATP binding, (3) *St*-Rtc in complex with Mg^{2+} ion and ATP under acidic condition (*St*-Rtc+AMP) to see the structure just before the first reaction, and (4) *St*-Rtc in complex with Mg^{2+} ion and ATP under basic conditions (*St*-Rtc*-AMP) to see the structure after the first reaction.

The crystal structures of *St*-Rtc*, *St*-Rtc, *St*-Rtc*+Mn, *St*-Rtc*+ATP, *St*-Rtc+AMP and *St*-Rtc*-AMP have been successfully determined at 2.25, 2.2, 3.2, 2.4 2.9 and 2.25

Å resolutions, respectively. The two subunits forming a dimer have been found in every crystal obtained under different conditions (Figure 1). The subunit is composed of a large domain and a small domain. In the large domain, the βαβαββ motifs are repeated by three times to fold into a pseudo three-fold symmetric structure. On the other hand, the small domain contains a βαββββα motif, which is inserted into the third motif of the large domain.

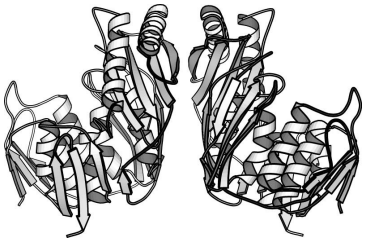


Figure 1. Two subunits forming a dimer found in the *St*-Rtc crystals.

In the *St*-Rtc*+Mn crystal, it was difficult to find the density of added Mn^{2+} ions. It may be ascribed to low Mn^{2+} concentration in crystallization and/or to no ATP. The latter means that Mg^{2+} binding requires a help of the bound ATP. In the *St*-Rtc*+ATP crystal, the adenine ring of ATP is stacked between the side chains of Pro126 and His283. In addition, Asp250 forms a hydrogen bond with the adenine base and Asp286 forms double hydrogen bonds with the ribose. These structural features are also found in the *St*-Rtc+AMP and *St*-Rtc*-AMP crystals. However, only the α-phosphate group is disordered. To fix it in a correct orientation, a Mg^{2+} ion will be bound to nearby Glu10. In the *St*-Rtc+AMP crystal, ATP is hydrolyzed in acidic condition into two parts, pyrophosphate and AMP, and they are bound separately. The α-phosphate group of AMP may be allowed to disorder or further hydrolyzed to adenosine. In the *St*-Rtc*-AMP crystal, the α-phosphate group of AMP is covalently bound to the Nε atom of His307 (Figure 2) as the end product of the first reaction. Based on these structures, a plausible reaction mechanism of the first reaction has been proposed, and possible molecular movement in the second reaction has been speculated.

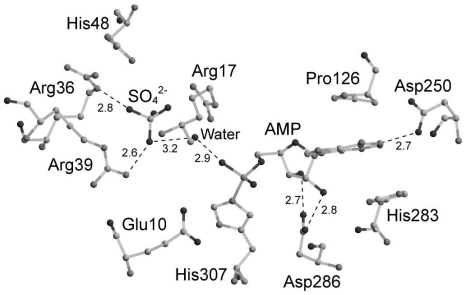


Figure 2. The AMP binding sites in the St-Rtc+AMP* crystal

2. X-ray analysis of two threonyl-tRNA synthetases with complementary functions from *Aeropyrum pernix* and *Sulfolobus tokodaii*

To maintain the highest fidelity in protein synthesis, twenty kinds of aminoacyl-tRNA synthetases (aaRSs) exist in general for twenty kinds of amino acids (aa). Each aaRS is specialized to recognize only the cognate amino acid and the cognate $tRNA^{aa}$. It was reported that threonyl-tRNA synthetase (ThrRS) from *E. coli* (*Ec*ThrRS) is composed of the three domains, editing, catalytic and anticodon-binding. The editing domain is essential to hydrolyze the misaminoacylated Ser-$tRNA^{Thr}$, because threonine as well as serine is acylated to $tRNA^{Thr}$ in the catalytic domain. In archaea *Pyrococcus*

abyssi ThrRS, the X-ray structure of the truncated editing domain is quite different from that of *Ec*ThrRS.

Recently two genes were found for ThrRS in the genomes of some organisms. It was proposed that two proteins have different functions complementary to each other, one is only for catalysis and the other is only for editing. In this study, it has been surveyed that other creanarchaeal organisms also contain the two ThrRS genses in their genomes. Sequence alignment analyses of the two gene products (ThrRS-1 and ThrRS-2), have shown that ThrRS-1s possess the highly conserved catalytic and anticodon-binding domains, but no editing domain, while ThrRS-2s are composed of the archaeal type editing domain and the anticodon-binding domain but no catalytic domain. Furthermore, they contain an additional unknown domain. In order to clarify the three-dimensional structures by the X-ray method, ThrRS-1 and ThrRS-2 from *Aeropyrum pernix* (*Ap*) and those from *Sulfolobus tokodaii* (*St*) have been overexpressed in *E. coli*, purified and crystallized.

The crystal structure of *Ap*ThrRS-1 has been successfully determined at 2.3 Å resolution, as the first example. *Ap*ThrRS-1 is a dimeric enzyme composed of the two identical subunits. The subunit consists of two domains. The *N*-terminal domain has the three motifs characteristic as the catalytic domain of the class II aaRS. In this domain, the zinc atom is bound to Cys112, His166 and His310 (Figure 3). By comparing with *Ec*ThrRS structure (Figure 4), it is easily identified that the second domain is for anticodon-binding. Therefore, it is concluded that the editing domain is completely missing as expected. Superimposition of the *Ap*ThrRS-1 structure onto that of *Ec*ThrRS in complex with $tRNA^{Thr}$ shows that their overall peptide foldings are similar to each other, but different in the peripheral regions. In the anticodon-binding sites, the amino acid residues interacting with G35 and U36 of $tRNA^{Thr}$ are conserved. These features are consistent to that *Ap*ThrRS-1 cannot distinguish threonine from serine without the editing domain, suggesting the necessity of *Ap*ThrRS-2 for removing serine moiety of the misacylated Ser-$tRNA^{Thr}$. Finally, possible structure features of ThrRS-2 have been discussed.

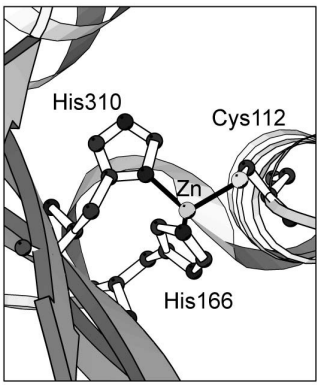


Figure 3. The zinc finger of *Ap*ThrRS-1

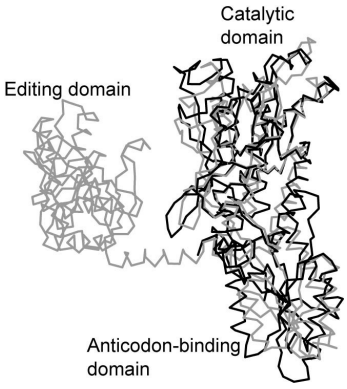


Figure 4. Superimposed Cα chains of *Ap*ThrRS-1 (black) onto *Ec*ThrRS (gray, PDB-ID=1QF6).

Contents

Contents

List of abbreviations used in the thesis

A

Aa; Aquifex aeolicus

Aa-Rtc; Rtc from *Aa*

Ap; *Aeropyrum pernix*

*Ap*ThrRS-1; one of two ThrRS gene product from *Ap*, catalytic emzyme

*Ap*ThrRS-2; one of two ThrRS gene product from *Ap*, editing emzyme

*Ap*ThRS-1*; selenomethionine derivertive of *Ap*TrhRS-1

Ap-Rtc; Rtc from *Ap*

D

Dm; *Drosophila melanogaster* (fruit fly)

Dm-Rtc; Rtc from *Dm*

E

Ec-Rtc; Rtc from *Escherichia coli*

H

HEPES; 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, $C_8H_{18}N_2O_4S$, titrated using NaOH

Hs; *Homo sapience*

*Hs*ThrRS; ThrRS from *Hs*

Hs-Rtc; Rtc from *Hs*

M

Mm; *Methanosarcina mazei*

*Mm*ThrRS; ThrRS from *Mm*

Ms; *Metallosphaera sedula*

*Ms*ThrRS-1; one of two ThrRS gene product from *Ms*, catalytic emzyme

*Ms*ThrRS-2; one of two ThrRS gene product from *Ms*, editing emzyme

P

Pa; *Pyrococcus abyssi*

*Pa*ThrRS; ThrRS from *Pa*

Ph; Pyrococcus horikoshii

Ph-Rtc; Rtc from *Ph*

R

Rtc; RNA 3′-terminal phosphate cyclase

S

Sa; *Staphylococcus aureus*

*Sa*ThrRS; ThrRS from *Sa*

Sc; *Saccharomyces cerevisiae*

*Sc*ThrRS; ThrRS from *Sc*

SDS-PAGE; poly-acrylamide gel electrophoresis

*S*f; *Sulfolobaceae* family (containing *St*, *Ss*, *Si*, and *Ms*)

*Sf*ThrRS-1; one of two ThrRS gene product from *S*f, catalytic emzyme

*S*fThrRS-2; one of two ThrRS gene product from *S*f, editing emzyme

Si; *Sulfolobus acidocaldarius*

*Si*ThrRS-1; one of two ThrRS gene product from *Si*, catalytic emzyme

*Si*ThrRS-2; one of two ThrRS gene product from *Si*, editing emzyme

Ss; *Sulfolobus solfataricus*

*Ss*ThrRS-1; one of two ThrRS gene product from *Ss*, catalytic emzyme (ThrRS-cat)

*Ss*ThrRS-2; one of two ThrRS gene product from *Ss*, editing emzyme (ThrRS-ed)

St; *Sulfolobus tokodaii*

St-Rtc; Rtc from *St* (*apo*-form)

St-Rtc*; selenomethionine derivative of *St*-Rtc

St-Rtc*+ATP; *St*-Rtc* in complex with ATP and no Mg^{2+} ion, which will provide an insight of ATP binding

St-Rtc*+Mn; *St*-Rtc* in complex with Mn^{2+} ion to identify the Mg^{2+} ion binding site

St-Rtc*-AMP; *St*-Rtc* in complex with Mg^{2+} and ATP under basic condition

St-Rtc+AMP; *St*-Rtc in complex with Mg^{2+} and ATP under acidic condition

*St*ThrRS-1*; selenomethionine derivertive of *St*ThrRS-1

*St*ThrRS-2*; selenomethionine derivertive of *St*ThrRS-2

*St*ThrRS-2; one of two ThrRS gene product from *St*, editing emzyme

*St*ThRS-1; one of two ThrRS gene product from *St*, catalytic emzyme

T

Tk; *Thermococcus kodakaraensis*

Tk-Rtc; Rtc from *Tk*

ThrRS; threonyl-tRNA synthetase

ThrRS-1; one of two ThrRS gene products containing catalytic and anticodon-binding domains and lacking editing domain

ThrRS-2; one of two ThrRS gene products containing editing and anticodon-binding domains and lacking catalytic domain

ThrRS-cat; *Ss*ThrRS-1

ThrRS-ed; *Ss*ThrRS-2

Tris-HCl; tris(hydroxymethyl)aminomethane, $NH_2C(CH_2OH)_3$, titrated using HCl

X

XAFS; X-ray absorption fine structure

< Chapter 1; Introduction >

1-1. RNA related enzymes

RNAs perform so many different functions essential for maintaining the cell system. Syntheses of RNAs are catalyzed by an RNA polymerase in transcription of DNA strand [1]. The well known RNAs are messenger RNAs (mRNA), transfer RNAs (tRNA), ribosome RNA [2], the spliceosome RNA [3] and some ribozymes [4], some of which have catalytic activities. Many protein enzymes are involved in generating these functional RNAs. For example, a poly-A tail and a 5'-cap are added to eukaryotic pre-mRNA by poly-A polymerase and three enzymes containing phosphatase, guanylyl transferase and methyl transferase (reviewed in [5]). Aminoacylation of tRNAs are catalyzed by the cognate aminoacyl-tRNA synthetases [6], which are used in the translation of the genetic codes during protein synthesis. Ribonucleoprotein particles (RNPs) [7] regulate mRNAs in posttranscription of genes. These enzymes, related closely to generate the functional RNAs, have been well studied for understanding the molecular mechanisms of their functions, but the details are still unclear.

In order to reveal the mechanisms of RNA modifying enzymes, two interesting enzymes; RNA 3′-terminal phosphate cyclase (Rtc) and threonyl-tRNA synthetase (ThrRS) from crenarchaea are suggested for structural studies in this thesis as described below.

1-2. RNA 3'-terminal phosphate cyclase (Rtc)

1-2-a. RNA splicing

Splicing is a modification of an RNA, in which introns are removed and exons are joined. Several years ago, it was considered that RNA splicing is necessary to make a mature mRNA in the nuclear envelop of eukaryotes. The process contains a series of reactions which are catalyzed by a splicesome composed of small nuclear ribonucleoproteins (snRNPs). This idea was expanded to other organisms after discovery of self-spliced introns; group I intron [8] and group II introns [9] in bacteria cell, and identification of putative introns in tRNA genes of archaea [10].

Recently, six different classes of introns have been identified, and they are grouped in four different splicing mechanisms that can be categorized as follows; (1) group I introns; (2) group II, group III introns; (3) nuclear mRNA introns; (4) archaeal and eukaryotic nuclear tRNA introns (reviewed in [11-12] and see also Figure 1-2-a). The first two mechanisms are self-splicing without requiring any proteins. The 3′-OH of the 5′-exon triggers the transesterification at the 3′-spliced site to join the exons together

(see Figure 1-2-a (1), (2)). In the third mechanism, nuclear mRNA introns are removed by spliseosome, of which the catalytic mechanism is similar to those of group II and III introns [13]. This splicing is presumed to be mainly catalyzed by several snRNAs built into snRNPs.

The fourth mechanism is completely different from the reaction catalyzed by protein enzymes. Archaeal and eukaryotic tRNA introns are excised by a splicing endnuclease that generates 5′-OH and 2′,3′-cyclic phosphate (see Figure 1-2-a(3)). Then the eukaryotic ligase possesses 2′,3′-phosphodiesterase, 5′-kinase and ligase activity to produce 3′,5′-linkage and 2′-phosphate group. The third enzyme, 2′-phosphatetransferase, removes the 2′-phosphate at the ligation junction. In archaea, the ligation product has no 2′-phosphate at the ligation junction despite the presence of the 2′-phosphotransferase homolog in all archaeal genomes [14-15]. Rtc is considered to be related to the forth splicing mechanism before the ligation because a ligase requires 2′, 3′-cyclic phosphate group for its activity.

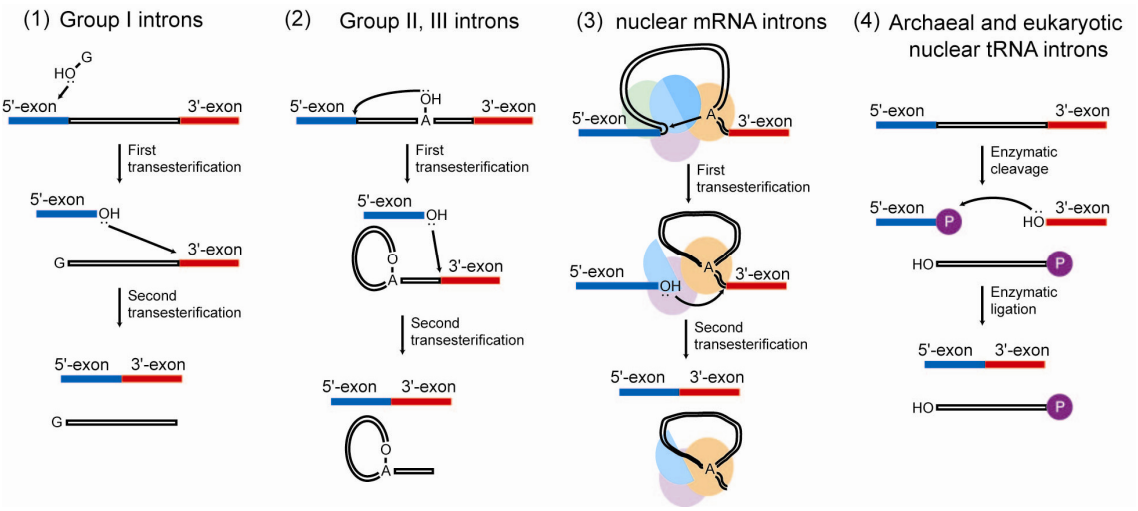


Figure 1-2-a. Four different mechanisms for intron removal. 5'-exons, 3'-exons and introns are shown by blue, red and white bars, respectively. (1) The mechanism of Group I introns containing the two transesterification reaction. (2) and (3) That of Group II, Group III and nuclear mRNA introns with producing an intron lariat. Blue, light blue, green, orange and purple circles indicate snRNPs. (4) The mechanism of archaeal and nuclear tRNA introns intermediated with some protein enzymes that generate 5′-OH and 2′, 3′ -cyclic phosphate in the first step. Purple circles indicate phosphate groups.

1-2-b. Background of Rtc

Rtc was initially identified in extracts of HeLa cell and *Xenopus* oocytes [16, 17]. This enzyme is localized in nucleoplasm, consistent with its postulated role in RNA processing.

The enzyme catalyzes the ATP-dependent conversion of 3′-phosphate to

2′,3′-terminal phosphodiester at the end of various RNA substrates. It is proposed that the cyclization of the 3′-phosphate occurs in the following three steps:

(i) Rtc+ATP→Rtc-AMP+PP*i*

(ii) RNA-$N^{3'}$p+Rtc-AMP→RNA-$N^{3'}$p$p^{5'}$A+Rtc

(iii) RNA-$N^{3'}$p$p^{5'}$A→RNA-N>p+AMP.

In the first step, Rtc and ATP react with each other to form a covalent bonded Rtc-AMP intermediate. The second step is proposed by detecting the RNA-$N^{3'}$p$p^{5'}$A formation when the ribose at the RNA 3′-terminus is replaced with the 2′deoxy- or 2′-*O*-methylribose [18]. The third step probably takes place non-enzymatically [19-21].

It is interesting to note a similar formation of covalent protein-NMP intermediate followed a transfer of NMP to the terminal phosphate also occurs by RNA and DNA ligases, or capping enzymes (reviewed in [22]). In these ligases, the nucleotidyl group is transferred to the 5′-terminal phosphate or pyrophosphate and a covalent lysyl-NMP intermediate is formed at the active-site lysine residue of a conserved sequence motif, KxDG (in single-letter amino acid code where x denotes any amino acid). However, this motif is not present in the sequence of Rtc, and it was reported that the adenylyl group is transferred not to a lysine but to a histidine residue of Rtc of *E. coli*. [23].

Rtc genes were identified in all organisms of eukaryote, bacteria and archaea [20]. The three dimensional structure of Rtc of *E. coli* was determined by X-ray crystallography [24]. The structure consists of two domains, large and small domains. The large domain consists of three repeats of a folding unit constructed with two α helices and a four-stranded β sheet. This motif of the folding unit appears in several nucleic acid binding proteins, for example the C-terminal domain of the bacterial initiation factor 3 (IF3) [25], and the folding of the large domain is similar to the partial structure of UDP–N-acetylglucosamine enolpyruvyl transferase (MurA) [26-27] and 5-enol-pyruvylshikimate-3-phosphate synthase (EPSP synthase) [28]. However the catalytic mechanism of Rtc is not clarified yet.

Furthermore, the biological role of Rtc is also unclear. The splicing endonucleases directly generate 5'-tRNA halves carrying 2', 3'-cyclic phosphate [29] in tRNA splicing reactions. The cellular RNA ligase identified in bacteria also requires 2', 3'-cyclic ends for ligation [30, 31]. It is possible that Rtc may regenerate the cyclic ends for the subsequent ligation step with 3′-monophospate-terminated RNA formed by decyclizing with phosphodiesterase or hydrolysis after the above reactions.

1-2-c. Purpose and significance of the X-ray analysis of Rtc from *Sulfolobus tokodaii*

Although the structure of Rtc from *E. coli* was determined by X-ray analysis, the substrate binding site and the catalytic mechanism is not clear. To reveal the detail mechanism of Rtc catalysis, X-ray analyses of Rtc from *Sulfolobus tokodaii* and its complex with some substrate were performed. *Sulfolobus tokodaii* is an aerobic thermoacidophilic crenarchaeon and proteins are generally stable and easy for biochemistry chemical experiments. In order to determine the *apo*-form crystal structure of *Sulfolobus tokodaii* in the pre-reaction state, Rtc of *Sulfolobus tokodaii* (*St*-Rtc) was overexpressed in *E. coli* and crystallized. In order to reveal the detailed interactions at the substrate binding site and the reaction mechanism of the first step, *St*-Rtc was crystallized with ATP in several states. Based on these structures, a possible reaction mechanism will be proposed.

1-3. Threonyl-tRNA synthetase (ThrRS)

1-3-a. Protein biosynthesis

A flow of information for gene expression is processed from DNA to protein through RNA using many enzymes. The first step is called transcription, in which the corresponding messenger RNA (mRNA) is produced by an RNA polymerase from one strand of DNA duplex. In the second step called translation, according to the gene information, the three nucleotides (codon) on mRNA is converted to the cognate amino acid of protein on a ribosome.

For beginning translation, an initiator tRNA is bound to the start codon of mRNA positioned at the P site of a ribosome with help of an initiation complex. A cognate aminoacyl-tRNA adaptable to the next codon is bound to the empty A site, and the formyl-methionine moiety of the initiator tRNA is transferred to the neighboring amino acid moiety to make a peptic bond. The remaining initiator tRNA is moved with the mRNA from the P site to the E site to be released. At the same time, the peptidyl-tRNA is shifted from the A site to the P site to make the A site free. Then the opened A site is filled again by the new aminoacyl-tRNA corresponding to the next codon of mRNA for a peptide elongation. By repeating these steps, the peptide is extended to become the protein. At the end of the elongation, a release factor binds to the stop codon on tRNA to terminate the translation and release the produced protein from the ribosome.

In this mechanism, the proper attachment of an amino acid to the cognate tRNA has to perform by an aminoacyl tRNA syntetase (aaRS) specific to the tRNA prior to and independent to the protein synthesis. AaRS catalyzes such an aminoacylation by the

following two step reactions. At first, aaRS synthesizes aminoacyl-AMP from ATP and amino acid. Second, the activated aminoacyl group of the aminoacyl-AMP is transferred to the cognate tRNA to form the aminoacyl-tRNA. To maintain high fidelity in translation, twenty kinds of aaRSs exist in general for twenty kinds of amino acids. And they are divided into two classes, class I and class II, each containing equally ten aaRSs for ten kinds of amino acids (see Figure 1-3-a).

Class I aaRSs are generally monomeric and composed of Rossmann fold, containing the highly conserved motifs; KMSKS and HIGH. These enzymes approach the minor groove side of the cognate tRNAs, and are bound to the tRNA accepter stem to animoacylate the 2′-OH group of the terminal adenosine residue. On the other hand, class II aaRSs are generally dimeric or tetrameric enzymes that approach the major groove side of their respective tRNAs to charge the amino acid to the 3′-OH group of the terminal adenosine residue [32].

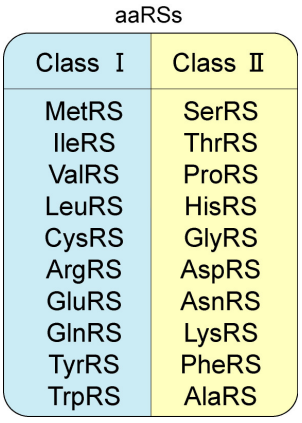


Figure 1-3-a Two classes of twenty aaRSs. Class I and Class II have ten kinds of aaRS each other.

Some aaRSs are allowed to misactivate and misaminoacylate noncognate amino acids that are isosterically and chemically related to the cognate amino acids [33]. For example, aspartyl-tRNA synthetase (AspRS) in some organism [34] and grutamyl-tRNA (GluRS) in all archaea, most bacteria and some eukaryotic organisms [35] are known as the nondiscriminating enzymes. These organisms are missing the genes of AsnRS and GlnRS. The nondiscriminating AspRS catalyzes aspartylation of both $tRNA^{Asp}$ and $tRNA^{Asn}$. The nondiscriminating GluRS catalyzes gultamylation for both $tRNA^{Glu}$ and $tRNA^{Gln}$. These misaminoacylated tRNAs are later converted to the collect amino acids by the following enzymes [35-36]. GatCAB converts the carbonyl group to the carbamoly group for both Asp-$tRNA^{Asn}$ and Glu-$tRNA^{Gln}$. GatDE converts Glu-$tRNA^{Gln}$

to Gln-tRNA^{Gln}. I determined the X-ray structure of the nondiscriminating AspRS from *Aeropyrum pernix*. The details have been published in [34] (see the attached manuscript).

Several aminoacyl-tRNA mutants similar to the above case are also reported for the biosyntheses of selenocystein-tRNA^{Sec} [37] and Cys-tRNA^{cys}[38]. The most interesting case is threonyl-tRNA synthetase (ThrRS) of creanarchaea, as described below.

1-3-b. Complementary ThrRSs of crenarchaea

ThrRS generally consists of three domains: an N-terminal editing domain, a catalytic domain and a C-terminal anticodon-binding domain [39-40]. ThrRS must discriminate threonine from valine and serine. A conserved zinc ion in the catalytic domain precludes the misactivation of noncognate valine, resulting in the discrimination of valine [39]. It is known in *Escherichia coli* that the N-terminal editing domain of ThrRS removes the mischarged serine moiety from Ser-tRNA^{Thr} [41]. However, the N-terminal domains of archaeal ThrRSs have no sequence homology to those of eukaryotic and bacterial ThrRSs. Nevertheless, it has been reported that the archaeal ThrRS from *Methanosarcina mazei* is likewise able to remove serine mischarged onto tRNA^{Thr} [42]. The structure of the truncated N-terminal domain of *Pyrococcus abyssi* ThrRS [43] shows a similarity to that of D-aminoacyl-tRNA deacylase of *E. coli*, which specifically removes a D-amino acid mischarged on the tRNA.

Furthermore, several crenarchaeal species have been found to possess two different genes encoding ThrRSs [44]. In *Sulfolobus solfataricus* harboring also two ThrRS genes, the phenotype products were named as ThrRS-cat and ThrRS-ed [45], Korencic *et al*. reported that ThrRS-cat could synthesize both Thr-tRNA^{Thr} and Ser-tRNA^{Thr} without editing the misacylated Ser-tRNA^{Thr}, and that ThrRS-ed has no catalytic activity for aminoacylation, but it would deaminoacylate the misacylated Ser-tRNA^{Thr} produced by ThrRS-cat.

We have found that four other creanarchaea also contain the two ThrRS genses in their genomes, and performed similar sequence analysis of the two ThrRSs (ThrRS-1 and ThrRS-2) including the following creanarchaea; *Aeropyrum pernix* and *Sulfolobaceae* family (*Sulfolobus tokodaii*, *Sulfolobus solfataricus*, *Sulfolobus acidocaldarius* and *Metallosphaera sedula*). The detailed results will be described in Chapter 5. The sequences of ThrRS-1 and ThrRS-2 are similar to ThrRS-cat and ThrRS-ed, respectively. The amino acid sequences of ThrRS-1s possess the highly conserved catalytic and anticodon-binding domains, but no N-terminal editing domain. On the other hand, the sequence alignment shows that ThrRS-2s are composed of the

archaeal editing domain and the anticodon-binding domain but no catalytic domain. The sequence of editing domain of *Sulfolobus tokodaii* is extremely short as compared with other organisms.

1-3-c. Purpose and significance of X-ray analyses of the complementally ThrRSs from *Aeropyrum pernix* and *Sulfolobus tokodaii*

The sequence alignment has shown the domain components of ThrRS-1 and ThrRS-2. In order to prove the above speculations, it is necessary to clarify the tertiary strictures of ThrRS-1 and ThrRS-2. We have started X-ray analyses of ThrRS-1 and ThrRS-2 from *Aeropyrum pernix* (*Ap*). Furthermore, X-ray analyses of those from *Sulfolobus tokodaii* (*St*) have been also performed to compare the structures, from which the significant evolutionary aspects of the organisms would be revealed.

1-4. Composition of the thesis

This thesis is composed of nine chapters, which are grouped into two parts. The chapter 1 describes the introduction of both Part I and Part II. The Part I including chapters 2, 3 and 4 contains X-ray analyses of *St*-Rtc and its complexes, and the Part II including chapters 5, 6, 7 and 8 contains sequence alignments and X-ray analyses of ThrRSs. The discussion of both Part I and Part II are summarized in the last chapter 9.

1-5. Reference

1. J. Hurwitz (2005) “The discovery of RNA polymerase” *J. Biol. Chem.*, **280**, 42477-42485.
2. M. Selmer, C. M. Dunham, F. V. Murphy, A. Weixlbaumer, S. Petry, A. C. Kelley, J. R. Weir & V. Ramakrishnan (2006) “Structure of the 70S ribosome complexed with mRNA and tRNA” *Science*, **313**, 1935-1942
3. R. Luhrmann, B. Kastner & M. Bach (1990) “Structure of spliceosomal snRNP’s and their role in pre-mRNA splicing” *Biochim. Biophys. Acta*, **1087**, 265–292.
4. W. G. Scott, (1998) “RNA catalysis” *Curr. Opin. Struct. Biol.*, **8**, 720–726.
5. S. Meyer, C. Temme & E.Wahle (2004) “Messenger RNA turnover in eukaryotes: pathways and enzymes” *Crit Rev Biochem Mol Biol*., **39**, 197-216.
6. D. Moras (1992) “Aminoacyl-tRNA synthetases” *Curr. Opin. Struct. Biol.*, **2**, 138–142
7. G. Varani & K. Nagai (1998) “RNA recognition by RNP proteins during RNA processing” *Annu. Rev. Biophys. Biomol. Struct.*, **27**, 407–445
8. F. K. Chu, G. F. Maley, F. Maley & M. Belfort (1984) “Intervening sequence in the thymidylate synthase gene of bacteriophage T4” *PNAS*, **81**, 3049–3053.
9. J. Ferat & F. Michel (1993) “Group II self-splicing introns in bacteria” *Nature*, **364**, 358–361.
10. B. P. Kaine, R. Gupta & C. R. Woese (1983) “Putative introns in tRNA genes of prokaryotes” *PNAS*, **80**, 3309–3312.
11. J. Lykke-Andersen, C. Aagaard, M. Semionenkov & R. A. Garrett (1997) “Archaeal introns: splicing, intercellular mobility and evolution” *Trends Biochem Sci*., **9**, 326-331.
12. M. Belfort, M. E. Reaban, T. Coetzee & J. Z. Dalgaard (1995) “Prokaryotic introns and inteins: a panoply of form and function” *J. Bacteriol*., **177**, 3897-3903.
13. M. J. Schellenberg, D. B. Ritchie & A. M. MacMillan (2008) “Pre-mRNA splicing: a complex picture in higher definition” *Trends Biochem Sci*., 6, 243-246.
14. A. K. Hopper & E. M. Phizicky (2003) “tRNA transfers to the limelight” *Genes Dev*. 17, 162-180.
15. I. Gomes & R. Gupta (1997) “RNAsplicing ligase activity in the archaeon *Haloferax volcanii*” *Biochem. Biophys. Res. Commun*. **237**, 588–594.
16. O. Vicente & W. Filipowicz (1988) “Purification of RNA 3’-terminal phosphate cyclase from HeLa cells.

Covalent modification of the enzyme with different nucleotides" *Eur. J. Biochem*., **176**, 431−439.
17. W. Filipowicz & A. J. Shatkin (1983) "Origin of splice junction phosphate in tRNAs processed by HeLa cell extract" *Cell*, **32**, 547–557.
18. W. Filipowicz, K. Strugala, M. Konarska & A. J. Shatkin (1985) "Cyclization of RNA 39-terminal phosphate by cyclase from HeLacells proceeds via formation of N(3')pp(5')A activated intermediate" *PNAS.*, **82**, 1316–1320.
19. K. D. Genschik & W. Filipowicz (1998) "Characterization of the Escherichia coli RNA 3′ -terminal phosphate cyclase and its σ^{54}-regulated operon" *J. Biol. Chem*., **73**, 25516−25526.
20. P. Genschik, E. Billy, M. Swianiewicz & W. Filipowicz (1997) "The human RNA 3'-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea" *EMBO J*., **16**, 2955−2967.
21. W. Filipowicz & O. Vicente (1990) "RNA 3'-terminal phosphate cyclase from HeLa cells" *Methods Enzymol*., **181**, 499−510.
22. S. Shuman & B. Schwer (1995) "RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotidyl transferases" *Mol. Microbiol*., **17**, 405–410.
23. E. Billy, D. Hess, J. Hofsteenge & W. Filipowicz (1999) "Characterization of the adenylation site in the RNA 3'-terminal phosphate cyclase from *Escherichia coli*" *J. Biol. Chem*., **274**, 34955−34960.
24. G. J. Palm, E. Billy, W. Filipowicz, A. Wlodawer (2000) "Crystal structure of RNA 3'-terminal phosphate cyclase, a ubiquitous enzyme with unusual topology" *Structure*, **8**, 13-23.
25. V. Biou, F. Shu, & V. Ramakrishnan (1995) "X-ray crystallography shows that translational initiation factor IF3 consists of two compact alpha/beta domains linked by an alpha-helix" *EMBO J*., **14**, 4056-4064.
26. E. Schönbrunn, S. Sack, S. Eschenburg, A. Perrakis, F. Krekel, N. Amrhein & E. Mandelkow (1996) "Crystal structure of UDP-N-acetylglucosamine enolpyruvyltransferase, the target of the antibiotic fosfomycin" *Structure*, **4**, 1065-1075.
27. T. Skarzynski, A. Mistry, A. Wonacott, S. E. Hutchinson, V. A. Kelly & K. Duncan (1996) "Structure of UDP-N-acetylglucosamine enolpyruvyl transferase, an enzyme essential for the synthesis of bacterial peptidoglycan, complexed with substrate UDP-N-acetylglucosamine and the drug fosfomycin" *Structure*, **4**, 1465-1474.
28. W. C. Stallings, S. S. Abdel-Meguid, L. W. Lim, H. S. Shieh, H. E. Dayringer, N. K. Leimgruber, R. A. Stegeman, K. S. Anderson, J. A. Sikorski, S. R. Padgette & G. M. Kishore (1991) "Structure and topological symmetry of the glyphosate target 5-enol-pyruvylshikimate-3-phosphate synthase: a distinctive protein fold" *PNAS*, **88**, 5046-5050.
29. R. Rauhut, P. R. Green & J. Abelson (1990) "Yeast tRNA-splicing endonuclease is a heterotrimeric enzyme" *J Biol Chem*, **265**, 18180–18184.
30. C. L. Greer, B. Javor & J. Abelson (1983) "RNA ligase in Bacteria: formation of a 2',5' linkage by an *E. coli* extract" *Cell*, **33**, 899–906.
31. E. A. Arn & J. N. Abelson (1996) "The 2'-5' RNA ligase of *Escherichia coli*. Purification, cloning and genomic disruption" *J. Biol. Chem.*, **269**, 31145–31153.
32. G. Eriani, M. Delarue, O. Poch, J. Gangloff & D. Moras (1990) "Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs" Nature, **347**, 203-206.
33. E. Schmidt & P. Schimmel (1994) "Mutational isolation of a sieve for editing in a transfer RNA synthetase" *Science*, **264**, 265-267.
34. Y. Sato, Y. Maeda, S. Shimizu, M. T. Hossain, S. Ubukata, K. Suzuki, T. Sekiguchi & A. Takénaka (2007) "Structure of the nondiscriminating aspartyl-tRNA synthetase from the crenarchaeon *Sulfolobus tokodaii* strain 7 reveals the recognition mechanism for two different tRNA anticodons" *Acta Crystallogr*, **D63**, 1042-1047.
35. A.W. Curnow, K. W. Hong, R. Yuan, S. I. Kim, O. Martins, W. Winkler, T. M. Henkin, & D. Söll (1997) "Glu-tRNA^{Gln} amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation" *PNAS*., **94**, 11819-11826.
36. D. L. Tumbula, H. D. Becker, W. Z. Chang & D. Söll (2000) "Domain-specific recruitment of amide amino acids for protein synthesis" *Nature*, **407**, 106–110.
37. A. Böck (2000) "Biosynthesis of selenoproteins-an overview" *Biofactors*, **11**, 77–78.
38. A. Sauerwald, W. Zhu, T. A. Major, H. Roy, S. Palioura, D. Jahn, W. B. Whitman, J. R. Yates 3rd, M. Ibba & D. Söll (2005) "RNAdependent cysteine biosynthesis in archaea" *Science*, **307**, 1969–1972.
39. R. Sankaranarayanan, A. C. Dock-Bregeon, P. Romby, J. Caillet, M. Springer, B. Rees, C. Ehresmann, B. Ehresmann, & D. Moras (1999) "The structure of threonyl-tRNA synthetase-tRNA^{Thr} complex enlightens its repressor activity and reveals an essential zinc ion in the active site" *Cell*, **97**, 371–381.
40. A. Torres-Larios, A. C. Dock-Bregeon, P. Romby, B. Rees, R. Sankaranarayanan, J. Caillet, M. Springer, C. Ehresmann, B. Ehresmann & D. Moras (2002) "Structural basis of translational control by Escherichia coli threonyl tRNA synthetase" *Nature Struct. Biol*., **9**, 343–347.
41. A. C. Dock-Bregeon, B.Rees, A. Torres-Larios, G. Bey, J. Caillet & D. Moras (2004) "Achieving error-free translation; the mechanism of proofreading of threonyl-tRNA synthetase at atomic resolution" *Mol. Cell*, **16**, 375–386.
42. Beebe, E. Merriman, L. Ribas de Pouplana & P. Schimmel (2004) "A domain for editing by an archaebacterial tRNA synthetase" *Proc. Natl Acad. Sci. USA*, **101**, 5958–5963.
43. Hussain, S. P. Kruparani, B. Pal, A. C. Dock-Bregeon, S. Dwivedi, M. R. Shekar, K. Sureshbabu & R.

Sankaranarayanan (2006) "Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea" *EMBO J.*, **25**, 4152-4162.
44. C. R. Woese, G. J. Olsen, M. Ibba & D. Söll (2000) *Microbiol.Mol.Biol.Rev.*, **64**, 202–236.
45. D. Korencic, I. Ahel, J. Schelert, M. Sacher, B. Ruan, C. Stathopoulos, P. Blum, M. Ibba & D. Söll (2004) "A freestanding proofreading domain is required for protein synthesis quality control in Archaea" *Proc. Natl Acad. Sci. USA*, **101**, 10260-10265.

Part I. RNA 3'-terminal phosphate cyclase

< Chapter 2. X-ray analyses of *St*-Rtc and its complexes >

2-1. Introduction

2-1-a. Characterization of Rtc

Rtc is an enzyme which catalyzes the cyclization of the 3'-terminal phosphate groups of truncated RNAs by the following two-step reactions [1,2]:

(i) Rtc+ATP→Rtc-AMP + PP*i*

(ii) RNA-$N^{3'}$p + Rtc-AMP→ RNA-N>p + Rtc + AMP

In the first reaction, Rtc complex with the covalently bound AMP (Rtc-AMP) is formed in the pH range of 8.0-8.5 in the presence of truncated RNA with a 3′-terminal phosphate group, ATP and divalent cations, either Mg^{2+} or Mn^{2+} [3-6]. In the second reaction, it is considered that once an RNA-$N^{3'}pp^{5'}$A molecule is found as the reaction intermediate [3], the cyclization may occur non-enzymatically. It is shown that Rtc exists as a dimer composed of the identical two subunits with a molecular weight of about 30000 Da, but it is easy dissociated into the two subunits [7]. Rtc is localized in nucleus to perform its postulated role in RNA processing [8]. Genes encoding proteins similar to Rtc in sequence were identified in three domains, Eukaryota, Bacteria and Archaeota. Although a three dimensional structure of Rtc of *E. coli* (*Ec*-Rtc) (*apo*-form) was reported [9], the catalytic mechanism is not clarified. In Rtc from *Sulfolobus tokodaii* (*St*-Rt), the subunit contains 339 amino acid residues. The estimated molecular weight and pI are 37505 and 9.89, respectively.

2-1-b. Strategy for X-ray analyses

At first we succeeded in crystallization of Rtc (*St*-Rtc) from *Sulfolobus tokodaii*. In preliminary trials of structure analysis, the molecular replacement method was applied using the *Ec*-Rtc structure (PDB-ID 1QMH) [9], but it was failed. Therefore, we will start to solve the crystal structure of selenomethionine-labeled Rtc, *St*-Rtc*, by the multiple-wavelength anomalous dispersion (MAD) method. And then the native structure of *St*-Rtc (*apo*-form) will be determined by the molecular replacement (MR) method using the *St*-Rtc* structure as a probe. In order to reveal the structural features of the first reaction, *St*-Rtc will be crystallized in the following states, *St*-Rtc* in complex with Mn^{2+} ion (*St*-Rtc*+Mn) to identify the Mg^{2+} ion binding site. Since the number of electrons on Mg^{2+} ion is close that of water oxygen atom, it is difficult to

distinguish between Mg^{2+} ion and water oxygen atom. Therefore, Mn^{2+} ion is often used as a probe of Mg^{2+}. The crystal structure of *St*-Rtc in complex with ATP and no Mg^{2+} ion (*St*-Rtc*+ATP) will provide an insight of ATP binding. Furthermore, the crystal structures of *St*-Rtc in complex with Mg^{2+} and ATP under acidic condition (*St*-Rtc+AMP) and under basic conditions (*St*-Rtc*-AMP) will be much interesting to see before and after the first reaction, because it is reported that the optimal pH values are 8-9 [5-6] for Rtcs from Hela cell and *E. coli*.

2-2. Methodology

2-2-a. Crystallization

In order to see the molecular structure of a protein at atomic resolution, it is necessary to fix the molecule during experiments. Another difficult problem is the scattering power. The intensity of molecular scattering is too weak to detect the scattered light from a single molecule. To resolve these conditions, the simplest way is to crystallize the molecules. In crystalline state, molecules are fixed within an allowed range such a frozen state and a huge amount of molecules (in mol scale) are arranged in the same orientations by periodical regularity including symmetric operations.

The most convenient method for protein crystallizations is based on a hanging-drop vapor-diffusion technique, in which a droplet containing purified protein in a buffer solution with precipitants at low concentrations is equilibrated to a reservoir solution containing precipitants at high concentrations in a sealed container. During condensation of the droplet, crystalline nuclei will appear at the over-saturation of protein, and then they will grow up to a certain size. For X-ray analysis, these grown crystals should be as perfect as no significant internal imperfections such as cracks or twinning and be typically larger than several dozen μm in all dimensions.

2-2-b. X-ray diffraction experiments

Although X-rays are lights with wavelength in the range of 10~0.1 nm, the typical wavelength used for X-ray analysis of protein crystals is around 1 Å (0.1 nm), which is nearly a same scale as atomic radius. When the parallel light is irradiated to a crystal, X-ray radiations generated from all electrons cause diffraction by interference of the scattered lights in a certain direction. This process can be expressed by the following Fourier transformation equation,

$$F(\boldsymbol{h}) = \int \rho(\boldsymbol{r})^{\mathrm{xtal}} \exp(2\pi i \boldsymbol{h}\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \qquad (2\text{-}1)$$

where $\rho(\boldsymbol{r})^{xtal}$ is the electron density distribution of the unit cell and $\boldsymbol{h}$ is a vector in the reciprocal space which is equivalent to the observed direction. The distribution of $F(\boldsymbol{h})$ is called a diffraction pattern, and all diffractions in the three-dimensional space are obtained by rotating the crystal.

The unit-cell dimensions (a, b, c, α, β and γ) and the symmetry of the crystal are derived by analyzing several diffraction patterns in different angular rotations. The diffracted beam in equation (2-1) is equivalent to the following equation,

$$F(\boldsymbol{h}) = |F(\boldsymbol{h})|\exp[i\phi(\boldsymbol{h})] \qquad (2\text{-}2)$$

where $F(\boldsymbol{h})$ is the diffracted beam with the amplitude $|F(\boldsymbol{h})|$ and the phase $\phi(\boldsymbol{h})$. On the other hand, the reversed Fourier transformation of the equation (2-1) expresses the electron density distribution of the unit cell $\rho(\boldsymbol{r})^{xtal}$, as shown below.

$$\rho(\boldsymbol{r})^{xtal} = (1/V)\sum_{\boldsymbol{h}} F(\boldsymbol{h})\exp(-2\pi i\boldsymbol{hr}) \qquad (2\text{-}3)$$

It is noted that when we observe the diffracted beams of equation (2-2), the electron distribution of the crystal can be calculated by equation (2-3). After that, it is easily derived the three-dimensional structure of the protein by putting atoms at the maximum densities. However, the situation is practically not easy to calculate the equation (2-3), because we can observe only the amplitudes of the diffracted beams, but not the phases.

In real X-ray experiments, we have to collect all the diffracted patterns of the targeted crystal, and then process them by computer programs, precisely and accurately, to get the diffracted intensity data, which is then converted to a set of $|F(\boldsymbol{h})|^{obs}$. The phase information can be derived by the different methods as described below.

2-2-c. Phase problem

In order to reconstruct the electron density distribution $\rho(\boldsymbol{r})$ from the observed scattering, it is necessary to know the phase angle $\phi(\boldsymbol{h})$ of each reflection. The phase information can be estimated by the several methods; the multiple-wavelength anomalous dispersion (MAD) method, the single wavelength anomalous dispersion (SAD) method, the multiple isomorphous replacement (MIR) method and the molecular replacement (MR) method [10]. Among them, the following two methods will be applied to solve the crystal structures of *St*-Rtc and its complexes.

Multiple-wavelength anomalous dispersion (MAD) method

A powerful solution for phase problem is the multiple wavelength anomalous diffraction (MAD) method (reviewed in [11]). In X-ray diffraction experiment, X-ray photons vibrate all electrons of atoms. If the incident photons have enough energy to bring an atom to an excited state, an electron of the atom is thrown out of the orbital. Because some energy are used to the electron transition (atomic absorption), the amplitude of diffraction beam is slightly reduced, called anomalous scattering, and the wavelength at the drastic change in atomic absorption is called absorption edge. As the result of the anomalous scattering, the phase angles are varied, as well as the intensities are changed. Therefore the atomic scattering factor f is simply expressed as a function of wavelength λ.

$$f(\boldsymbol{\lambda})=f'(\boldsymbol{\lambda})+if''(\boldsymbol{\lambda}) \qquad (2\text{-}4)$$

where f' contains the conventional atomic scattering factor and the real part of the anomalous scattering at the wavelength $\boldsymbol{\lambda}$, and f'' is the imaginably part of the anomalous scattering factor. The profiles of these factors can be estimated by an extended X-ray absorption fine structure (EXAFS) analysis. These wavelengths with the largest f'' value, with the lowest f' value and with a normal f' value is called "peak", "edge" and "remote", respectively. The anomalous effects at three are applied to estimate the phase angles.

In practical application of the MAD method, the protein crystal is required to be contained atoms with high anomalous effect. For the purpose, a most convenient way is to introcduce selenomethionine into protein, because this amino acid has chemical properties similar to methionine. The absorption edge of K-shell of Se atom is around 0.98 Å. To generate X-ray with this wavelength, a synchrotron radiation is used in X-ray experiments.

Molecular replacement (MR) method

If a protein structure which is similar to that of the target protein to be determined is known, it can be used as a search probe in molecular replacement to estimate the orientation and the positions of the probe molecules in the unit cell. When the amino acid sequence homology between the two proteins is high, it can be assumed that their polypeptide foldings are, in general, similar to each other.

2-2-d. Structure refinement

When approximate phases are obtained by any of the above methods, an electron

density distribution can be calculated according to the equation (2-3). By putting atoms or residues on the electron density map, an initial protein structure can be constructed. Given a structure of protein, its atom positions and thermal parameters are refined by the least-squares method to get better fitting between the observed diffraction amplitudes $|F(\boldsymbol{h})^{obs}|$, and the calculated amplitude $|F(\boldsymbol{h})^{calc}|$, ideally yielding a better set of phases. An improved structure can then be derived by fitting the current structure to the newly calculated electron density map, and a further round of least-squares atomic refinement is carried out. These processes are repeated until the correlation between the observed $|F(\boldsymbol{h})|$ data and the structure is maximized. The agreement is measured by an R factor defined as

$$R = \Sigma_{\mathbf{h}}[|F(\boldsymbol{h})^{obs}|\text{-Scale}|F(\boldsymbol{h})^{cal}|]/\Sigma_{\mathbf{h}}|F(\boldsymbol{h})^{obs}| \qquad (2\text{-}4)$$

A similar quality criterion is R_{free}, which is calculated from a subset (~10%) of reflections that were not included in the structure refinement [12].

2-3. Sample prepareation

2-3-a. Overexpression and purification of *St*-Rtc

The gene encoded *St*-Rtc (gene ID; ST0570, 339 amino acids, 37505 Da, pI 9.89) were inserted into the pET11a vector (Novagen). The recombinant plasmids were transformed in *E. coli* Rosetta-gami (DE3) (Novagen), and the cells were grown in LB culture at 310 K. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Resource Iso column (Amersham Bioscience). Bound proteins were eluted at 2 ml/min flow rate with 20 mM Tris-HCl (pH 8.0) and a linear gradient of 1.75-0 M ammonium sulfate. The pooled fractions were loaded onto a Resource S column (Amersham Bioscience), and the bound proteins were eluted at 4ml/min flow rate with 50 mM MES (pH6.0), 5 mM β-mercaptoethanol and 0-0.4 M NaCl linear gradient. Moreover, the eluted fractions were applied onto a HiLoad 16/60 Superdex75 column (Amersham Bioscience) with 20 mM Tris-HCl (pH8.0), 5mM β-mercaptoethanol and 150 mM NaCl at 1 ml/min flow rate. The eluted fractions were finally applied onto a BioScale CH51 column (Bio-Rad) with 10-400 mM potassium phosphate (pH 7.0) linear gradient and 5 mM β-mercaptoethanol at 2ml/min flow rate.

The proteins were concentrated to final concentrations of 5 mg/ml in 20 mM

Tris-HCl (pH 8.0) with centrifugal filter devices (Vivaspin 500, Sartorius Stedim Biotech). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

2-3-b. Overexpression and purification of *St*-Rtc*

The selenomethionine derivative was expressed in *E. coli* B834 (DE3) (Novagen) using minimal M9 medium containing L-selenomethionine. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Resource Iso column (Amersham Bioscience). Bound proteins were eluted at 4 ml/min flow rate with 20 mM Tris-HCl (Tris hydroxymethyl aminomethane titrated using HCl) (pH 8.0) and a linear gradient of 1.75-0 M ammonium sulfate. The pooled fractions were loaded onto a Resource S column (Amersham Bioscience), and the bound proteins were eluted at 4 ml/min flow rate with 50 mM MES (2-Morpholinoethanesulfonic acid monohydrate titrated using NaOH) (pH6.0), 5 mM β-mercaptoethanol and 0-0.4 M NaCl linear gradient. Moreover, the eluted fractions were applied onto a BioScale CH51 column (Bio-Rad) at 2 ml/min flow rate with 10-400 mM potassium phosphate (pH 7.0) linear gradient and 5 mM β-mercaptoethanol. The eluted fractions were finally applied onto a HiLoad 16/60 Superdex75 column (Amersham Bioscience) with 20 mM Tris-HCl (pH8.0), 5mM β-mercaptoethanol and 150 mM NaCl at 1ml/min flow rate.

The proteins were concentrated to final concentrations of 10 mg/ml in 20 mM Tris-HCl (pH 8.0) with centrifugal filter devices (Vivaspin 500, Sartorius Stedim Biotech). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

2-4. X-Ray analyses of *St*-Rtc*

2-4-a. Crystallization of *St*-Rtc*

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further

optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

2-4-b. X-Ray data collection and data processing of *St*-Rtc*

To apply a Multiple Anomalous Dispersion (MAD) method, it is necessary to measure the data at the exact wavelengths with and without anomalous effects. X-Ray experiments were performed at the BL-17A beamline of the Photon Factory (PF; Ibaraki, Japan). Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). These wavelengths were estimated by X-ray Absorption Fine Structure (XAFS) measurement, four wavelengths being chosen as the peak, the edge, the high remote and the low remote. Their diffraction beams were detected at 95 K with 1˚-360˚ oscillation range using the Quantum 4R CCD detector (ADSC). Exposure time was 2 s for 1° oscillation angle per image. The detector was positioned from the crystal at 210.8mm, 210.7mm, 209.3mm and 215.0 mm for the four wavelengths, respectively. Diffraction spots were indexed, and their intensities were integrated and scaled using the *HKL*2000 package [13] and the intensity data were converted to the amplitudes using programs in the *CCP4* suite [14].

2-4-c. Preliminary structure analysis of *St*-Rtc*

The MAD method was applied to solve the phase problems of the *St*-Rtc* crystal with the program *SHARP/autoSHARP* [15]. The program *autoSHARP* is the automated structure solution pipeline, built around the program *SHELXD* [16] for location of heavy atoms, the program *SHARP* for the heavy atom refinement, the program *SOLOMON* [17] for density modification and the program *ARP/wARP* [18] for automated structural model building and refinement using the program *REFMAC 5* [19].

In order to refine the atomic parameters of the constructed protein structures, the following two steps were repeatedly applied. One is refinement of the atomic parameters in the reciprocal space. The positional parameters (x, y and z) and the temperature factors (B) for all atoms without hydrogen in the asymmetric unit are refined using the program *REFMAC 5* [19]. And those of other atoms and assigned in the real space on an electron density map. The most effective way to correct their position and to add several atoms was manual adjustments using the program *COOT* [20]. The structure was validated with a Ramachandran plot using the program *PROCHECK* [21].

2-5. X-ray analyses of *St*-Rtc

2-5-a. Crystallization of *St*-Rtc

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

2-5-b. X-ray data collection and data processing of *St*-Rtc

Crystal obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 5 s for 1° oscillation angle per image. Detector distance from the crystal and the used wavelength were 189.6mm and 1.00 Å, respectively. Diffraction spots were indexed, and the intensities were integrated and scaled using the *HKL*2000 package [13] and the intensity data were converted to the amplitudes using programs in the *CCP4* suite [14].

2-5-c. Structure determination

The initial structure of the crystal was derived by the molecular replacement method with the program *AMoRe* [23] using the *St*-Rtc* structure as a probe. All the atomic parameters were refined with the program *REFMAC 5* [19]. The structural modification was performed with the program *COOT* [20]. These two operations were complementally repeated. The structure was validated with a Ramachandran plot using the program *PROCHECK* [21].

2-6. X-ray analyses of *St*-Rtc*+Mn

2-6-a. Crystallization of *St*-Rtc*+Mn

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein with Mn^{2+} ion and the reservoir solutions, and then equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential

crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

2-6-b. X-ray data collection and data processing of *St*-Rtc*+Mn

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 20 s for 1° oscillation angle per image. The detector distance and the used wavelength are fixed with 212.5mm and 1.00 Å, respectively. Diffraction spots were indexed, and the intensities were integrated and scaled using the *HKL*2000 package [13] and the intensity data were converted to the amplitudes using programs in the *CCP4* suite [14].

2-6-c. Structure determination of *St*-Rtc*+Mn

The initial structures of the crystal was derived by the molecular replacement method with the program *Phaser* [22] using the *St*-Rtc* structure as a probe. All the atomic parameters were refined with the program *REFMAC 5* [19]. The structural modification was performed with the program *COOT* [20]. These two operations were complementally repeated. The structure was validated with a Ramachandran plot using the program *PROCHECK* [21].

2-7 X-ray analyses of *St*-Rtc*+ATP

2-7-a. Crystallization of *St*-Rtc*+ATP

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein containing ATP and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

2-7-b. X-ray data collection and data processing of *St*-Rtc*+ATP

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 10 s for 1° oscillation angle per image. The detector distance and the used wavelength are fixed with 169.9mm and 0.98 Å, respectively. Diffraction spots were indexed, and the intensities integrated and scaled using the *HKL*2000 package [13]. The intensity data were converted to amplitudes using programs in the *CCP4* suite [14].

2-7-c. Structure determination of *St*-Rtc*+ATP

The initial structures of the crystal was derived by the molecular replacement method with the program *Phaser* [22] using the *St*-Rtc* structure as a probe. All the atomic parameters were refined with the program *REFMAC 5* [19]. The structural modification was performed with the program *COOT* [20]. These two operations were complementally repeated. The structure was validated with a Ramachandran plot using the program *PROCHECK* [21].

2-8. X-Ray analyses of *St*-Rtc+AMP†

2-8-a. Crystallization of *St*-Rtc+AMP†

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. We were prepared the protein solution containing *St*-Rtc with Mg^{2+} and ATP under acidic condition. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates with acidic conditions appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

2-8-b. X-ray data collection and data processing of *St*-Rtc+AMP†

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction

† The obtained crystals will be refered due to the reaction that the additive ATP is charged to AMP or adenosine as described in the reaction 3.6 of the chapter 3.

measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 4R CCD detector (ADSC) at the BL-6A beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 10 s for 1° oscillation angle per image. Detector distance from the crystal and the used wavelength are fixed with 195.5mm and 0.978 Å, respectively. Diffraction spots were indexed, and the intensities were integrated and scaled using the *HKL*2000 package [13]. The intensity data were converted to amplitudes using programs in the *CCP4* suite [14].

2-8-c. Structure determination of *St*-Rtc+AMP†

The initial structures of the crystal were derived by the molecular replacement method with the program *Phaser* [22] using the *St*-Rtc* structure as a probe. All the atomic parameters were refined with the program *REFMAC 5* [19]. The structural modification was performed with the program *COOT* [20]. These two operations were complementally repeated. The structure was validated with a Ramachandran plot using the program *PROCHECK* [21].

2-9. X-Ray analyses of *St*-Rtc*-AMP

2-9-a. Crystallization of *St*-Rtc*-AMP

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. The protein solutions containing *St*-Rtc with Mg^{2+} and ATP under basic conditions were prepared to obtain its complex crystals. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

2-9-b. X-ray data collection and data processing of *St*-Rtc*-AMP

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 5 s for 1° oscillation angle per image. Detector distance from the crystal and the used wavelength are fixed with 212.5mm and 1.00 Å, respectively. Diffraction spots were indexed, the intensities were integrated and scaled

using the *HKL*2000 package [13] and the intensity data were converted to amplitudes using programs in the *CCP4* suite [14].

2-9-c. Structure determination of *St*-Rtc*-AMP

The initial structures of the crystal was derived by the molecular replacement method with the program *Molrep* [24] using the *St*-Rtc* structure as a probe. All the atomic parameters were refined with the program *REFMAC 5* [19]. The structural modification was performed with the program *COOT* [20]. These two operations were complementally repeated. The structure was validated with a Ramachandran plot using the program *PROCHECK* [21].

2-10. Reference

1. W. Filipowicz & A. J. Shatkin (1983) "Origin of splice junction phosphate in tRNAs processed by HeLa cell extract" *Cell*, **32**, 547–557.
2. W. Filipowicz, M. Konarska, H. J. Gross & A. J. Shatkin (1983) "RNA 3'-terminal phosphate cyclase activity and RNA ligation in HeLa cell extract" *Nucleic Acids Res*, **11**, 1405–1418.
3. W. Filipowicz, K. Strugala, M. Konarska & A. J. Shatkin (1985) Cyclization of RNA 3'-terminal phosphate by cyclase from HeLa cells proceeds via formation of N(3')pp(5')A activated intermediate. *Proc Natl Acad Sci USA*, **82**, 1316–1320.
4. D. Reinberg, J. Arenas & J. Hurwitz (1985) "The enzymatic conversion of 3'-phosphate terminated RNA chains to 2',3'-cyclic phosphate derivatives" *J Biol Chem*, **260**, 6088–6097.
5. O. Vicente & W. Filipowicz (1988) "Purification of RNA 3'-terminal phosphate cyclase from HeLa cells" *Eur J Biochem*, **176**, 431–439.
6. P. Genschik, K. Drabikowski & W. Filipowicz (1998) "Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its σ^{54}-regulated operon." *J. Biol. Chem.*, **273**, 25516-25526.
7. Filipowicz W, Strugala K, Konarska M, Shatkin AJ (1985) "Cyclization of RNA 3'-terminal phosphate by cyclase from HeLa cells proceeds via formation of N(3')pp(5')A activated intermediate" *PNAS*, 82, 1316-1320.
8. G. Pascal, B. Eric, S. Michal & F. Witold (1997) "The human RNA 3'-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea" *EMBO*, **16**, 2955–2967.
9. G. J. Palm, E. Billy, W. Filipowicz & A. Wlodawer (2000) "Crystal structure of RNA 3'-terminal phosphate cyclase, a ubiquitous enzyme with unusual topology" *Structure*., **8**, 13-23.
10. G. Taylor (2003) "The phase problem" *Acta Crystallogr.*, **D59**, 1881-1890.
11. S. E. Ealick (2000) "Advances in multiple wavelength anomalous diffraction crystallography" *Current opinion in chemical biology*, **4**, 495-499.
12. A. T. Brünger (1992) **"**Free *R* value: a novel statistical quantity for assessing the accuracy of crystal structures" *Nature*, **355**, 472-475.
13. Z. Otwinowski & W. Minor, (1997) "Processing of X-ray Diffraction Data Collected in Oscillation Mode" *Methods in Enzymology*, **276**, 307-326.
14. Collaborative Computational Project, Number 4, 1994.
15. C. Vonrhein, E. Blanc, P. Roversi & G. Bricogne, (2007) "Automated structure solution with autoSHARP" *Methods Mol. Biol*., **364**, 215-230.
16. T. R. Schneider and G. M. Sheldrick, (2002) "Substructure solution with *SHELXD* 2002" *Acta Cryst*. **D58**, 1772-1779.
17. J. P. Abrahams & A. G. W. Leslie, (1996) "Methods used in the structure determination of bovine mitochondrial F_1 ATPase" *Acta Cryst.*, **D52**, 30-42.
18. A. Perrakis, R. M. Morris, & V. S. Lamzin, (1999) "Automated protein model building combined with iterative structure refinement" *Nature Struct.Biol*., **6**, 458-463.
19. G. N. Murshudov, A. A.Vagin & E. J.Dodson, (1997) "Refinement of Macromolecular Structures by the Maximum-Likelihood Method" *Acta Cryst.* **D53**, 240-255.
20. P. Emsley & K. Cowtan (2004) "Coot: model-building tools for molecular graphics" *Acta Cryst.*, **D60**, 2126-2132.
21. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, (1993) "PROCHECK: a program to check the stereochemical quality of protein structures" *J. App. Cryst.*, **26**, 283-291.
22. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L.C. Storoni & R.J. Read. (2007) "*Phaser*

crystallographic software" *J. Appl. Cryst.*, **40**, 658-674.
23. J.Navaza (1994) "*AMoRe*: an automated package for molecular replacement" *Acta Cryst.* **A50**, 157-163.
24. A.Vagin,A.Teplyakov, (1997) "*MOLREP*: an automated program for molecular replacement" *J. Appl. Cryst*., **30**, 1022-1025.

<Chapter 3 Results of X-ray analyses of *St*-Rtc and its complexes>

3-1. Overexpression and purification of *St*-Rtc and *St*-Rtc*

3-1-a. Purified quality of *St*-Rtc

After the final purification, the molecular weight of the target protein was estimated by an SDS-PAGE. The result is shown in Figure 3-1 (*a*). The target protein appears as a main band at a position corresponding to the molecular weight (37505 Da) calculated from the amino acid sequence of *St*-Rtc. Thus it is considered the protein was highly purified for crystallization.

3-1-b. Purified quality of *St*-Rtc*

Figure 3-1 (*b*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (37786 Da) calculated from the amino acid sequence of *St*-Rtc*. Thus it is considered the protein was highly purified for crystallization.

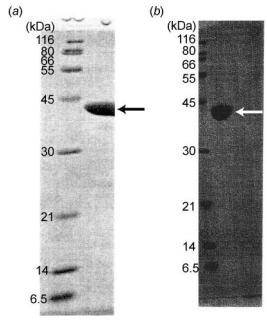


Figure 3-1. SDS–PAGEs of the target proteins of *St*-Rtc (*a*) and *St*-Rtc* (*b*). The left and right lanes contain proteins of molecular weight markers and the protein, respectively. An arrow indicates the protein. Values indicate approximate molecular weights (kDa) at the maker bands.

3-2. X-ray analysis of *St*-Rtc*

3-2-a. *St*-Rtc* crystals

The optimized condition, under which suitable crystals for X-ray analysis were obtained after three days, is that the protein solution containing 10 mg/ml *St*-Rtc* in 20 mM Tris-HCl (pH8.0) buffer and the reservoir solution (pH4.2) containing 2.0 M ammonium dihydrogen phosphate in 0.1 M Tris-HCl (pH8.5) buffer were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The crystal shape is like a needle or a plate, as shown in Figure 3-2-a. The largest crystal indicated by an arrow in Figure 3-2-a has a size of approximately 200×100×50 μm. This crystal was used for X-ray diffraction experiment.

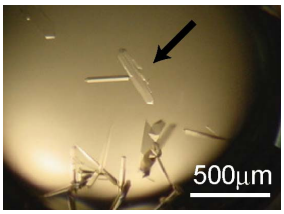


Figure 3-2-a. A photo of the *St*-Rct* crystals. An arrow indicates the crystal used for X-ray diffraction experiment

3-2-b. X-ray data of the *St*-Rtc* crystal

A spectrum of f' and f'' around the absorption edge wavelength of Se atoms is drawn based on XAFS measurement of the crystal (Figure 3-2-b-1). In this spectrum, the four wavelengths are assigned as the peak (0.97898Å), the edge (0.97931Å), the high remote (0.96408Å) and the low remote (0.98319Å). The same crystal was used throughout X-ray experiments.

In the measurement of the peak wavelength, the observed intensity data were processed to 2.00 Å resolution. The data completeness and an R_{marge} value are 99.7% and 6.5%, respectively. The numbers of observed and unique reflections are 657599 and 58120, respectively. The crystal belongs to the $P3_1$ space group. The cell constants are $a = b = 84.1$ and $c = 109.2$ Å.

In the measurement at the edge wavelength, the observed intensity data were processed to 2.10 Å resolution. The data completeness and an R_{marge} value are 100% and 6.3%, respectively. The numbers of observed and unique reflections are 584648 and 50413, respectively.

In the measurement at the high remote wavelength, the observed intensity data were processed to 2.20 Å resolution. The data completeness and an R_{marge} value are 100% and 6.2%, respectively. The numbers of observed and unique reflections are 511916 and 43942, respectively.

In the measurement at the low remote wavelength, the observed intensity data were processed to 2.20 Å resolution. The data completeness and an R_{marge} value are 100% and 6.0%, respectively. The numbers of observed and unique reflections are 512338 and 43942, respectively.

A statistics of the observed intensity data and the crystallographic data are given in Table 3-2-b. Figure 3-2-b-2 shows an example of diffraction images taken using the peak wavelength.

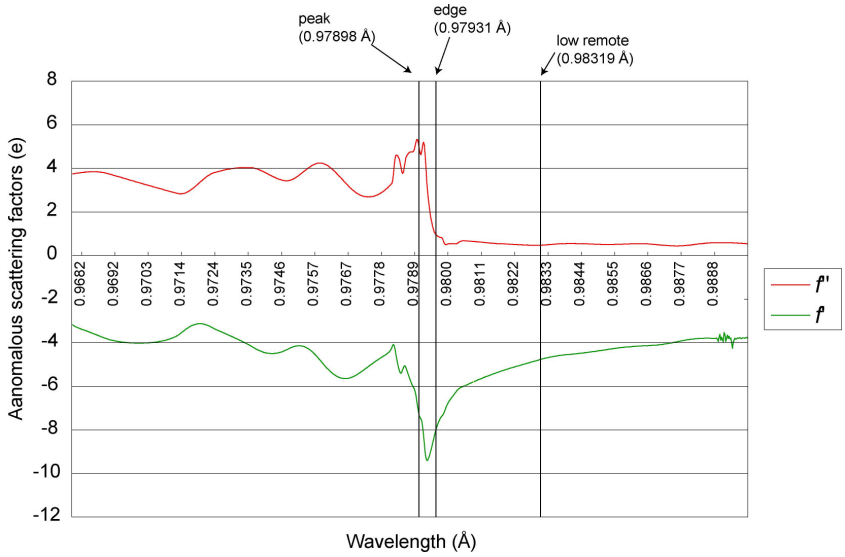


Figure 3-2-b-1. Variation of real (*f*′; green line) and imaginary parts (*f*″; red line) of anomalous scattering factors of selenium atoms in the *St*-Rtc* crystal as a function of wavelength.

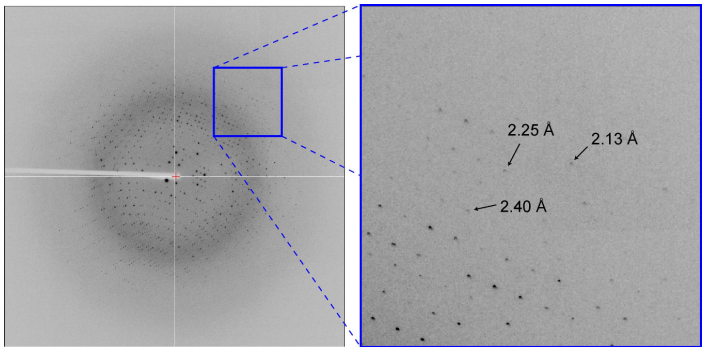


Figure 3-2-b-2. An X-ray diffraction image of the *St*-Rtc* crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 3-2-b. A statistics of the observed intensity data and the crystallographic data of the *St*-Rtc* crystal at four wavelengths

	peak	edge	high remote	low remote
Wavelength (Å)	0.97898	0.97931	0.96408	0.98319
Resolution (Å)	50-2.00 (2.07-2.00)	50-2.10 (2.18-2.10)	50-2.20 (2.28-2.20)	50-2.20 (2.28-2.20)
Space group	$P3_1$	$P3_1$	$P3_1$	$P3_1$
Unit cell $a = b$ (Å) c (Å)	 84.0 109.2	 84.1 109.2	 84.2 109.3	 84.2 109.3
Observed reflections	657599	584648	511916	512338
Unique reflections	58120	50413	43942	43942
Completeness (%)	99.7 (96.9)	100 (100)	100 (100)	100 (100)
Redundancy	5.7 (4.6)	5.8 (5.5)	5.8 (5.7)	5.8 (5.7)
I/σ	31.4 (2.8)	34.4 (4.6)	34.3 (5.4)	36.8 (5.8)
R_{merge} (%)	6.5 (37.9)	6.3 (31.2)	6.2 (27.9)	6.0 (27.0)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl) - I(hkl)| / \sum_{hkl}\sum_i I(hkl)$

3-2-c. Phase angle determination and structural refinement of *St*-Rtc*

The number of protein molecules in the asymmetric unit (Z) was estimated from a V_M value calculated with the unit cell volume and the protein molecular weight. In general, the value should be in a range between 1.7 and 3.5 Å^3/Da [1]. The space group was kept to be $P3_1$ during the X-ray experiments. The average cell constants are $a = b =$ 84.15 and c = 109.25 Å. An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 2, the V_M value to be 2.96 Å^3/Da and the solvent content to be 58.4%. Twelve Se atoms were found for the two molecules in the asymmetric unit, because *St*-Rtc* contains six selenomethionine residues.

Using the two data sets measured at the peak and the edge wavelengths, the best solution was obtained using six Se atoms found in the asymmetric unit. The correlation coefficient between F_o and F_c was 0.292. After the heavy atom position refinements, the additional seven Se atoms were found automatically. Phase angles were calculated from a total of thirteen Se atoms. An electron density map calculated with the estimated phase angles were modified by solvent flattering technique. The solvent content was fixed to be 45% to include solvent molecules surrounding proteins. The final correlation coefficient in $|E^2|$ was 0.822. After structure building with the program *ARP/wARP*, more than 96% residues of the two proteins in the asymmetric unit were successfully assigned.

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. In the asymmetric unit, 338 of the 339 residues per each molecule are assigned except for the C-terminal one residue, one phosphate ion and 276 water molecules are assigned. The final R factor and R_{free} are 18.4% and 24.3%, respectively, indicating that the structure was reasonably refined. A statistic of the structural refinement is given in Table 3-2-c. The final $2|F_o|-|F_c|$ map, contoured at 1.5 σ level, is shown in Figure 3-2-c-1. The refined structure is well fitted on the map. A Ramachandran plot is shown in Figure 3-2-c-2. In the plot, 524 (89.3%) residues are in the most favored geometry, 58 (9.9%) in the allowed geometry and 5 (0.9%) in generally allowed geometry, respectively. There is no residue in disallowed geometry. One phosphate ion interacts with the side chains of Arg202 and Lys205 of the small domain of the A-subunit, as shown in Figure 3-2-c-3.

Table 3-2-c-1. A statistics of the structural refinement of *St*-Rtc*

Resolution (Å)	43.8-2.20
Used reflections	43931
R (%)	18.4
R_{free} (%)	24.3
r.m.s bond length (Å)	0.02
r.ms. angles (˚)	2.0
Protein atoms	5262
Water molecules	278
PO_4^{3-}	1

*$R = 100 \times \sum||F_o|-|F_c|| / \sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

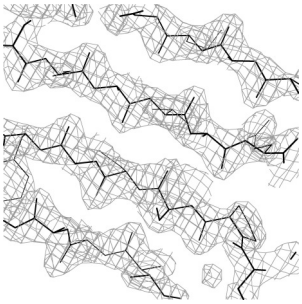


Figure 3-2-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *St*-Rtc* crystal, contoured at 1.5 σ level (colored in gray).

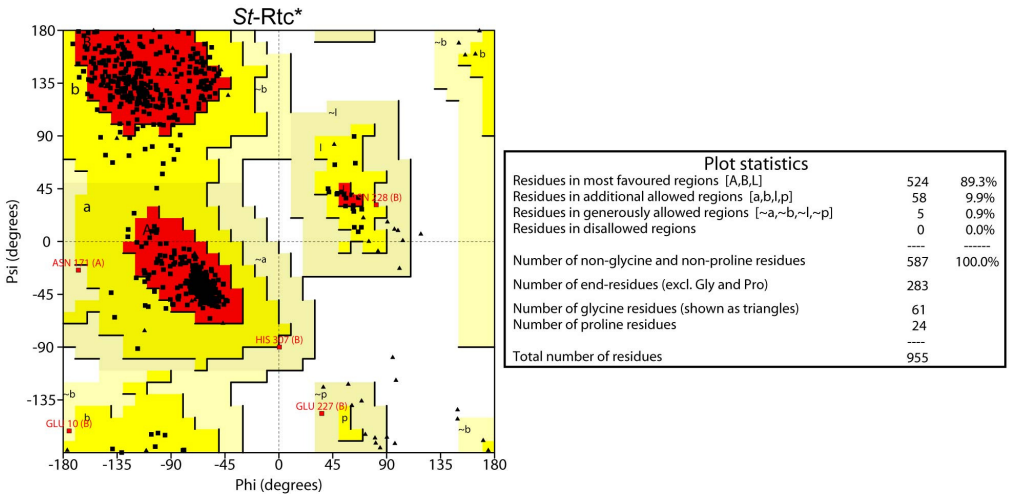


Figure 3-2-c-2. A Ramachandran plot of *St*-Rtc*. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

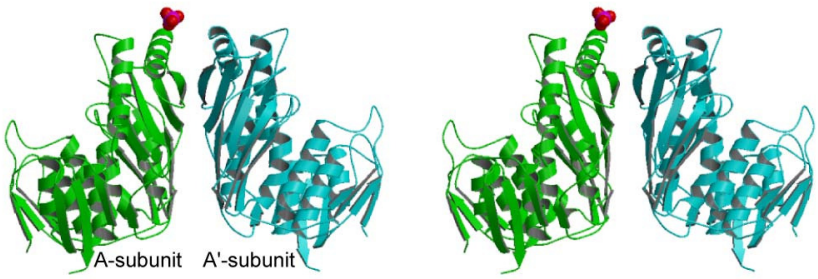


Figure 3-2-c-3. A stereo-pair drawing of the two molecules (subunits) of a dimer and one phosphate ion in the *St*-Rtc* crystal, indicated in ribbon models and pink (phosphorus atom) and red (oxygen atoms) sphere models. Green and cyan wire models indicate the A-subunit and the A′-subunit of the dimer, as described below.

3-2-d. Dimer structure of *St*-Rtc*

The two molecules, A and A′, in the asymmetric unit are shown in Figure 3-2-d-1. When they are superimposed on each other, the estimated r.m.s deviation between the corresponding Cα atoms is 0.66 Å, suggesting that their folding are similar to each other, as shown in wire models in Figure 3-2-d-2. Figure 3-2-d-1 also shows that the two molecules are associated by a non-crystallographic two-fold symmetry to form a dimer. The dimer formation is stabilized by forming a β-sheet through hydrogen bonds between the two main chains, Ile223 and Ile223′, as shown in Figure 3-2-d-1. An apostrophe indicates the residue of the counter subunit A′. The atomic distances given in Table 3-2-d-1 show that there are no abnormal contacts between the two subunits. Hereafter the two molecules are referred to the subunits to form a dimer of the protein. The detailed molecular structure of the dimer will be described in the next section (3-3) and discussed in the next chapter4.

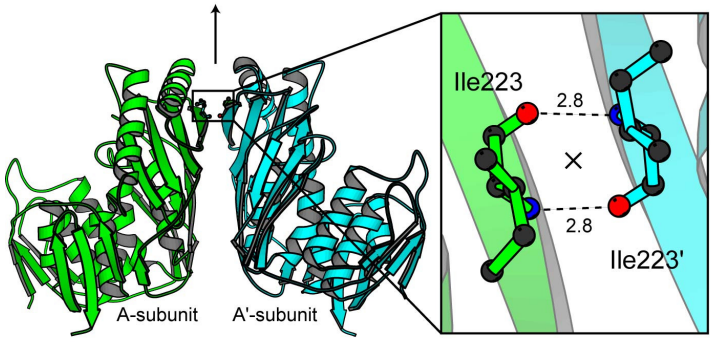


Figure 3-2-d-1. The two protein molecules in the asymmetric unit of the *St*-Rtc* crystal. The two subunits are associated to each other to form a dimer by a non-crystallographic two-fold symmetry shown by an arrow or a symbol x. The right figure shows a zoomed up view at the Ile223 residues. The residue indicated with an apostrophe is the related counter one. Black, blue and red spheres indicate carbon, nitrogen and oxygen atoms respectively. Broken lines indicate the hydrogen bonds with the distances (Å) between the two antiparallel β-strands.

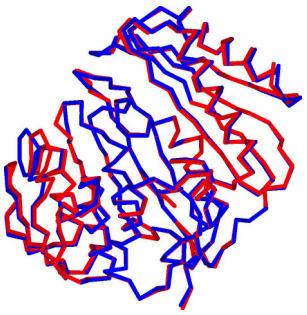


Figure 3-2-d-2. A superimposition of the Cα chains between the two molecules (subunits of the dimer) in the asymmetric unit of the *St*-Rtc* crystal. The A-subunit and the A′-subunits are colored in blue and red, respectively.

Table 3-2-d-1. Interactions between the two subunits of the dimer in the asymmetric unit of the *St*-Rtc* crystal.

Interactions between the small domains				
Direct hydrogen bonds				
A-subunit (Small domain)	Distance (Å)		A′-subunit (Small domain)	
N- Ile223	2.8		O-Ile223	
O-Ile223	2.8		N-Ile223	
Nη-Arg226	2.6		Oε-Glu220	
Hydrogen bond Interactions intermediated by a water molecule				
A-subunit (Small domain)	Distance (Å)	Water	Distance (Å)	A′-subunit (Small domain)
Oδ-Asp224	2.6	W288	2.5	Oε-Glu220
O-Ile223	3.1	W204	2.7	O-Ile221
N-Val225	2.8			

Interactions between the small domain and the opposite large domain.	
Hydrophobic contacts	
A-subunit (Small domain)	A′-subunit (Large domain)
Ile184 Met185 Pro218	Ile149 Ile151

Hydrogen bond interactions intermediated by a water molecule				
A-subunit (Large domain)	Distance (Å)	Water	Distance (Å)	A′-subunit (Small domain)
Nζ-Lys164	2.7	W147	2.9	O-Leu184
Hydrophobic contacts				
A-subunit (Large domain)			A′-subunit (Small domain)	
Ile149 Ile151			Ile184 Met185 Pro218	

*The domain structures (small domain and large domain) of the *St*-Rtc* subunit will be described latter in the next section 3-3.

3-2-e. Dimer packing of the St-Rtc* crystal

The dimer packing of the *St*-Rtc* crystal is shown in Figures 3-2-e-1 and 3-2-e-2. The dimers are in contact with a head-and-tail manner around the crystallographic 3_1 axis, so that a screwed column is formed by stacking the A subunit on the A′-subunit of

the subsequent dimer. Along with the *a* and *b* axes, the columns are interacted to the adjacent columns. The atomic distances in the contact regions along the *c* axis are listed in Table 3-2-e-1, and those along the *a* and *b* axes are listed in Table 3-2-e-2.

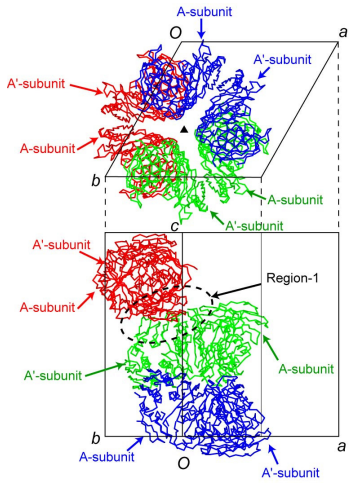


Figure 3-2-e-1. The dimers around the crystallographic three-fold screw axis in the *St*-Rtc* crystal, viewed down the *c* axis (upper) and down the *b** axis (bottom)

Table 3-2-e-1. Interactions between the *St*-Rtc* dimers related by the crystallographic 3_1 symmetry (Region-1) (see Figure 3-2-e-1).

Region-1				
Direct hydrogen bonds				
A′-subunit	Distance (Å)			A-subunit
Oε-Glu179	3.2			Oγ-Ser176
Oγ-Ser176	2.7			Oε-Glu179
Oε-Gln324	3.1			Nδ-Asn277
Oδ-Asn277	3.1			Nε-Gln324
	3.2			Oγ-Ser326
Oδ-Asn277	3.3			Nη-Arg334
Oγ-Ser326	2.6			O-Ser276
Hydrogen bond interactions intermediated by a water molecule				
A′-subunit	Distance (Å)	Water	Distance (Å)	A-subunit
O-Pro330	2.7	W141	2.7	O-Pro330
Oε-Glu298	2.8	W134	2.7	Nδ-Asn277
Oγ-Ser326	3.3	W98	2.7	Oδ-Asn277
O-Ile325	2.7	W233	3.3	O-Ile275
Oε-Gln324	3.0	W93	2.7	Oγ-Ser326
Nη-Arg180	2.9	W277	3.0	Nη-Arg334
Nη-Arg334	2.9	W22	3.3	Nη-Arg180
	2.8	W202	3.2	
			3.1	Nε-Gln324
Hydrophobic contacts				
A′-subunit		A-subunit		
Met178 Ile300 Ile332 Pro330		Met178 Ile300 Ile332 Pro330		

Table 3-2-e-2. Interactions between the *St*-Rtc* dimers in the two different screwed columns related by a periodic shift along the *a* axis (Region 2) and *b* axis (Region 3) (see Figure 3-2-e-2).

Region-2						
Direct hydrogen bond						
A-subunit		Distance (Å)		A-subunit		
O-Pro40		2.8		Nζ-Lys214		
Hydrogen bond interactions intermediated by a water molecule						
A-subunit	Distance (Å)	Water	Distance (Å)	A-subunit		
Oγ-Thr71	2.7	W219	2.6	O-Glu209		
O-Gly69	2.7	W30	3.0	O-Ser206		
Oγ-Ser70	2.7	W107	3.1	Oε-Glu209		
N-Val68	3.1	W123	2.4	Oε-Glu260		
Hydrogen bond interactions intermediated by solvents						
A-subunit	Distance (Å)	Water	Distance (Å)	Ion	Distance (Å)	A-subunit
N-Gly-65	3.0	W42	2.1	PO_4^{3-}	2.8	Nζ-Lys205
					3.3	Nη-Arg202
Hydrogen bond interactions intermediated by two water molecules						
A-subunit	Distance (Å)	Water	Distance (Å)	Water	Distance (Å)	A-subunit
O-Pro42	2.8	W63	2.7	W30	3.0	O-Ser206

Region-3				
Hydrogen bond interactions intermediated by a water molecule				
A′-subunit	Distance (Å)	Water	Distance (Å)	A-subunit
O-Val68	2.7	W130	3.2	Oε-Glu227
			3.0	Oγ-Ser194
O-Pro42	3.1	W185	3.1	Nδ-Asn228
			2.5	Oδ-Asn228
N-Thr71	3.0	W74	2.7	O-Arg230
Oγ-Ser70	3.0	W5	3.2	Oε-Glu158

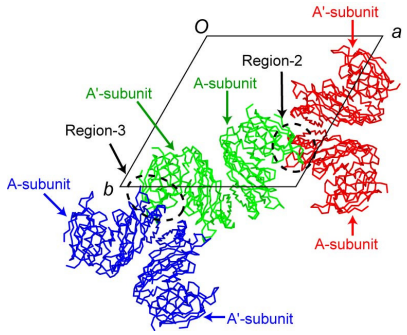


Figure 3-2-e-2. Interaction regions of the dimers between the two screwed columns related by a periodic shift along the *a* axis (Region-2) and along the *b* axis (Region-3)

3-3. X-ray analysis of *St*-Rtc

3-3-a. *St*-Rtc crystals

Crystallization of *St*-Rtc was performed to determine as the *apo*-form structure. The optimized condition, under which suitable crystals for X-ray analysis appeared after one week, is that the protein solution containing 5 mg/ml *St*-Rtc in 0.02 M Tris-HCl (pH 8.0) buffer and the reservoir solution (pH4.6) containing 2.0 M ammonium sulfate in

0.1 M sodium acetate (pH 4.6) buffer were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The largest crystal indicated by an arrow in Figure 3-3-a has a size of approximately 500×100×100 μm. This crystal was used for X-ray diffraction experiment.

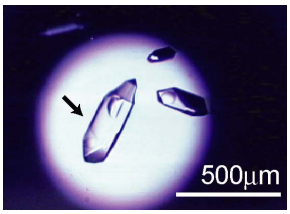


Figure 3-3-a. A photo of the *St*-Rct crystals. An arrow indicates the crystal used for X-ray diffraction experiment

3-3-b. X-ray data of the *St*-Rtc crystal

A diffraction image of the *St*-Rtc crystal is shown in Figure 3-3-b. Although the reflection spots were observed till 2.17 Å resolution, their intensity data were processed to 2.25 Å resolution. The crystal belongs to the space group $P3_1$ with cell constants $a = b = 84.4$ and $c = 109.4$ Å. The numbers of observed and unique reflections are 238521 and 41221, respectively. The data completeness and an R_{marge} value are 99.9% and 5.5%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 3-3-b.

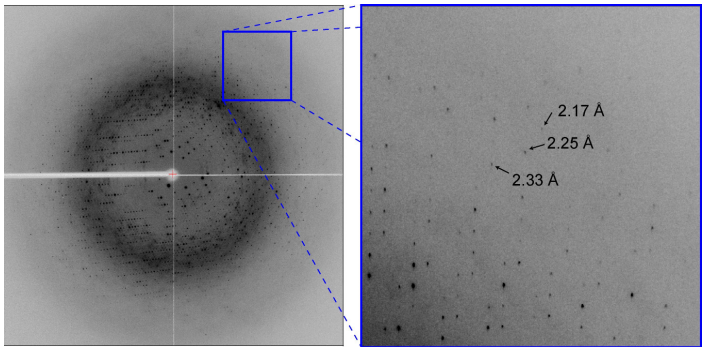


Figure 3-3-b. An X-ray diffraction image of the *St*-Rtc crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 3-3-b. A statistics of the observed intensity data and the crystallographic data of the *St*-Rtc crystal

Wavelength (Å)	1.00
Resolution (Å)	50-2.25 (2.33-2.25)
Space group	$P3_1$
Unit cell (Å)	$a = b$ =84.4, c = 109.4
Observed reflections	239382
Unique reflections	41221
Completeness (%)	99.9 (99.6)
Redundancy	5.8 (5.8)
I/σ	43.2 (6.3)
R_{merge} (%)	5.5 (30.2)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i|I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

3-3-c. Phase angle determination and structural refinement of *St*-Rtc

An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 2, the V_M value to be 2.99 Å^3/Da and the solvent content to be 58.9%. Despite the crystallization conditions are quite different from those of the SeMet derivative, these crystal habits are similar to those of the SeMet crystal. In the *St*-Rtc* crystal, two molecules (subunits) in the asymmetric unit are associated to each other by a non-crystallographic two-fold symmetry to form a dimer. To confirm this, the molecular replacement method was applied using one of the two subunit structures of *St*-Rtc* as a probe. The program *AMoRe* was used for computations. A unique solution containing two subunits in the asymmetric unit has been obtained with the 15-2.5 Å resolution data (Table 3-3-c-1). The two modules (subunits) are associated to each other to form a dimer, as expected. The crystal structure shows a reasonable dimer packing as described in the next subsection (3-3-d).

Table 3-3-c-1. A unique solution from molecular replacement of a subunit of *St*-Rtc*.

Subunit	α	β	γ	x	y	z	Cc	R
1	29.29	16.57	354.38	0.3048	0.4238	0.0000	76.7	37.3
2	93.20	163.31	191.18	0.4524	0.3919	0.0000	61.2	46.9

*α, β, γ, Eulerian angles (º), x, y, z; translations,
Cc; Correlation (%), R; Crystallographic R-factor (%)

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. In the asymmetric unit, 338 residues of the 339 residues per each molecule (subunit) have been assigned except for the C-terminal one residue. 364 water molecules and five sulfate ions are also assigned. The final R factor and R_{free} are 20.0% and 27.7%, respectively, indicating that the structure was reasonably refined. A statistics of the refinement is given in Table 3-3-c-2. The final $2|F_o|-|F_c|$ electron density map, contoured at 1.5 σ level, is shown in Figure 3-3-c-1. The refined structure is well fitted on the map. A Ramachandran plot is shown in Figure 3-3-c-2. In the Ramachandran plot,

510 (86.7%) residues are in the most favored geometry, 75 (12.8%) in the allowed geometry, 3 (0.5%) in generally allowed geometry, respectively. There is no residue in disallowed geometry. The sulfate ion 1 interacts with the side chain of Lys17 of the A-subunit. The sulfate 2 interacts with the side chains of Arg202 and Lys205 of the A′-subunit. The sulfate 3 interacts with the side chains of Lys117, Arg152 and Arg230 of the A′-subunit. The sulfate 4 interacts with the main chain of Lue183 of A′-subunit. The sulfate ion 5 interacts with the main chain of Asp5 and the side chain of Lys317 of the A′-subunit. A stereo-pair drawing of the subunits and the ions is shown in Figure 3-3-c-3.

Table 3-3-c-2. A statistics of the structural refinement of *St*-Rtc

Resolution (Å)	42.2-2.25
Used reflections	41206
R (%)	20.0
R_{free} (%)	27.7
r.m.s bond length (Å)	0.027
r.ms. angles (˚)	2.415
Protein atoms	5245
Water molecules	364
SO_4^{2-}	5

*R = 100 $\times\sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

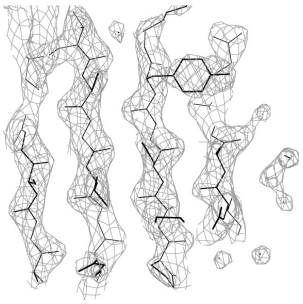


Figure 3-3-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *St*-Rtc crystal, contoured at 1.5 σ level (colored in gray).

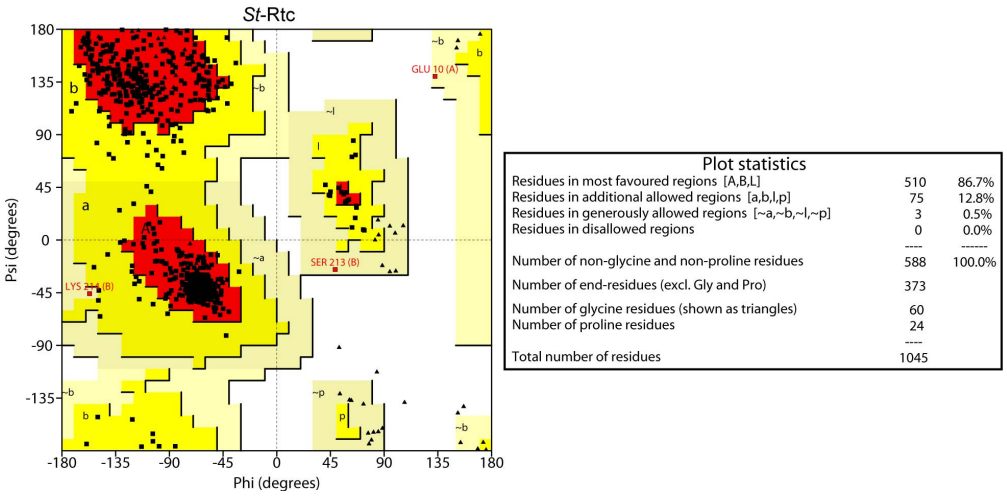


Figure 3-3-c-2. A Ramachandran plot of *St*-Rtc. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

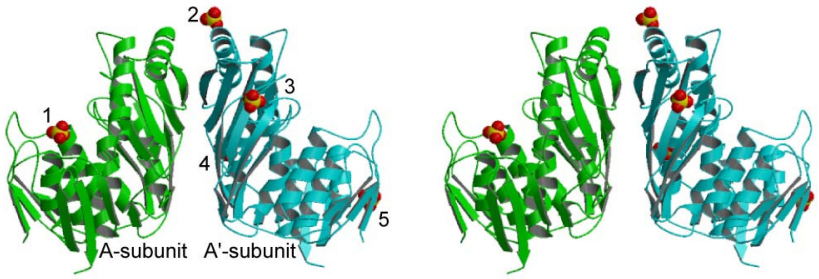


Figure 3-3-c-3. A stereo pair drawing of the A-subunit and A′-subunits of the *St*-Rtc dimer and the five sulfate ions indicated by ribbon models and yellow (sulfur atoms) and red (oxygen atoms) sphere models, respectively. The A-subunit and the A′-subunits are colored in green and cyan, respectively.

3-3-d. Dimer structure of *St*-Rtc

In the asymmetric unit, the two subunits, A and A′ are associated by a non-crystallographic two-fold symmetry to form a dimer similar to that of the SeMet derivartive. Superimposed Cα chains between two dimers in the *St*-Rtc and the *St*-Rtc* crystals are shown in wire models in Figure 3-3-d-1. The estimated r.m.s.d of Cα atoms between the dimers in the *St*-Rtc and *St*-Rtc* crystals are given in Table 3-3-d-1. It should be noted that the dimer formations maintained even under the different crystallization conditions (see subsection 3-3-a and 3-2-a).

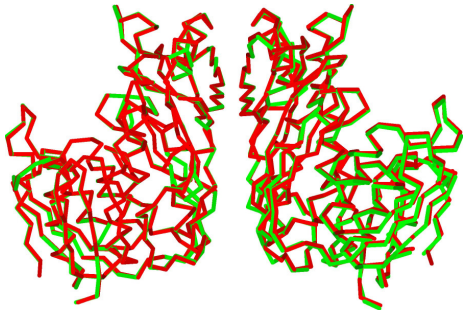


Figure 3-3-d-1. A superimposition of Cα chains of the *St*-Rtc* dimer (green line) onto the *St*-Rtc dimer (red line).

Table 3-3-d-1. The estimated r.m.s.d (Å) of Cα atoms of the dimers in the *St*-Rtc crystal with those in the *St*-Rtc* crystal.

		St-Rtc*
		A-A′ dimer
St-Rtc	A-A′ dimer	0.25
	A′-A dimer	0.36

The two subunits in the asymmetric unit are shown in Figure 3-3-d-2. When the subunits are superimposed on each other, the estimated r.m.s deviation between the corresponding Cα atoms is 0.63 Å, suggests that their folding are similar to each other, as shown in wire models in Figure 3-3-d-3. The dimer formation is stabilized by forming a β-sheet through hydrogen bonds between the two main chains, Ile223 and Ile223′. An apostrophe indicates the residue of the counter subunit. When the hydrogen bond interactions between the two subunits of the dimers in the *St*-Rtc and *St*-Rct* crystals are compared using Table 3-3-d-1 and Table 3-2-d-1, the directi hydrogen bonds are formed by atoms of different residues except for those of the main chains of Ile223 residues. The atomic distances between two subunits of the *St*-Rtc dimer are given in Table 3-3-d-2.

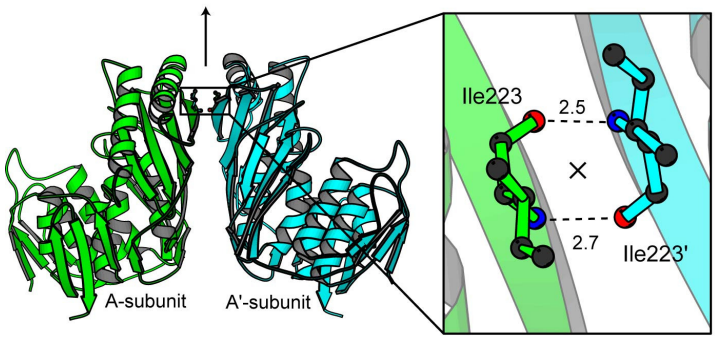


Figure 3-3-d-2. The two subunits in the asymmetric unit of the *St*-Rtc crystal. The subunits are associated to each other to form a dimer by a non-crystallographic two-fold symmetry shown with an arrow and a symbol x. The residue indicated with an apostrophe is the related counter one. The right figure shows a zoomed up view at the Ile223 residues. Black, blue and red spheres indicate carbon, nitrogen and oxygen atoms respectively. Broken lines indicate the hydrogen bonds with the distances (Å) between the two β-strands.

Table 3-3-d-2. Interactions of the two subunits of the dimer in the asymmetric unit of the *St*-Rtc crystal.

Between the small domains		
Direct hydrogen bonds		
A-subunit (Small domain)	Distance (Å)	A′-subunit (Small domain)
N-Ile223	2.5	O-Ile223
O-Ile223	2.7	N-Ile223

Between the small domain and the large domain of the opposite molecule		
Direct hydrogen bond		
A-subunit (Small domain)	Distance (Å)	A′-subunit (Large domain)
Oε-Glu220	2.8	Nη-Arg132
Hydrophobic contacts		
Chain A (Small domain)		Chain A′ (Large domain)
Ile184 Met185 Pro218		Ile149 Ile151
Salt bridge		
A-subunit (Large domain)	Distance (Å)	A′-subunit (Small domain)
Nη-Arg132	3.2	Oε-Glu220
Nη-Arg132	3.1	Oε-Glu220
Hydrophobic contacts		
A-subunit (Large domain)		A′-subunit (Small domain)
Ile149 Ile151		Ile184 Met185 Pro218

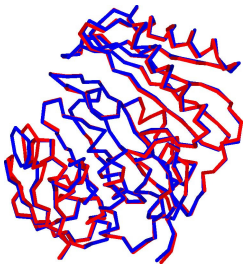


Figure 3-3-d-3. A superimposition of the Cα chains between the two subunits of the dimer of the *St*-Rtc crystal. The A-subunit and A′-subunit are colored in blue and red, respectively.

3-3-e. Subunit structure of *St*-Rtc

The subunit structure is described below. The main chain is folded into the two domains. The large domain (residue numbers 1-177 and 283-339) is composed of three repeated βαβαββ motifs, as shown in Figure 3-3-e-1 (*a*). Each motif contains a β-sheet of four β-strands and two α-helices. The repeated three motifs are arranged by a pseudo three-fold symmetry (Figure 3-3-e-1 (*b*)) so that the assembled six α-helices are surrounded by three β-sheets. On the other hand, the small domain (residue numbers 178-282) is formed a βαββββα motif inserted into the third motif of the large domain. This small domain is also composed of a β-sheet and two α-helices and their folded

structure seems to be the same as that of each motif of the large subunit, but the order of the secondary structures or their topology is interestingly different, as shown in Figure 3-3-e-2. The β11 strand of the small domain builds a β-sheet with another β11 strand of the counter subunit to form a dimer (see Figure 3-3-d-2). These structural features of the subunit are quite the same as those of the SeMet derivative.

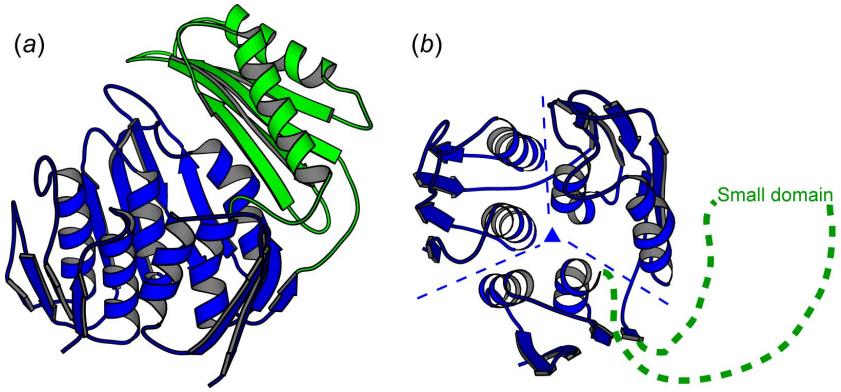


Figure 3-3-e-1. A ribbon drawing of the overall subunit *St*-Rtc (*a*). The large and small domains are colored in blue and green, respectively.
Three repeated βαβαββ motifs of the large domain, arranged by a pseudo three-fold symmetry indicated with a blue triangle (*b*). Green broken lines suggest the protruded small domain.

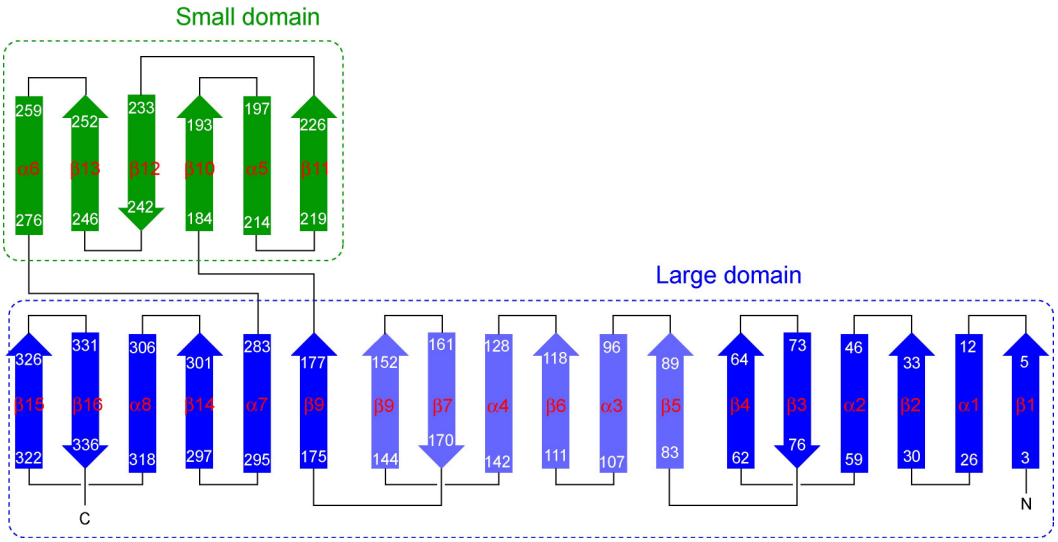


Figure 3-3-e-2. A topological diagram of the secondary structures in *St*-Rtc. Large and small domains are indicated blue and green, respectively. Arrows and bars indicate β-strands and α-helices, respectively. White and red characters on the arrows or on the bars indicate the residue number of the secondary structure code, respectively.

3-3-f. Dimer packing of the *St*-Rtc crystal

A stereo-pair drawing with Cα models of the dimers in the *St*-Rtc crystal is shown in Figure 3-3-f-1. The dimer packing in the unit cell of the *St*-Rtc crystal is shown in Figure 3-3-f-2~3. The dimers are in contact with a head-and-tail manner around the crystallographic 3_1 axis, so that a screwed column is formed by stacking the A-subunit

on the A′-subunit of the subsequent dimer. Along with the *a* and *b* axes, the columns are interacted to the adjacent columns. When the dimer packings in the *St*-Rtc* and the *St*-Rtc crystals are compared (see Figures 3-3-f-3 and 3-2-e-2), it is interesting to note that the *St*-Rtc crystal is similar to the *St*-Rtc* crystal, but these two are not isomorphous. The *St*-Rtc column is rotated by about 15º around the c axis. As seen in Figure 3-2-f-4, the contacting face of the triangle is changed to the other side (face) of the same triangle. Therefore, in the *St*-Rtc* crystal, the A′-subunit gets in contact with the A-subunit, while the A′-subunit is in contact with the A′-subunit. The atomic distances in the contacting regions along the *c* axis are listed in Table 3-3-f-1, and those along the *a* and *b* axes are listed in Table 3-3-f-2.

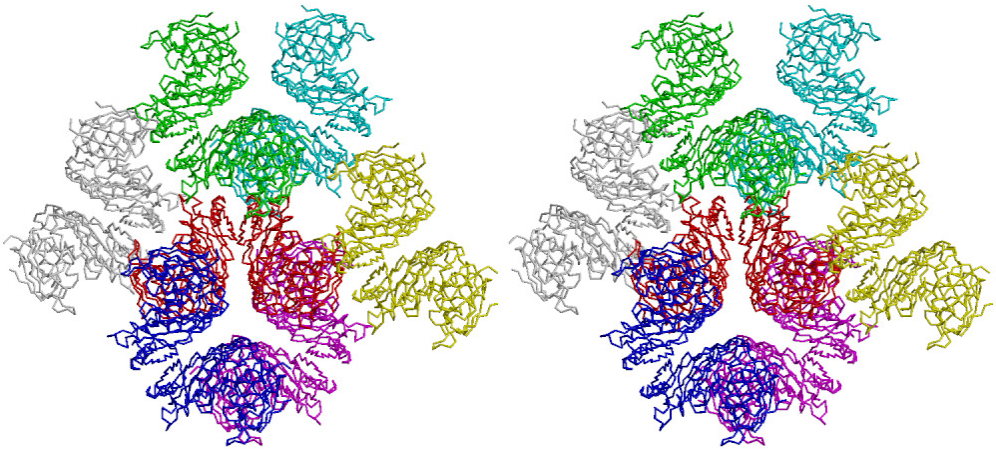


Figure 3-3-f-1. A stereo-pair drawing with Cα models of the dimers in the *St*-Rtc crystal, viewed down the *c* axis. Three dimers form a screwed column around a crystallographic 3_1 axis. Lateral interactions of the dimers are also included.

Table 3-3-f-1. Interactions of the *St*-Rtc dimmers related by the crystallographic 3_1 symmetry (Region-1) (see Figure 3-3-f-2).

Region-1				
Direct hydrogen bonds				
A′-subunit	Distance (Å)			A-subunit
Oε-Glu179	2.8			Oγ-Ser176
Oγ-Ser176	3.2			Oε-Glu179
Oε-Gln324	2.6			Nδ-Asn277
Nδ-Asn277	2.8			Oε-Gln321
O-Ser276	2.6			Oγ-Ser326
Hydrogen bond Interactions intermediated by a water molecule				
A′-subunit	Distance(Å)	Water	Distance (Å)	A-subunit
O-Pro330	2.7	W181	2.8	Pro330-O
Glu179-Oε	3.3	W359	2.7	N-Glu179
Nη-Arg180	3.3	W136	2.5	Nη-Arg334
Nδ-Asn277	2.8		3.0	Oε-Glu324
	2.6	W120	2.8	Oε-Glu298
Hydrophobic contacts				
A′-subunit		A-subunit		

Met178 Ile300 Ile332 Pro330	Met178 Ile300 Ile332 Pro330

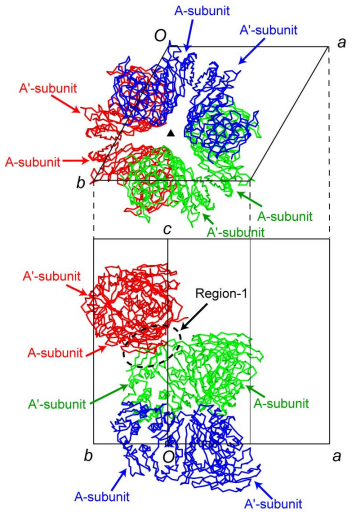


Figure 3-3-f-2. The dimers around the crystallographic three-fold screw axis in the *St*-Rtc crystal, viewed down the *c* axis (upper) and down the *b** axis (bottom)

Table 3-3-f-2. Interactions of the *St*-Rtc dimers between the two screwed column related by a periodic shift along the *a* axis (region-3) and along the *b* axis (Region-2) (See Figure 3-3-f-3).

Region-2				
Direct hydrogen bonds				
A′-subunit	Distance (Å)		A′-subunit	
O-Arg39	3.2		Nζ-Lys214	
N-Thr71	3.0		O-Glu209	
Hydrogen bond Interactions intermediated by a water molecule				
A′-subunit	Distance (Å)	Water	Distance (Å)	A′-subunit
N-Val168	3.0	W5	2.7	Oε-Glu260
O-Ala-37	3.0	W355	2.4	Nζ-Lys-214
O-Arg39	2.8			
Oγ-Ser70	3.1	W349	2.4	Oε-Glu209
O-Gly69	2.4	W104	3.3	Oε-Glu264
Hydrogen bond interactions intermediated by solvent				
A′-subunit	Distance (Å)	Ion	Distance (Å)	A′-subunit
N-Gly65	2.9	SO_4^{2-}	2.9	Nζ-Lys205
			3.1	Nη-Arg202
			3.2	Oε-Glu201
			3.1	Oε-Glu201
			3.0	Nε-Arg202

Region-3						
Hydrogen bond Interactions intermediated by a water molecule						
A-subunit	Distance (Å)	Water	Distance (Å)	A′-subunit		
Oγ-Thr71	3.3	W99	2.7	O-Arg320		
O-Val168	2.6	W115	3.0	Oγ-Ser194		
			3.1	N-Asn228		
Hydrogen bond Interactions intermediated by two water molecules						
A-subunit	Distance (Å)	Water	Distance (Å)	Water	Distance (Å)	A′-subunit

Oγ-Ser70	2.8	W224	3.0	W282	3.3	Oε-Glu158

Salt bridge		
A-subunit	Distance(Å)	A′-subunit
Nε-Gln45	3.0	Oδ-Asn228
Oε-Gln45	2.6	Nδ-Asn228

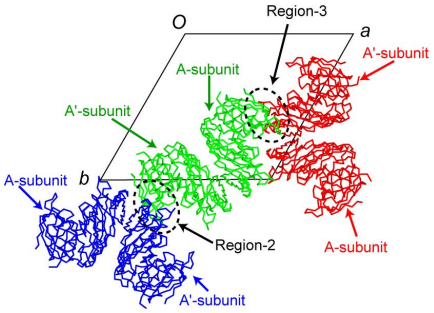


Figure 3-3-f-3. Interaction regions of the dimers between the two screwed columns related by a periodic shift along the *a* axis (Region-3) and along the *b* axis (Region-2).

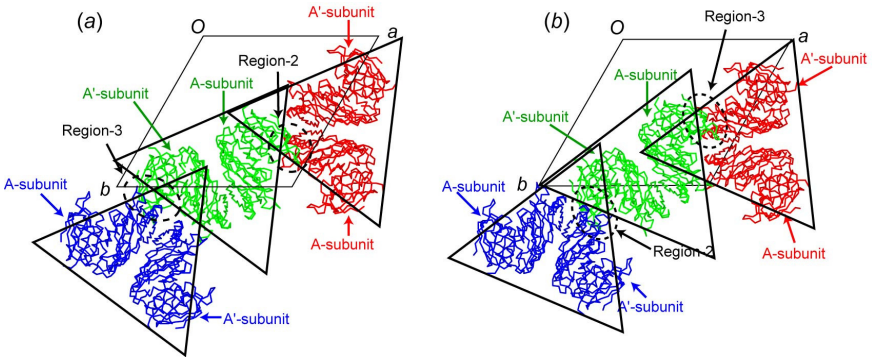


Figure 3-3-f-4. Interaction regions between the dimers in the *St*-Rtc crystal (*a*) and the *St*-Rtc* crystal (*b*). Each dimer is surrounded by a triangle to resolve the difference of the packing formation and the contacting regions.

3-4. X-ray analysis of *St*-Rtc*+Mn

3-4-a. *St*-Rtc*+Mn crystals

In crystallization of *St*-Rtc*+Mn, 2mM $MnCl_2$ was added to 0.02 M Tris-HCl buffer solution (pH8.0) containing 5mg/ml *St*-Rtc*. The condition was that the protein solution and the reservoir solution (pH4.6) containing 2.0 M ammonium sulfate in 0.1 M sodium acetate (pH 4.6) buffer were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293K. Under this condition, crystals for X-ray analysis appeared after one week. The crystal indicated by an arrow in Figure 3-4-a has a size of approximately 50×50×100 μm. This crystal was used for X-ray diffraction experiment. Crystallization of *St*-Rtc+Mn was also preformed, but more suitable

crystals of *St*-Rtc+Mn for X-ray diffraction experiment could not appeared than the *St*-Rtc*+Mn crystals.

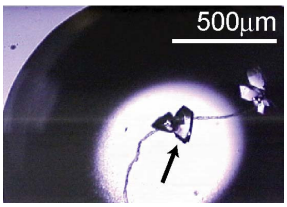


Figure 3-4-a. A photo of the *St*-Rtc*+Mn crystals. An arrow indicates the crystal used for X-ray diffraction experiment

3-4-b. X-ray data of the *St*-Rtc*+Mn crystal

A diffraction image of the *St*-Rtc*+Mn crystal is shown in Figure 3-4-b. Although reflection spots were observed till 2.96 Å resolution, their intensity data were processed to 3.20 Å resolution. The crystal belongs to the space group $P3_1$ with cell constants a = b = 83.5 and c = 109.7 Å. The numbers of the observed and unique reflections are 82064 and 14154, respectively. The data completeness and an R_{marge} value are 99.7% and 6.5%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 3-4-b.

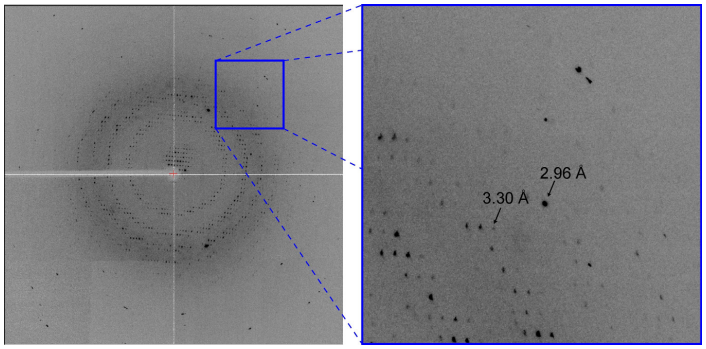


Figure 3-4-b. An X-ray diffraction image of the *St*-Rtc*+Mn crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with the values.

Table 3-4-b. A statistics of the observed intensity data and the crystallographic data of the *St*-Rtc*+Mn crystal

Wavelength (Å)	1.00
Resolution (Å)	50-3.20 (3.31-3.20)
Space group	$P3_1$
Unit cell (Å)	$a = b = 83.5$, $c = 109.7$
Observed reflections	82064
Unique reflections	14154
Completeness (%)	99.7 (99.5)
Redundancy	5.8 (5.8)
I/σ	35.3 (11.9)
R_{merge} (%)	6.5 (17.3)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl) - I(hkl)| / \sum_{hkl}\sum_i I(hkl)$

3-4-c. Phase angle determination and structural refinement of *St*-Rtc*+Mn

An acceptable number of the protein molecules in the asymmetric unit was estimated to be 2, the V_M value to be 2.94 Å^3/Da and the solvent content to be 58.2%. Despite the crystallization conditions are quite different from those of the *St*-Rtc and its SeMet derivative, these crystal habits are similar to them. In the *St*-Rtc and its SeMet derivative crystals, two molecules (subunits) in the asymmetric unit are associated to each other by a non-crystallographic two-fold symmetry to form a dimer. To confirm this again, the molecular replacement was applied using one of the two subunit structures of *St*-Rtc* as a probe. The program *Phaser* was used for computations. A unique solution containing two subunits in the asymmetric unit has been obtained with the 43.7-3.2 Å resolution data (Table 3-4-c-1). The two molecules (subunits) are associated to each other to form a dimer, as expected. The crystal structure shows a reasonable dimer packing as described in the subsection (3-4-e).

Table 3-4-c-1. A unique solution from molecular replacement of a subunit of *St*-Rtc*

Subunit	α	β	γ	x	y	z	RFZ	TFZ
1	52.6	1.1	307.4	-0.005	-0.006	-0.001	21.1	27.0
2	110.9	174.9	37.3	0.797	0.461	0.189	18.6	52.7

*α, β, γ, Eulerian angles (º), x, y, z; Translations,
RFZ; Z-score (number of standard deviations above the mean value of peaks) of rotation
TFZ; Z-score (number of standard deviations above the mean value of peaks) of translation

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. A statistics of the structural refinement is given in Table 3-4-c-2. In the asymmetric unit, 338 residues of the 339 residues are assigned except for the C-terminal one residue in one subunit, and all 339 residues are assigned in another subunit. The final R factor and R_{free} are 25.0% and 25.8%, respectively. The final $2|F_o|-|F_c|$ electron density map, contoured at 1.5 σ level, is shown in Figure 3-4-c-1. The refined structure is well fitted on the map. A Ramachandran plot is shown in Figure 3-4-c-2. In the Ramachandran plot, 485 (82.3%) residues are in the most favored geometry, 84 (14.3%) in the allowed geometry, 14 (2.4%) in generously allowed geometry and 6 (1.0%) in disallowed geometry, respectively.

However, it was difficult to find the density of added Mn atoms. It may be ascribed to low resolution data or to low Mn^{2+} concentration in crystallization. The latter reason is more plausible, because it was reported that the optimal Mg^{2+} concentration for the catalytic reaction was 4~5 mM [3]. Another reason is that Mg^{2+} or Mn^{2+} ion is bound the enzyme with the help of ATP binding.

Table 3-4-c-2. A statistics of the structural refinement of *St*-Rtc*+Mn

Resolution (Å)	41.8-3.2
Used reflections	13445
R (%)	25.0
R_{free} (%)	25.8
r.m.s bond length (Å)	0.025
r.ms. angles (˚)	2.509
Protein atoms	5269

*R = 100 ×$\sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

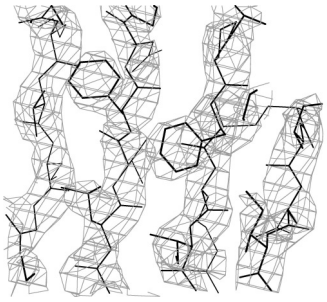


Figure 3-4-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *St*-Rtc*+Mn crystal, contoured at 1.5 σ level (colored in gray).

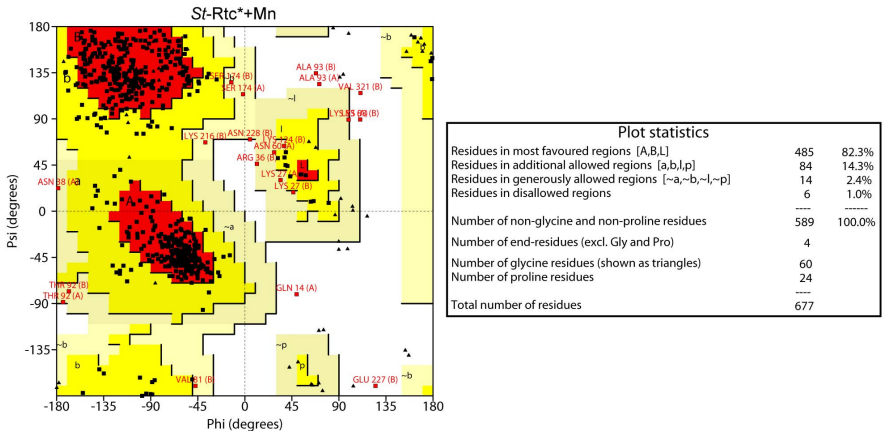


Figure 3-4-c-2. A Ramachandran plot of *St*-Rtc*+Mn. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

3-4-d. Dimer structure of *St*-Rtc*+Mn

In the asymmetric unit, the two subunits, A and A′ are associated by a non-crystallographic two-fold symmetry to form a dimer similar to that in the *St*-Rtc and the *St*-Rtc* crystals. Superimposed Cα chains between three dimers in the *St*-Rtc*,

St-Rtc and the *St*-Rtc*+Mn crystals are shown in wire models in Figure 3-4-d-1. The estimated r.m.s.d of Cα atoms between the dimers in the *St*-Rtc and the *St*-Rtc* crystals are given in Table 3-4-d-1. It should be noted that the dimer formation is maintained even under the three different crystallization conditions (see subsection (3-2-a), (3-3-a) and (3-4-a)).

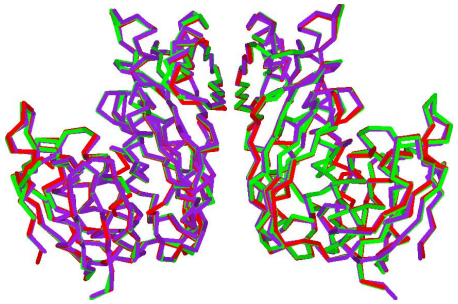


Figure 3-4-d-1. A superimposition of Cα chains with the dimers in the *St*-Rtc* (green line), the *St*-Rtc (red line) and the *St*-Rtc*+Mn crystals (purple line).

Table 3-4-d-1. The estimated r.m.s.d (Å) of Cα atoms between the dimers of the *St*-Rtc*, the *St*-Rtc and the *St*-Rtc*+Mn crystals.

		St-Rtc*	*St*-Rtc
		A-A′ dimer	A-A′ dimer
St-Rtc*+Mn	A-A′ dimer	0.51	0.57
	A′-A dimer	0.96	1.00

The two subunits, A and A′ in the asymmetric unit are shown in Figure 3-4-d-2. When they are superimposed on each other, the estimated r.m.s deviation between the corresponding Cα atoms is 0.32 Å, suggests that their folding are similar to each other, as shown in wire models in Figure 3-4-d-3. Figure 3-4-d-2 also shows that the two subunits are associated by a non-crystallographic two-fold symmetry to form a dimer, similar to those in the *St*-Rtc* and the *St*-Rtc crystals. The dimer formation is stabilized by forming a β-sheet through hydrogen bonds between the two main chains of Ile223 and Ile223′. An apostrophe indicates the residue of the counter subunit. The atomic distances given in Table 3-4-d-2 show that there are no abnormal contacts between the two subunits.

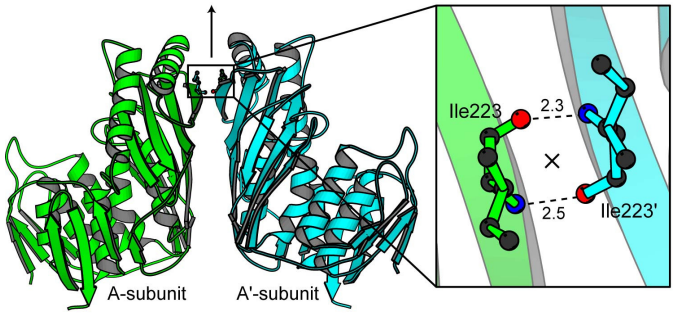


Figure 3-4-d-2. The two subunits in the asymmetric unit of the *St*-Rtc*+Mn crystal. The subunits are associated to each other to form a dimer by a non-crystallographic two-fold symmetry shown with an arrow or a symbol x. The residue indicated with an apostrophe is the related counter one. The right figure shows a zoomed up view at the Ile223 residues. Black, blue and red spheres indicate carbon, nitrogen and oxygen atoms respectively. Broken lines indicate the hydrogen bonds with the distances (Å) between the two β-strands.

Table 3-4-d-2. Interactions of the two subunits of the dimer in the asymmetric unit of the *St*-Rtc*+Mn crystal.

Between the small domains		
Direct hydrogen bonds		
A-subunit (Small domain)	Distance (Å)	A′-subunit (Small domain)
N-Ile223	2.5	O-Ile223
O-Ile223	2.3	N-Ile223
Oε-Glu220	3.3	Nη-Arg226
Nη-Arg226	3.3	Oε-Glu220

Between the small domain and the opposite large domain	
Hydrophobic contacts	
A-subunit (Small domain)	A′-subunit (Large domain)
Ile184 Met185 Pro218	Ile149 Ile151
Hydrophobic interactions	
A-subunit (Large domain)	A′-subunit (Small domain)
Ile149 Ile151	Ile184 Met185 Pro218

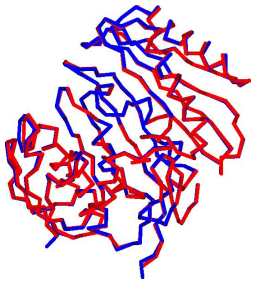


Figure 3-4-d-3. A superimposition of the Cα chains between the two subunits of the dimer in the asymmetric unit of the *St*-Rtc*+Mn crystal, drawn by wire models. The A-subunit and the A′-subunit are colored in blue and red, respectively.

3-4-e. Dimer packing of the *St*-Rtc*+Mn crystal

The dimer packing in the unit cell of the *St*-Rtc*+Mn crystal is shown in Figure 3-4-e-1~2. The dimers are in contact with a head-and-tail manner around the crystallographic 3_1 axis, so that a screwed column is formed by stacking the A-subunit on the A′-subunit of the subsequent dimer. Along with the *a* and *b* axes, the columns are interacted to the adjacent columns. As seen in Figures 3-2-e-2 and 3-4-e-2, the dimer packings are similar to each other between the St-Rtc*+Mn crystal and the St-Rtc* crystal. The atomic distances in the contact regions along the *c* axis are listed in Table 3-4-e-1, and those along the *a* and *b* axes are listed in Table 3-4-e-2.

Table 3-4-e-1. Interactions of the *St*-Rtc*+Mn dimmers related by the crystallographic 3_1 symmetry (Region-1) (see Figure 3-4-e-1).

Region-1		
Direct hydrogen bonds		
A′-subunit	Distance (Å)	A-subunit
Oε-Glu179	3.0	Oγ-Ser176
Oε-Gln324	3.1	Nδ-Asn277
Oδ-Asn277	3.2	Nε-Glu298
Oε-Glu182	2.9	Nη-Arg334
O-Ser176	2.5	Oγ-Ser326

Hydrophobic contacts	
A′-subunit	A-subunit
Met178 Ile300 Ile332 Pro330	Met178 Ile300 Ile332 Pro330

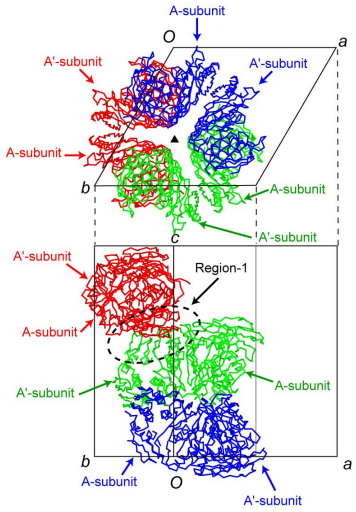


Figure 3-4-e-1. The dimers around the crystallographic three-fold screw axis in the *St*-Rtc*+Mn crystal, viewed down the *c* axis (upper) and down the *b** axis (bottom)

Table 3-4-e-2. Interactions of the *St*-Rtc*+Mn dimers between the two screwed column related by a periodic shift along the *a* axis (region-2) and along the *b* axis (Region-3) (See Figure 3-4-e-2).

Region-2		
Direct hydrogen bonds		
A-subunit	Distance (Å)	A-subunit
N-Thr71	2.7	Oε-Glu209
O-Pro40	2.6	Nζ-Lys214

Region-3		
Direct hydrogen bond		
A′-subunit	Distance (Å)	A-subunit
O-Val68	2.3	Nδ-Asn228

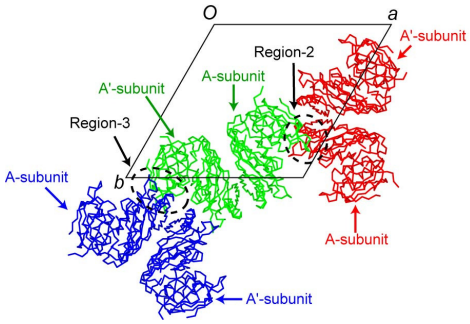


Figure 3-4-e-2. Interaction regions of the dimers between the two screwed columns related by a periodic shift along the *a* axis (Region-2) and along the *b* axis (Region-3).

3-5. X-ray analysis of *St*-Rtc*+ATP

3-5-a. *St*-Rtc*+ATP crystals

Crystallization of *St*-Rtc*+ATP was performed to determine the ATP binding site with the absence of Mg^{2+} ion to suppress the first step reaction. The optimized condition, under which suitable crystals for X-ray analysis were obtained after one week, is that the protein solution containing 5 mg/ml *St*-Rtc* and 0.5 mM ATP in 0.02 M Tris-HCl (pH 8.0) buffer, and the reservoir solution (pH8.2) containing 2.1 M ammonium sulfate in 0.1 M Tris-HCl (pH8.0) buffer were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293K. The crystal indicated by an arrow in Figure 3-5-a, has a size of approximately 50×50×200 μm. This crystal was used for X-ray diffraction experiment.

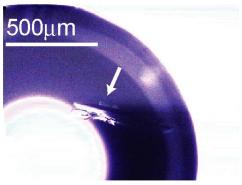


Figure 3-5-a. A photo of *St*-Rtc*+ATP crystals. An arrow indicates the crystal used for X-ray diffraction experiment

3-5-b. X-ray data of the *St*-Rtc*+ATP crystal

A diffraction image of the *St*-Rtc*+ATP crystal is shown in Figure 3-5-b. Although the reflection spots were observed till 2.35 Å resolution, their intensity data were processed to 2.40 Å resolution. The crystal belongs to the space group $P4_2$ with cell constants $a = b = 91.1$ and $c = 89.9$ Å. The space group and the cell constants are different from the other three crystals; the *St*-Rtc*, the *St*-Rtc and the *St*-Rtc*+Mn crystals. The numbers of the observed and unique reflections are 211829 and 28513, respectively. The data completeness and an R_{marge} value are 99.8% and 8.7%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 3-5-b.

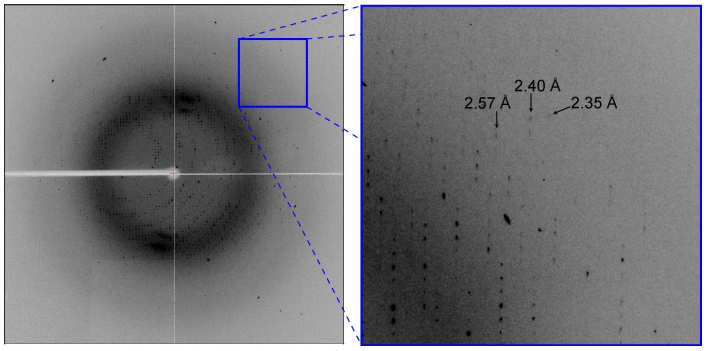


Figure 3-5-b. An X-ray diffraction image of the *St*-Rtc*+ATP crystal, taken with an oscillation angle of 1º (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 3-5-b. A statistics of the observed intensity data and the crystallographic data of the *St*-Rtc*+ATP crystal

Wavelength (Å)	0.98
Resolution (Å)	50-2.40 (2.49-2.40)
Space group	$P4_2$
Unit cell (Å)	$a = b = 91.1$, $c = 89.9$
Observed reflections	211829
Unique reflections	28513
Completeness (%)	99.8 (99.8)
Redundancy	7.4 (7.1)
I/σ	36.6 (7.6)
R_{merge} (%)	8.7 (25.8)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

3-5-c. Phase angle determination and structural refinement of *St*-Rtc*+ATP

An acceptable number of the protein molecules in the asymmetric unit was estimated to be 2, the V_M value to be 2.44 Å^3/Da and the solvent content to be 49.6%. These crystal habits are different from those of the *St*-Rtc*, the *St*-Rtc and the *St*-Rtc*+Mn crystals. To estimate initial phase angles, the molecular replacement method was applied using one of the two subunits of *St*-Rtc* as a probe. The program *Phaser* was used for computations. A unique solution containing two molecules in the

asymmetric unit has been obtained (Table 3-5-c-1).

Table 3-5-c-1. A unique solution from the molecular replacement of a subunit of *St*-Rtc*

Molecule	α	β	γ	x	y	z	RFZ	TFZ
1	61.4	87.3	233.0	0.159	-0.453	0.001	19.9	24.8
2	28.6	92.7	53.0	0.047	-0.340	-0.052	16.5	57.2

*α, β, γ, Eulerian angles, x, y, z; Translations,
RFZ; *Z*-score (number of standard deviations above the mean value of peaks) of rotation function,
TFZ; *Z*-score (number of standard deviations above the mean value of peaks) of translation function

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. 337 residues of the 339 residues per each molecule are assigned except for the C-terminal two residues. 219 water molecules, two ATP molecules and four sulfate ions are also assigned. The final *R* factor and R_{free} are 20.7% and 26.4%, respectively indicating the structures were reasonably refined. A statistics of the refinement is given in Table 3-5-c-2. The final $2|F_o|-|F_c|$ electron density map, contoured at 1.5 σ level, is shown in Figure 3-5-c-1. The refined structure is well fitted on the map. A Ramachandran plot is shown in Figure 3-5-c-2. In the Ramachandran plot, 526 (89.8%) residues in an asymmetric unit are in the most favored geometry, 55 (9.4%) in the allowed geometry, 3 (0.5%) in generously allowed geometry and 2 (0.3%) in disallowed geometry, respectively. One sulfate ion (see Figure 3-5-c-3 (*a*)) interacts with the main chain of Asp5 and the side chain of Lys317 of the large domain of the A-subunit. The second sulfate ion (see Figure 3-5-c-3 (*b*)) interacts with the side chain of Arg202 and the main chain of Ala259 of the small domain of the B-subunit. The third sulfate ion is found between the B and B′ subunits. Stereo-pair drawings of the ions and the dimers are shown in Figure 3-5-c-3

Table 3-5-c-2. A statistics of the structural refinement of *St*-Rtc*+ATP

Resolution (Å)	32.0-2.4
Used reflections	28197
R (%)	20.7
R_{free} (%)	26.4
r.m.s bond length (Å)	0.02
r.ms. angles (°)	2.3
Protein atoms	5246
Water molecules	220
ATP	2
SO_4^{2-}	3

*$R = 100 \times \sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

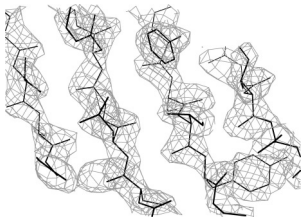
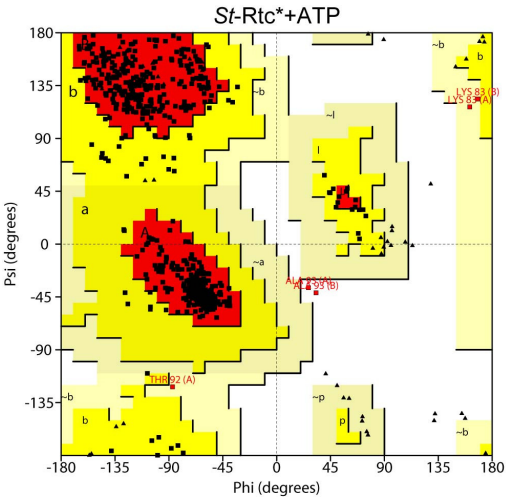


Figure 3-5-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *St*-Rtc*+ATP crystal, contoured at 1.5 σ level (colored in gray).



Plot statistics

Residues in most favoured regions [A,B,L]	526	89.8%
Residues in additional allowed regions [a,b,l,p]	55	9.4%
Residues in generously allowed regions [~a,~b,~l,~p]	3	0.5%
Residues in disallowed regions	2	0.3%
Number of non-glycine and non-proline residues	586	100.0%
Number of end-residues (excl. Gly and Pro)	228	
Number of glycine residues (shown as triangles)	60	
Number of proline residues	24	
Total number of residues	898	

Figure 3-5-c-2. A Ramachandran plot of *St*-Rtc*+ATP. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

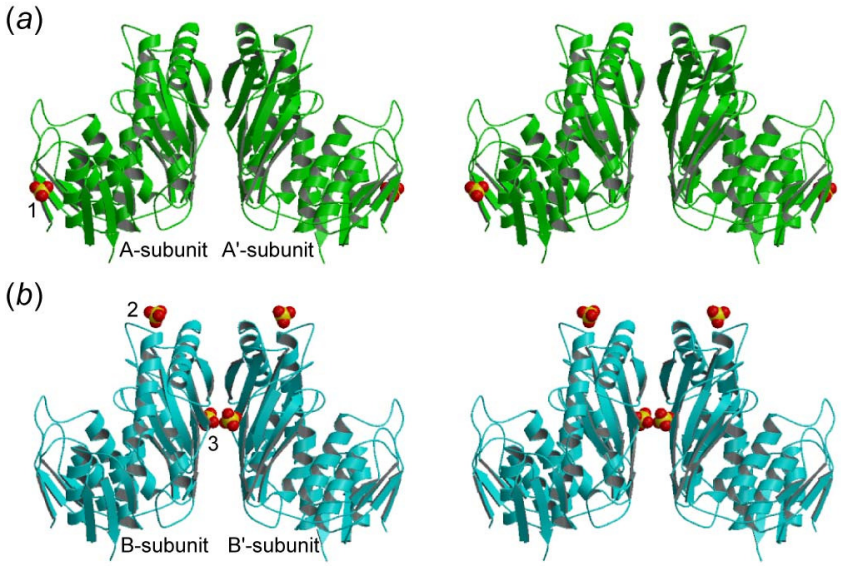


Figure 3-5-c-3. Stereo-pair drawings of A-A′ (*a*) and B-B′ dimers (*b*) with the bound three sulfate ions. They are indicated in ribbon models and yellow and red sphere models, respectively. The A-A′ and the B-B′ dimers are colored in green and cyan, respectively.

3-5-d. Dimer structure of *St*-Rtc*+ATP

In the *St*-Rtc*+ATP crystal, the two found molecules A and B in the asymmetric unit separately form dimers with those related by the crystallographic two-fold symmetries passing the origin and passing the center of the *a*-*b* plane (see Figure 3-5-e in the next subsection 3-5-e). In other words, there are two dimers in the unit cell, one between A and A′ subunits related by an exact two-fold symmetry, and other between B and B′ subunits related by another exact two-fold symmetry. This is contrast to those found in the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn crystals. In the latter cases, the dimers are formed by non-crystallographic (approximate) two-fold symmetries.

The two dimers, A-A′ and B-B′, in the *St*-Rtc*+ATP crystal are similar to each other with an r.m.s.d of 0.14 Å. Superimposed Cα chains between the two *St*-Rtc*+ATP dimers and those of the three dimers in the *St*-Rtc*, the *St*-Rtc and the *St*-Rtc*+Mn crystals is shown by wire models in Figure 3-5-d-1. The estimated r.m.s deviations of the Cα atoms are shown in Table 3-5-d-1. It is important to note that the dimer formation is maintained not only under the different crystallization conditions but also through the different molecular packing, as described in the next subsection.

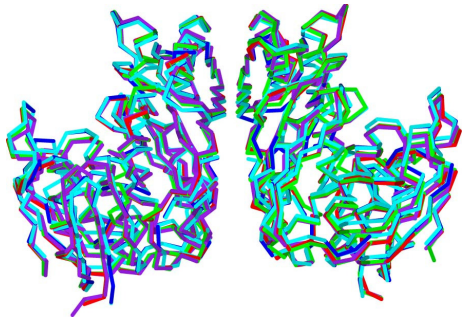


Figure 3-5-d-1. A superimposition of Cα chains between the dimers in the *St*-Rtc* (green), the *St*-Rtc (red), *St*-Rtc*+Mn (purple), and the A-A′ (blue) and B-B′ dimer (cyan) in the *St*-Rtc*+ATP crystal.

Table 3-5-d-1. The estimated r.m.s.d (Å) of Cα atoms of the two dimers in the *St*-Rtc*+ATP, with the three dimers in the *St*-Rtc*, *St*-Rtc, *St*-Rtc*+Mn crystals.

		St-Rtc*	*St*-Rtc	*St*-Rtc*+Mn
		A-A′ dimer	A-A′ dimer	A-A′ dimer
St-Rtc*+ATP	A-A′ dimer	1.57	1.57	1.77
	B-B′ dimer	1.58	1.58	1.77

The dimer formations are stabilized by a forming β-sheet through hydrogen bonds between the two main chains, Ile223 and Ile223′, as shown in Figure 3-5-d-2. An apostrophe indicates the residue of the counter subunit. They are similar to those in the *St*-Rtc*, the *St*-Rtc and the *St*-Rtc*+Mn crystals. The hydrogen bond interactions

between the small domains of the A-A′ dimer and of the B-B′ dimer are given in Table 3-5-d-2 and Table 3-5-d-3, respectively. The hydrogen bond, Oε (Glu227) ⋯ Nζ (Lys205) is formed in both dimers. Several salt bridges are formed between the side chains of Arg132 in the large domain and the side chain of Glu220 in the opposite small domain of B-B′ dimer. Two sulfate ions, each bound to the B or B′ subunit are bound near these salt bridges (see Figure 3-5-c-3). The atomic distances between two subunits of A-A′ dimer are given in Table 3-5-d-2 and those between two subunits of B-B′ dimer are given in Table 3-5-d-3.

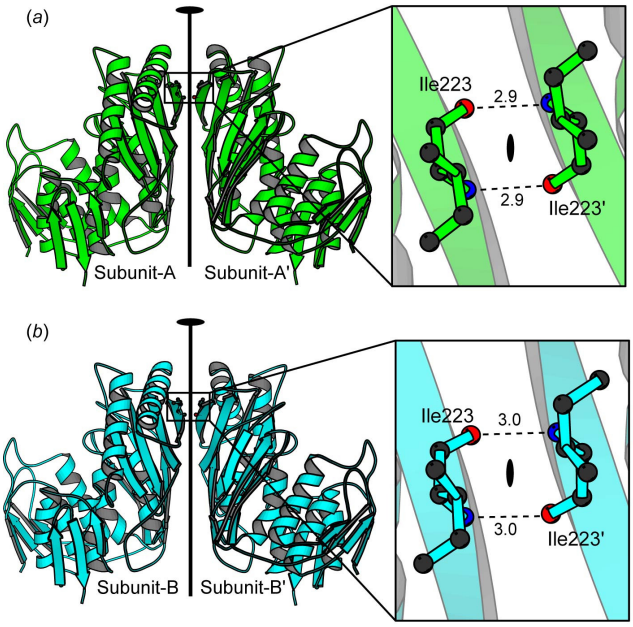


Figure 3-5-d-2. The A-A′ (*a*) and the B-B′ dimers (*b*) in the *St*-Rtc*-AMP crystal. A-subunit is associated to the A′-subunit, related by a crystallographic two-fold symmetry shown with a thick line and an elliptical symbol. The B-B′ dimers is also formed by another crystallographic two-fold symmetry. The residue indicated with an asterisk is the related counter one. The right figure shows a zoomed up view at the Ile223 residues. Black, blue and red spheres indicate carbon, nitrogen and oxygen atoms respectively. Broken lines indicate the hydrogen bonds with the distances (Å) between the two β-strands.

Table 3-5-d-2. Interactions between the A and A′ subunits related by a two-fold crystallographic symmetry passing the center of *a*-*b* plane of the unit cell (see Figure 3-5-e).

Between the small domains		
Direct hydrogen bonds		
A-subunit (Small domain)	Distance (Å)	A′-subunit (Small domain)
O-Ile223	2.9	N-Ile223
Oε-Glu227	2.6	Nζ-Lys205
Oε-Glu201	3.2	Oγ-Ser197
Oε-Glu220	3.2	Nη-Arg226
N-Ile223	2.9	O-Ile223
Nζ-Lys205	2.6	Oε-Glu227
Oγ-Ser197	3.2	Oε-Glu201
Nη-Arg226	3.2	Oε-Glu220
Between the small domain, and the large domain of the opposite chain.		

Direct hydrogen bonds				
A-subunit (Small domain)	Distance (Å)			A′-subunit (Large domain)
Oε-Glu182	3.0			Nζ-Lys113
Nζ-Lys113	3.0			Oε-Glu182
Hydrogen bond interactions intermediated by a water molecule				
A-subunit (Small domain)	Distance(Å)	Water	Distance (Å)	A′-subunit (Large domain)
Oε-Glu220	2.6	W126	3.1	N-Ile150
N-Ile150	3.1	W126	2.6	Oε-Glu220
Hydrophobic contacts				
A-subunit (Small domain)		A′-subunit (Small domain)		
Ile184 Met185 Pro218		Ile149 Ile151		
Ile149 Ile151		Ile184 Met185 Pro218		
Between the large domains				
Hydrogen bond interactions intermediated by a water molecule				
A-Subunit (Large domain)	Distance(Å)	Water	Distance (Å)	A′-subunit (Large domain)
Oε-Glu140	3.4	W172	3.4	Oε-Glu140
Oε-Glu140	2.9		2.9	Oε-Glu140

Table 3-5-d-3. Interactions between the B and B′ subunits related by a two-fold crystallographic symmetry passing the origin of the unit cell (see Figure 3-5-e).

Between the small domains				
Direct hydrogen bonds				
B-subunit (Small domain)	Distance (Å)			B′-subunit (Small domain)
O-Ile 223	3.0			N-Ile223
Oε-Glu227	3.2			Nζ-Lys205
Oε-Glu227	2.9			
Oε-Glu201	3.2			Oγ-Ser197
N-Ile223	3.0			O-Ile 223
Nζ-Lys205	3.2			Oε-Glu227
	2.9			Oε-Glu227
Oγ-Ser197	3.2			Oε-Glu201
Hydrogen bond interactions intermediated by a water molecule				
B-subunit (Small domain)	Distance(Å)	Water	Distance (Å)	B′-subunit (Small domain)
O-Ile221	2.6	W53	2.9	N-Val225
N-Val225	2.9	W53	2.6	O-Ile221
Between the small domain, and the large domain of the opposite chain.				
Direct hydrogen bond				
B-subunit (Small domain)	Distance (Å)			B′-subunit (Large domain)
Oε-Glu167	2.9			Nε-Gln243
Nε-Gln243	2.9			Oε-Glu167
Hydrogen bond interactions intermediated by a water molecule				
B-Subunit (Small domain)	Distance(Å)	Water	Distance (Å)	B′-subunit (Large domain)
N-Leu150	3.0	W136	2.7	Oε-Glu220
Oε-Glu220	2.7	W136	3.0	N-Leu150
Salt bridge				
B-subunit (Small domain)	Distance (Å)			B′-subunit (Large domain)

Oε-Glu220	2.7	Nη-Arg132
Oε-Glu220	2.8	Nη-Arg132
Nη-Arg132	2.7	Oε-Glu220
Nη-Arg132	2.8	Oε-Glu220
Hydrophobic contacts		
B-subunit (Small domain)		B′-subunit (Large domain)
Ile184 Met184 Pro218		Ile149 Ile151
Ile149 Ile151		Ile184 Met184 Pro218

3-5-e. Dimer packing in the *St*-Rtc*+ATP crystal

Packing diagrams of the dimers in the *St*-Rtc*+ATP crystal is shown in Figure 3-5-e. The two columns, one is composed of the A-A′ dimers and the other is composed of B-B′ dimers, are arranged in an opposite direction to each other to deny the helical polarity. In each column, the dimers are stacked on each other on the c-axis by rotating 90º successively. The dimer-dimer interaction regions in the A-A′ dimer column and those in the B-B′ dimer column are referred to Region-A1, Region-A2, Region-B1 and Region-B2, respectively (see Figure 3-5-e). Hydrogen bond interactions between Oε (Glu82), Nε (Arg254) and Nη (Arg258) atoms are found in both Region-A1 and Region-B1. In Region-A2, however, every hydrogen bond, formed though water molecules. The atomic distances of these interactions are listed in Table 3-5-e-1 and Table 3-5-e-2.

The two interaction regions between the adjacent columns are referred to Region-AB1 and Region-AB2 (see Figure 3-5-e). The two hydrogen bonds, Oε (Glu243) ⋯ Nζ (Lys34), and Oη (Tyr271) ⋯ N (Met1) are formed in both Region-AB1 and Region-AB2. The atomic distances of these interactions in those regions are given in Table 3-5-e-3.

The A and B subunits are in contact to each other in Region-AB3 and Region-AB4 in the asymmetric unit. Atomic distances in the regions are given in Table 3-5-e-4.

These interactions are quite different from those in the other three crystals, *St*-Rtc*, *St*-Rtc and *St*-Rtc*+Mn, which were crystallized under acidic conditions. The basic crystallization condition changes the dimers packing.

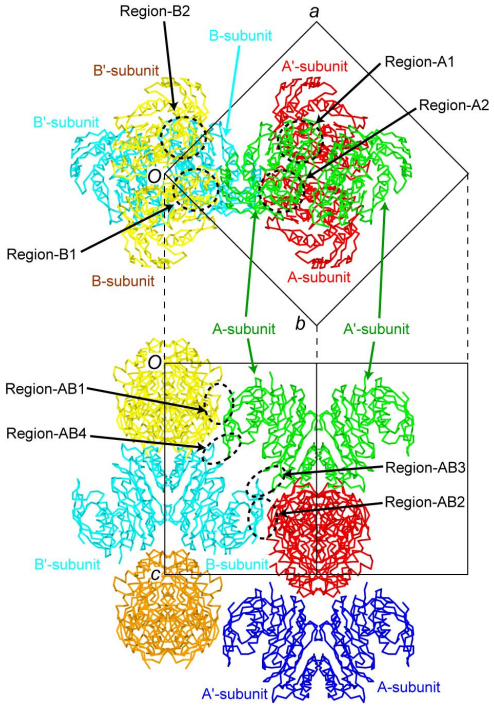


Figure 3-5-e. Packing diagrams in the *St*-Rtc*+ATP unit cell, viewed down the *c* axis (upper) and along the (-1 1 0) direction (bottom). The A-A′ dimers indicated by wire models are colored in green, red and blue. And the B-B′ dimers are colored in yellow, cyan and orange.

Table 3-5-e-1. Interactions between A and A′ subunits related by a periodic shift along the *c* axis (Region-A1) (See Figure 3-5-e).

Region-A1				
Salt bridge				
A-subunit	Distance (Å)			A′-subunit
Oε-Glu82	3.2			Nε-Arg254
Oε-Glu82	2.8			Nη-Arg258
Nε-Arg254	3.2			Oε-Glu82
Nη-Arg258	2.8			Oε-Glu82
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance(Å)	Water	Distance (Å)	A′-subunit
Nη-Arg145	3.3	W121	3.2	Nε-His198
Nε-His198	3.2	W121	3.3	Nη-Arg145

Table 3-5-e-2. Interactions between the B and B′ subunits related by a periodic shift along the *c* axis (Region-B1 and Region-B2) (See Figure 3-5-e).

Region-B1		
Direct hydrogen bonds		
B-subunit	Distance (Å)	B′-subunit
Oε-Glu82	2.9	Nε-Arg258
Oε-Glu82	2.6	Nη-Arg258
Nε-Arg258	2.9	Oε-Glu82
Nη-Arg258	2.6	Oε-Glu82

Region-B2		
Hydrogen bonds		
B-subunit	Distance (Å)	B′-subunit
Oγ-Ser174	3.2	Oε-Glu227

Oε-Glu227	3.2	Oγ-Ser174

Table 3-5-e-3. Interactions between two columns of the A-A′ and the B-B′ dimers in the *St*-Rtc*+ATP crystal (Region-AB1 and Region-AB2) (See Figure 3-5-e).

Region-AB1		
Direct hydrogen bonds		
A-subunit	Distance (Å)	B-subunit
N-Arg180	3.1	Oε-Glu72
N-Gly181	2.8	Oε-Glu72
Oε-Glu243	2.8	Nζ-Lys34
Oη-Tyr271	2.8	N-Met1
O-Lys329	3.2	Nε-Lys64

Region-AB2		
Direct hydrogen bonds		
B-subunit	Distance (Å)	A-subunit
Oε-Glu179	3.4	Oγ-Thr71
O-Lys216	3.4	Oε-Glu3
Oε-Glu243	2.8	Nζ-Lys34
Oη-Tyr271	2.9	N-Met1
O-Ile275	2.7	Nζ-Lys64

Table 3-5-e-4. Interactions between the A and the B subunits in the asymmetric unit of the *St*-Rtc*+ATP crystal (Region-AB3 and Region-AB4) (See Figure 3-5-e).

Region-AB3		
Hydrogen bonds		
A-subunit	Distance (Å)	B-subunit
Oδ-Asn38	3.2	Nζ-Lys255

Region-AB4		
Hydrogen bonds		
B-subunit	Distance (Å)	A-subunit
Oδ-Asn38	3.0	Nζ-Lys255

3-5-f. ATP binding site of *St*-Rtc*+ATP

The structure of *St*-Rtc in complex with ATP is shown in Figure 3-5-f-1. The substrate ATP is bound to the one end of the large domain so as to position the ribose at the center of six-helix assembly. The adenine ring contacts to α7 helix and the loop between β6 strand and α4 helix in the large domain, and to β13 strand of the small domain. The β- and γ- phosphates are contacts to α1 helix and the loop between β2 strand and α2 helix.

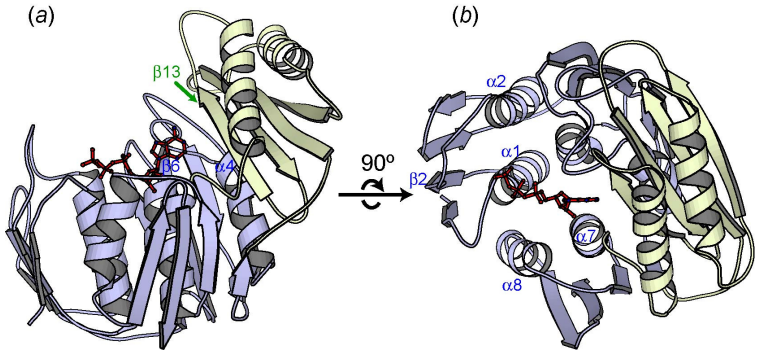


Figure 3-5-f-1. A ribbon drawing of the *St*-Rtc*+ATP A-subunit (*a*), and its top view (*b*). The structure of the B-subunit is omitted, because it is similar to that of the A-subunit. The large and small domains are colored in light purple and lemon, respectively. The substrate ATP molecule is drawn by red sticks. The colored characters indicate the code of the secondary structures referred to Figure 3-3-c-9.

Figure 3-5-f-2 shows the final $|F_o|$-$|F_c|$ omit maps of the bound ATPs. The phosphate groups and the ribose of ATP are partially disordered. The adenosine moiety is stacked between the side chains of Pro126 and His283 in both subunits. Figure 3-5-f-3~4 show the details of the stacked adenine bases. The carboxyl group of Asp250 interacts with the N1 atom of ATP though a hydrogen bond (Figure 3-5-f-4). The 2′- and 3′-hydroxyl groups of the ribose are bound to the carboxyl group of Asp286 though two hydrogen bonds (Figure 3-5-f-5).

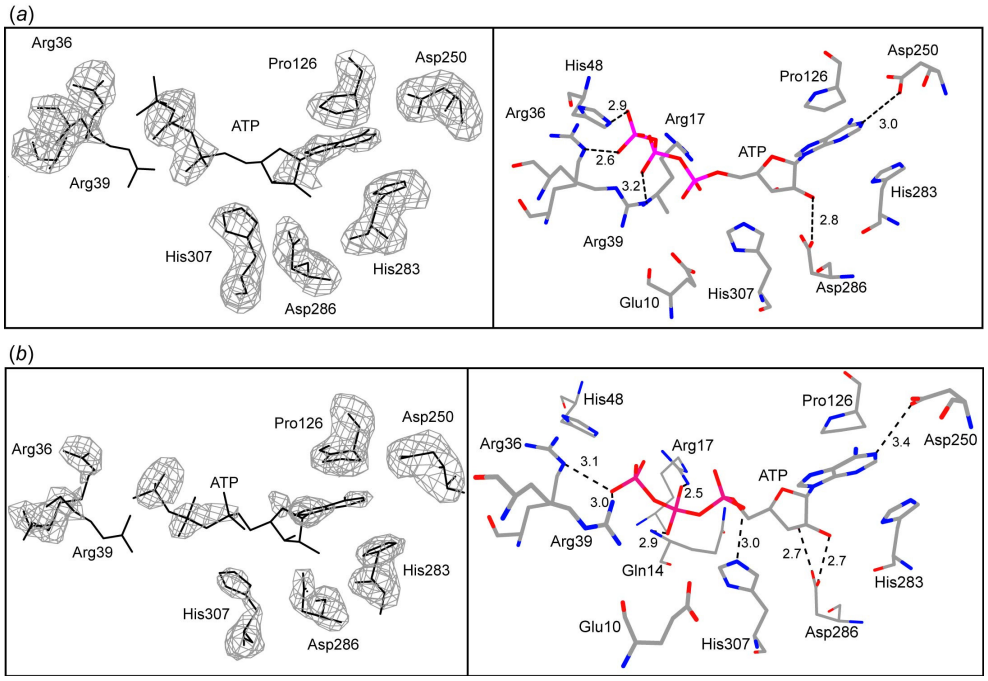


Figure 3-5-f-2. The ATP binding sites in the A-subunit (*a*) and in the B-subunit (*b*). The final $|F_o|$-$|F_c|$ omitted map of ATP and the surrounding residues is contoured at 2.5 σ level, are shown in the left column. The right column shows the possible hydrogen bonds between ATP and the surrounding residues. Carbon, nitrogen, oxygen, phosphorus and sulfur atoms are colored in gray, blue, red, pink and yellow, respectively.

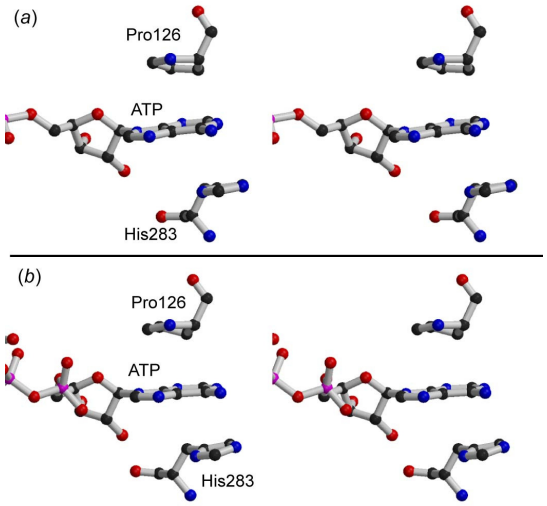


Figure 3-5-f-3. Stereo-pair drawings of ATP stacked between Pro126 and His283 of the A-subunit (*a*) and of the B-subunit (*b*).

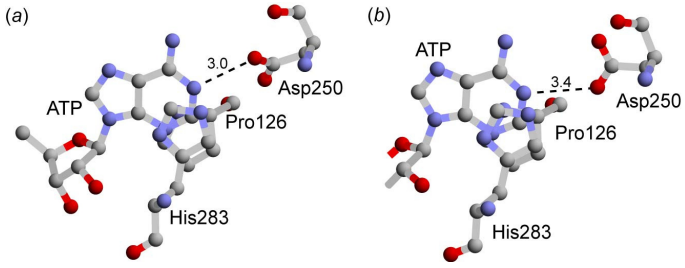


Figure 3-5-f-4. A hydrogen bond formation of the adenine rings with Asp250 of the A-subunit (*a*) and of the B-subunit (*b*).

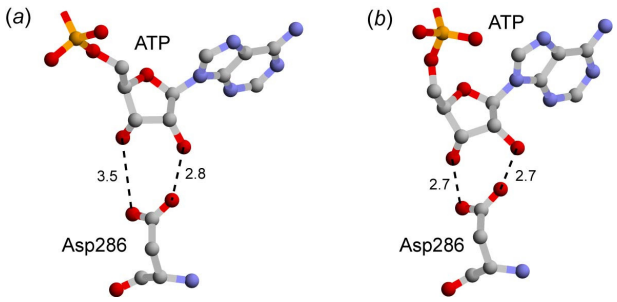


Figure 3-5-f-5. The double hydrogen bonds to trap ATP in the A-subunit (*a*) and in the B-subunit (*b*).

3-6. X-ray analysis of *St*-Rtc+AMP

3-6-a. Crystals of *St*-Rtc+AMP

Crystallization was performed to determine the interactions between the ATP

molecule and Mg^{2+} ion under acidic condition to suppress the reaction. The optimized condition, under which suitable crystals for X-ray analysis appeared after one week, is that the protein solution containing 5 mg/ml *St*-Rtc, 2mM $MgCl_2$ and 0.5 mM ATP in 0.02 M Tris-HCl (pH 8.0) buffer, and the reservoir solution (pH4.6) containing 2.0 M ammonium sulfate in 0.1 M sodium acetate (pH4.6) buffer were equally mixed into the droplet which was then equilibrated to the reservoir solution at 293K. The crystal indicated by an arrow in Figure 3-6-a has a size of approximately 100×100×200 μm. This crystal was used for X-ray diffraction experiment. As described in the latter careful discussion (section 3-6-f), the crystal structure has been successfully solved, and the final $|F_o|$-$|F_c|$ omit map of the ATP binding site suggests that the bound molecule is AMP, but not ATP. For this reason, the obtained crystal is named as *St*-Rtc+AMP to prevent confusion.

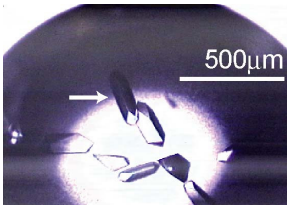


Figure 3-6-a. A photo of *St*-Rtc+AMP crystals. An arrow indicates the crystal used for X-ray diffraction experiment

3-6-b. X-ray data of the *St*-Rtc+AMP crystal

A diffraction image of the crystal is shown in Figure 3-6-b. Although the reflection spots were observed till 2.85 Å resolution, their intensity data were processed to 2.90 Å resolution. The crystal belongs to the space group $P3_1$ with cell constants $a = b = 84.4$ and $c = 108.4$ Å. The numbers of the observed and unique reflections are 110711 and 19249, respectively. The data completeness and an R_{marge} value are 100% and 8.1%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 3-6-b.

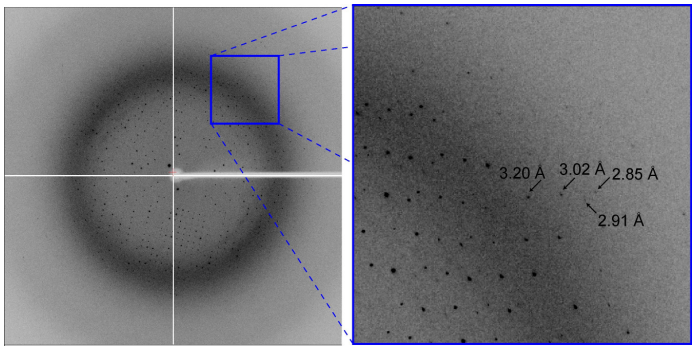


Figure 3-6-b. An X-ray diffraction image of the *St*-Rtc+AMP crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 3-6-b. A statistics of the observed intensity data and the crystallographic data of the *St*-Rtc+AMP crystal.

Wavelength (Å)	0.98
Resolution (Å)	50-2.90 (3.00-2.90)
Space group	$P3_1$
Unit cell (Å)	a=b=84.4, c=108.4
Observed reflections	110711
Unique reflections	19249
Completeness (%)	100 (100)
Redundancy	5.8 (5.7)
I/σ	30.1 (7.2)
R_{merge} (%)	8.1 (31.8)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

3-6-c. Phase angle determination and structural refinement of *St*-Rtc+AMP

An acceptable number of the protein molecules in asymmetric unit was estimated to be 2, the V_M value to be 2.97 Å^3/Da and the solvent content to be 58.6%. Despite the crystallization conditions are quite different from those of the *St*-Rtc*, *St*-Rtc and *St*-Rtc*+Mn crystals, these crystal habits are similar to them. The *St*-Rtc*, *St*-Rtc and *St*-Rtc*+Mn crystals show that two subunits in the asymmetric unit are associated to each other by a non-crystallographic two-fold symmetry to form a dimer. Therefore, molecular replacement method was applied using the dimer of *St*-Rtc* as a probe. The program *Phaser* was used for computations. A unique solution containing one dimer in the asymmetric unit has been obtained (Table 3-6-c-1). The crystal structure shows a reasonable dimer packing as described in the subsection 3-6-e.

Table 3-6-c-1. A unique solution from molecular replacement of a dimer of *St*-Rtc*.

Subunit	α	β	γ	x	y	z	RFZ	TFZ
1	79.8	162.4	163.7	0.0872	-0.3632	0.0001	37.4	37.5

*α, β, γ, Eulerian angles, x, y, z; Translations,
RFZ; Z-score (number of standard deviations above the mean value of peaks) of rotation function
TFZ; Z-score (number of standard deviations above the mean value of peaks) of translation function

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. In the asymmetric unit, 338 residues of the 339 residues per each subunit have been assigned except for the C-terminal one residue. 53 water molecules, two ATP molecules and two sulfate ions are assigned. The final R factor and R_{free} are 22.8% and 28.4%, respectively, indicating the structure was reasonably refines. A statistics of the refinement is given in Table 3-6-c-2. The final $2|F_o|-|F_c|$ electron density map, contoured at 1.5 σ level, is shown in Figure 3-6-c-1. The refined structure is well fitted on the map. A Ramachandran plot is shown in Figure 3-6-c-2. In the

Ramachandran plot, 487 (83.0%) residues are in the most favored, 86 (14.7%) in the allowed and 12 (2.0%) in generously allowed geometry, 12 (2.0%) in disallowed geometry, respectively. One sulfate ion (see Figure 3-6-c-3) interacts with the main chain of Gly85 and the side chains of Arg202 and Lys205 of the A-subunit. The second sulfate ion interacts with main chain of Lue183 of the A-subunit.

Table 3-6-c-2. A statistics of the structural refinement of *St*-Rtc+AMP.

Resolution (Å)	36.5-2.9
Used reflections	19217
R (%)	22.8
R_{free} (%)	28.4
r.m.s bond length (Å)	0.04
r.ms. angles (˚)	3.2
Protein atoms	5262
Water molecules	53
Adenine	2
Pyrophosphate	2
SO_4^{2-}	2

*R = 100 ×$\sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

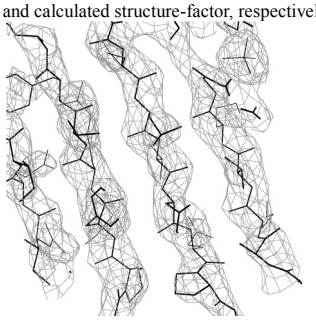


Figure 3-6-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *St*-Rtc+AMP crystal, contoured at 1.5 σ level (colored in gray).

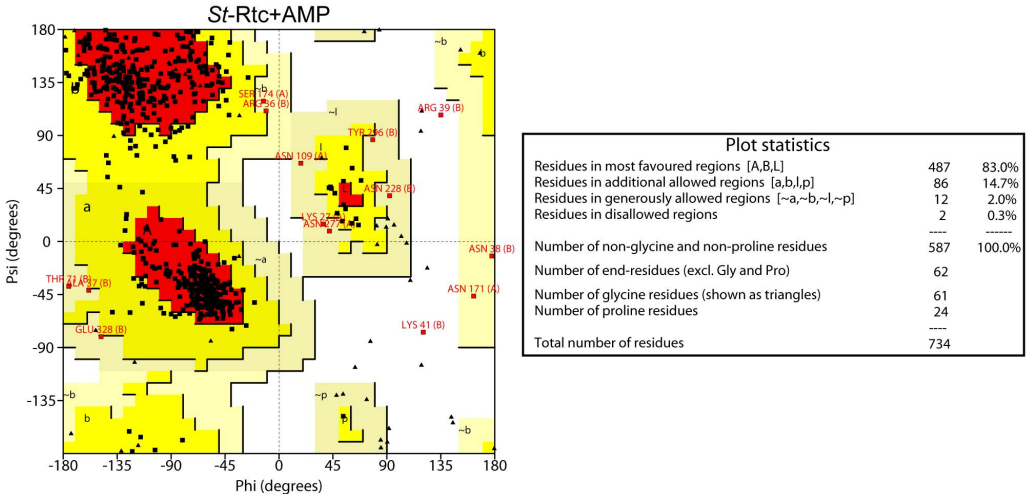


Figure 3-6-c-2. A Ramachandran plot of *St*-Rtc+AMP. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

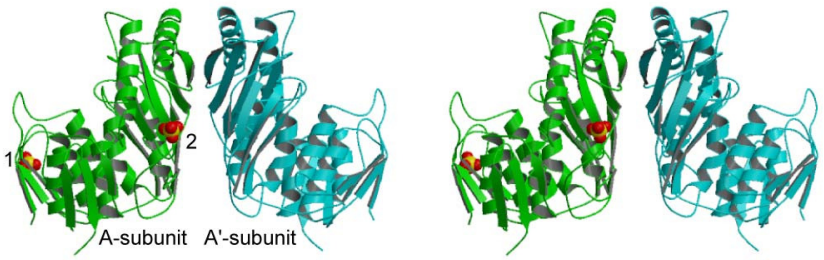


Figure 3-6-c-3. A stereo-pair drawing of the subunit structures with the bound two sulfate ions. They are indicated in ribbon models and yellow (sulfur atoms) and red (oxygen atoms) sphere models, respectively. The A-subunit and the A′-subunit are colored in green and cyan, respectively.

3-6-d. Dimer structure of *St*-Rtc+AMP

In the asymmetric unit, the two subunits, A and A′ are associated by a non-crystallographic two-fold symmetry to form a dimer. Superimposed Cα chains between the five dimers already found in the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn, the *St*-Rtc*+ATP and the *St*-Rtc+AMP crystals are shown in wire models in Figure 3-6-d-1. The dimer formation in the *St*-Rtc+AMP crystal is similar to those of the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn and the *St*-Rtc*+ATP crystals. The estimated r.m.s deviations of Cα atoms between them are given in Table 3-6-d-1. It is interesting to note that the dimer formations are still maintained even under the deferent crystallization conditions.

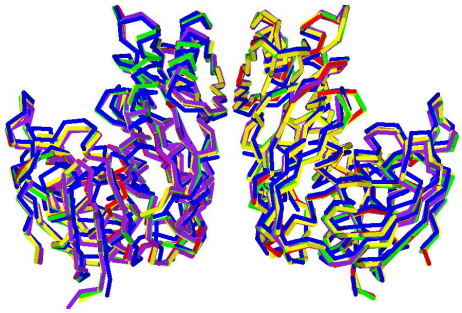


Figure 3-6-d-1. A superimposition of Cα chains of and the dimers in the *St*-Rtc* (green), the *St*-Rtc (red), *St*-Rtc*+Mn (purple) and the *St*-Rtc*+ATP (blue) crystals onto that in the *St*-Rtc+AMP crystal (yellow).

Table 3-6-d-1. The estimated r.m.s.d (Å) of Cα atoms between the dimers of the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn, *St*-Rtc*+ATP and *St*-Rtc+AMP the crystals.

		St-Rtc*	*St*-Rtc	*St*-Rtc*+Mn	*St*-Rtc*+ATP	
		A-A′ dimer	A-A′ dimer	A-A′ dimer	A-A′ dimer	B-B′ dimer
St-Rtc+AMP	A-A′ dimer	0.52	0.43	0.63	1.52	1.56
	A′-A dimer	0.94	0.92	1.01		

The two subunits in the asymmetric unit are shown in Figure 3-6-d-2. When they

are superimposed on each other, the estimated r.m.s deviation between the corresponding Cα atoms is 0.12 Å, suggested that their folding are similar to each other, as shown in Figure 3-6-d-3. The dimerization between the two subunits interface is stabilized by forming a β-sheet through hydrogen bonds between the two main chains, Ile223 and Ile223′. An apostrophe indicates the residue of the counter subunit. The atomic distances given in Table 3-6-d-2 show that there are no abnormal contacts between the two subunits.

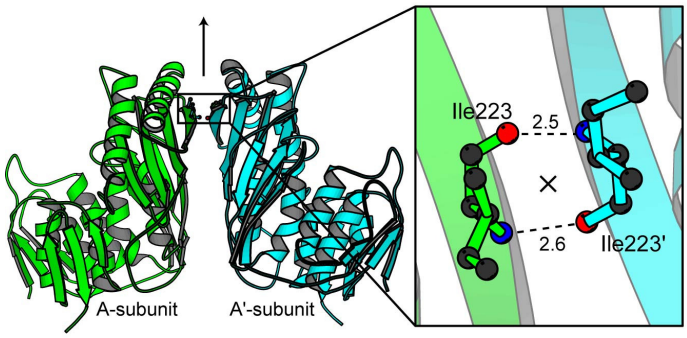


Figure 3-6-d-2. The dimer in the asymmetric unit of the *St*-Rtc+AMP crystal. The two subunits are associated to each other to form a dimer by a non-crystallographic two-fold symmetry, shown with an arrow and a symbol x. The residue indicated with an apostrophe is the related counter one. The right figure shows a zoomed up view at the Ile223 residues. Black, blue and red spheres indicate carbon, nitrogen and oxygen atoms respectively. Broken lines indicate the hydrogen bonds with the distances (Å) between the two β-strands.

Table 3-6-d-2. Interactions between the two subunits to form the dimer in the asymmetric unit of the *St*-Rtc+AMP crystal.

Between the small domains		
Direct hydrogen bonds		
A-subunit	Distance (Å)	A′-subunit
N-Ile223	2.6	O-Ile223
O-Ile223	2.5	N-Ile223
Oε-Glu227	2.8	Nζ-Lys208
Oε-Glu201	3.1	Oγ-Ser197
Nζ-Lys208	3.0	Oε-Glu227

Between the small domain and the large domain of the opposite molecule		
Direct hydrogen bonds		
A-subunit (Small domain)	Distance (Å)	A′-subunit (Large domain)
Oε-Glu220	2.7	Nη-Arg132
Hydrophobic contacts		
A-subunit (Small domain)		A′-subunit (Large domain)
Ile184 Met185 Pro218		Ile149 Ile151
Direct hydrogen bonds		
A-subunit (Large domain)	Distance (Å)	A′-subunit (Small domain)
Oε-Glu220	3.1	Nη-Arg132
Hydrophobic contacts		
A-subunit (Large domain)		A′-subunit (Small domain)
Ile184 Met185		Ile149 Ile151

Pro218	

Between the large domains		
Direct hydrogen bonds		
A-subunit (Large domain)	Distance (Å)	A′-subunit (Large domain)
Nη-Arg145	2.8	Oε-Glu140

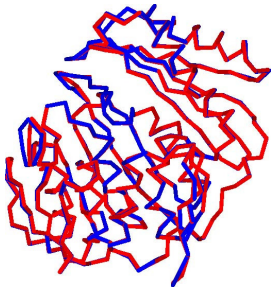


Figure 3-6-d-3. A superimposition of the Cα chains between the two subunits of the dimer in the *St*-Rtc+AMP crystal. The A-subunit and A′-subunit are indicated by wire models with blue and red colors, respectively.

3-6-e. Dimer packing in the *St*-Rtc+AMP crystal

The dimers are in contact with a head-and-tail manner around the crystallographic 3_1 axis, so that a screwed column is formed by stacking the A-subunit on the A′-subunit of the subsequent dimer. Along with the *a* and *b* axes, the columns are interacted to the adjacent columns. The atomic distances in the contact regions along the *c* axis are listed in Table 3-6-e-1, and those along the *a* and *b* axes are listed in Table 3-6-e-2. As seen in Figures 3-2-e-2, 3-4-e-2 and 3-6-e-2, the dimers are similarly packed.

Table 3-6-e-1. Interactions of the *St*-Rtc+AMP dimmers related by the crystallographic 3_1 symmetry (Region-1) (see Figure 3-6-e-1).

Region-1		
Direct hydrogen bonds		
A-subunit	Distance (Å)	A′-subunit
Oε-Glu179	2.8	Oγ-Ser176
Oε-Gln324	3.1	Nδ-Asn277
Oδ-Asn277	3.1	Nη-Arg334
O-Ser276	2.5	Oγ-Ser326
O-Lys329	3.3	N-Gly327
Oγ-Ser176	2.9	Oε-Gln179
	3.3	Oε-Gln179
Nε-Gln324	3.2	O-Ile275
Hydrophobic contacts		
A-subunit	A′-subunit	
Met178 Ile300 Ile332 Pro330	Met178 Ile300 Ile332 Pro330	

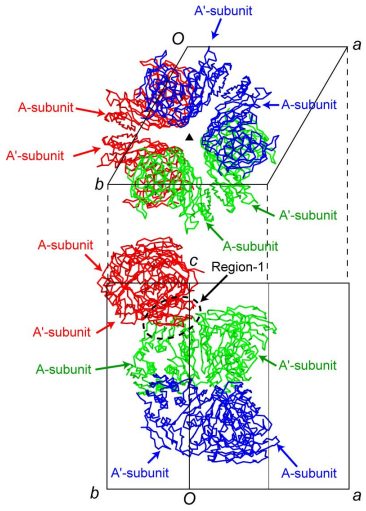


Figure 3-6-e-1. The *St*-Rtc+AMP dimers around the crystallographic three-fold screw axis, viewed along the *c*-axis (upper) and along the *b**-axis (bottom)

Table 3-6-e-2. Interactions of the *St*-Rtc+AMP dimers between the two screwed columns related by a periodic shift along the *a* axis (region-2) and along the *b* axis (Region-3) (See Figure 3-6-e-2).

Region-2				
Direct hydrogen bond				
A-subunit	Distance (Å)		A-subunit	
O-Pro40	2.8		Nζ-Lys214	
O-Val68	2.6		Oε-Glu264	
Hydrogen bond interactions intermediated by a sulfate ion				
A-subunit	Distance (Å)	Ion	Distance (Å)	A-subunit
N-Gly65	3.2	SO_4^{2-}	3.0	Nζ-Lys205
			3.3	Nε-Arg202
			2.7	Nη-Arg202

Region-3		
Hydrogen bond interactions intermediated by a water molecule		
A′-subunit	Distance (Å)	A-subunit
Oε-Gln45	2.9	Oδ-Asn228
Nε-Gln45	2.6	

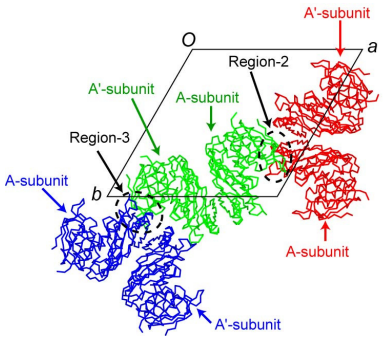


Figure 3-6-e-2. Interacting regions of the dimers between the two screwed columns related by a periodic shift along the *a* axis (Region-2) and *b* axis (Region-3).

3-6-f. AMP binding site of *St*-Rtc+AMP

The final $|F_o|$-$|F_c|$ omit map of the ATP binding site is shown in Figure3-6-f-1. The β- and γ-phosphate groups are easily assigned. However, there are no density for α-phosphate group. Except this phosphate, the remaining parts of AMP are visible on the map. There are two possibilities to explain this situation. One is simply that ATP is hydrolyzed in acidic condition into two parts, pyrophosphates and AMP, and the two components are bound independently to the ATP binding site. Only the α-phosphate group of AMP is allowed to be disordered. The other in that once ATP is reacted to produce an AMP-His307 covalent intermediate, then AMP is hydrolyzed in acidic condition. These formations are commonly observed in both subunits.

The adenine moiety is stacked between the side chains of Pro126 and His283 in both subunits, as shown in Figure 3-6-f-2 and in Figure 3-6-f-3. The binding mode is similar to those found in the *St*-Rtc*+ATP crystal (see Figure 3-5-f-4). The carboxyl group of Asp250 forms a hydrogen bond with the N1 atom of AMP (Figure 3-6-f-3). The two hydroxyl groups of the ribose are bound to the carboxyl group of Asp286 through double hydrogen bonds (Figure 3-6-f-4). These hydrogen bond formations are also observed in the *St*-Rtc*+ATP crystal (see Figure 3-5-f-4~5).

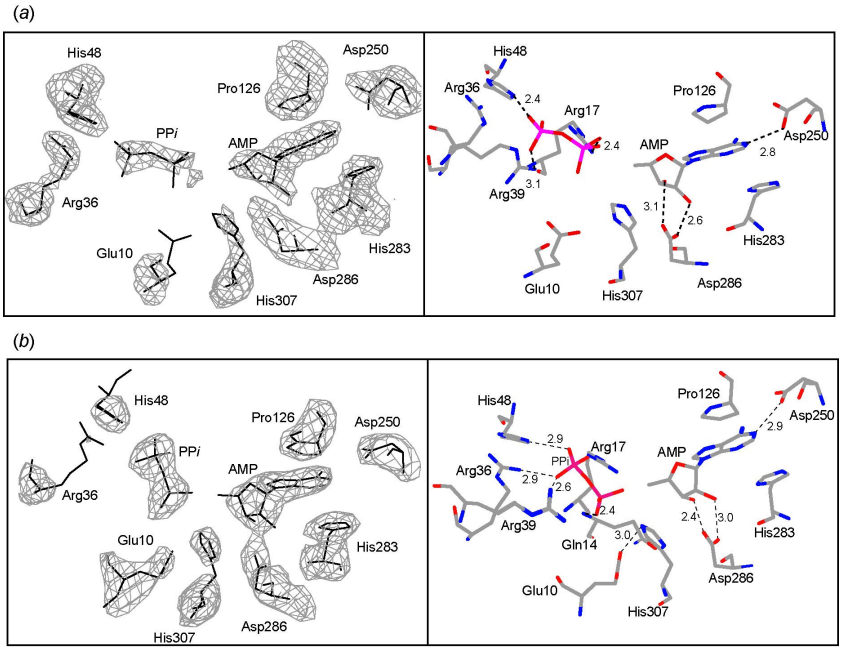


Figure 3-6-f-1. The AMP binding sites of the A-subunit (*a*) and the A′-subunit (*b*). The final $|F_o|$-$|F_c|$ omit map of AMP and the surrounding residues contoured at 2.5 σ level, in colored in gray in the left column. The right column shows possible hydrogen bonds between AMP and the surrounding residues. Carbon, nitrogen, oxygen, phosphorus and sulfur atoms are colored in gray, blue, red, pink and yellow, respectively.

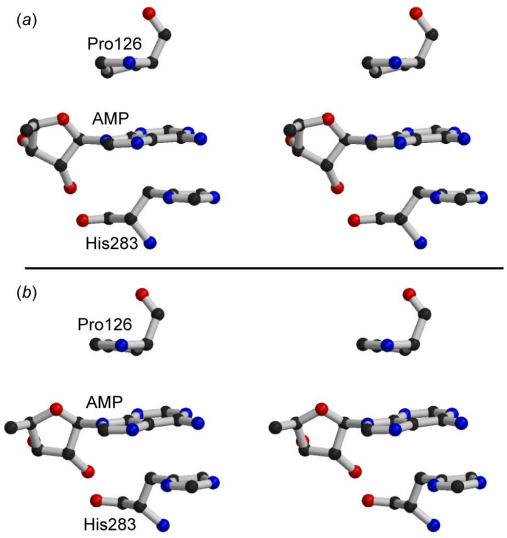


Figure 3-6-f-2. A stereo-pair drawing of AMP trapped between Pro126 and His283 of the A-subunit (*a*) and of the A′-subunit (*b*).

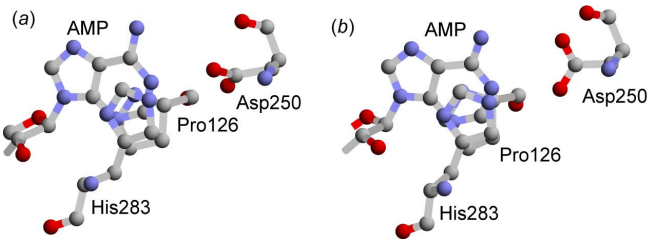


Figure 3-6-f-3. Adenine base hydrogen-bonded to Asp250 in the A-subunit (*a*) and in the A′-subunit (*b*).

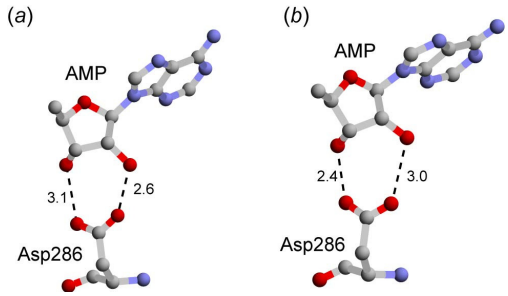


Figure 3-6-f-4. The two hydroxyl group of the ribose hydrogen-bonded to the carboxyl group of Asp286 in the A-subunit (*a*) and in the A′-subunit (*b*).

3-7. X-ray analysis of *St*-Rtc*-AMP

3-7-a. Crystals *St*-Rtc*-AMP

Crystallization of *St*-Rtc*-AMP was performed to determine the structure after the first step reaction. The preparation of the protein solution mixing with *St*-Rtc*, Mg^{2+} ion and ATP under basic condition near optimal pH of the cyclization, was preformed 30 minutes before crystallization. The optimized condition, under which suitable crystals

appeared after one week, is that the protein solution containing the 10 mg/ml *St*-Rtc*, 2mM $MgCl_2$ and 0.5 mM ATP in 0.02 M Tris-HCl (pH8.0) buffer and reservoir solution (pH8.2) containing 2.0 M ammonium sulfate in 0.1 M Tris-HCl (pH8.0) buffer were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293K. The crystal is indicated by an arrow in Figure 3-7-a. The size of the rod crystal was approximately 50×50×400 μm. This crystal was used for X-ray diffraction experiment.

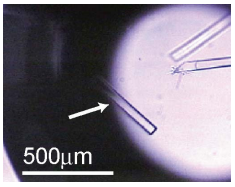


Figure 3-7-a. A photo of the *St*-Rtc*-AMP crystals. An arrow indicates the crystal used for X-ray diffraction experiment

3-7-b. X-ray data of the *St*-Rtc*-AMP crystal

A diffraction image of the *St*-Rtc*-AMP crystal is shown in Figure 3-7-b. Although the reflection spots were observed till 2.12 Å resolution, their intensity data were processed to 2.25 Å resolution. The crystal belongs to the space group $P4_2$ with cell constants $a = b = 90.8$ and $c = 90.5$ Å. The numbers of the observed and unique reflections are 265661 and 34783, respectively. The data completeness and an R_{marge} value are 99.7% and 8.0%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 3-7-b.

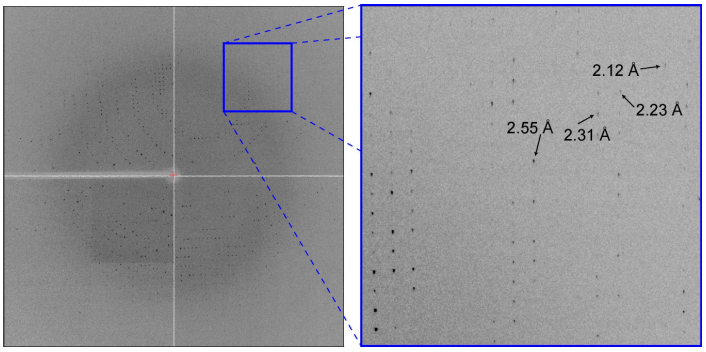


Figure 3-7-b. An X-ray diffraction image of the *St*-Rtc*-AMP crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 3-7-b. A statistics of the observed intensity data and the crystallographic data of the *St*-Rtc*-AMP crystal.

Wavelength (Å)	1.00
Resolution (Å)	50-2.25 (2.33-2.25)
Space group	$P4_2$
Unit cell (Å)	a=b=90.8, c=90.3
Observed reflections	265661
Unique reflections	34783
Completeness	99.7 (99.1)
Redundancy	7.6 (7.4)
I/σ	23.1 (2.8)
R_{merge} (%)	8.0 (28.5)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i|I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

3-7-c. Phase angle determination and structural refinement of *St*-Rtc*-AMP

An acceptable number of the protein molecules in asymmetric unit was estimated to be 2, the V_M value to be 2.48 Å^3/Da and the solvent content to be 50.4 %. Despite the crystallization conditions are different from those for crystals, the crystal habits are similar to those of the *St*-Rtc*+ATP crystal. In the latter crystal, the two subunits are independently associated to from the dimers with the respective subunits related by a crystallographic two-fold symmetry. To conform this, the molecular replacement method was applied using one subunit of *St*-Rtc* as a probe. The program *Phaser* was used for computations. A unique solution containing two molecules in the asymmetric unit has been obtained (Table 3-7-c-1). As expected, the arrangement of the two subunits in the symmetric unit is similar to that of the *St*-Rtc*+ATP crystal. Each subunit is reasonably packed in the unit cell as described in subsection 3-7-e.

Table 3-7-c-1. A unique solution from molecular replacement of a subunit of *St*-Rtc*.

Subunit	α	β	γ	x	y	z	RFZ	TFZ
1	61.2	87.4	232.8	0.159	-0.451	0.0543	20.0	48.2
2	28.8	92.6	52.8	0.0486	-0.341	0.00275	18.0	26.1

*α, β, γ, Eulerian angles, x, y, z; Translations,
RFZ; *Z*-score (number of standard deviations above the mean value of peaks) of rotation funciton
TFZ; *Z*-score (number of standard deviations above the mean value of peaks) of translation function

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. In the asymmetric unit, 338 residues of the 339 residue pre each molecule have been assigned except for the C-terminal one residue. 452 water molecules, two AMP, six sulfate ions are assigned. The final R factor and R_{free} are 18.1% and 24.5%, respectively, indicating the structure was reasonably refined. A statistics of the refinement is given in Table 3-7-c-2. The final $2|F_o|-|F_c|$ electron density map, contoured at 1.5 σ level is shown in Figure 3-7-c-1. The refined structures are well fitted on the map. A Ramachandran plot is shown in and Figure 3-7-c-2. In the

Ramachandran plot, 524 (89.1%) residues in the asymmetric unit are in the most favored geometry, 61 (10.4%) in the allowed geometry and 3 (0.5%) in generously allowed geometry, respectively. There is no residue in disallowed geometry. Two sulfate ions are bound near the AMP binding sites (see Figure 3-7-c-3). The sulfate ions 1 and 2 interact with the side chains of Arg36 and Arg39 and the main chain of Gly13 of the A- and B-subunits. The sulfate ions 3 and 4 interact with the main chain of Asp5 of the A- and B-subunits. The phosphate 4 and 5 interact with the main chain of Ala259 of the A- and B-subunit. Stereo-pair drawings of the dimers and the ions are shown in Figure 3-7-c-3.

Table 3-7-c-2. A statistics of the structural refinement of *St*-Rtc*-AMP

Resolution (Å)	40.4-2.25
Used reflections	34417
R (%)	18.1
R_{free} (%)	24.5
r.m.s bond length (Å)	0.02
r.ms. angles (˚)	2.0
Protein atoms	5262
Water molecules	452
AMP	2
SO_4^{2-}	6

*$R = 100 \times \sum||F_o|-|F_c|| / \sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

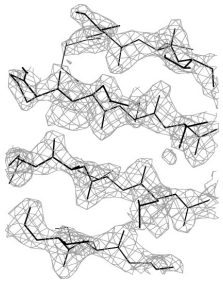


Figure 3-7-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *St*-Rtc*-AMP crystal, contoured at 1.5 σ level (colored in gray).

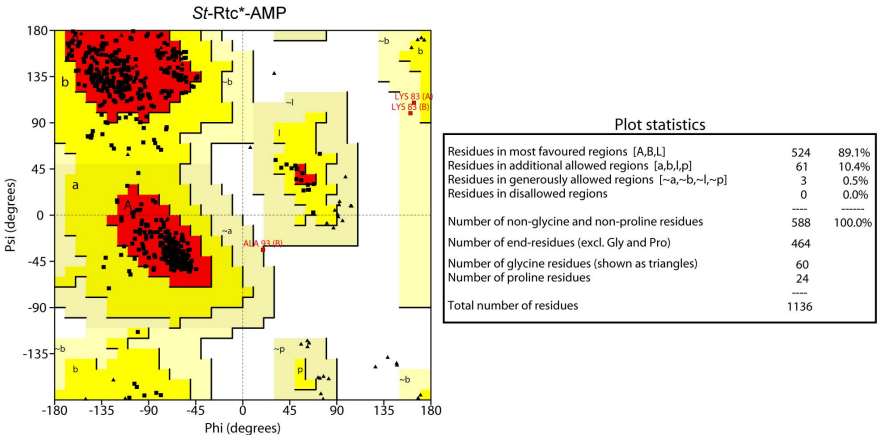


Figure 3-7-c-2. A Ramachandran plot of *St*-Rtc*-AMP. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

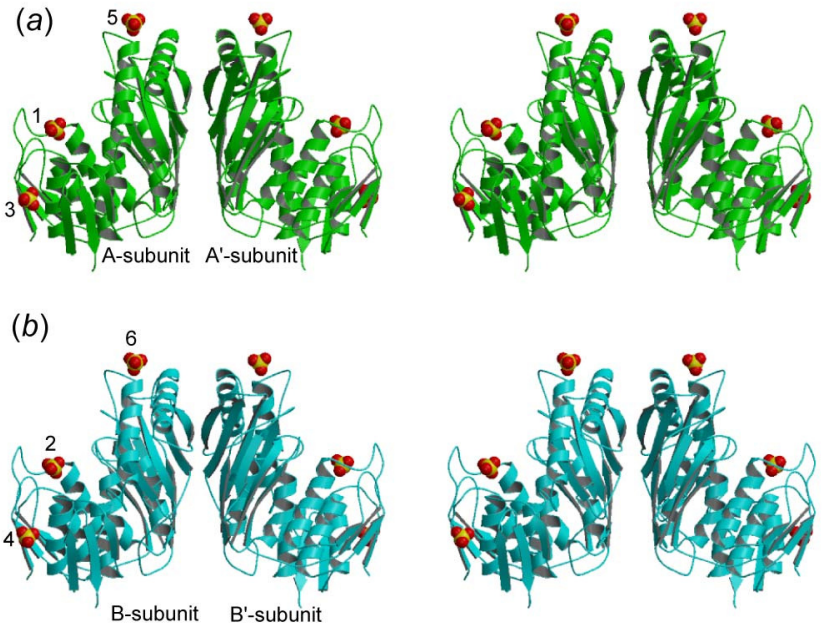


Figure 3-7-c-3. Stereo-pair drawings of the A-A′ (*a*) and B-B′ dimers (*b*) with six sulfate ions, indicated in ribbon models and yellow (sulfur atoms) and red (oxygen atoms) sphere models, respectively. The A-A′ and B-B′ dimers are colored in green and cyan, respectively.

3-7-d. Dimer structure of *St*-Rtc*-AMP

As like in the *St*-Rtc*+ATP crystal, the two subunits, A and B, in the *St*-Rtc*-AMP crystal are separately form dimers with A′ and B′ related by the crystallographic two-fold symmetries passing the origin and passing the center of the *a*-*b* plane, respectively (see Figure 3-7-e in the next subsection). The dimer structures are similar

to those found in the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn, the *St*-Rtc*+ATP and the *St*-Rtc+AMP crystals. Every dimer is superimposed to identify the similarity as shown in Figure 3-7-d-1. The r.m.s.d values indicate that the dimer structures in the *St*-Rtc*-AMP crystal closer to those in the *St*-Rtc*+ATP crystal. This similarity seems to be ascribed to the crystallization conditions, because they are basic, while the crystallization conditions for *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn and the *St*-Rtc+AMP crystals are acidic.

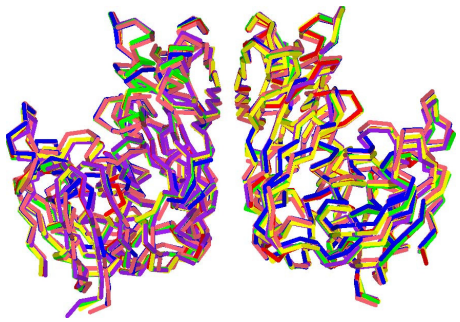


Figure 3-7-d-1. A superimposition of Cα chains between the dimers in the *St*-Rtc* (green), the *St*-Rtc (red), the *St*-Rtc*+Mn (purple), the *St*-Rtc*+ATP (blue), the *St*-Rtc+AMP (yellow) and the *St*-Rtc*-AMP crystals (pink).

Table 3-7-d-1. The estimated r.m.s.d (Å) of Cα atoms between the dimers in the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn, *St*-Rtc*+ATP, *St*-Rtc+AMP and *St*-Rtc*-AMP crystals.

		St-Rtc*	*St*-Rtc	*St*-Rtc*+Mn	*St*-Rtc+ATP		*St*-Rtc+AMP
		A-A′ dimer	A-A′ dimer	A-A′ dimer	A-A′ dimer	B-B′ dimer	A-A′ dimer
St-Rtc*-AMP	A-A′ dimer	1.54	1.54	1.73	0.20	0.20	1.48
	B-B′ dimer	1.55	1.55	1.73	0.20	0.20	1.49

The dimerization interfaces in the *St*-Rtc*-AMP crystal are stabilized by forming a β-sheet through hydrogen bonds between the two main chains, Ile223 and Ile223′, as shown in Figure 3-7-d-2, similar to those in the *St*-Rtc*, the *St*-Rtc, the *St*-Rtc*+Mn, *St*-Rtc*+ATP and *St*-Rtc+AMP crystals. An apostrophe indicates the residue of the counter subunit. Hydrogen bond interactions between the small domains of the A-A′ and of the B-B′ dimers are given in Table 3-7-d-2 and Table 3-7-d-3. The hydrogen bonds, Oε (Glu227) ⋯ Nζ (Lys205) and Oε (Glu201) ⋯ Oγ (Ser197), are commonly observed in both dimers. The atomic distances between the two subunits of the A-A′ dimer are given in Table 3-7-d-2 and those between the two subunits of B-B′ dimers are given in Table 3-7-d-3.

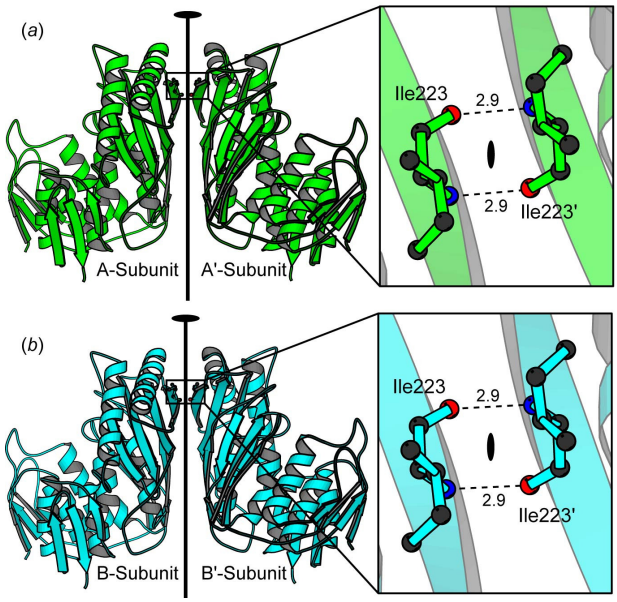


Figure 3-7-d-2. The A-A′ (*a*) and the B-B′ dimers (*b*) in the *St*-Rtc*-AMP crystal. A-subunit is associated to the A′-subunit, related by a crystallographic two-fold symmetry shown with a thick line and an elliptical symbol. The B-B′ dimer is also formed by another crystallographic two-fold symmetry. The residue indicated with an apostrophe is the related counter one. The right figure shows a zoomed up view at the Ile223 residues. Black, blue and red spheres indicate carbon, nitrogen and oxygen atoms respectively. Broken lines indicate the hydrogen bonds with the distances (Å) between the two β-strands.

Table 3-7-d-2. Interactions between the A and A′ subunits, related by a two-fold crystallographic symmetry passing the center of *a*-*b* plane of the unit cell (see Figure 3-7-e).

Between Chain two A-subunits		
Between the small domains		
Direct hydrogen bonds		
A-subunits (Small domain)	Distance (Å)	A′-subunits (Small domain)
O-Ile 223	2.8	N-Ile223
Oδ-Asp224	3.1	Nη-Arg226
Oε-Glu227	2.8	Nζ-Lys205
Oε-Glu227	3.4	
Oε-Glu201	3.4	Oγ-Ser197
Oε-Glu201	2.9	
N-Ile223	2.8	O-Ile 223
Nη-Arg226	3.1	Oδ-Asp224
Nζ-Lys205	2.8	Oε-Glu227
	3.4	Oε-Glu227
Oγ-Ser197	3.4	Oε-Glu201
	2.9	Oε-Glu201

Between the small domain, and the large domain of the opposite chain.				
Hydrogen bond interactions intermediated by a water molecule				
A-subunits (Small domain)	Distance(Å)	Water	Distance (Å)	A′-subunits (Large domain)
N-Leu150	3.0	W136	2.7	Oε-Glu220
Oε-Glu140	2.9	W264	2.9	Oε-Glu140
Oε-Glu220	2.7	W136	3.0	N-Leu150

Salt bridge		
A-subunits (Small domain)	Distance(Å)	A′-subunits (Large domain)
Oε-Glu220	2.8	Nη-Arg132
Oε-Glu220	2.8	Nη-Arg132
Nη-Arg132	2.8	Oε-Glu220
Nη-Arg132	2.8	Oε-Glu220

Hydrophobic contacts	
A-subunits (Small domain)	A′-subunits (Large domain)
Ile184 Met185 Pro218	Ile149 Ile150
Ile149 Ile150	Ile184 Met185 Pro218

Table 3-7-d-3. Interactions between the B and B′ subunits, related by a two-fold crystallographic symmetry passing the origin of the unit cell (see Figure 3-7-e).

Between two B-subunits				
Between the small domains				
Direct hydrogen bonds				
B-subunit (Small domain)	Distance (Å)		B′-subunit (Small domain)	
O-Ile223	2.9		N-Ile223	
Oε-Glu227	2.8		Nζ-Lys205	
Oε-Glu227	3.4			
Oε-Glu201	3.4		Oγ-Ser197	
Oε-Glu201	3.1			
Oε-Glu220	3.0		Nη-Arg226	
N-Ile223	2.9		O-Ile223	
Nζ-Lys205	2.8		Oε-Glu227	
	3.4		Oε-Glu227	
Oγ-Ser197	3.4		Oε-Glu201	
	3.1		Oε-Glu201	
Nη-Arg226	3.0		Oε-Glu220	
Hydrogen bond interactions intermediated by a water molecule				
B-subunits (Small domain)	Distance(Å)	Water	Distance(Å)	B′-subunits (Small domain)
Oε-Glu201	2.7	W194	3.1	Oε-Glu227
O-Ile274	2.8	W236	2.8	N-Val225
Oε-Glu227	3.1	W194	2.7	Oε-Glu201
N-Val225	2.8	W236	2.8	O-Ile274
Between the small domain, and the large domain of the opposite chain.				
Hydrogen bond interactions intermediated by a water molecule				
B-subunits (Small domain)	Distance(Å)	Water	Distance (Å)	B′-subunits (Large domain)
Oε-Glu220	2.7	W300	3.1	N-Ile150
O-Phe242	3.0	W29	3.3	Oδ-Asn147
N-Ile150	3.1	W300	2.7	Oε-Glu220
Oδ-Asn147	3.3	W29	3.0	O-Phe242
Salt bridge				
B-subunits (Small domain)	Distance (Å)		B′-subunits (Large domain)	
Oε-Glu220	2.8		Nη-Arg132	
Oε-Glu220	2.7		Nη-Arg132	
Nη-Arg132	2.8		Oε-Glu220	
Nη-Arg132	2.7		Oε-Glu220	
Hydrophobic contacts				

B-subunits (Small domain)	B-subunits (Large domain)
Ile184 Met185 Pro218	Ile149 Ile151
Ile149 Ile151	Ile184 Met185 Pro218

Between the large domains				
Hydrogen bond interactions intermediated by a water molecule				
B-subunits	Distance(Å)	Water	Distance (Å)	B′-subunits
Oε-Glu140	2.8	W260	2.8	Oε-Glu140
Nζ-Lys141	2.9	W29	3.3	Oδ-Asn147
Oδ-Asn147	3.3	W29	2.9	Nζ-Lys141

3-7-e. Dimer packing in the *St*-Rtc*-AMP crystal

Figure 3-7-e shows the dimers in the *St*-Rtc*-AMP crystal. The two columns composed of the A-A′ dimers and composed of B-B′ dimers are arranged in an opposite direction to each other to deny the helical polarity. In each column, the dimers are stacked on each other on the *c* axis by rotating 90º successively. This packing is similar to that in the *St*-Rtc*+ATP crystal (see Figure 3-5-e). The dimer-dimer interaction regions in the A-A′ dimer column and those in the B-B′ dimer column are referred to Region-A1, Region-A2, Region-B1 and Region-B2, respectively (see Figure 3-7-e). The atomic distances of these interactions are listed in Table 3-7-e-1 and Table 3-7-e-2.

The two interaction regions between the adjacent columns are referred to Region-AB1 and Region-AB2 (see Figure 3-7-e). The three hydrogen bonds, Oε (Glu243) ⋯ Nζ (Lys34), Oη (Tyr271) ⋯ N (Met1) and N (Arg180) ⋯ Oε (Glu72) are found in both Region-AB1 and Region-AB2. The atomic distances of these interactions are given in Table 3-7-e-3.

The A and B subunits are in contact to each other in Region-AB3 and Region-AB4 in the asymmetric unit. In these regions, hydrogen bonds are not found.

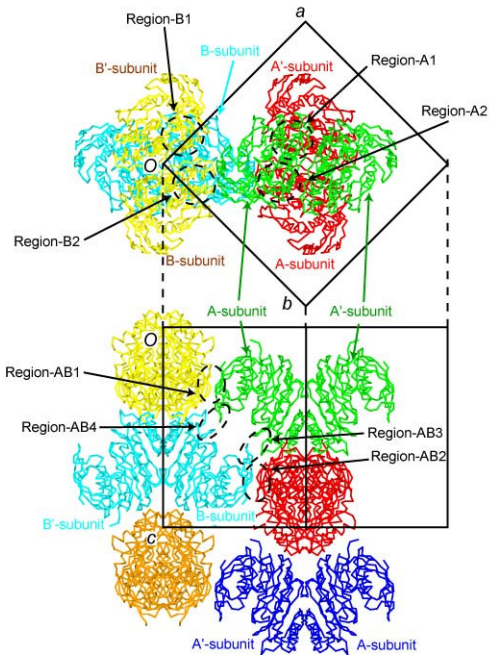


Figure 3-7-e. The dimers in the *St*-Rtc*-AMP unit cell, viewed along the *c* axis (upper) and along the (-1 1 0) direction (bottom). Dimers of the A-A′ subunits indicated by wire models are colored in yellow, cyan and orange. Those of the B-B′ subunits are colored in green, red and blue.

Table 3-7-e-1. Interactions between A and A′ subunits related by a periodic shift along the *c* axis (Region-A1 and Region-A2) (See Figure 3-7-e).

Region-A1				
Direct hydrogen bonds				
A-subunit	Distance (Å)		A′-subunit	
Oδ-Asn171	3.2		Oε-Glu201	
Oγ-Ser173	3.4		Nζ-Lys205	
O-Pro172	3.0			
Salt bridge				
A-subunit	Distance (Å)		A′-subunit	
Oε-Glu82	2.9		Nε-Arg254	
Oε-Glu82	2.8		Nη-Arg258	
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance(Å)	Water	Distance (Å)	A′-subunit
Nζ-Lys168	2.4	W432	3.2	Nε-His198

Region-A2		
Direct hydrogen bonds		
A-subunit	Distance (Å)	A′-subunit
O-Tyr229	3.1	Nδ-Asn228
Oγ-Ser174	2.8	Oδ-Glu227

Table 3-7-e-2. Interactions between the B and B′ subunits related by a periodic shift along the *c* axis (Region BB′-1 and Region BB′-2) (See Figure 3-7-e).

Region-B1		
Direct hydrogen bonds		
B-subunit	Distance (Å)	B′-subunit
Oε-Glu82	2.9	Nε-Arg258

Oε-Glu82	2.5	Nη-Arg258
Oδ-Asn171	3.2	Nζ-Lys205

Region-B2		
Direct hydrogen bonds		
B-subunit	Distance (Å)	B′-subunit
Oγ-Ser173	3.4	Oε-Glu227
Oγ-Ser174	3.0	O-Glu227
O-Tyr296	3.2	Nδ -Asn228

Table 3-7-e-3. Interactions between the two subunits of two columns with the A-A′ and the B-B′ dimers in the *St*-Rtc*+ATP crystal along the diagonal line of the γ angle (Region-AB1 and Region-AB2) (See Figure 3-7-e-1).

Region-AB1		
Direct hydrogen bonds		
B-subunit	Distance (Å)	A-subunit
N-Gly181	2.8	Oε-Glu72
N-Arg180	3.0	
	3.1	Oε-Glu72
Oε-Glu243	2.8	Nζ-Lys34
Oη-Tyr271	2.8	Met1-N
O-Lys329	3.4	Nζ-Lys64
Oε-Glu179	3.3	Oγ-Thr71
Nδ-Asn227	3.1	O-Glu72
N-Leu183	3.4	Oε-Glu3

Region-AB2				
Direct hydrogen bonds				
A-subunit	Distance (Å)		B-subunit	
Oε-Glu179	2.7		Oγ-Thr71	
N-Arg180	3.1		Oε-Glu72	
N-Ile183	3.4		Oε-Glu3	
O-Lys214	3.2		N-Met1	
Nδ-Asn227	3.2		O-Glu72	
Oε-Glu243	3.1		Nζ-Lys34	
Oη-Tyr271	3.0		N-Met1	
O-Ile275	3.3		Nζ-Lys64	
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance (Å)	Water	Distance (Å)	B-subunit
O-Gly181	2.5	W244	3.1	Oε-Glu72
O-Ile275	2.7	W397	3.4	OεG-lu76

3-7-f. AMP binding site of *St*-Rtc*-AMP.

As seen in Figure 3-7-f-1, the final $|F_o|$-$|F_c|$ omit map of AMP and the surrounding residues definitely shows that the α-phosphate group of AMP is covalently bound to the Nε atom of His307. This feature is commonly found in both dimers. The adenine moiety is stacked between the side chains of Pro126 and His283 in both subunits, as shown in Figure 3-7-f-2~3. This binding is similar to those of AMP in the *St*-Rtc+AMP and ATP in the *St*-Rtc*+ATP crystals (see also Figures 3-5-f-3 and 3-6-f-2). The carboxyl group of Asp250 interacts with the N1 atom of the AMP through a hydrogen bond (Figure 3-7-f-3). The 2′- and 3′- hydroxyl groups of the ribose are bound to the carboxyl group

of Asp286 through double hydrogen bonds (Figure 3-6-f-4). These hydrogen bond interactions are also found in the *St*-Rtc*+ATP and the *St*-Rtc+AMP crystals (see Figures 3-5-f-4~5 and 3-6-f-3~4).

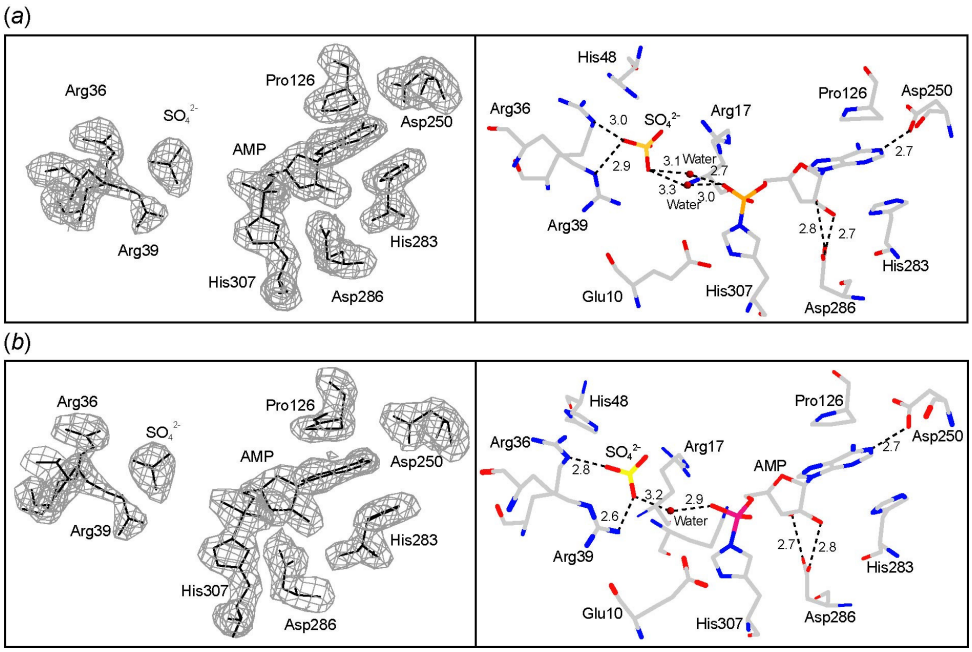


Figure 3-7-f-1. The ATP binding sites in the A-subunit (*a*) and in the B-subunit (*b*). The final $|F_o|$-$|F_c|$ omit map of AMP and the surrounding residues is contoured at 2.5 σ level, the left columns. The right column shows possible hydrogen bond interactions. Carbon, nitrogen, oxygen, phosphorus and sulfur atoms are colored in gray, blue, red, pink and yellow, respectively.

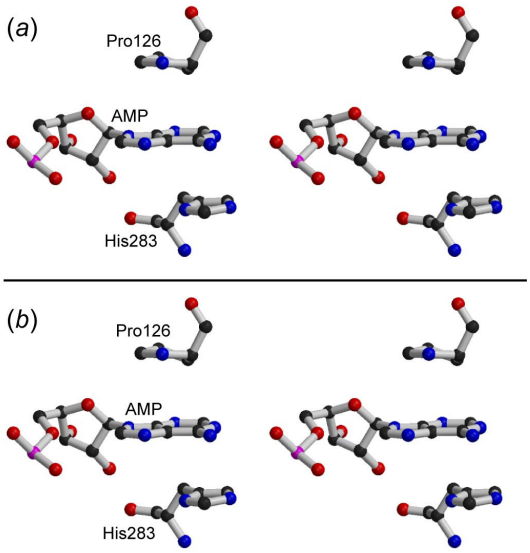


Figure 3-7-f-2. Stereo-pair drawings of the bound AMP between Pro126 and His283 in the A-subunit (*a*) and in the B-subunit (*b*).

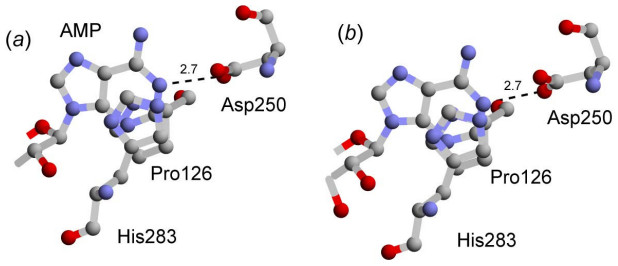


Figure 3-7-f-3. The adenine rings trapped between Pro126 and His283 and hydrogen-bonded to with Asp250 in the A-subunit (*a*) and in the B-subunit (*b*).

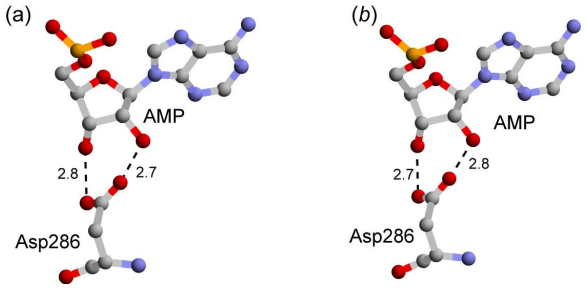


Figure 3-7-f-4. The two hydroxyl groups of the ribose hydrogen-bonded to Asp250 in the A-subunit (*a*) and in the B-subunit (*b*).

Reference

1. B. W. Matthews (1968) “Solvent content of protein crystals” *J.Mol.Biol.*, **33**, 491-497.
2. G. J. Palm, E. Billy, W. Filipowicz & A. Wlodawer (2000) “Crystal structure of RNA 3'-terminal phosphate cyclase, a ubiquitous enzyme with unusual topology” *Structure*, **8**, 13-23.
3. O. Vicente & W. Filipowicz (1988) “Purification of RNA 3'-terminal phosphate cyclase from HeLa cells. Covalent modification of the enzyme with different nucleotides” *Eur. J. Biochem.*, **176**, 431-439.

< Chapter 4; Discussion of *St*-Rtc and its complexes >

4-1. Dimers of *St*-Rtc and its complexes

In this study, the six different crystals, *St*-Rtc*, *St*-Rtc, *St*-Rtc*+Mn, *St*-Rtc*+ATP, *St*-Rtc+AMP, *St*-Rtc*-AMP, were obtained by changing ionic species and pH (acidic and basic), and with or without ATP. In every crystal, the dimer formations have been found as shown in Figure 4-1-1 (*a*). The two *St*-Rtc subunits are associated to each other by the crystallographic regular two-fold symmetry or a non-crystallographic two-fold symmetry. As shown in Figure 4-1-1 (*b*), the dimer is formed through the β-sheet extension between the two subunits. Therefore, it is concluded that *St*-Rtc exists as a dimer composed of the two identical subunits as the biological unit. This is quite different from the case of *Ec*-Rtc [1]. The original authors reported that the dimerization occurred with an S-S linkage formed between the two subunits, in which a cysteine residue near the catalytic site is involved. However it is far from the dimerization interface found in this study. The cysteine residue is replaced with serine residue at the corresponding site in *St*-Rtc, so that such a linkage formation is impossible. In the *Ec*-Rtc crystal, the S-S linked dimers are further associated together through a β-sheet formation on the other side, which is quite similar to the dimerization interface found in this study. Therefore, the S-S linkage might be an artifact in a process of purification or crystallization of *Ec*-Rtc.

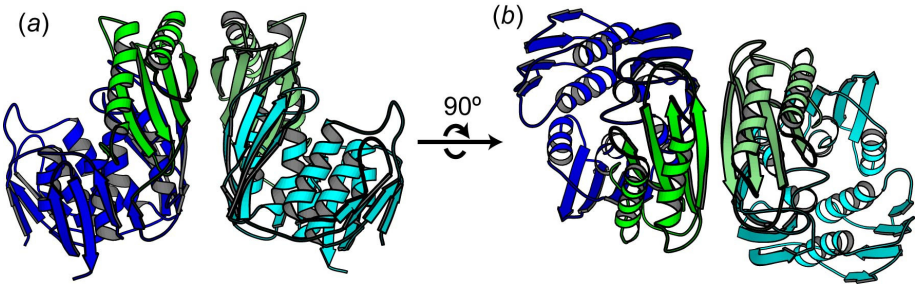


Figure 4-1-1. *St*-Rtc dimer (a) and the top view (b) shown in ribbon model. Small domains are colored in green and light green. Large domains are colored in blue and cyan.

The found dimers are grouped into two forms. Ones are the dimers in the the *St*-Rtc*, the *St*-Rtc, *St*-Rtc*+Mn and *St*-Rtc+AMP crystals obtained under acidic conditions, and the others are those in the *St*-Rtc*+ATP and *St*-Rtc*-AMP crystals obtained under basic conditions. The dimers in the basic form are just on the crystallographic 4_2 axis and stacked on each other. In the acidic form, the dimers are stacked on each other around the crystallographic 3_1 axis, and these screwed columns are arranged in parallel at the periodic distance along the *a* and *b* axes. The parallel

columns are in contact to each other at their triangular edges. Three of the four crystals, *St*-Rtc*, *St*-Rtc*+Mn and *St*-Rtc+AMP, are isomorphous. It is interesting to note that the *St*-Rtc crystal is similar to the others, but is not isomorphous. The column is rotated by 15º around the *c* axis, so that the contacting face of the edge is changed to the other side (face) of the same edge.

Here we can compare the two types of the dimers, one is an acidic form and the other is a basic form. The estimated r.m.s deviations of Cα atoms between the acidic dimers and between basic those are given in Table 4-1-1 and Table 4-1-2, respectively, and between the acidic and basic forms in Table 4-1-3. The values between the two forms are higher than those between the same forms. This suggests that the dimer structures are slightly different between the acidic and the basic forms. The surface charges, as shown in Figure 4-1-2, determine the dimer-dimer interaction to be crystallized in the two different space groups, $P3_1$ and $P4_2$.

Table 4-1-1. The estimated r.m.s. deviations (Å) of Cα atoms between the dimers of the *St*-Rtc, *St*-Rtc*, *St*-Rtc*+Mn and *St*-Rtc+AMP crystals appearing in acidic conditions.

		St-Rtc*	*St*-Rtc*+Mn	*St*-Rtc+AMP
		A-A′ dimer	A-A′ dimer	A-A′ dimer
St-Rtc	A-A′ dimer	0.36	0.57	0.43
	A′-A dimer	0.97	1.01	0.99

Table 4-1-2. The estimated r.m.s. deviations (Å) of Cα atoms between the dimers of the *St*-Rtc*+ATP and *St*-Rtc*-AMP crystals appearing in basic conditions.

		St-Rtc*-AMP	
		A-A′ dimer	B-B′ dimer
St-Rtc*+ATP	A-A′ dimer	0.22	0.20
	B-B′ dimer	0.21	0.20

Table. The estimated r.m.s. deviations (Å) of Cα atoms between the dimers of the *St*-Rtc, *St*-Rtc*, *St*-Rtc*+Mn and *St*-Rtc+AMP crystals appearing in acidic conditions, and those of the *St*-Rtc*+ATP and *St*-Rtc*-AMP crystals appearing in basic conditions.

			Crystals appearing in acidic conditions			
			St-Rtc	*St*-Rtc*	*St*-Rtc*+Mn	*St*-Rtc+AMP
			A-A′ dimer	A-A′ dimer	A-A′dimer	A-A′ dimer
Crystals appearing in acidic condition	*St*-Rtc*+ATP	A-A′ dimer	1.58	1.57	1.78	1.54
		B-B′ dimer	1.58	1.58	1.77	1.54
	St-Rtc*-AMP	A-A′ dimer	1.54	1.54	1.73	1.51
		B-B′ dimer	1.56	1.55	1.74	1.52

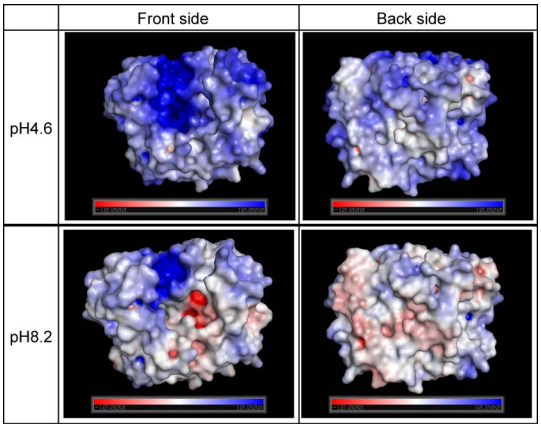


Figure 4-1-2. Electrostatic potential surfaces of *St*-Rtc at pH4.6 (acidic) and pH8.2 (basic). Blue and red colors indicate positively and negatively charge, respectively. The bottom bar indicates correspondence between colors and charge values (-10~10) (kT/e, k; Boltzmann factor, T; template (K) e; elementary charge). Poisson-Boltzmann electrostatics was caluculated by the program *pdb2pqr* [3] and the figures were drawn by the program *Pymol* [4] with plugin of the program *APBS* [5].

4-2. ATP binding sites of Rtc

The residues for the ATP binding in the *St*-Rtc*+ATP crystal are superimposed on those in the *St*-Rtc crystal in Figure 4-2-1. The residues interacting to the adenine ring of ATP are Pro126, Asp250 and His283. They adopt conformations similar to each other between the two crystals without and with ATP. On the other hand, the side chain conformation of Glu10, Arg17, Arg36, Arg39, Asp286 and His307 are different between the two states, depending on without and with ATP. Especially, the side chain of Asp286 moves to form a double hydrogen bonds with the 2′ and 3′ hydroxyl groups of the ribose. The side chains of Glu10 and His307 also move to interact with the phosphate groups of ATP. These changes suggest that once the adenine moiety is bound, the surrounding residues move to trap the ribose and the phosphate groups of ATP. However, only the α-phosphate group is disordered. Between the postulated α-phosphate group and the carboxyl group of Glu10, there is a large space, in which a Mg^{2+} ion might be bound. It is expected that when a Mg^{2+} ion is bound, the α-phosphate group is forced to have a fixed position.

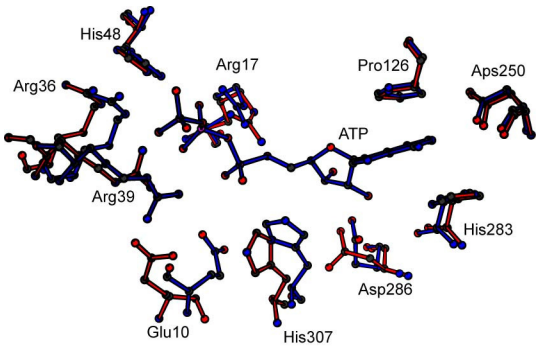


Figure 4-2-1. A superimposition of the residues of the ATP binding site in the *St*-Rtc*+ATP crystal (blue) onto those in the *St*-Rtc crystal (red), drawn by ball-and-stick models. Black, blue, pink and red spheres indicate carbon, nitrogen, phosphorus and oxygen atoms, respectively.

The AMP binding site in the *St*-Rtc+AMP crystal is compared to the corresponding site in the *St*-Rtc crystal in Figure 4-2-2. The residues interacting with the adenine ring, Pro126, Asp250 and His283, adopt the conformations similar to those without AMP. The side chain of Asp286 is also kept a conformation to interact with the ribose hydroxyl group through hydrogen bonds. The side chains of Arg17, Arg36 and Arg39 slightly move to interact to the pyrophosphate. The side chain of His307 adopts the conformation similar to that of *St*-Rtc*+ATP, nearby there are no electron density for the α-phosphate group. They have to interact to each other, if ATP is bound. These features suggest that ATP is hydrolyzed in acidic condition into two parts, pyrophosphates and AMP or is reacted at first to produce AMP-His307 linked state, and then hydrolyzed in acidic condition.

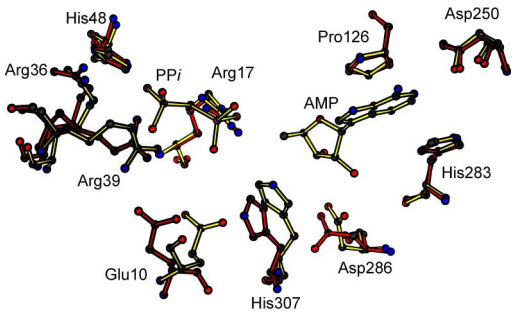


Figure 4-2-2. A superimposition of the residues of the AMP binding site in the *St*-Rtc+AMP crystal (yellow) onto those in the *St*-Rtc crystal (red), drawn by ball-and-stick models. Black, blue, pink and red spheres indicate carbon, nitrogen, phosphorus and oxygen atoms, respectively.

In the *St*-Rtc*-AMP crystal (Figure 4-2-3), Glu10 and His307 change their conformations because the covalent bond is formed between the Nε atom of His307 and the α-phosphorus of AMP. The side chains of Arg17, Arg36 and Arg39 interact with a

sulfate ion by hydrogen bonds. These argeinine residues might be involved in RNA binding in the second step reaction.

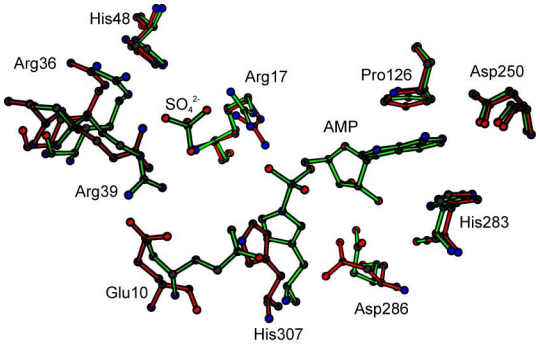


Figure 4-2-3. A superimposition of the residues of the ATP binding site in the *St*-Rtc*-AMP crystal (green) onto those in the *St*-Rtc crystal (red), drawn by ball-and-stick models. Black, blue, pink and red spheres indicate carbon, nitrogen, phosphorus and oxygen atoms, respectively.

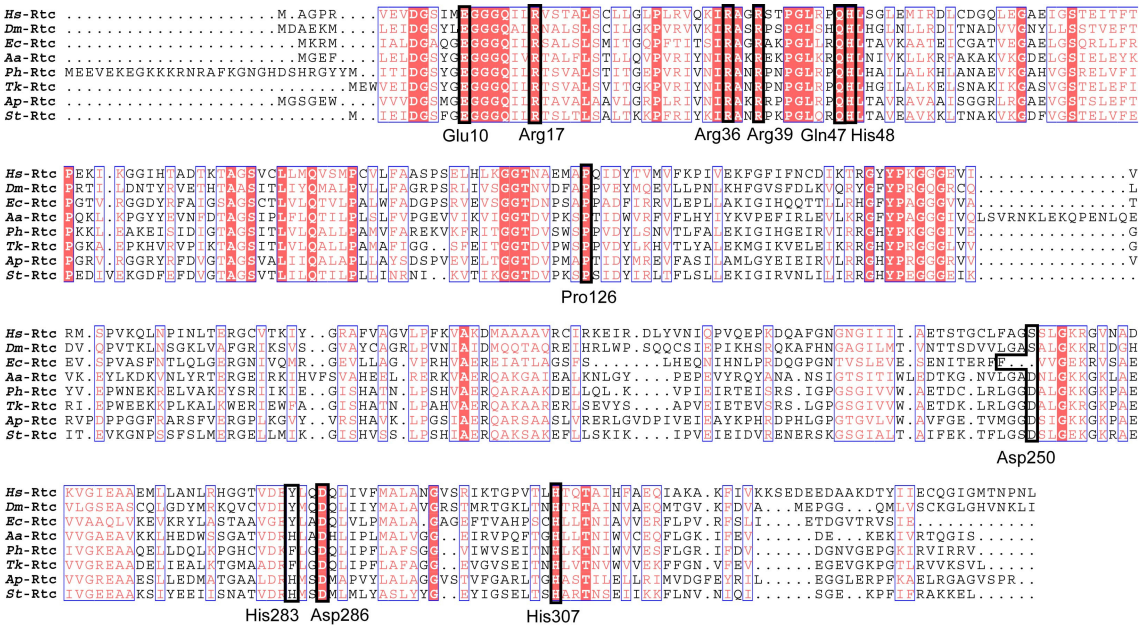


Figure 4-2-4. Sequence alignment of Rtcs from eukaryote; *Homo sapiens* (human) (*Hs*-Rtc) and *Drosophila melanogaster* (fruit fly) (*Dm*-Rtc), from bacteria; *Escherichia coli* (*Ec*-Rtc) and *Aquifex aeolicus* (*Aa*-Rtc), from euryarchaea; *Pyrococcus horikoshii* (*Ph*-Rtc) and *Thermococcus kodakaraensis* (*Tk*-Rtc) and from crenarchaea; *Aeropyrum pernix* (*Ap*-Rtc) and *Sulfolobus tokodaii* (*St*-Rtc). Characters surrounding the thick lines indicate the corresponding residues in the substrate binding site of *St*-Rtc. Red and white characters indicate homology and identity residues among the Rtcs, respectively. The alignment parameters are 11.0% for the gap opening penalty which is a gap for opening and 0.85% for the gap extending penalty which is for extending a gap by one residue.

To verify the residues concerning to the ATP binding of Rtc in other organisms, a multiple sequence arraignment of Rtcs from several species of three kingdoms are performed using the program *Kalign* [2] (Figure 4-2-4). The sequences from eukaryote;

Homo sapiens (human) (*Hs*-Rtc; gene ID 8634) and *Drosophila melanogaster* (fruit fly) (*Dm*-Rtc; gene ID Dmel_CG4061), from bacteria; *Escherichia coli* (*Ec*-Rtc; gene ID b4475) and *Aquifex aeolicus* (*Aa*-Rtc; gene ID aq_642), from euryarchaea; *Pyrococcus horikoshii* (*Ph*-Rtc; gene ID PH1529) and *Thermococcus kodakaraensis* (*Tk*-Rtc; gene ID TK1615) and from crenarchaea; *Aeropyrum pernix* (*Ap*-Rtc; gene ID APE_0511.1) and *Sulfolobus tokodaii* (*St*-Rtc; gene ID ST0570) were downloaded from the KEGG database (http://www.genome.jp/kegg/).

The histidine residues corresponding to His307, which is covalently bound to the α-phosphorus atom of AMP in the *St*-Rtc*-AMP crystal, are conserved in every Rtc of other organisms.

The proline residues corresponding to Pro126, which is involved in the adenine ring of ATP and AMP in the *St*-Rtc*+ATP, the *St*-Rtc+AMP and the *St*-Rtc*-AMP crystals, are highly conserved in other Rtcs. However, the partner residues corresponding to Asp250 and His283 in *St*-Rtc, which also interact with the adenine ring, are slightly changed. Asp250 of *St*-Rtc is conserved in *Aa*-Rtc *Ph*-Rtc, *Tk*-Rtc and *Ap*-Rtc, but replaced with serine in *Hs*-Rtc and *Dm*-Rtc and with phenylalanine in *Ec*-Rtc. His283 of *St*-Rtc is conserved by the replacing with other aromatic amino acid, tyrosine in *Hs*-Rtc, *Dm*-Rtc and *Ec*-Rtc and phenylalanine in *Ph*-Rtc and *Tk*-Rtc. The aromatic rings of these residues can capture on the adenine ring of ATP and AMP in a similar trapping manner instead of His283 in *St*-Rtc.

The residues corresponding to Glu10, Arg17, Arg36, Arg39, Glu47, His48 and Asp286, which interact with the phosphate groups and the ribose of ATP and AMP in *St*-Rtc are also highly conserved in every organism.

4-3. Reaction mechanism of the first reaction

Based on the structures solved in this study, a reaction mechanism at the first step can be figured out as shown in Figure 4-3-1. At first, ATP enters the binding site of *St*-Rtc (Figure 4-3-1 (*a*)). The adenine ring is staked between the side chains of Pro126 and His283 (Figure 4-3-1 (*b*)), and its orientation fixed by hydrogen bonds between Asp250 and the adenine ring, and between Asp286 and the ribose hydroxyl groups. The basic residues, Arg17, Arg36, Arg39 and His48, form positively charged surface, as shown in Figure 4-3-2. The β- and γ-phosphate groups of the ATP are bound to these residues. The essential Mg^{2+} ion might be bound to Glu10 so that the α- and β-phosphate groups of ATP can be trapped as shown in Figure 4-3-1 (*c*). The bound Mg^{2+} ion changes the conformations of the α-phosphate groups of ATP, so that the lone-paired electron on the Nε atom (His307) might be forced to attack the phosphorus

atom of the α-phosphate group. After a phosphoamide bond is formed between the Nε atom and the phosphorus atom, a pyrophosphate ion is released (Figure 4-3-1 (*d*)). This covalently bonded *St*-Rtc-AMP intermediate (Figure 4-3-1 (*e*)) is used in the second reaction.

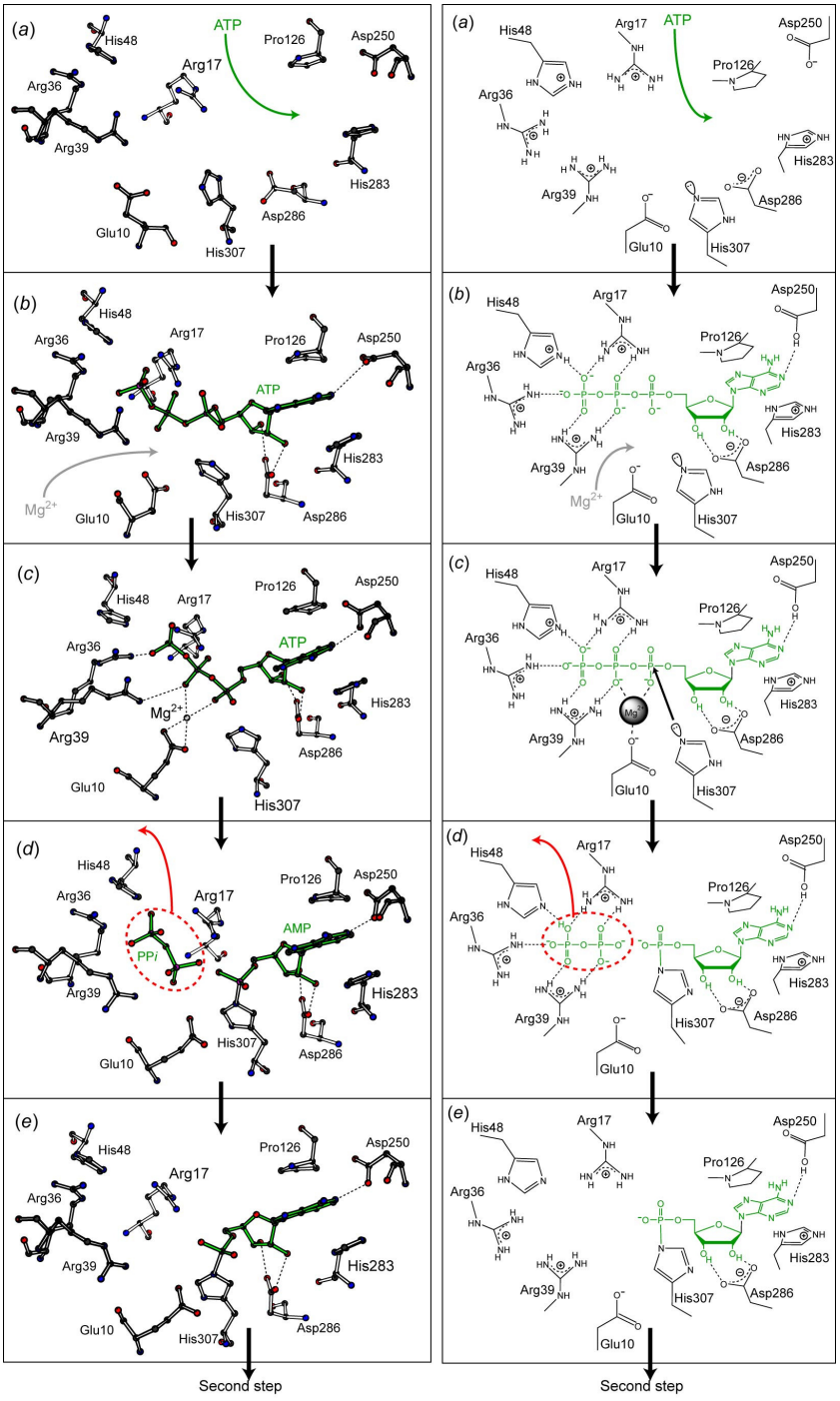


Figure 4-3-1. Models (left column) and schemes (right column) of a proposed mechanism of the first reaction of *St*-Rtc. In the left figures, black, blue, pink and red spheres indicate carbon, nitrogen, phosphorus and oxygen atoms, respectively.

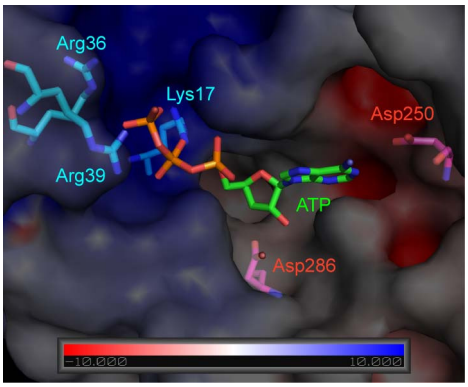


Figure 4-3-2. An electrostatic potential surface of the ATP binding site of *St*-Rtc*+ATP subunit at pH8.2. Blue and red colors indicate positively and negatively charge, respectively. The bottom bar indicates correspondence between colors and charge values (-10~10) (kT/e, k; Boltzmann factor, T; template (K) e; elementary charge). Poisson-Boltzmann electrostatics was caluculated by the program *pdb2pqr* [3] and the figures were drawn by the program *Pymol* [4] with plugin of the program *APBS* [5].

4-4 Reaction mechanism of the second reaction

The second reaction is speculated as described below. A single stranded RNA with a phosphate group at the 3′-terminus will be bound in *St*-Rtc*-AMP in a way as shown in Figure 4-4-1. The region indicated by a white arrow shows negative charge. The eight basic residues, Arg17, Arg36, Arg39, Lys41, Arg46, His48 and Lys124 are assembled in the region. These suggest that RNA may be bound to the region.

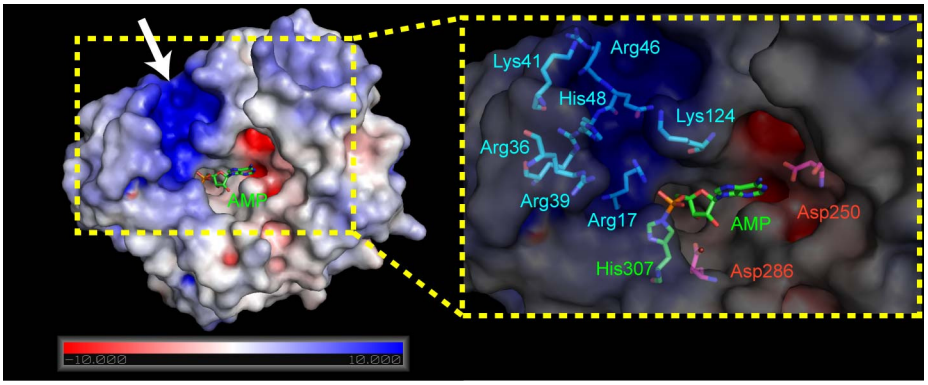


Figure 4-4-1. An electrostatic potential of surface of *St*-Rtc*-AMP subunit at pH8.2. A white arrow indicates a possible RNA binding site. The right figure shows a zoom up view around the binding site. Blue and red colors indicate positively and negatively charge, respectively. The bottom bar indicates correspondence between colors and charge values (-10~10) (kT/e, k; Boltzmann factor, T; template (K) e; elementary charge). Poisson-Boltzmann electrostatics was caluculated by the program *pdb2pqr* and the figures were drawn by the program *Pymol* with plugin of the program *APBS*.

Figure 4-4-2 and 3 show a reaction mechanism of the second step, proposed from this study. A spliced RNA with the 3′-terminal phosphate will be bound in a possible region, as shown in Figure 4-4-1 (Figure 4-4-2 (*a*)). His48 and Gln47 may form hydrogen bonds with the 2′-hydroxyl group of the 3′-end of RNA. The side chains of Arg17 and Arg38 interact with the phosphate groups of RNA. The oxygen atom of the

3′-phosphate group might be forced to attack the α-phosphorus atom bound to His307 (Figure 4-4-2 (*b*)). The P-N bond is cleaved to produce an intermediate linked by a pyrophosphate bond between the 3′-phosphate group of RNA and AMP (Figure 4-4-2 (*c*)). As a Schiff base, His48 withdraws a proton from the 2′-hydroxyl group of the RNA and the 2′-oxygen atom attacks the 3′-phosphorus atom to form a free AMP and the 2′,3′-cyclic phosphate group at the 3′ end of RNA (Figure 4-4-2 (*d*)). This reaction scheme must be proved by further structural studies.

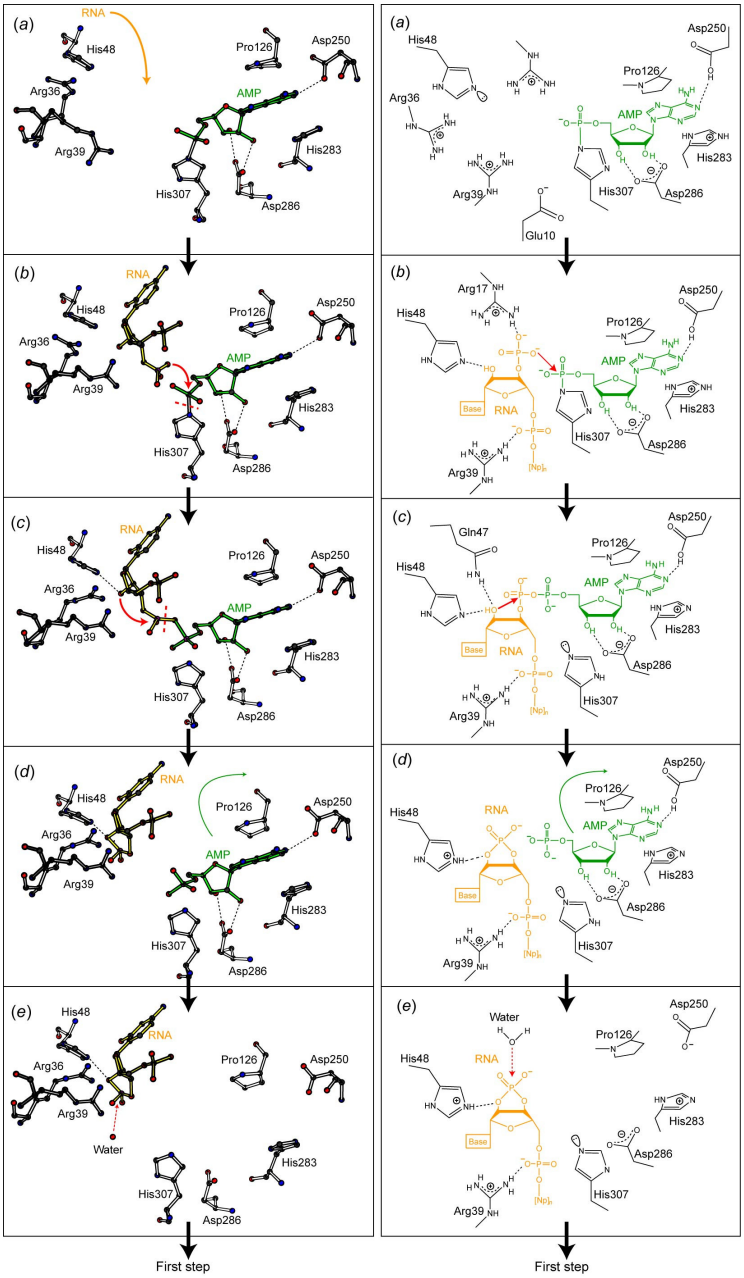


Figure 4-4-2. Models (left column) and a scheme (right column) of proposed mechanism of the second reaction of *St*-Rtc. In the left figures, black, blue, pink and red spheres indicate carbon, nitrogen, phosphorus and oxygen atoms, respectively.

4-5. Reference

1. G. J. Palm, E. Billy, W. Filipowicz & A. Wlodawer (2000) "Crystal structure of RNA 3'-terminal phosphate cyclase, a ubiquitous enzyme with unusual topology" *Structure*. **8**, 13-23.
2. T. Lassmann & E. L. Sonnhammer (2005) "Kalign--an accurate and fast multiple sequence alignment algorithm" *Bioinformatics*. **6**, 298.
3. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon & N. A. Baker (2004) "PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations" *Nucleic Acids Research*, **32**, W665-W667.
4. DeLano Scientific LLC.(http://www.pymol.org/)
5. N. A. Baker, D. Sept, S. Joseph, M. J. Holst & J. A. McCammon (2001) "Electrostatics of nanosystems: application to microtubules and the ribosome" *Proc. Natl. Acad. Sci*. USA **98**, 10037-10041.

Part II: Threonyl-tRNA synthetases

< Chapter 5; Architectural analyses of ThrRSs from crenarchaea >

5-1. Introduction

The X-ray structures of ThrRSs from bacteria, *E. coli* [1] (*Ec*ThrRS) and *Staphylococcus aureus* [2] (*Sa*ThrRS) were reported. Both are dimeric enzyme of the two subunits. As shown in Figure 5-1-1, the subunit is composed of three domains; the N-terminal editing domain, the catalytic domain and the C-terminal anticodon-binding domain. In this study, the amino acid sequences of two eukaryotic ThrRSs from *Homo sapience* (*Hs*ThrRS) and *Saccharomyces cerevisiae* (*Sc*ThrRS) have been compared with those of the above two ThrRSs. As shown in Figure 5-1-2, the three domains are conserved even in eukaryote except for additional regions at the N-terminus which contain about 70 residues with no homology to *Hs*ThrRS and *Sc*ThrRS. It has been found that the two essential sequence motifs of HxxxH and GxCRPH [3] on the editing domains of bacterial ThrRSs are almost conserved in the eukaryotic editing domains. A cystain and two histidine residues are consistently conserved for the zinc finger at the active site of the catalytic domain. In the anticodon-binding domain, a lysine and an argeinine residues are conserved for interactions with the anticodon G35 and U36 residues of $tRNA^{Thr}$, respectively.

Furthermore the Fig 5-1-2 also suggests that the archaeal ThrRSs of *Pyrococcus abyssi* contains three domains, but only the N-terminal domain has no sequence similarity. The truncated N-terminal domain was confirmed to have an activity as a editing domain [4]. The X-ray structure of only the editing domain revealed that the tertiary structure is quite different from those of *Ec*ThrRS and *Sa*ThrRS as shown in Figure 5-1-3 [4].

Recently, two different genes were assigned for ThrRS in *Sulfolobus solfataricus* genome. One is the catalytic enzyme (ThrRS-cat) and the other is the editing enzyme (ThrRS-ed) [5]. Namely, ThrRS-cat synthesizes both Thr-$tRNA^{Thr}$ and Ser-$tRNA^{Thr}$, but lacks the editing activity of Ser-$tRNA^{Thr}$. In contrast, ThrRS-ed lacks the aminoacylation activity but removes the serine from the misacylated Ser-$tRNA^{Thr}$. Therefore, it is considered that the two enzymes may work complementally to each other.

In the present study, we have found that four crenarchaea, *Aeropyrum pernix* (*Ap*),

Sulfolobus tokodaii (*St*), *Sulfolobus acidocaldarius* (*Si*) and *Metallosphaera sedula* (*Ms*), similarly contain two ThrRS genes*. In order to speculate their structural features, their amino acid sequences were compared to those of other organisms, including *Ss*ThrRS.

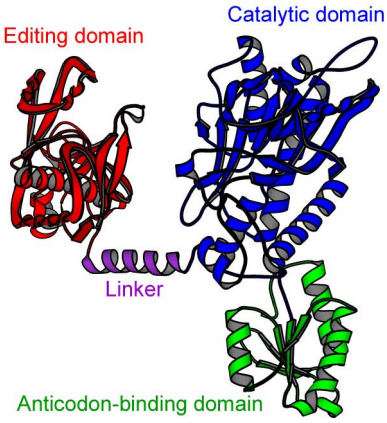


Figure 5-1-1. Structural composition of ThrRS from *E.coli* (PDB ID 1QF6 [1]). The editing, the catalytic and the anticodon-binding domains are colored in red, blue and green, respectively. The linker region is show in purple color.

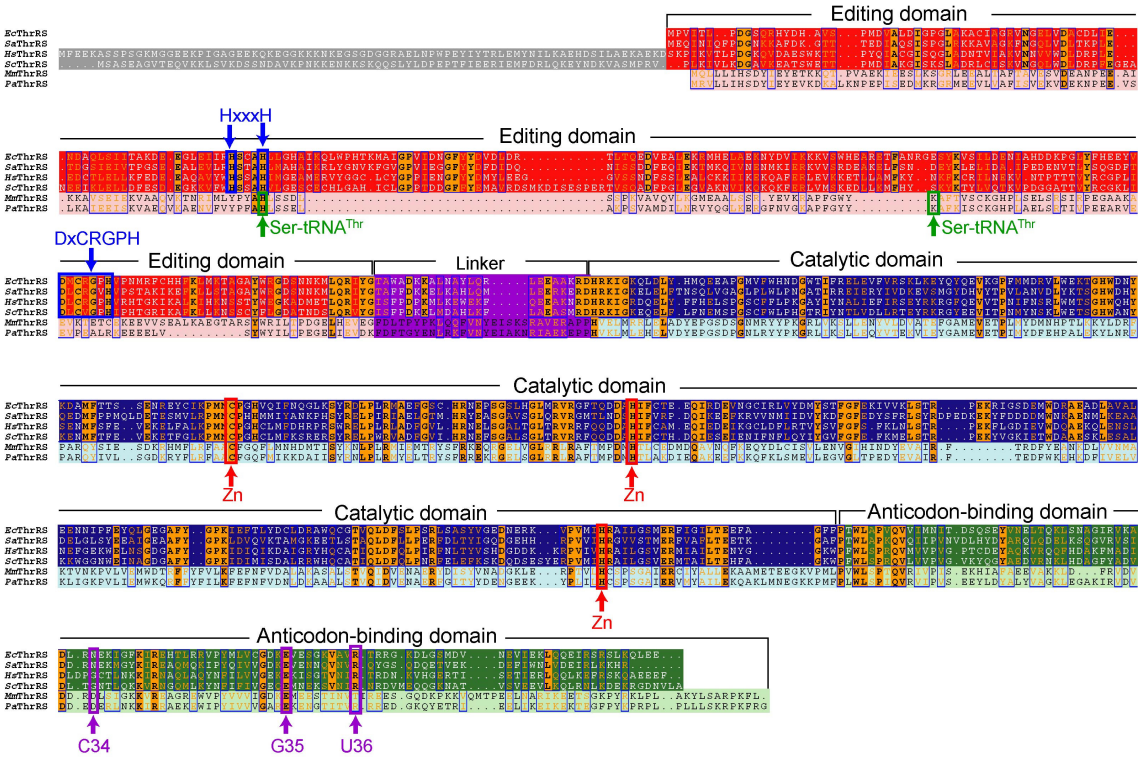


Figure 5-1-2. A sequence alignment of bacterial; *Ec*ThrRS and *Sa*ThrRS, eukaryotic; *Hs*ThrRS and *Sc*ThrRS, and archaeal; *Mm*ThrRS and *Pa*ThrRS. Orange boxes and orange characters indicate identical and homologous residues, respectively. Red arrows indicate the residues constructing zinc finger. The characters surrounded with red and purple boxes indicate the estimated residues interacting with a zinc atom in the catalytic site and with anticodon residues of $tRNA^{Thr}$, rescpectively. Blue and green arrows indicate the essential residues of the editing function of *Ec*ThrRS and *Pa*ThrRS, respectively.

*In the KEGG database, there are found two different ThrRS genes of in some bacteria, eukaryote and archaea. They are not composed yet in this study.

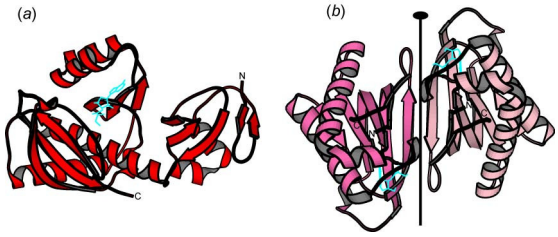


Figure 5-1-3. (*a*) Monomeric structure of editing domain of ThrRS from *E. coli* (PDBID 1TKG [3]). (*b*) Dimeric structure of the editing domains of ThrRS from *Pyrococcus abyssi* (PDBID 2HL0 [5]). A black line with a black ellipse indicates a crystallographic 2-fold symmetry. The analog molecules, serine-3'-aminoadenosine are shown in cyan color. The characters "N" and "C" indicate N- and C-terminus, respectively.

5-2 Multiple sequence alignment of creanarchaeal ThrRSs

5-2-1. Sequence composition and alignment

The amino acid sequences of ThrRS from *E.coli* (gene ID b1719) and *Pyrococcus abyssi* (gene ID PAB1490) (*Pa*ThrRS) were downloaded from the KEGG database. The amino acid sequences of ThrRSs of *Aeropyrum pernix* (gene IDs APE0809.1 and APE0117.1), *Sulfolobus tokodaii* (gene IDs ST0966 and ST2187), *Sulfolobus solfataricus* (gene IDs SSO2486 and SSO0384), *Sulfolobus acidocaldarius* (gene IDs SACI1261 and SACI2363) and *Metallosphaera sedula* (gene IDs MSED1585 and MSED1601) were also downloaded from the KEGG database. Their multiple alignments were performed using the program *clustalX* [6]. The following multiple alignment parameters were used; 10.0% for the gap opening penalty which is a gap for opening and 0.2% for the gap extending penalty which is for extending a gap by one residue. Increasing the gap opening penalty will make gaps less frequent and the gap extending penalty will make gaps shorter. For easy understanding of the output diagram, unhomologous parts were moved on a line so as to separate to each other at different positions, and the modified diagram was colored by program *ESPript* [7]. Similar sequence comparison was performed on the FASTA server (http://fasta.ddbj.nig.ac.jp/top-j.html) [8].

5-2-2. Results

Depending on domain arrangement with sequence homology, the two ThrRS genes of crenarchaea organisms are divided into two groups, ThrRS-1 and ThrRS-2. As shown in Figure 5-5-2-1, ThrRS-1 contains APE0809.1, ST0966, SSO2486, SACI1261 and MSED1585, and as shown in Figure 5-5-2-2, ThrRS-2 contains APE0117.1, ST2187, SSO0384, SACI2363 and MSED1601. However it is interesting to note that both ThrRS-1 and ThrRS-2 commonly possess the C-terminal anticodon-binding domain.

In ThrRS-1, there are no sequences corresponding to the editing domain such as found in *Ec*ThrRS. Closer examination shows that the sequences of the catalytic and anticodon-binding domains of ThrRS-1s are more similar to those of *Ec*ThrRS than those of *Pa*ThrRS. The residues constructing a zinc finger at the active site of *Ec*ThrRS (Cys334, His385 and His511) and interacting to the 35th and 36th residues of the anticodon tRNA^{Thr} are conserved in all ThrRS-1. These residues are indicated with red and purple arrows. This feature suggests that ThrRS-1 can bind to tRNA^{Thr} to aminoacylate it at the zinc finger motif as like *Ec*ThrRS [1]. An unknown region in *Sf*ThrRS-1 contains 110 residues, which are well conserved among *Sulfolobaceae* family (*St*, *Ss*, *Si*, and *Ms*). These sequences cannot be found in any other ThrRS, when surveyed on the FASTA server.

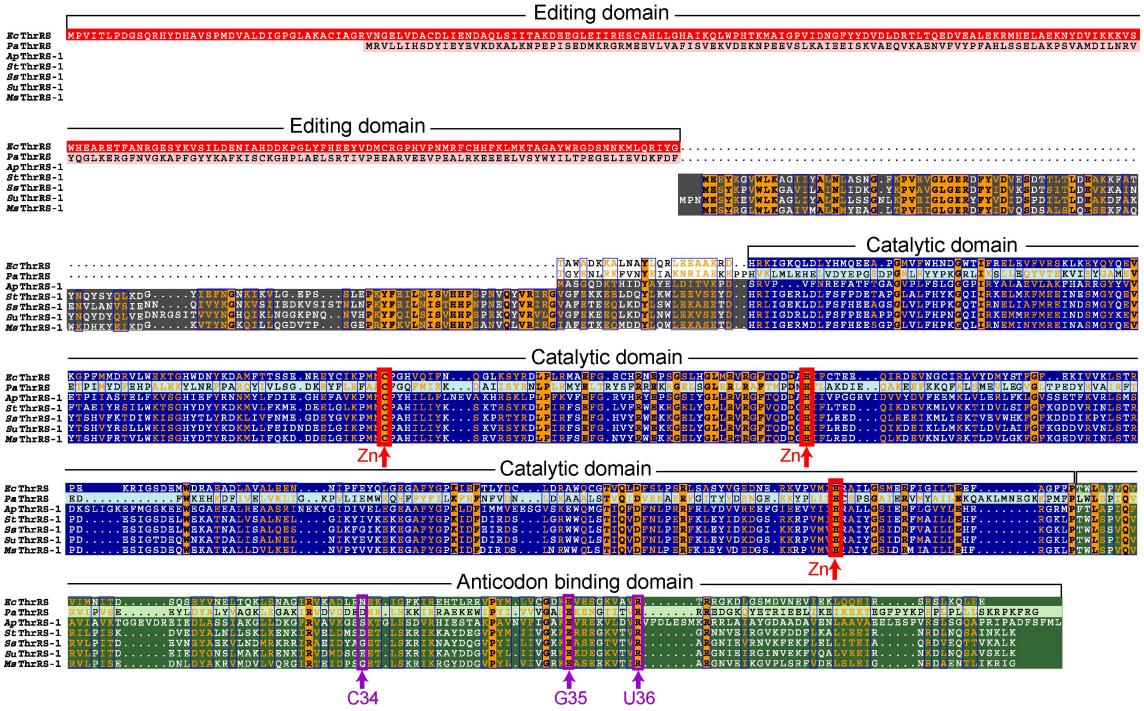


Figure 5-2-2-1. A sequence alignment of *Ec*ThrRS, *Pa*ThrRS, *Ap*ThrRS-1 and ThrRS-1 from several *Sulfolobaceae* family, *Sulfolobus tokodaii* (*St*), *Sulfolobus solfataricus* (*Ss*), *Sulfolobus acidocaldarius* (*Si*) and *Metallosphaera sedula* (*Ms*). Orange boxes and orange characters indicate identical and homologous residues, respectively. Red arrows indicate the residues constructing zinc finger. The characters surrounded with red and purple boxes indicate the estimated residues interacting with a zinc atom in the catalytic site and with anticodon residues of tRNA^{Thr}, rescpectively.

The remarkable difference is that ThrRS-2 completely lacks the catalytic domain as shown in Figure 5-2-2-2. In the anticodon-binding domain, a lysine residue interacting with G35 of the anticodon in *Ec*ThrRS is conserved in ThrRS-2. At the position of the Arg609 residue interacting with G36 in *Ec*ThrRS, the basic arginine or lysine residue is conserved. These residues suggest a possibility that ThrRS-2 can bind to the tRNA^{Thr}. The sequences of the editing domains are similar to that of *Pa*ThrRS. The proposed

residues for cleaving Ser-tRNA^{Thr}[5], hystidine and lysine, are well conserved in ThrRS-2 except for *St*ThrRS-2 which lacks a sequence containing these residues. In the case of *St*ThrRS-2, however, the editing domain also displays high sequence conservation, but its sequence is externally short. Only the 13 amino acid residues are linked to at the N-terminus. It is too short to have an editing function. Instead, the subsequent 130 residues are well conserved in every *Sf*ThrRS-2. They have no sequence homology with any other sequences found on the FASTA server. *Ap*ThrRS-2 also has such an inserted sequence but it is different from those of *Sulfolobaceae* family.

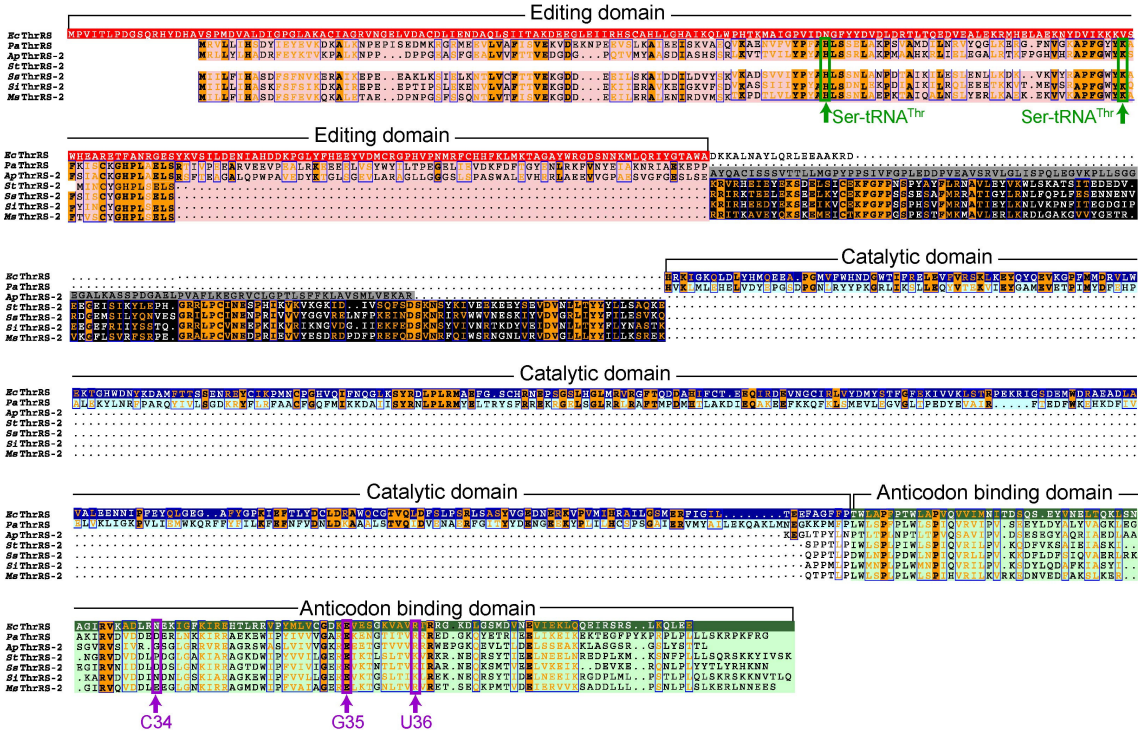


Figure 5-2-2-2. A sequence alignment of *Ec*ThrRS, *Pa*ThrRS, *Ap*ThrRS-2 and ThrRS-2 from several *Sulfolobaceae* family, *Sulfolobus tokodaii* (*St*), *Sulfolobus solfataricus* (*Ss*), *Sulfolobus acidocaldarius* (*Si*) and *Metallosphaera sedula* (*Ms*). Orange boxes and orange characters indicate identical and homologous residues, respectively. The characters surrounded with green and purple boxes indicate the estimated residues editing Ser-tRNA^{Thr} and interacting with anticodon residues of tRNA^{Thr}

5-2-3. Discussion

A schematic summary of the above results is shown in Figure 5-2-3-1, from which it is easy to see the difference in domain arrangements. ThrRS-1 is completely missing the editing domain. Every ThrRS-1 has an additional sequence at the N-terminus except for *Ap*ThrRS-1. These sequences contain about 140 residues. On the other hand, every ThrRS-2 is completely missing the catalytic domain. The additional sequence at the N-terminus of *Ap*ThrRS-2 is similar to those of the editing domain of *Pa*ThrRS. The corresponding sequences of *Sf*ThrRS-2 except for *St*ThrRS-2 are also similar to those of

*Pa*ThrRS, but the sizes are shorter by 70 residues. It is the most interesting to note that the size in *St*ThrRS-2 is too short, but its sequence of 13 residues is similar to the C-terminal sequence of the *Pa*ThrRS editing domain. The subsequent sequences with about 130 residues are commonly found in all *Sf*ThrRS-2. Their sequences have high homology to each other. The corresponding sequence is found in *Ap*ThrRS-2. The length is a little bit short and its sequence has no homology.

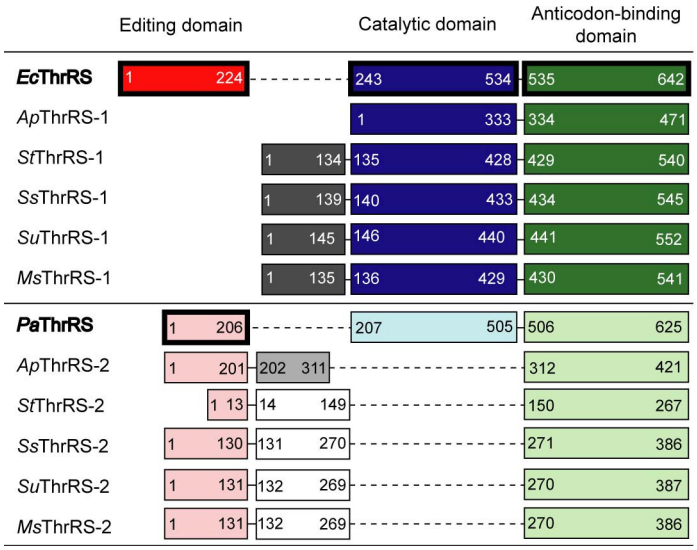


Figure 5-2-3-1. Domain arrangement in various ThrRS, based on sequence alignment. Red and pink bars show the bacterial and archaeal editing domains, dark and light blue bars show the bacterial and archaeal catalytic domains, and dark and light green bars show the bacterial and archaeal anticodon-binding domains, respectively. The numeric characters in bars indicate the residue number assigned from N-terminus. The bold-framed boxes indicate the domains whose three dimensional structures were solved by X-ray crystallography [1, 4].

The domain arrangement in Figure 5-2-3-1 stimulates us to speculate their tertiary structures. ThrRS-1 looks rather shorter, because the N-terminal editing domain is missing. In the *Ec*ThrRS structure, the anticodon-binding and the catalytic domains participate in stabilization of the dimer formation with a two-fold symmetry (Figure 5-2-3-2). The editing domain is linked to the catalytic domain with short linker sequence at the opposite side of the dimerization interface as a member of class II aaRS. This feature is possible to maintain in ThrRS-1.

On the other hand, it is interesting to know how ThrRS-2 can form a dimer without the catalytic domain. An alternative possibility arises from the editing domain, because the two editing domains can form a dimer in the case of *Pa*ThrRS as shown in Figure 5-2-3-3. The interfacing sequence contains the highly conserved 13 amino acid residues in every *Sf*ThrRS-2. Therefore the sequences might occupy the center of the most essential part of the interface. It is necessary to clarify the structural basis for these interests in ThrRS-1 and ThrRS-2 from two crenarchaea, *Aeropyrum pernix* and

Sulfolobus tokodaii. Their X-ray analyses will be described in the subsequence chapter.

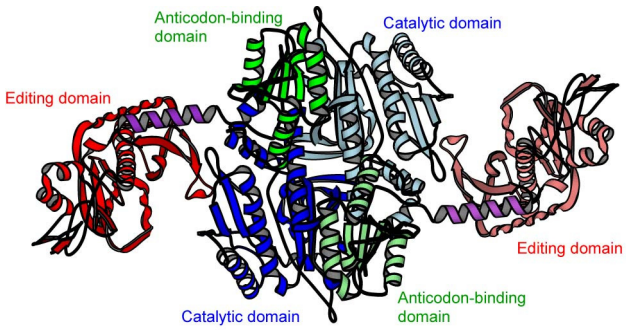


Figure 5-2-3-2. Dimeric structure of *Ec*ThrRS [1]. The editing, the catalytic and the anticodon-binding domains are colored in red, blue and green, respectively. The linker region is show in purple color. Each light colors indicate crystallographic equivalent molecule forming a dimer.

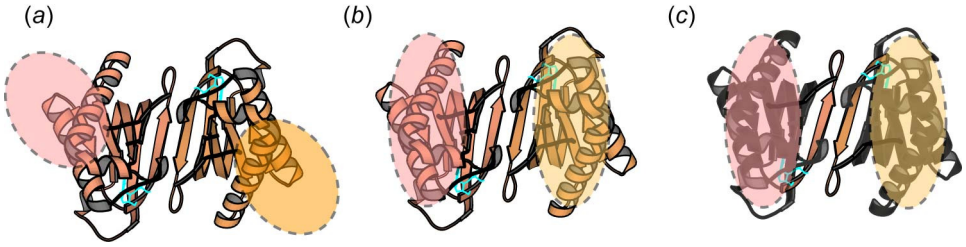


Figure 5-2-3-3. Possible structures of the editing domains of ThrRS-2s; *Ap*ThrRS-2 (*a*), *St*ThrRS-2 (*b*) and the other *Sf*ThrRS-2 (*c*). The colored circles with dot line are illustrated as remaining unassigned amino acids and black cartoon in (*b*) indicates inserted regions based on the sequence alignments.

5-3. Reference

1. R. Sankaranarayanan, A. C. Dock-Bregeon, P. Romby, J. Caillet, M. Springer, B. Rees, C. Ehresmann, B. Ehresmann & D. Moras (1999) "The structure of threonyl-tRNA synthetase-tRNA^{Thr} complex enlightens its repressor activity and reveals an essential zinc ion in the active site" *Cell*, **97**, 371-381.
2. A. Torres-Larios, R. Sankaranarayanan, B. Rees, A. C. Dock-Bregeon & D. Moras (2003) "Conformational movements and cooperativity upon amino acid, ATP and tRNA binding in threonyl-tRNA synthetase" *J.Mol.Biol*., **331**, 201-211.
3. A. C. Dock-Bregeon, B. Rees, A. Torres-Larios, G. Bey, J. Caillet & D. Moras (2004) "achieving error-Free translation; The mechanism of proofreading of threonyl-tRNA synthetase at atomic resolution" *Mol. Cell*, **16**, 375-386.
4. T. Hussain, S. P. Kruparani, B. Pal, A. C. Dock-Bregeon, S. Dwivedi, M. R. Shekar, Sureshbabu, K. & Sankaranarayanan, R. (2006) "Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea" *EMBO J*., **25**, 4152-4162.
5. D. Korencic, I. Ahel, J. Schelert, M. Sacher, B. Ruan, C. Stathopoulos, P. Blum, M. Ibba & D. Söll "A freestanding proofreading domain is required for protein synthesis quality control in archaea" *PNAS*., **101** 10260-10265.
6. J. D. Thompson, T. J. Gibson, F. Plewniak, F. Jeanmougin & D. G. Higgins (1997) "The *ClustalX* windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools" *Nucleic Acids Research*, **24** 4876-4882.
7. P. Gouet, E. Courcelle, D. I. Stuart & F. Metoz (1999) "*ESPript*: multiple sequence alignments in PostScript" *Bioinformatics*. **15** 305-308.
8. W. R. Pearson & D. J. Lipman (1988) "Improved tools for biological sequence comparison" *PNAS*, **85**, 2444–2448.

< Chapter 6. X-ray analyses of ThrRSs from crenarchaeota >

6-1. Introduction

In order to elucidate the structural basis of the editing and catalytic mechanisms in the two crenarchaea *A. pernix* and *S. tokodaii*, the two ThrRSs from both organisms were overexpressed in *E. coli* and purified. Crystals of the native and selenomethionine-labelled ThrRSs were obtained and X-ray diffraction data were collected.

6-2. Overexpression and purification

6-2-a. Overexpression and purification of *Ap*ThrRS-1*

The gene encoding *Ap*ThrRS-1 (gene ID; APE0809.1, 471 amino acids, 53122 Da, pI 5.89) were inserted into the pET11a vector (Novagen) introduced into *E. coli* B834 (DE3) (Novagen). The selenomethionine derivative was expressed using minimal M9 medium containing L-selenomethionine. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Resource Iso column (Amersham Bioscience). Bound proteins were eluted at 5 ml/min flow rate with 20 mM Tris-HCl (Tris hydroxymethyl aminomethane titrated using HCl) (pH 8.0), 5 mM β-mercaptoethanol and a linear gradient of 1.2-0 M ammonium sulfate. The pooled fractions were loaded onto a Resource Q column (Amersham Bioscience), and the bound proteins were eluted at 4 ml/min flow rate with 20 mM Tris-HCl (pH8.0), 5 mM β-mercaptoethanol and 0-0.6 M NaCl linear gradient. The eluted fractions were finally applied onto a HiLoad 16/60 Superdex75 column (Amersham Bioscience) with 20 mM Tris-HCl (pH8.0), 5mM β-mercaptoethanol and 150 mM NaCl at 1ml/min flow rate.

The protein was concentrated to a final concentration of 6 mg/ml in 20 mM Tris-HCl (pH 8.0) with centrifugal filter devices (Vivaspin 500, Sartorius Stedim Biotech). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-2-b. Overexpression and purification of *Ap*ThrRS-1

The recombinant plasmids containing the gene of *Ap*-ThrRS-1 were transformed in *E. coli* Rosetta-gami (DE3) (Novagen), and the cells were grown in LB culture at 310

K. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Q-Sepharose Fast Flow column (Amersham Bioscience). Bound proteins were eluted at 1 ml/min flow rate with 50 mM potassium phosphate (pH 7.0) and a linear gradient of 0-200 mM NaCl. The pooled fractions were loaded onto a Hydroxyapatite Type 1 column (Bio-Rad), and the bound proteins were eluted at 2ml/min flow rate with 10-250 mM potassium phosphate (pH7.0) linear gradient. The eluted fractions were finally applied onto a Superdex 200 column (Amersham Bioscience) with 50 mM potassium phosphate (pH 7.0) and 150 mM NaCl at 2ml/min flow rate.

The protein was concentrated to a final concentration of 6 mg/ml in 20 mM potassium phosphate (pH 7.0) with centrifugal filter devices (Vivaspin 500, Sartorius Stedim Biotech). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-2-c. Overexpression and purification of *Ap*ThrRS-2

The gene encoding *Ap*ThrRS-2 (gene ID; APE0117.1, 421 amino acids, 45003 Da, pI 6.74) were inserted into the pET11a vector (Novagen). The recombinant plasmids were transformed in *E. coli* Rosetta-gami (DE3) (Novagen), and the cells were grown in LB culture at 310 K. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Q-Sepharose Fast Flow column (Amersham Bioscience). Bound proteins were eluted at 1 ml/min flow rate with 50 mM potassium phosphate (pH 7.0) and a linear gradient of 0-500 mM NaCl. The pooled fractions were loaded onto a Phenyl-Toyopeal Pak650S column (Tosoh), and the bound proteins were eluted at 3ml/min flow rate with 50 mM potassium phosphate (pH7.0) and 0-0.8 M ammonium sulfate linear gradient. Moreover, the eluted fractions were applied onto a DEAE-cellulose A500-sf column (Seikagaku Biobusiness Corp.) with 50 mM potassium phosphate (pH 7.0) and 0-500 mM NaCl at 2ml/min flow rate. The eluted fractions were finally applied onto a Superdex200 column (Amersham Bioscience) with 50 mM potassium phosphate (pH7.0) and 500-250 mM NaCl linear gradient.

The protein was concentrated to a final concentration of 6 mg/ml in 20 mM

potassium phosphate (pH 7.0) with centrifugal filter devices (Vivaspin 500, Sartorius Stedim Biotech). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-2-d. Overexpression and purification of *St*ThrRS-1

The gene encoding *St*ThrRS-1 (gene ID; ST0966, 540 amino acids, 63110 Da, pI 6.49) were inserted into the pET11a vector (Novagen). The recombinant plasmids were transformed in *E. coli* Rosetta-gami (DE3) (Novagen), and the cells were grown in LB culture at 310 K. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Resource Iso column (Amersham Bioscience). Bound proteins were eluted at 6 ml/min flow rate with 20 mM Tris-HCl (Tris hydroxymethyl aminomethane titrated using HCl) (pH 8.0), 5 mM β-mercaptoethanol and a linear gradient of 1.35-0 M ammonium sulfate. The pooled fractions were loaded onto a Resource Q column (Amersham Bioscience), and the bound proteins were eluted at 4 ml/min flow rate with 20 mM Tris-HCl (pH8.0), 5 mM β-mercaptoethanol and 0-2 M NaCl linear gradient. Moreover, the eluted fractions were applied onto a BioScale CH5I column (Bio-Rad) with 10-500 mM potassium phosphate (pH7.0) linear gradient, 5 mM β-mercaptoethanol and 0.5 M KCl at 1.5 ml/min flow rate. The eluted fractions were finally applied onto a HiLoad 16/60 Superdex200 column (Amersham Bioscience) with 20 mM Tris-HCl (pH8.0), 5mM β-mercaptoethanol and 150 mM NaCl at 1ml/min flow rate.

The protein was concentrated to a final concentration of 5 mg/ml in 20 mM Tris-HCl (pH 7.0) with centrifugal filter devices (Centricon Ultracel YM-30, Millipore Co.). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-2-e. Overexpression and purification of *St*ThrRS-1*

The recombinant plasmids were introduced in *E. coli* B834 (DE3) (Novagen). The selenomethionine derivative was expressed using minimal M9 medium containing L-selenomethionine. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were

applied onto a Resource Iso column (Amersham Bioscience). Bound proteins were eluted at 6 ml/min flow rate with 20 mM Tris-HCl (Tris hydroxymethyl aminomethane titrated using HCl) (pH 8.0), 5 mM β-mercaptoethanol and a linear gradient of 1.5-0 M ammonium sulfate. The pooled fractions were loaded onto a Resource Q column (Amersham Bioscience), and the bound proteins were eluted at 4 ml/min flow rate with 20 mM Tris-HCl (pH8.0), 5 mM β-mercaptoethanol and 0-1 M NaCl linear gradient. Moreover, the eluted fractions were applied onto a BioScale CH5I column (Bio-Rad) with 10-500 mM potassium phosphate (pH7.0) linear gradient, 5 mM β-mercaptoethanol and 0.5 M KCl at 2 ml/min flow rate. The eluted fractions were finally applied onto a HiLoad 16/60 Superdex75 column (Amersham Bioscience) with 20 mM Tris-HCl (pH8.0), 5mM β-mercaptoethanol and 150 mM NaCl at 1ml/min flow rate.

The protein was concentrated to a final concentration of 5 mg/ml in 20 mM Tris-HCl (pH 7.0) with centrifugal filter devices (Centricon Ultracel YM-30, Millipore Co.). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-2-f. Overexpression and purification of *St*ThrRS-2

The gene encoding *St*ThrRS-2 (gene ID; ST2187, 267 amino acids, 30855 Da, pI 6.93) were inserted into the pET11a vector (Novagen). The recombinant plasmids were transformed in *E. coli* Rosetta-gami (DE3) (Novagen), and the cells were grown in LB culture at 310 K. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Q-Sepharose Fast Flow column (Amersham Bioscience). Bound proteins were eluted at 5 ml/min flow rate with 20 mM Tris-HCl (pH 8.0), 5 mM β-mercaptoethanol and a linear gradient of 50-600 mM NaCl. The pooled fractions were loaded onto a HiTrap Heparin HP column (Amersham Bioscience), and the bound proteins were eluted at 3ml/min flow rate with 20 mM MES (pH6.0), 5 mM β-mercaptoethanol and 0-700 mM NaCl linear gradient. Moreover, the eluted fractions were applied onto a BioScale CH10I (Bio-Rad) with 10-500 mM potassium phosphate (pH7.0) and 5 mM β-mercaptoethanol at 3 ml/min flow rate. The eluted fractions were finally applied onto a HiLoad 16/60 Seperdex75 column (Amersham Bioscience) with 20 mM Tris-HCl (pH 8.0), 5 mM β-mercaptoethanol and 150 mM NaCl at 0.9 ml/min flow rate.

The protein was concentrated to a final concentration of 3 mg/ml in 20 mM Tris-HCl (pH 7.0) with centrifugal filter devices (Centricon Ultracel YM-30, Millipore Co.). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-2-g. Overexpression and purification of *St*ThrRS-2*

The recombinant plasmids were introduced in *E. coli* B834 (DE3) (Novagen). The selenomethionine derivative was expressed using minimal M9 medium containing L-selenomethionine. Following overnight incubation, the cells were harvested by centrifugation at 6000 rpm for 10 min at 277 K, and disrupted by sonication. The cell lysates were incubated at 343 K for 30 min to denature the *E. coli* proteins.

After centrifugation at 18000 rpm for 20 min at 277 K, the supernatants were applied onto a Resource Q column (Amersham Bioscience). Bound proteins were eluted at 5 ml/min flow rate with 20 mM Tris-HCl (Tris hydroxymethyl aminomethane titrated using HCl) (pH 8.0), 5 mM β-mercaptoethanol and a linear gradient of 0-0.6 M NaCl. The pooled fractions were loaded onto a BioScale CH5I column (Bio-Rad), and the bound proteins were eluted at 1.5 ml/min flow rate with 10-300 mM potassium phosphate (pH7.0) linear gradient and 5 mM β-mercaptoethanol. Moreover, the eluted fractions were applied onto a HiTrap Heparin HP column (Amersham Bioscience) with 50 mM MES (pH6.0), 5 mM β-mercaptoethanol and 0-0.6 M NaCl linear gradient at 2 ml/min. The eluted fractions were finally applied onto a HiLoad 16/60 Superdex75 column (Amersham Bioscience) with 20 mM Tris-HCl (pH8.0), 5mM β-mercaptoethanol and 150 mM NaCl at 1ml/min flow rate.

The protein was concentrated to a final concentration of 3 mg/ml in 20 mM Tris-HCl (pH 7.0) with centrifugal filter devices (Centricon Ultracel YM-30, Millipore Co.). The fractions obtained in all the purification steps, as well as the concentrated proteins were analyzed by SDS-PAGE.

6-3. X-ray analysis of *Ap*ThrRS-1*

6-3-a. Crystallization

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton

Research). The conditions under which crystalline precipitates appeared were optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

6-3-b. X-ray data collection and data processing

The crystal obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). X-Ray experiments were performed at the BL-17A beamline of the Photon Factory (PF; Ibaraki, Japan). To apply a Multiple Anomalous Dispersion (MAD) method, it is necessary to measure the data at the exact wavelengths with and without anomalous effects. These wavelengths were estimated by X-ray Absorption Fine Structure (XAFS) measurement, there wavelengths being chosen as the peak, the edge and the remote with the crystal. Their diffraction beams were detected at 95 K with 1˚-360˚ oscillation range using the Quantum 4R CCD detector (ADSC). Exposure time was 5 s for 1° oscillation angle per image. The detector was positioned from the crystal at 167.6mm, 167.4mm and 163.2mm for the three wavelengths, respectively. Diffraction spots were indexed, and their intensity were integrated and scaled using the *HKL*2000 package [1] and the intensity data were converted to the amplitudes using programs in the *CCP4* suite [2].

6-3-c. Structure determination

The MAD method was applied to solve the phase problems of the RtcSe crystal with the program *SHARP/autoSHARP* [3]. The program *autoSHARP* is the automated structure solution pipeline, built around the program *SHELXD* [4] for location of heavy atoms, the program *SHARP* for the heavy atom refinement, the program *SOLOMON* [5] for density modification and the program *ARP/wARP* [6] for automated structural model building and refinement using the program *REFMAC 5* [7].

6-4. X-ray analysis of *Ap*ThrRS-1

6-4-a. Crystallization

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by

changing the pH of the buffer.

6-4-b. X-ray data collection and data processing

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the BL-6A beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 2 s for 1° oscillation angle per image. Detector distance from the crystal and the using wavelength were 148.9 mm and 1.00 Å, respectively. Diffraction spots were indexed, and the intensities were integrated and scaled using the *HKL*2000 package [1]. The intensity data were converted to amplitudes using programs in the *CCP4* suite [2].

6-4-c. Structure determination

The initial structures of the native crystals were derived by the molecular replacement method with the program *AMoRe* [11] using the *Ap*ThrRS-1* structure as a probe.

All the atomic parameters were refined with the program *REFMAC 5* [7]. The structural modification was performed with the program *COOT* [8]. These two operations were complementally repeated. The structure was validated with a Ramachandran plot using the program *PROCHECK* [9].

6-5. X-ray analysis of *Ap*ThrRS-2

6-5-a. Crystallization

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen and Crystal Screen 2 kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

6-5-b. X-ray data collection and data processing

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The

diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 315 CCD detector (ADSC) at the BL-5A beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 2 s for 1° oscillation angle per image. Detector distance from the crystal and the used wavelength were 451.9 mm and 1.00 Å, respectively. Diffraction spots were indexed, and the intensity data integrated and scaled using the *HKL*2000 package [1].

6-5-c. Structure determination

The initial structures of the native crystals were derived by the molecular replacement method with the program *AMoRe* [11], *MOLREP* [12] or *Phaser* [10]. The ThrRS structures from *E. coli* [13, 14] and *Staphylococcus aureus* [15], truncated according to the sequence alignment, and the editing domain structure from *Pyrococcus abyssi* [16] were used as searched model.

6-6 X-ray analysis of *St*ThrRS-1

6-6-a. Crystallization

Crystallization trials were performed using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen, Crystal Screen 2 and PEG/Ion Screen kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

6-6-b. X-ray data collection and data processing

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 2 s for 1° oscillation angle per image. Detector distance from the crystal and the using wavelength were 301.4 mm and 1.00 Å, respectively. Diffraction spots were indexed, and the intensities integrated and scaled using the *HKL*2000 package [1]. The intensity data were converted to amplitudes using programs in the *CCP4* suite [2].

6-6-c. Structure determination

The initial structures of the native crystals were derived by the molecular replacement method with the program *AMoRe* [11], *MOLREP* [12] or *Phaser* [10]. The ThrRS structures from *E. coli* [13, 14] and *Staphylococcus aureus* [15], truncated according to the sequence alignment, were used as searched model.

6-7 X-ray analysis of *St*ThrRS-1*

6-7-a. Crystallization

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen, Crystal Screen 2 and PEG/Ion Screen kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

6-7-b. X-ray data collection and data processing

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). X-Ray experiments were performed at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan).To apply a Multiple Anomalous Dispersion (MAD) method, it is necessary to measure the data at the exact wavelengths with and without anomalous effects. These wavelengths were estimated by X-ray Absorption Fine Structure (XAFS) measurement, there wavelengths being chosen as the peak, the edge and the remoteH and remoteL. Their diffraction beams were detected at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC). Exposure time was 2 s for 1° oscillation angle per image. The detector was positioned from the crystal at 286.0mm, 285.9mm 290.8mm and 284.6mm for the four wavelengths, respectively. Diffraction spots were indexed, and their intensities were integrated and scaled using the *HKL*2000 package [1]. The intensity data were converted to the amplitudes using programs in the *CCP4* suite [2].

6-7-c. Structure determination

The MAD method was applied to solve the phase problems of the RtcSe crystal with the program *SHARP/autoSHARP* [3]. The program *autoSHARP* is the automated structure solution pipeline, built around the program *SHELXD* [4] for location of heavy

atoms, the program *SHARP* for the heavy atom refinement, the program *SOLOMON* [5] for density modification and the program *ARP/wARP* [6] for automated structural model building and refinement using the program *REFMAC 5* [7].

6-8. X-ray analysis of *St*ThrRS-2

6-8-a. Crystallization

Crystallization trials were carried out using the hanging-drop vapor-diffusion method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen, Crystal Screen 2 and PEG/Ion Screen kits (Hampton Research). The conditions under which crystalline precipitates appeared were optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

6-8-b. X-ray data collection and data processing

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). The diffraction measurements were performed at 95 K with 1˚-360˚ oscillation range using the Quantum 210 CCD detector (ADSC) at the BL-5A beamline of the Photon Factory (PF; Ibaraki, Japan). Exposure time was 2 s for 1° oscillation angle per image. Detector distance from the crystal and the used wavelength were 191.1 mm and 1.00 Å, respectively. Diffraction spots were indexed, and the intensities integrated and scaled using the *HKL*2000 package [1]. The intensity data was converted to amplitudes using programs in the *CCP4* suite [2].

6-8-c. Structure determination

The initial structures of the native crystals were derived by the molecular replacement method with the program *AMoRe* [11], *MOLREP* [12] or *Phaser* [10]. The ThrRS structures from *E. coli* [13, 14] and *Staphylococcus aureus* [15] and the editing domain structure from *Pyrococcus abyssi* [16], truncated according to the sequence alignment, were used as searched model.

6-9. X-ray analysis of *St*ThrRS-2*

6-9-a. Crystallization

Crystallization trials were carried out using the hanging-drop vapor-diffusion

method by mixing equal volumes (1 μl) of the protein and the reservoir solutions, and then by equilibrating the mixed solutions against 700 μl reservoir solution in a 24-well plate (Stem Corporation) at 293 K. Initial screening for potential crystallization conditions was performed with the Crystal Screen, Crystal Screen 2 and PEG/Ion Screen kits (Hampton Research). The conditions under which crystalline precipitates appeared were further optimized by changing the concentrations of the protein, precipitant and salt, and by changing the pH of the buffer.

6-9-b. X-ray data collection and data processing

Crystals obtained were soaked in the respective reservoir solutions containing 40% glycerol for 30 s, and mounted on a CryoLoop (Hampton Research). X-Ray experiments were performed at the NW12 beamline of the Photon Factory (PF; Ibaraki, Japan). To apply a Multiple Anomalous Dispersion (MAD) method, it is necessary to measure the data at the exact wavelengths with and without anomalous effects. These wavelengths were estimated with X-ray Absorption Fine Structure (XAFS) measurement, there wavelengths being chosen as the peak, the edge and the remoteH and remoteL. The diffraction measurements were performed at 95 K with 1˚-180˚ oscillation range using the Quantum 210 CCD detector (ADSC). Exposure time was 2 s for 1° oscillation angle per image. The detector was positioned from the crystal 232.0mm, 231.9mm 236.1mm and 236.9mm for four wavelengths, respectively. Diffraction spots were indexed, the intensities were integrated and scaled using the *HKL*2000 package [1]. The intensity data were converted to amplitudes using programs in the *CCP4* suite [2].

6-9-c. Structure determination

The MAD method was applied to solve the phase problems of the RtcSe crystal with the program *SHARP/autoSHARP* [3]. The program *autoSHARP* is the automated structure solution pipeline, built around the program *SHELXD* [4] for location of heavy atoms, the program *SHARP* for the heavy atom refinement, the program *SOLOMON* [5] for density modification and the program *ARP/wARP* [6] for automated structural model building and refinement using the program *REFMAC 5* [7].

6-10. Reference

1. Z. Otwinowski & W. Minor, (1997) “Processing of X-ray Diffraction Data Collected in Oscillation Mode” *Methods in Enzymology*, **276**, 307-326.
2. Collaborative Computational Project, Number 4, 1994.
3. C. Vonrhein, E. Blanc, P. Roversi & G. Bricogne, (2007) “Automated structure solution with

autoSHARP" *Methods Mol. Biol*., **364**, 215-230.
4. T. R. Schneider and G. M. Sheldrick, (2002) "Substructure solution with *SHELXD* 2002" *Acta Cryst*. **D58**, 1772-1779.
5. J. P. Abrahams & A. G. W. Leslie, (1996) "Methods used in the structure determination of bovine mitochondrial F_1 ATPase" *Acta Cryst.*, **D52**, 30-42.
6. A. Perrakis, R. M. Morris, & V. S. Lamzin, (1999) "Automated protein model building combined with iterative structure refinement" *Nature Struct.Biol*., **6**, 458-463.
7. G. N. Murshudov, A. A.Vagin & E. J.Dodson, (1997) "Refinement of Macromolecular Structures by the Maximum-Likelihood Method" *Acta Cryst.* **D53**, 240-255.
8. P. Emsley & K. Cowtan (2004) "Coot: model-building tools for molecular graphics" *Acta Cryst.*, **D60**, 2126-2132.
9. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, (1993) "PROCHECK: a program to check the stereochemical quality of protein structures" *J. App. Cryst.*, **26**, 283-291.
10. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L.C. Storoni & R.J. Read. (2007) "*Phaser* crystallographic software" *J. Appl. Cryst.*, **40**, 658-674.
11. J. Navaza, (1994) "AMoRe: an automated package for molecular replacement" *Acta Cryst*. **A50**, 157-163.
12. A. Vagin, A. Teplyakov, (1997) "MOLREP: an automated program for molecular replacement" *J. Appl. Cryst*., **30**, 1022-1025.
13. R. Sankaranarayanan, A. C. Dock-Bregeon, P. Romby, J. Caillet, M. Springer, B. Rees, C. Ehresmann, B. Ehresmann, D. Moras (1999) "The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site" *Cell*, **97**, 371-81.
14. R. Sankaranarayanan, A. C. Dock-Bregeon, B. Rees, M. Bovee, J. Caillet, P. Romby, C. S. Francklyn, D. Moras (2000) "Zinc ion mediated amino acid discrimination by threonyl-tRNA synthetase" *Nature Struct. Biol*., 7, 461-465.
15. A. Torres-Larios, R. Sankaranarayanan, B. Rees, A. C. Dock-Bregeon, D. Moras (2003) "Conformational movements and cooperativity upon amino acid, ATP and tRNA binding in threonyl-tRNA synthetase" *J. Mol. Biol*., **331**, 201-211.
16. T. Hussain, S. P. Kruparani, B. Pal, A. C. Dock-Bregeon, S. Dwivedi, M. R. Shekar, K. Sureshbabu, R. Sankaranarayanan (2006) "Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea" *EMBO J*., **25**, 4152-4162.

< Chapter 7. Results of X-ray analyses >

7-1. Overexpression and purification

7-1-a. Purified quality of *Ap*ThrRS-1*

After the final purification, the molecular weight of *Ap*-ThrRS-1* was estimated by an SDS-PAGE. The result is shown in Figure 7-1 (*a*). The target protein appears as a main band at a position corresponding to the molecular weight (53684 Da) calculated from the amino acid sequence of *Ap*ThrRS-1*. Thus it is considered the protein was highly purified for crystallization.

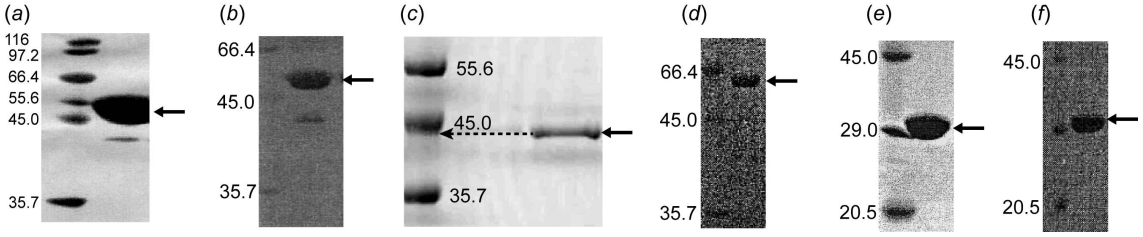


Figure 7-1. SDS-PAGEs of the target protein. The left and right lanes contain proteins of molecular weight markers and the protein, respectively. An arrow indicates the protein. Values indicate approximate molecular weights (kDa) at the maker bands. (*a*), (*b*), (*c*), (*d*), (*e*) and (*f*) indicate SDS-PAGEs for *Ap*ThrRS-1*, ApThrRS-1, *Ap*ThrRS-2, *St*-ThrRS-1, *St*-ThrRS-1*, *St*-ThrRS-2 and *St*-ThrRS-2*, respectively.

7-1-b. Purified quality of *Ap*ThrRS-1

Figure 7-1 (*b*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (53122Da) calculated from the amino acid sequence of *Ap*ThrRS-1. Thus it is considered the protein was highly purified for crystallization.

7-1-c. Purified quality of *Ap*ThrRS-2

Figure 7-1 (*c*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (45003 Da) calculated from the amino acid sequence of *Ap*ThrRS-2. Thus it is considered the protein was highly purified for crystallization.

7-1-d. Purified quality of *St*ThrRS-1

Figure 7-1(*d*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (63110 Da) calculated from the amino acid sequence of *St*-ThrRS-1. Thus it is considered the protein was highly purified for crystallization.

7-1-e. Purified quality of *St*ThrRS-1*

Figure 7-1 (*e*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (63673 Da) calculated from the amino acid sequence of *St*-ThrRS-1*. Thus it is considered the protein was highly purified for crystallization.

7-1-f. Purified quality of *St*ThrRS-2

Figure 7-1 (*f*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (30855 Da) calculated from the amino acid sequence of *St*-ThrRS-2. Thus it is considered the protein was highly purified for crystallization.

7-1-g. Prified quality of *St*ThrRS-2*

Figure 7-1 (*g*) shows an SDS-PAGE of the purified target protein. The main band appears at a position corresponding to the molecular weight (30949 Da) calculated from the amino acid sequence of *St*-ThrRS-2*. Thus it is considered the protein was highly purified for crystallization.

7-2. X-ray analysis of *Ap*ThrRS-1*

7-2-a. *Ap*ThrRS-1* Crystals

The optimized condition, under which suitable crystals for X-ray analysis appeared after one week, is that the protein solution contains with 6 mg/ml *Ap*ThrRS-1* in 20 mM potassium phosphate (pH7.0) buffer and the reservoir solution contains 1.5 M ammonium sulfate in 0.1 M HEPES (pH7.5) buffer and were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293K. The crystal has a size of approximately 300×150×50 μm, shown in Figure 7-2-a. This crystal was used for X-ray diffraction experiment.

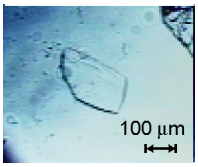


Figure 7-2-a. A photo of the *Ap*ThrRS-1* crystal.

7-2-b. X-ray data of the *Ap*ThrRS-1* crystal

A spectrum of f' and f'' near the absorption edge wavelength of Se atoms (Figure 7-2-b-1) is drawn based on XAFS measurement of the crystal. In this spectrum, the four wavelengths are assigned as the peak (0.97898Å), the edge (0.97966Å) and the remote

(1.00Å). The same crystal was used throughout X-ray experiments. A statistics of the observed intensity data and the crystallographic data are given in Table 3-2-b. Figure 3-2-b-2 shows an example of diffraction images taken using the "peak" wavelength.

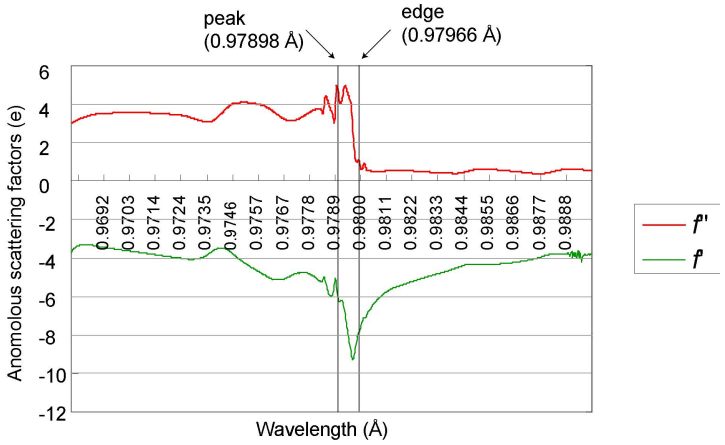


Figure 7-2-b-1 Variation of real (*f'*; green line) and imaginary parts (*f''*; red line) of anomalous scattering factors of selenium atoms in the *Ap*ThrRS-1* crystal as a function of wavelength.

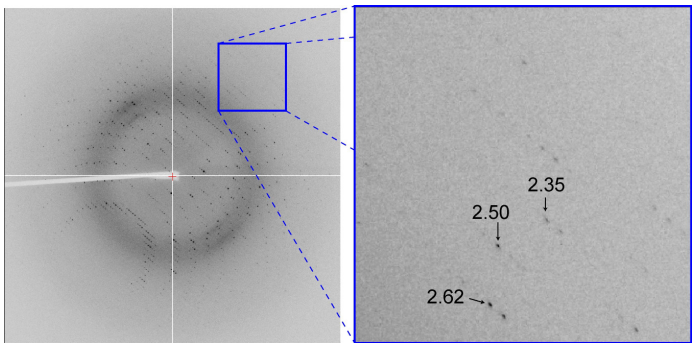


Figure 7-2-b-2. An X-ray diffraction image of the *Ap*-ThrRS-1* crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

In the measurement at the "peak" wavelength, the observed intensity data were processed to 2.5 Å resolution. The data completeness and an R_{marge} value are 100% and 8.3%, respectively. The numbers of observed and unique reflections are 244125 and 16854, respectively. The crystal belongs to the $C222_1$ space group. The cell constants are a=81.8, b=103.6 and c=112.5 Å.

In the measurement at the "edge" wavelength, the observed intensity data were processed to 2.5 Å resolution. The data completeness and an R_{marge} value are 100% and 8.4%, respectively. The numbers of observed and unique reflections are 246544 and 17017, respectively.

In the measurement at the "remote" wavelength, the observed intensity data were

processed to 2.5 Å resolution. The data completeness and an R_{marge} value are 100% and 8.3%, respectively. The numbers of observed and unique reflections are 240627 and 16636, respectively.

Table 7-2-b. A statistics of the observed intensity data and the crystallographic data of the *Ap*ThrRS-1* crystal at the three wavelengths

	peak	edge	remote
Wavelength (Å)	0.97898	0.97966	1.00
Resolution (Å)	50-2.5 (2.59-2.50)	50-2.5 (2.59-2.50)	50-2.5 (2.59-2.50)
Space group	$C222_1$	$C222_1$	$C222_1$
Unit cell (Å)	a = 81.8, b = 103.6, c = 112.5	a = 81.9, b = 103.6, c = 112.6	a = 82.0, b = 103.8, c = 112.7
Observed reflections	244125	246544	240627
Unique reflections	16854	17017	16636
Completeness (%)	100 (100)	100 (100)	100 (100)
Redundancy	14.5 (14.7)	14.5 (14.7)	14.5 (14.6)
I/σ	49.3 (17.3)	58.0 (16.0)	49.5 (14.4)
R_{merge} (%)	8.3 (20,9)	8.4 (23.3)	8.0 (25.8)

$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

7-2-c. Phase determination and structural refinement of *Ap*ThrRS-1*

An acceptable Z value (the number of the proteins in the asymmetric unit of the crystal) was estimated to be 1, the V_M value to be 2.23 Å^3/Da and the solvent content to be 44.8%. Twelve Se atoms were found for the one molecule in the asymmetric unit, because *Ap*ThrRS-1* contains twelve selenomethionine residues.

Using the data sets measured at three wavelengths, seven Se atoms were found in the asymmetric unit, and then the best solution showed that the correlation coefficient between F_o and F_c was 0.357. After the heavy atom position refinements, a total of eleven selenium atoms with occupancies greater than 50% were found finally. An electron density map calculated with the estimated phase angles was modified by solvent flattering technique. The solvent content was fixed to be 34% to include solvent molecules in the asymmetric unit. The final correlation coefficient in $|E^2|$ was 0.771. After structure model building, more than 96% residues of the protein were successfully assigned.

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. In asymmetric unit, 465 residues of the 471residues are assigned except for the N-terminal six residues. 305 water molecules, one zinc atom and two sulfate ions are also assigned. The final R factor and R_{free} values are 17.1% and 19.3%, respectively, indicating that the structure was refined reasonably. A statistics of the structural refinement is given in Table 7-2-c-1. The refined structure is well fitted on the map. The final $2|F_o|-|F_c|$ map, contoured at 1.5 σ level, is shown in Figure 7-2-c-2. A

Ramachandran plot is shown in Figure 7-2-c-1. In the Ramachandran plot, 378 (92.6%) residues are in the most favored geometry, 25 (6.1%) in the allowed geometry, 2 (0.5%) in generally allowed geometry and 3 (0.7%) in disallowed geometry, respectively. The wire models with the zinc atom and two sulfate ions are shown in Figure 7-2-c-3.

Table 7-2-c-1. A statistics of the structural refinement of *Ap*ThrRS-1*

Resolution (Å)	47.1-2.50
Used reflections	16761
R (%)	17.1
R_{free} (%)	19.3
r.m.s bond length (Å)	0.02
r.ms. angles (˚)	2.1
Protein atoms	3708
Water molecules	305
SO_4^{2-}	2
Zn^{2+}	1

$R = 100 \times \sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

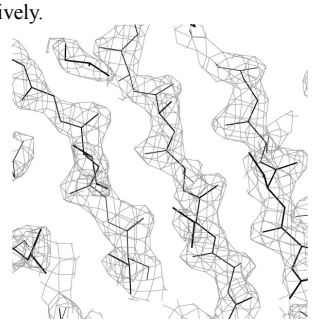


Figure 7-2-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *Ap*ThrRS-1* crystal, contoured at 1.5 σ level (colored in gray).

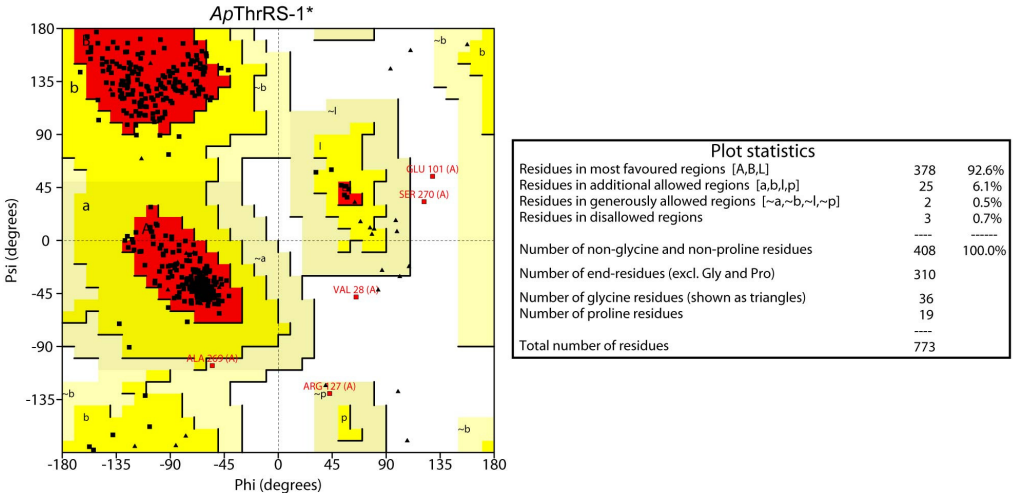


Figure 7-2-c-2. A Ramachandran plot of *Ap*-ThrRS-1*. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

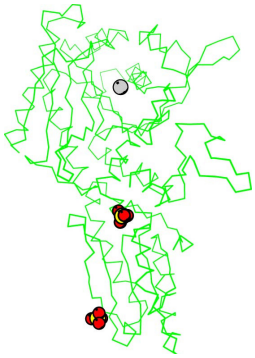


Figure 7-2-c-3. The molecule in the asymmetric unit of the *Ap*ThrRS-1* crystal with the one zinc atom (gray sphere) and two sulfate ions (yellow and red sphere indicating sulfur and oxygen atoms). Green wire models indicate the molecule.

7-2-d. Dimer structure of *Ap*ThrRS-1*

The two molecules (subunits) A and A′ related by a crystallographic two-fold symmetry along the *a* axis in the *Ap*ThrRS-1* crystal form a dimer, as shown in Figure 7-2-d-1. This dimer formation is of typical as a member of the Class II. The dimer has a size of approximately 70×80×80 Å. The atomic distances between the subunits of the dimer are listed in Table 7-2-d-1. The molecular structure of the dimer will be described in the next section (7-3) and discussed in the next chapter8.

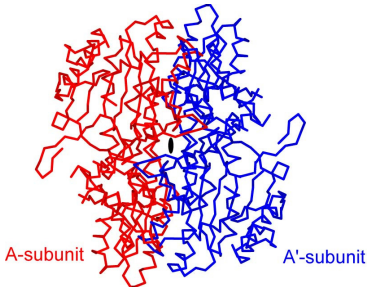


Figure 7-2-d-1. Two subunits of a dimer related by a crystallographic two-fold symmetry along the *a* axis in the *Ap*ThrRS-1* crystal. The each subunit is indicated by red or blue wire models. An ellipsoid indicates a crystallographic two-fold axis.

Table 7-2-d-1. Interactions of the dimer related a crystallographic two-fold symmetry along the *a* axis in the *Ap*ThrRS-1* crystal.

Between the anticodon binding domains				
Direct hydrogen bonds				
A-subunit	Distance(Å)		A′-subunit	
Nε-Glu459	2.7		O-Ala460	
Oε-Glu459	3.2		Arg462-N	
O-Ala460	2.7		Nε-Glu459	
Arg462-N	3.2		Oε-Glu459	
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance(Å)	Water	Distance (Å)	A′-subunit
O-Ser457	2.7	W10	2.7	O-Arg472
Oε-Glu459	2.9	W305	3.3	Oε-Glu459
Oε-Glu459	3.3		2.9	Oε-Glu459

O-Arg472	2.7	W10	2.7	O-Ser457

Between the catalytic domain and the opposite anticodon-binding domain				
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance(Å)	Water	Distance (Å)	A′-subunit
O- Tyr68	2.8	W10	2.9	O-Arg472
O-Arg472	2.9	W10	2.8	O- Tyr68

Between the catalytic domains		
Direct hydrogen bonds		
A-subunit	Distance(Å)	A′-subunit
Oη-Tyr53	3.4	N-Phe139
O- Thr37	3.1	N- Ser78
Oδ-Asp17	3.3	Nζ-Lys125
Oγ-Ser148	2.9	Oε-Glu101
N -Ser148	3.3	Oε-Glu101
Oε-Glu72	2.6	Nε-Arg141
Oε-Glu72	3.2	Nζ-Arg141
Nζ-Lys56	2.7	Oε-Glu57
	3.0	Oε-Glu57
N-Phe139	3.4	Oη-Tyr53
N- Ser78	3.1	O- Thr37
Nζ-Lys125	3.3	Oδ-Asp17
Oε-Glu101	2.9	Oγ-Ser148
Oε-Glu101	3.3	N -Ser148
Nε-Arg141	2.6	Oε-Glu72
Nζ-Arg141	3.2	Oε-Glu72
Oε-Glu57	2.7	Nζ-Lys56
Oε-Glu57	3.0	

Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance (Å)	Water	Distance (Å)	A′-subunit
Oδ-Asp17	3.4	W181	2.7	Oε-Glu122
Nε-His143	2.7	W21	2.8	N- Ile75
O-Val70	2.9	W7	3.1	O- Glu72
O-Phe44	2.9	W138	2.9	
Oε-Glu122	2.7	W181	3.4	Oδ-Asp17
N- Ile75	2.8	W21	2.7	Nε-His143
O- Glu72	3.1	W7	2.9	O-Val70
	2.9	W138	2.9	O-Phe44

Hydrophobic contacts	
A-subunit	A′-subunit
Val20 Val26 Phe29 Phe33 Phe36 Val41 Lue43 Lue46 Tyr96	Ile76 Tyr114 Lue117 Lue118 Lue120 Val184 Phe291
Ile76 Tyr114 Lue117 Lue118 Lue120 Val184 Phe291	Val20 Val26 Phe29 Phe33 Phe36 Val41 Lue43 Lue46

	Tyr96

* A′-subunit indicates the counter molecule of A-subunit related by a crystallographic two-fold symmetry along the *a* axis.

7-2-e. Dimer packing in the *Ap*ThrRS-1 crystal

The dimer packing in the unit cell of the *Ap*ThrRS-1* crystal is shown in Figure 7-2-e-1. The dimers are arranged by crystallographic two-fold screw symmetries along the *c* axis. These columns are arranged by crystallographic two-fold symmetries along *a* and *b* axes, respectively. The interacting regions of the two dimers in a column is referred to Region-1, and those related by a periodic shift along the *a* axis is Region-2, and along an arbitrary direction to resolve the overlapped dimers near the (-1 1 0) direction are Region-3 and Region-4. Their atomic interactions are listed in Table 7-2-e-1~2.The structural details will be described in next section (7-3).

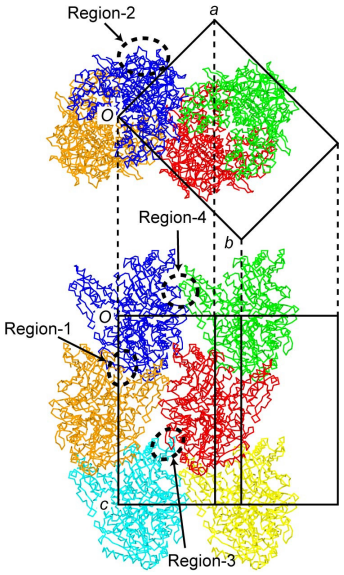


Figure 7-2-e-1. Dimer packing in the unit cell of the *Ap*ThrRS-1* crystal, viewed down the *c* axis (upper) and along an arbitrary direction to resolve the overlapped dimers near the (-1 1 0) direction (bottom).

Table 7-2-e-1. Interactions between the *Ap*ThrRS-1* subunits related by the crystallographic 2_1 symmetry along the *c* axis (Region-1) (Figure 7-2-e-1).

Region-1		
Direct hydrogen bond		
A-subunit	Distance (Å)	A′-subunit
Nη-Arg354	2.6	Oγ-Ser128

A′-subunit indicates the molecule related by a crystallographic 2_1 symmetry along the *c* axis.

Table 7-2-e-1. Interactions of the *Ap*ThrRS-1* subunits between the two screwed columns related by a periodic shift along the *a* axis (Region-2) and along along an arbitrary direction to resolve the overlapped dimers near the (-1 1 0) direction (Region-3 and Region-4) (See Figure 7-2-e-1).

Region-2		
Direct hydrogen bonds		
A-subunit	Distance (Å)	A′-subunit
O-Gly215	2.7	Nδ-His87
Nδ-His87	2.7	O-Gly215

Region 3		
Direct hydrogen bond		
A-subunit	Distance (Å)	A′-subunit
Nε-His390	3.2	Oε-Glu251

Region-4				
Direct hydrogen bonds				
A-subunit	Distance (Å)			A′-subunit
Nζ-Lys424	3.2			Oγ-Ser148
Nη-Arg416	3.3			O-Gly149
Nζ-Lys411	2.7			O-Ser273
O-Arg409	2.9			N-Glu275
O-Lys411	3.4			Nζ-Lys274
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance (Å)	Water	Distance (Å)	A′-subunit
N-Arg427	3.3	W65	2.7	O-Ser148
Oε-Glu410	2.7	W91	2.9	Nη-Arg319
			3.3	N-Gln277
Hydrogen bond interactions intermediated by an ion molecule				
A-subunit	Distance (Å)	Ion	Distance (Å)	A′-subunit
Nζ-Lys406	3.0	SO_4^{2-}	2.7	Nη-Arg156
Nη-Arg427	3.2		3.1	Nε-Arg156
Nη-Arg427	3.2			

7-3. X-ray analysis of *Ap*ThrRS-1

7-3-a. *Ap*ThrRS-1 Crystals

The optimization condition, under which suitable crystals for X-ray analysis appeared after one week, is that the protein solution containing 6 mg/ml *Ap*ThrRS-1 in 20 mM potassium phosphate (pH 7.0) buffer, and the reservoir solution containing 2.0 M ammonium sulfate in 0.1 M HEPES (pH 7.75) buffer were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293K. The crystal has a size of approximately 250×250×100 μm, as shown in Figure 7-3-a. This crystal was used for X-ray diffraction experiment.

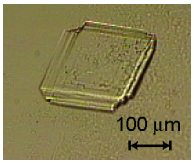


Figure 7-3-a. A photo of the *Ap*ThrRS-1 crystal.

7-3-b. X-ray data of the *Ap*ThrRS-1 crystal

A diffraction image of the *Ap*ThrRS-1 crystal is shown in Figure 7-3-b. Although the diffraction spots were observed till 2.26 Å resolution, their intensity data were processed to 2.3 Å resolution. The crystal belongs to the space group $C222_1$ with cell constants a = 81.5, b = 103.7 and c = 112.5 Å. The numbers of observed and unique reflections are 155186 and 21565, respectively. The data completeness and an R_{marge} value are 100% and 9.6%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 7-3-b.

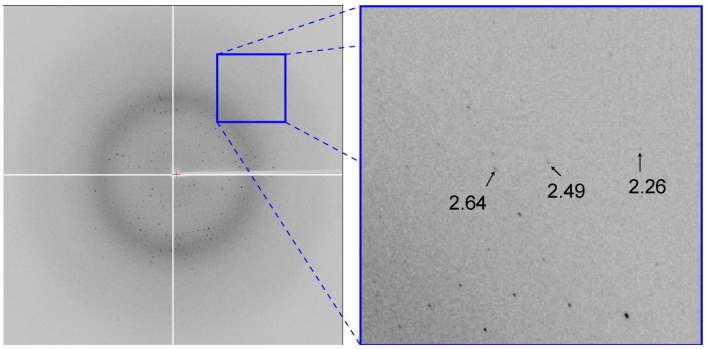


Figure 7-3-b. An X-ray diffraction image of *Ap*ThrRS-1 crystal, taken with an oscillation angle of 1º (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 3-7-b. A statistics of the observed intensity data and the crystallographic data of the *Ap*ThrRS-1 crystal.

Wavelength (Å)	1.00
Resolution (Å)	50-2.3 (2.38-2.3)
Space group	$C222_1$
Unit cell (Å)	a = 81.5, b = 103.7, c = 112.5
Observed reflections	155186
Unique reflections	21565
Completeness (%)	100 (99.9)
Redundancy	7.2 (7.0)
I/σ	19.9 (6.9)
R_{merge} (%)	9.6 (30.0)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

7-3-c. Phase determination and structural refinement of *Ap*ThrRS-1

An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 1, the V_M value to be 2.24 Å^3/Da and the solvent content to be 45.0%. These crystal habits are same to those of the SeMet crystal. Therefore it is considered that the *Ap*ThrRS-1 crystal is isomolphous with the *Ap*ThrRS-1* crystal. To confirm this, the molecular replacement method was applied using the subunit structure of *Ap*ThrRS-1* as a probe. The program *Molrep* was used for computations. A unique solution containing one subunit in the asymmetric unit has been obtained with the

15-2.3 Å resolution data (Table 7-3-c-1). The Solution 1 represents a correlation value of 0.70 and an *R* value of 38%. The other solutions represent correlation values of 0.10 ~0.19 and *R* values of 61~62%. A unique solution (Solution 1) which showed the highest value of correlation and the lowest *R* value is given in Table 3-7-c-1 together with other solutions.

Table 7-3-c-1. Top six solutions of molecular replacement of *Ap*ThrRS-1*

Solution	σ	ϕ	χ	*x*	*y*	*z*	*Cc*	*R*
1	180.00	0.00	179.84	0.188	0.099	0.426	0.701	38.0
2	91.33	-178.84	62.19	0.366	0.218	0.359	0.188	61.6
3	120.15	-138.46	72.17	0.370	0.332	0.052	0.159	62.1
4	166.34	-89.61	133.80	0.152	0.116	0.358	0.137	61.5
5	166.67	-90.97	131.83	0.422	0.357	0.104	0.136	61.9
6	155.85	-178.86	157.62	0.234	0.095	0.302	0.100	61.8

* σ, ϕ, χ; Polar angles (º), *x*, *y*, *z*; Translations, *Cc*; Correlation, *R*; Crystallographic *R*-facter (%)

The atomic parameters of the crystal structure were refined using the programs *Refmac* and *Coot*. In the asymmetric unit, 465 residues of the 471 residues in the asymmetric unit have been assigned except for the N-terminal six residues. 268 water molecules, one zinc atom and five sulfate ions are also assigned. The finel *R* factor and R_{free} are 18.3% and 24.5%, respectively, indicating that the structure was reasonably refined. A statistics of the structural refinement is given in Table 7-3-c-2. The final $2|F_o|-|F_c|$ electron density map, contoured at 1.5 σ level, is shown in Figure 7-3-c-1. The refined structure is well fitted on the map. A Ramachandran plot is shown in Figure 7-3-c-2. In the Ramachandran plot, 378 (92.6%) residues are in the most favored geometry, 22 (5.4%) in the allowed geometry, 3 (0.7%) in generally allowed geometry, and 5 (1.2%) in disallowed geometry, respectively. The wire models with the zinc atom and five sulfate ions are shown in Figure 7-3-c-3. The interacting regions of the zinc atom and two of five sulfate ions in the *Ap*ThrRS-1 crystal are similar to those in the selenomethionine derivative crystal (see Figure 7-2-c-3).

Table 7-3-c-2. A statistics of the structural refinement of *Ap*ThrRS-1

Resolution (Å)	40.7-2.3
Used reflections	21157
R (%)	18.3
R_{free} (%)	24.5
r.m.s bond length (Å)	0.02
r.ms. angles (˚)	19.2
Protein atoms	3711
Water molecules	268
SO_4^{2-}	5
Zn^{2+}	1

*$R = 100 \times \sum||F_o|-|F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor, respectively.

Figure 7-3-c-1. A part of the final $2|F_o|-|F_c|$ electron density map of the *Ap*ThrRS-1 crystal, contoured at 1.5 σ level (colored in gray).

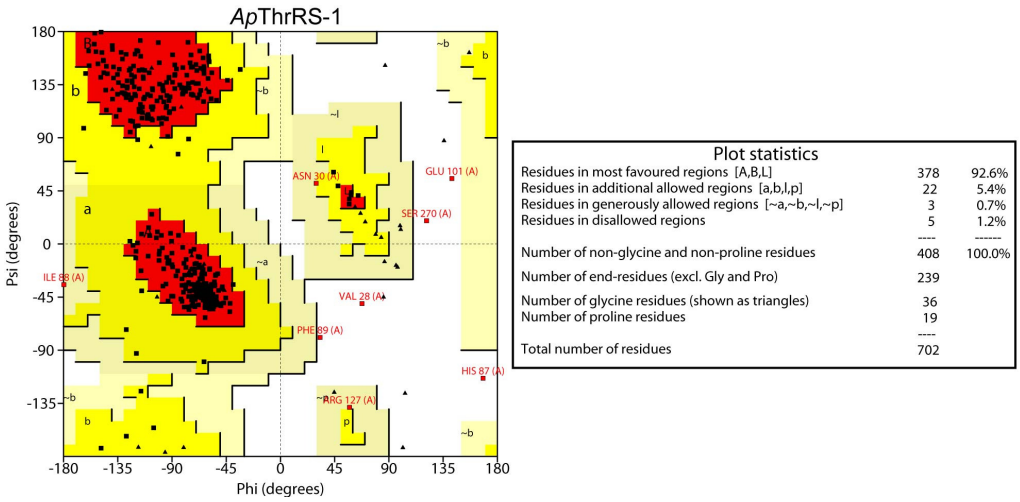


Figure 7-3-c-2. A Ramachandran plot of *Ap*ThrRS-1. Black triangles and squares indicate glycine and the other amino acid residues. Red, yellow, light yellow and white areas indicate the geometry of the main chains in the most favored, in the allowed, in generally allowed and in disallowed state, respectively. Red squares indicate the geometry in generally allowed or in disallowed state except for glycine and the adjacent characters show the residue name, the residue number and the chain name.

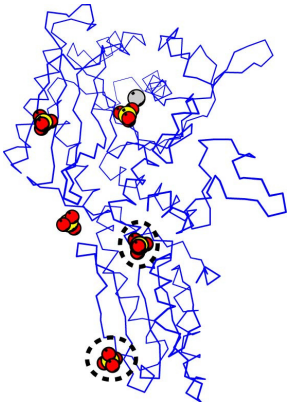


Figure 7-3-c-3. The molecule in the asymmetric unit of the *Ap*ThrRS-1 crystal with the one zinc atom (gray sphere) and five sulfate ions (yellow and red sphere indicating sulfur and oxygen atoms). Blue wire models indicate the subunit. Brocken circles indicate ions at the similar position compared with those in the *Ap*ThrRS-1* crystal (see Figure 7-2-c-3).

7-3-d. Dimer structure of *Ap*ThrRS-1

The two subunits related by a crystallographic two-fold symmetry along the *a* axis in the *Ap*ThrRS-1 crystal also form a typical dimer as a member of the Class II. Superimposed Cα atoms between the two dimers found in the *Ap*ThrRS-1 and in the *Ap*ThrRS-1* crystals shown in Figure 7-3-d-1. The dimer formations are also similar to each other with an r.m.s.d of 0.18 Å. The dimer interface represents 18.4% (3856.7 $Å^2$) of the total surface area (20848.7 $Å^2$) of each subunit. These values are calculated by the *PISA* server [2]. The two molecules are associated by a crystallographic two-fold symmetry along the *a* axis to form a dimer. The dimer formation interface is stabilized by the two catalytic domains. By comparison of the direct hydrogen bond interactions in Table 7-3-d-1 and Table 7-2-d-1, The hydrogen bond interactions between the two subunits of the dimer in the *Ap*ThrRS-1 crystal are formed by the atoms of different residues from those in the *Ap*ThrRS-1* crystal. The atomic distances between the two subunits of the dimer are listed in Table 7-3-d-1.

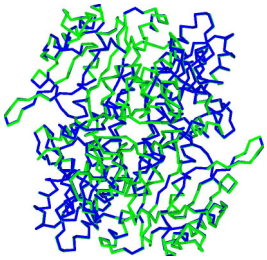


Figure 7-3-d-1. A superimposition of Cα chains of the *Ap*ThrRS-1 dimer and that of *Ap*ThrRS-1*

Table 7-3-d-1. Interactions of the dimer related a crystallographic two-fold symmetry in the *Ap*ThrRS-1 crystal

Between the anticodon binding domains				
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance(Å)	Water	Distance (Å)	A′-subunit
Oε-Glu459	3.3	W259	3.3	Oε-Glu459

Between the catalytic domains		
Direct hydrogen bonds		
A-subunit	Distance(Å)	A′-subunit
O-Tyr14	2.8	Nζ-Lys125
O-Thr30	3.3	N-Phe139
O-Thr37	3.3	N-Ser78
Oγ-Ser45	3.0	Oε-Glu122
Nε-Arg52	3.0	Oε-Glu72
Nη-Arg52	3.2	Oε-Glu72
Oη-Tyr145	2.5	Oε-Glu101
Oγ-Ser148	2.7	Oε-Glu101
Nζ-Lys125	2.8	O-Tyr14
N-Phe139	3.3	O-Thr30
N-Ser78	3.3	O-Thr37
Oε-Glu122	3.0	Oγ-Ser45
Oε-Glu72	3.0	Nε-Arg52
Oε-Glu72	3.2	Nη-Arg52
Oε-Glu101	2.5	Oη-Tyr145
Oε-Glu101	2.7	Oγ-Ser148

Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance (Å)	Water	Distance (Å)	A′-subunit
Oδ-Asp17	2.7	W123	2.9	Nδ-Asn121
Oγ-Ser24	2.8	W124	2.7	Oδ-Asn121
	3.4	W119	2.6	
O-Pro42	3.4	W7	2.7	O-Glu72
Nε-His143	2.8	W20	2.8	N-Ile75
Nδ-Asn121	2.9	W123	2.7	Oδ-Asp17
Oδ-Asn121	2.7	W124	2.8	Oγ-Ser24
	2.6	W119	3.4	
O-Glu72	2.7	W7	3.4	O-Pro42
N-Ile75	2.8	W20	2.8	Nε-His143

Hydrophobic contacts	
A-subunit	A′-subunit
Val20 Val26 Phe29 Phe33 Phe36 Val41 Lue43 Lue46 Tyr96	Ile76 Tyr114 Lue117 Lue118 Lue120 Val184 Phe291
Ile76 Tyr114 Lue117 Lue118 Lue120 Val184 Phe291	Val20 Val26 Phe29 Phe33 Phe36 Val41 Lue43 Lue46 Tyr96

7-3-e. Subunit structure of *Ap*ThrRS-1

A topological diagram of the secondary structure of the *Ap*ThrRS-1 subunit and the three dimensional structure shown in a ribbon model is given in Figure 7-3-e-1. The *Ap*ThrRS-1 structure is composed of the two domains. The N-terminal domain (residues 7-333) is formed with twelve α-helices (α1~α12) and ten β-strands (β1~β10). The three of ten β-strands form a small β-sheet (β1-β3) and the seven antiparallel strands (β4-β10) form a large β-sheet. The twelve α-helices are folded around the large β-sheets, as shown in Figure 3-7-e-1 (*b*). The large β-sheet contains the typical motifs of the Class II aaRSs, so called as motifs 1, 2 and 3 that each motif contains a conserved sequence: gϕxxϕxxPϕϕ for motif 1; (F/Y/H)Rx(E/D) followed 4 to 12 residues away by (R/H)xxxFxxx(D/E) for motif 2; and λxϕgϕgϕeRϕϕϕϕ for motif 3 (x stands for any residues, λ for small amino acids P, G, S, T; ϕ for hydrophobic residues F, Y, W, I, L, V, M, A) [1].

The C-terminal domain (residues 334-471) is made of the one parallel (β14) and four antiparallel β-strand (β11~β13, β15) and five α-helices (α13~α17) (see Figure 7-3-e-1 (*b*)). This folding is similar to the anticodon-binding domain of *Ec*ThrRS.

These domain structures will be discussed again in the next chapter 8 with comparing with those of the *Ec*ThrRS structure.

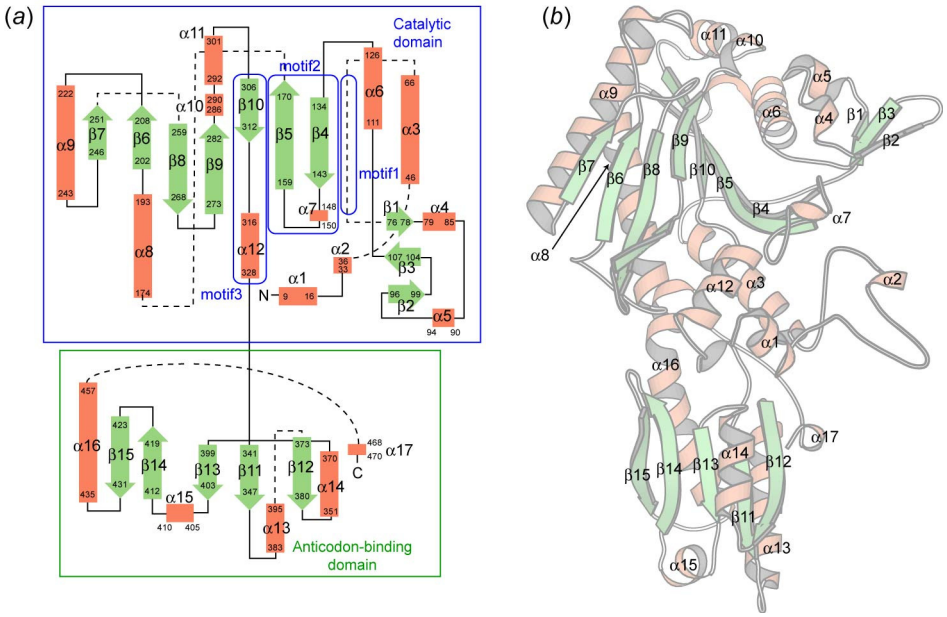


Figure 7-3-e-1. A topological diagram of the secondary structures of *Ap*ThrRS-1 (*a*). The catalytic domain and the anticodon-binding domain are surrounded by blue and green lines, respectively. Bars and arrows indicate α-helices and β-strands, respectively. Characters on or near the arrows and the bars indicate the residue number and codes of secondary structures. The three dimensional structure of *Ap*ThrRS-1 shown in ribbon model (b). α-helices and β-strands are colored in orange and green, respectively. The characters indicate the codes of the secondary structures corresponding to those in (*a*).

The catalytic site of the *Ap*ThrRS-1 subunit is shown in Figure 7-3-e-2. A zinc finger structure is formed with the Cys112, His166 and His310 residues, which is similar to the active site of the *Ec*ThRS in the catalytic domain. These residues are covalently bound to the zinc atom at the distances 2.3 Å (Zn-Sγ of Cys112), 2.0 Å (Zn-Nε of His166) and 2.1 Å (Zn-Nδ atom of His310) in a tetrahedral configuration. The fourth position of the tetrahedron is occupied by a water molecule, which is hydrogen-bonded to Asp164. Near by this site, a sulfate ion is bound through hydrogen bonds with Asp164. It is suggested that the substrate threonine is bound preferentially with high selectivity.

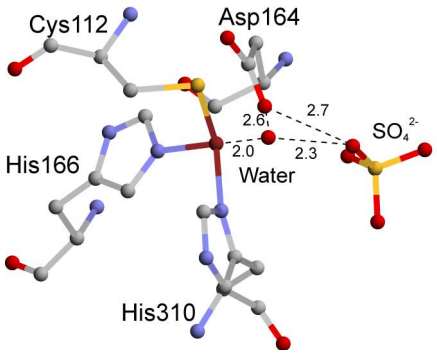


Figure 7-3-e-2. Catalytic site of *Ap*ThrRS-1, shown in ball-and-stick models.

7-3-f. Dimer packing of the *Ap*ThrRS-1 crystal

The dimer packing in the unit cell of the *Ap*ThrRS-1 crystal is shown in Figure 7-3-f-1. This crystal is isomolphous with the *Ap*ThrRS-1* crystal (see Figure 7-2-e-1). The interaction regions are similar to each other in the unit cells. Direct hydrogen bond interactions are not observed in the Region-4 of the *Ap*ThrRS-1 crystal. By comparison of direct hydrogen bonds in Table 7-3-f-1 and Table 7-2-e-1, a hydrogen bond interaction in Region-1 is observed between the Nη atom of Arg354 and the Oγ atom of Ser128 in both of the *Ap*ThrRS-1* and *Ap*ThrRS-1 crystals. The hydrogen interactions in Retion-2 and Region-3 of the *Ap*ThrRS-1 crystal are formed by the atoms of the different residues from those in the regions of the SeMet derivative crystal. The atomic distances of the contacting regions are listed in Table 7-3-f-1 and Table 7-3-f-2.

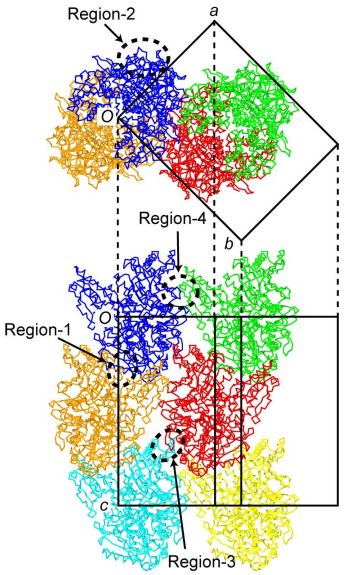


Figure 7-3-f-1. Dimer packing in the unit cell of the *Ap*ThrRS-1 crystal, viewed down the *c* axis (upper) and along along an arbitrary direction to resolve the overlapped dimers near the (-1 1 0) direction (bottom).

Table 7-3-f-1. Interactions between the *Ap*ThrRS-1 subunits related by the crystallographic 2_1 symmetry along the *c* axis (Region-1) (See Figure 7-3-f-1).

Region-1		
Direct hydrogen bond		
A-subunit	Distance (Å)	Chain A′
Nη-Arg354	2.5	Oγ-Ser128

Table 7-3-f-2. I Interactions of the *Ap*ThrRS-1 subunits between the two screwed columns related by a periodic shift along the *a* axis (Region-2) and along an arbitrary direction to resolve the overlapped dimers near the (-1 1 0) direction (Region-3~4) (See Figure 7-3-f-1).

Region-2				
Direct hydrogen bonds				
A-subunit	Distance (Å)			A′-subunit
O-Arg409	2.8			N-Glu275
Nη-Arg416	3.1			O-Gly149
Nζ-Lys424	3.1			Oγ-Ser148
Hydrogen bond interactions intermediated by an ion molecule				
A-subunit	Distance(Å)	Ion	Distance (Å)	A′-subunit
Nζ-Lys406	2.8	SO_4^{2-}	2.6	Nη-Arg156
Nη-Arg427	2.9		2.8	Nε-Arg156
Nη-Arg427	3.0			

Region-3				
Direct hydrogen bond				
A-subunit	Distance (Å)			A′-subunit
Oγ-Ser212	3.0			Oγ-Ser386
Oγ-Asp210	2.5			
Hydrogen bond interactions intermediated by a water molecule				
A-subunit	Distance(Å)	Water	Distance (Å)	A′-subunit
Nε-His390	2.2	W258	2.4	Oε-Glu251
			2.5	Oε-Glu251

7-4. X-ray analysis of *Ap*ThrRS-2

7-4-a. Crystals of *Ap*ThrRS-2

The condition, under which crystals used for X-ray analysis appeared after six months, is that the protein solution containing 5.8 mg/ml *Ap*ThrRS-2 in 20 mM potassium phosphate (pH 7.0) buffer and the reservoir solution of 2.0 M sodium formate in 0.1 M sodium acetate (pH 4.6) buffer and were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The crystal shown in Figure 7-4-a has a size of approximately 50×50×50 μm.

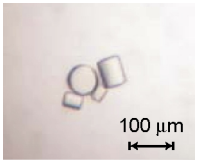


Figure 7-4-a. A photo of the *Ap*ThrRS-2 crystals.

7-4-b. X-ray data of the *Ap*ThrRS-2 crystal

A diffraction image of the *Ap*-ThrRS-2 crystal is shown in Figure 7-4-b. The reflection spots were observed till 3.73 Å resolution. The point groups and cell constants were indexed, but these could not be estimated using the programs *HKL2000* or *Crystal Clear*. These programs may not be able to extract each observed diffraction spots from back ground, because those are weak and broad. The reason is considered that the crystal is too small for X-ray diffraction experiment and the crystal mosaicity is not low. We are trying to get more suitable crystals using the different crystallization conditions.

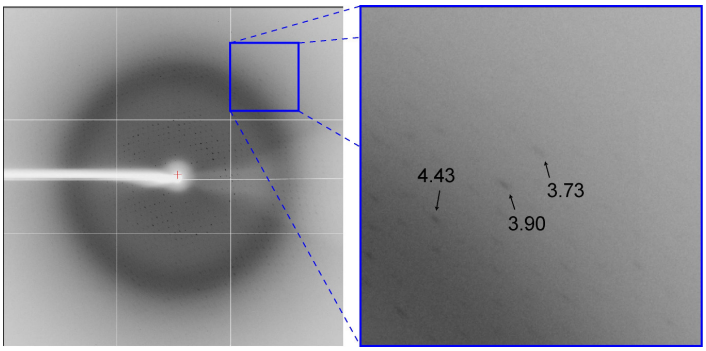


Figure 7-4-b. An X-ray diffraction image of *Ap*ThrRS-2 crystal , taken with an oscillation angle of 1º (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

7-5. X-ray analysis of *St*ThrRS-1

7-5-a. Crystals of *St*ThrRS-1

The condition, under which crystals used for X-ray analysis appeared after two months, is that the protein solution containing 5 mg/ml *St*ThrRS-1 in 20 mM Tris-HCl (pH 7.0) buffer and the reservoir solution containing 0.3 M ammonium chloride and 14% PEG3350 were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The crystal has a size of approximately 300×200×200 μm, as shown in Figure 7-5-a.

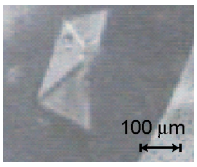


Figure 7-5-a. A photo of the *St*ThrRS-1 crystal.

7-5-b. X-ray data of the *St*ThrRS-1 crystal

A diffraction image of the *St*ThrRS-1 crystal is shown in Figure 7-5-b. Although the reflection spots were observed till 3.87 Å resolution, their intensity data were processed to 4.3 Å resolution. The crystal belongs to the space group $P2_1$ with cell constants a = 68.8, b = 99.7, c = 85.3 Å and β = 95.0°. The numbers of observed and unique reflections are 27253 and 7814, respectively. The data completeness and an R_{marge} value are 98.9% and 20.1%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 7-5-b.

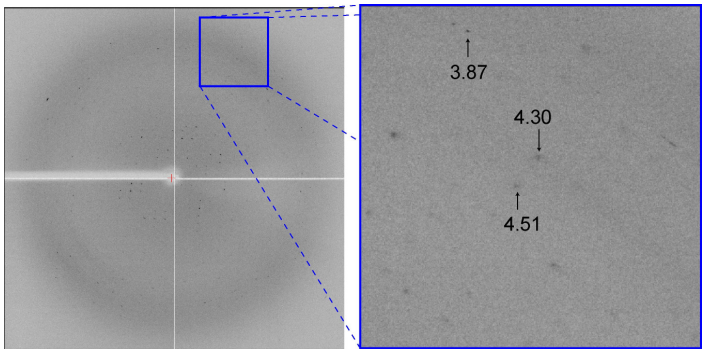


Figure 7-5-b. An X-ray diffraction image of *St*ThrRS-1 crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 7-5-b. A statistics of the observed intensity data and the crystallographic data of the *St*ThrRS-1 crystal

Wavelength (Å)	1.00
Resolution (Å)	50-4.3 (4.45-4.3)
Space group	$P2_1$
Unit cell	
a (Å)	68.8
b (Å)	99.7
c (Å)	85.3
β (º)	95.0
Observed reflections	27253
Unique reflections	7814
Completeness	98.9 (98.1)
Redundancy	3.5 (3.5)
I/σ	7.9 (4.5)
R_{merge} (%)	20.1 (31.1)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl) - I(hkl)| / \sum_{hkl}\sum_i I(hkl)$

7-5-c. Phase angle determination of *St*ThrRS-1

An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 2, the V_M value to be 2.31 Å^3/Da and the solvent content to be 46.7%. The molecular replacement method (*AMoRe*, *Molrep* and *Phaser*) was applied using two following models for the observed data in resolution range of 15-4.3 Å. A model containing the two domains, the catalytic domain and anticodon-binding domain, of *Ec*ThrRS (PDBID 1QF6) was used. And a model containing subunit of *Ap*ThrRS-1 is used as another probe.

In the case using the model containing the catalytic and anticodon-binding domains of *Ec*ThrRS as a probe, all solutions calculated by the program *AMoRe* represent correlation values of 0.32~0.33 and R values of 49~50 %. No unique solution was obtained. Using the program *Molrep*, all solutions represent correlation values of 0.16~0.18 and R values of 57~58%. A unique solution is not obtained. Using the program *Phaser*, two solutions represent RFZ (Z-score (number of standard deviations above the mean value) of rotation function) value of 5.0~5.4 and TFZ (Z-score (number of standard deviations above the mean value of peaks) of translation function) values of 10~12. For a rotation function, the correct solution may be with a Z-score under 4, and will not be found until a translation function is performed and picks out the correct solution. For a translation function the correct solution will generally have a Z-score over 5 and be well separated from the rest of the solutions. The RFZ values shows more than 5 and the TFZ values cannot be well separated to each other. The models form a dimer using one of the two solutions, but the structural refinement is difficult (R was 49.2% and R_{free} was 46.5%) because the maximum resolution of the structure factor data

is 4.3Å and the R_{marge} value is more than 20%.

In the case using the model containing the subunit of *Ap*ThrRS-1 as a probe, all solutions calculated by the program *AMoRe* represent correlation values of 0.26~0.27 and *R* values of 50~51%. No unique solution was obtained. Using the program *Molrep*, all solutions represent correlation values of 0.15 and *R* values of 57~58%. Using the program *Phaser*, all solutions represent RFZ value of 3.5~3.7 and TFZ values of 4.5. No unique solution was obtained.

The amino acid sequence of *St*ThrRS-1 is compared to those of the catalytic and anticodon-binding domains of *Ec*ThrRS with the identity of 45% and to those of *Ap*ThrRS-1 with the identity of 34% (These identity values are calculated by the program *Kalign* with the gap opening penalty of 11% and the gap extending penalty of 0.85%). The structure of *St*ThrRS-1 may be similar to those of *Ec*ThrRS and *Ap*ThrRS-1. However, it is difficult to estimate phase information using the molecular replacement method. One reason is considered the resolution is not so high and the R_{marge} value is more than 20% (see Table 7-5-b). We are trying to crystallize *St*ThrRS-1 using the other conditions and to get data with the higher resolution and lower R_{marge} value. In parallel, the selenomethionine derivative was crystallized and the MAD method were tried to solve the phase angle problem as described in the next section (7-6).

7-6. X-ray analysis of *St*ThrRS-1*

7-6-a. Crystals of *St*ThrRS-1*

The optimized condition, under which crystals for X-ray analysis appeared after two months, is that the protein solution containing 5 mg/ml *St*ThrRS-1* in 20 mM Tris-HCl (pH 7.0) buffer and the reservoir solution containing 0.08 M sodium dihydrogen phosphate monohydrate and 14% PEG4000 were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The crystal has a size of approximately 300×100×100 μm, as shown in Figure 7-6-a. This crystal was used for X-ray diffraction experiment.

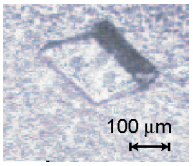


Figure 7-6-a. A photo of the *St*ThrRS-1* crystal.

7-6-b. X-ray data of the *St*ThrRS-1* crystal

A spectrum of f' and f'' around the absorption edge wavelength of Se atoms is drawn based on XAFS measurement of the crystal (Figure 7-6-b-1). In this spectrum, the four wavelengths are assigned as the peak (0.97910 Å), the edge (0.97942 Å), the high remote (0.96418 Å) and the low remote (0.98030 Å). The same crystal was used throughout X-ray experiments. A statistics of the observed intensity data and the crystallographic data are given in Table 6-2-b. Figure 6-2-b-2 shows an example of diffraction images taken using the peak wavelength.

In the measurement of the peak wavelength, the observed intensity data were processed to 4.0 Å resolution. The data completeness and an R_{marge} value are 98.8% and 13.4%, respectively. The numbers of observed and unique reflections are 78967 and 20429, respectively. The crystal belongs to the $P1$ space group. The cell constants are $a = 71.1$ Å, $b = 85.7$ Å, $c = 104.1$ Å, $\alpha = 89.9°$, $\beta = 92.3°$ and $\gamma = 95.4°$.

In the measurement at the edge wavelength, the observed intensity data were processed to 4.0 Å resolution. The data completeness and an R_{marge} value are 99.2% and 18.9%, respectively. The numbers of observed and unique reflections are 79910 and 20472, respectively.

In the measurement at the high remote wavelength, the observed intensity data were processed to 4.0 Å resolution. The data completeness and an R_{marge} value are 99.3% and 17.4%, respectively. The numbers of observed and unique reflections are 80233 and 20559, respectively.

In the measurement at the low remote wavelength, the observed intensity data were processed to 4.0 Å resolution. The data completeness and an R_{marge} value are 98.2% and 15.3%, respectively. The numbers of observed and unique reflections are 75493 and 20446, respectively.

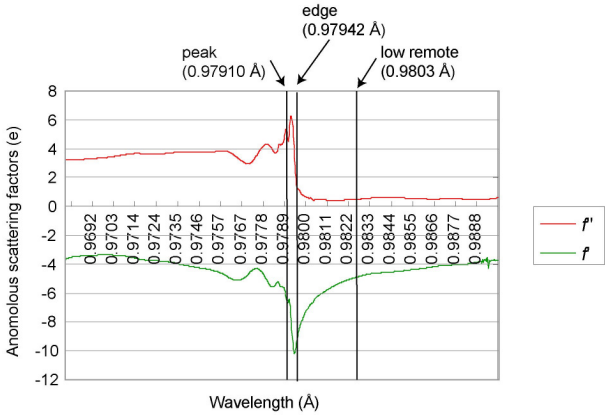


Figure 7-6-b-1. Variation of real (f'; green line) and imaginary parts (f''; red line) of anomalous scattering factors of selenium atoms in the *St*ThrRS-1* crystal as a function of wavelength.

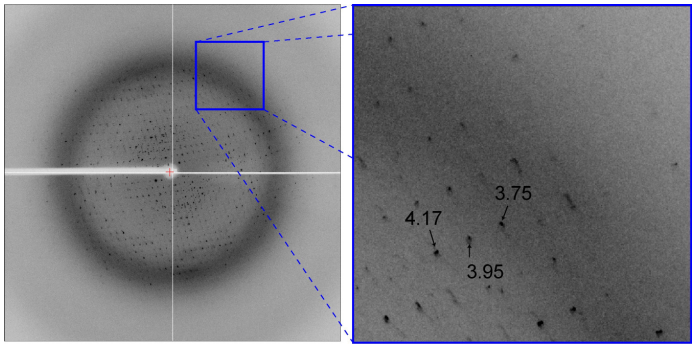


Figure 7-6-b-2. An X-ray diffraction image of *St*ThrRS-1* crystal, taken with an oscillation angle of 1º (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 7-6-b. A statistics of the observed intensity data of the *St*ThrRS-1* crystal and the crystallographic data at four wavelengths

	peak	edge	high remote	low remote
Wavelength (Å)	0.97910	0.97942	0.96418	0.98030
Resolution (Å)	50-4.0 (4.14-4.0)	50-4.0 (4.14-4.0)	50-4.0 (4.14-4.0)	50-4.0 (4.14-4.0)
Space group	*P*1	*P*1	*P*1	*P*1
Unit cell				
a (Å)	71.1	70.9	70.9	70.9
b (Å)	85.7	85.5	85.5	85.6
c (Å)	104.1	103.9	103.8	104.0
α (˚)	89.9	89.1	89.9	89.9
β (˚)	92.3	92.3	92.3	92.3
γ (˚)	95.4	95.2	95.2	95.2
Observed reflections	78967	79910	80233	75493
Unique reflections	20429	20472	20559	20446
Completeness (%)	98.8 (98.1)	99.2 (99.2)	99.3 (99.2)	98.2 (99.2)
Redundancy	3.9 (3.9)	3.9 (3.9)	3.9 (3.9)	3.7 (3.4)
I/σ	17.5 (5.3)	7.3 (2.4)	8.2 (2.5)	8.4 (2.5)
R_{merge} (%)	13.4 (32.7)	18.9 (50.0)	17.4 (47.1)	15.3 (38.4)

$R_{merge} = 100 \times \sum_{hkl}\sum_i |I_i(hkl) - I(hkl)| / \sum_{hkl}\sum_i I(hkl)$

7-6-c. Phase angle determination of *St*ThrRS-1*

The space group was kept to be *P*1 during the X-ray experiments. The average cell constants are a = 71.0 Å, b = 85.6 Å, c = 104.0 Å, α = 89.7º, β = 92.3º and γ = 95.3º. An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 4, the V_M value to be 2.47 Å^3/Da and the solvent content to be 50.2%. Fourty-eight Se atoms will be found for the two molecules in the asymmetric unit, because *St*ThrRS-1* contains twelve selenomethionine residues.

The MAD method using the four to two wavelength data sets and the SAD (single-wavelength anomalous dispersion) method of the peak data set were applied using the program *SHARP* to determine phase angles.

Using the data sets measured at four wavelengths with resolution range of

28.56 - 5.24 Å, fifty-four Se atoms were found in the asymmetric unit, and the correlation coefficient between F_o and F_c was 0.254. After the heavy atom position refinement, a total of ninety-six selenium atoms were found but only two atoms were with occupancies greater than 50% finally.

Using the peak, edge and remoteH data sets with resolution range of 28.56 - 6.0 Å, twelve Se atoms were found in the asymmetric unit, and the correlation coefficient between F_o and F_c was 0.126. After the heavy atom position refinement, a total of ninety-six selenium atoms were found and seventeen atoms were with occupancies greater than 50% finally.

Using the peak and edge data sets with resolution range of 28.56 - 6.0 Å, ninety-six Se atoms were found in the asymmetric unit, and then the solution showed that the correlation coefficient between F_o and F_c was 0.379. After the heavy atom position refinement, a total of one-hundred and twenty-six selenium atoms were found and fourteen atoms with occupancies greater than 50% finally.

A best result, chosen from three trials, has the maximum number of Se atoms with high occupancies. After heavy atom position refinement by maximum likelihood method, the final figure of merit was 0.305 and the final phasing power was 0.348. The figure of merit means the appropriate weight calculated from the phase probability distribution. It wound be 1 for a perfectly defined phase angle, tending to 0 as the error increase. The phasing power means the effect of the derivative in finding the phase angles. If the phasing power exceeds 1.5, the derivative is very useful to phase determination. It is useless if the value is less than 1.0. After solvent flattering technique with solvent content fixed to be 37.5%, the final correlation coefficient in $|E^2|$ was 0.811 and the final *R*-factor between the structure factor amplitudes from the modified map and the observed data was 27.9%. However it is difficult to estimate phase information using these MAD methods because peptide chains cannot be assigned on the density modified map. Therefore, the SAD method using the peak data set was also performed.

Using the peak data set with resolution range of 28.56 - 6.0 Å, forty-six Se atoms were found in the asymmetric unit, and then the solution showed that the correlation coefficient between F_o and F_c was 0.285. After heavy atom position refinement, a total of ninety-six selenium atoms were found but only one atom with occupancies greater than 50% finally. After heavy atom site refinement by maximum likelihood method, the final figure was merit of 0.239 and the final phasing power was 0.651. After solvent flattering technique with solvent content fixed to be 37.5%, the final correlation coefficient in $|E^2|$ was 0.811 and the final *R*-factor between the structure factor amplitudes from the modified map and the observed data was 27.9%. However it is

difficult to assign peptide chains on the density modified map.

It is difficult to estimate phase information using the MAD and SAD method. One reason is considered that the resolution is not so high and the R_{marge} value is more than 10% (see Table 7-6-b). We are crystallizing *St*ThrRS-1* using the other conditions to get data resulting in the higher resolution and lower R_{marge} value.

7-7. X-ray analysis of *St*ThrRS-2

7-7-a. Crystals of *St*ThrRS-2

The optimized condition, under which crystals for X-ray analysis appeared after two months, is that the protein solution containing 3 mg/ml *St*ThrRS-1 in 20 mM HEPES (pH 7.0) buffer and the reservoir solution containing 0.08 M ammonium acetate and 14% PEG4000 were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The crystal has a size of approximately 300×50×50 μm, as shown in Figure 7-7-a. This crystal was used for X-ray diffraction experiment.

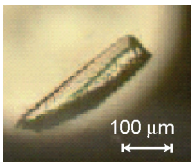


Figure 7-7-a. A photo of the *St*ThrRS-2 crystal.

7-7-b. X-ray data of the *St*ThrRS-2 crystal

A diffraction image of the *St*ThrRS-2 crystal is shown in Figure 7-7-b. Although the reflection spots were observed till 2.25 Å resolution, intensity data were processed to 2.3 Å resolution. The crystal belongs to the space group $P2_12_12_1$ with cell constants a = 60.3, b = 68.0 and c = 134.4 Å. The numbers of observed and unique reflections are 328094 and 25374, respectively. The data completeness and an R_{marge} value are 99.7% and 10.2%, respectively. A statistics of the observed intensity data and the crystallographic data are summarized in Table 7-7-b.

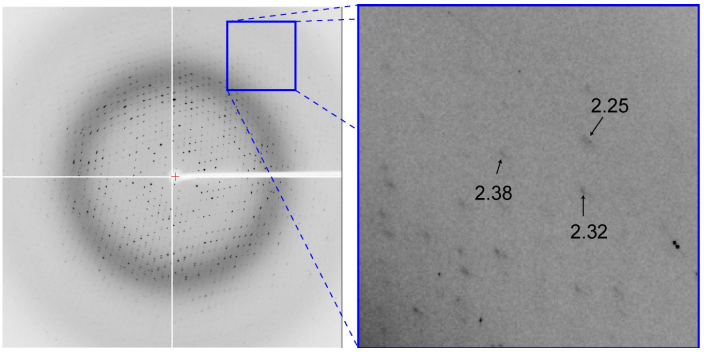


Figure 7-7-b. An X-ray diffraction image of *St*ThrRS-2 crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 7-7-b. A statistics of the observed intensity data and the crystallographic data of the *St*ThrRS-2 crystal

Wavelength (Å)	1.00
Resolution (Å)	50-2.3 (2.38-2.3)
Space group	$P2_12_12_1$
Unit cell (Å)	a = 60.3, b = 68.0, c = 134.4
Observed reflections	328094
Unique reflections	25374
Completeness	99.7 (100)
Redundancy	12.9 (13.1)
I/σ	63.0 (15.4)
R_{merge} (%)	10.2 (30.3)

*$R_{merge} = 100 \times \sum_{hkl}\sum_i|I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

7-7-c. Phase angle determination of *St*ThrRS-2

An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 2, the V_M value to be 2.23 Å^3/Da and the solvent content to be 44.9%. The molecular replacement method (*AMoRe*, *Molrep* and *Phaser*) was applied using three following models for the observed data in resolution range of 15-3.5 Å. A model containing the anticodon-binding domain of *Ec*ThrRS (PDBID 1QF6) was used. And the subunit of *Ap*ThrRS-1 and the dimer of *Pa*ThrRS (PDBID 2HL0) are used as another probes.

In the case using the model containing the anticodon-binding domain of *Ec*ThrRS as a probe, all solutions calculated by the program *AMoRe* represent correlation values of 0.16~0.19 and R values of 60~61%. No unique solution was obtained. In the case using the model the anticodon-binding domain of *Ap*ThrRS-1 as a probe, all solutions represent correlation values of 0.17~0.19 and R values of 52~54%. No unique solution was obtained. In the case using the model containing two editing domain forming dimer of *Pa*ThrRS as a probe, all solutions represent correlation values of 0.14~0.15 and R values of 52~53%. No unique solution was obtained.

Using the program *Molrep*, all solutions represent correlation values of 0.17~0.21 and *R* values of 61~63% in the case using the model containing anticodon-binding domain of *Ec*ThrRS as a probe. A unique solution is not obtained. In the case using the model containing the anticodon-binding domain of *Ap*ThrRS-1 as a probe, all solutions represent correlation values of 0.14~0.16 and *R* values of 60~62%. No unique solution was obtained. In the case using the model containing the two editing domains of *Pa*ThrRS forming a dimer as a probe, all solutions represent correlation values of 0.11~0.12 and *R* values of 61~63%. No unique solution was obtained.

Using the program *Phaser*, all solutions represent RFZ value of 3.5~3.6 and TFZ values of 4.5~5.8 in the case using the model containing the anticodon-binding domain of *Ec*ThrRS as a probe. A unique solution is not obtained. In the case using the model containing the anticodon-binding domain of *Ap*ThrRS-1 as a probe, all solutions represent all RFZ values of 3.7 and TFZ values of 4.6~5.8%. No unique solution was obtained. In the case using the model containing the two editing domains of *Pa*ThrRS forming a dimer as a probe, all solutions represent RFZ values of 3.6~4.4 and TFZ values of 5.4~6.6. No unique solution was obtained.

The amino acid sequence of the anticodon-binding domain of *St*ThrRS-2 is compared to that of the anticodon-binding domains of *Ec*ThrRS with the identity of 26% and is compared to that of *Ap*ThrRS-1 with the identity of 30% (These identity values are calculated by the program *Kalign* with the gap opening penalty of 11% and the gap extending penalty of 0.85%). The structure of the anticodon binding domain of *St*ThrRS-1 may not be similar to those of *Ec*ThrRS and *Ap*ThrRS-1. The sequence of the editing domain of *Pa*ThrRS is compared with that of *St*ThrRS-2 with the identity of 20%. The well fitted sequences between the editing domains are constructed with only 13 amino acid residues (see chapter 5). Therefore, it is difficult to estimate phase angle informaiton using the molecular replacement method. In parallel, the selenomethionine derivative was crystallized and the MAD method were tried to solve the phase angle problem as described in the next section (7-8).

7-8. X-ray analysis of *St*ThrRS-2*

7-8-a. Crystals of *St*ThrRS-2*

The optimized condition, under which suitable crystals for X-ray analysis appeared after two months, is that the protein solution containing 3 mg/ml *St*ThrRS-2* in 20 mM HEPES (pH 7.0) buffer and the reservoir solution containing 0.08 M ammonium acetate and 14% PEG4000 were equally mixed into a droplet, which was then equilibrated to the reservoir solution at 293 K. The crystal has a size of approximately 300×100×50 μm,

as shown in Figure 7-8-a. This crystal was used for X-ray diffraction experiment.

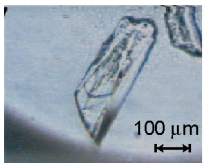


Figure 7-8-a. A photo of the *St*ThrRS-2* crystal.

7-8-b. X-ray data of the *St*ThrRS-2* crystal

A spectrum of f' and f'' around the absorption edge wavelength of Se atoms is drawn based on XAFS measurement of the crystal (Figure 7-8-b-1). In this spectrum, the four wavelengths are assigned as the peak (0.97893 Å), the edge (0.97907 Å), the high remote (0.96384 Å) and the low remote (0.98295 Å). The same crystal was used throughout X-ray experiments. A statistics of the observed intensity data and the crystallographic data are given in Table 7-8-b-1. Figure 7-8-b-2 shows an example of diffraction images taken using the peak wavelength.

In the measurement of the peak wavelength, the observed intensity data were processed to 2.9 Å resolution. The data completeness and an R_{marge} value are 99.0% and 11.2%, respectively. The numbers of observed and unique reflections are 83055 and 22735, respectively. The crystal belongs to the $P2_12_12_1$ space group. The cell constants are a = 60.5 b = 67.9 and c = 134.2 Å.

In the measurement at the edge wavelength, the observed intensity data were processed to 2.9 Å resolution. The data completeness and an R_{marge} value are 99.0% and 11.2%, respectively. The numbers of observed and unique reflections are 83857 and 22803, respectively.

In the measurement at the high remote wavelength, the observed intensity data were processed to 2.9 Å resolution. The data completeness and an R_{marge} value are 99.0% and 11.5%, respectively. The numbers of observed and unique reflections are 84524 and 23010, respectively.

In the measurement at the low remote wavelength, the observed intensity data were processed to 2.9 Å resolution. The data completeness and an R_{marge} value are 98.1% and 14.7%, respectively. The numbers of observed and unique reflections are 73735 and 22915, respectively.

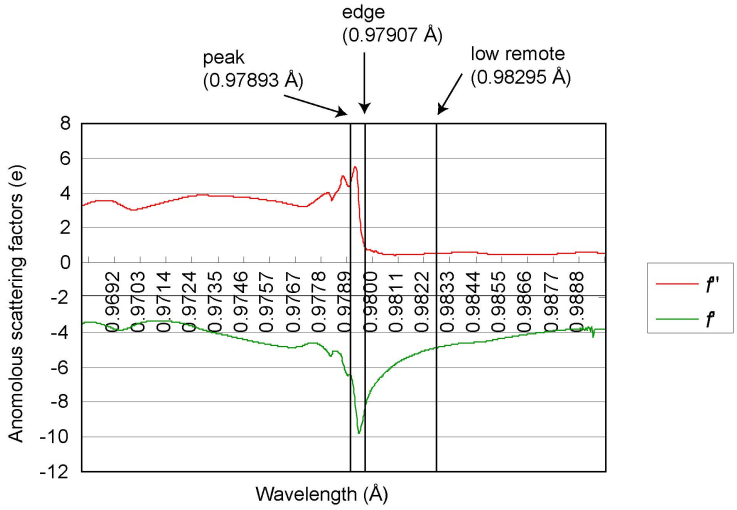


Figure 7-8-b-1. Variation of real (*f*′; green line) and imaginary parts (*f*″; red line) of anomalous scattering factors of selenium atoms in the *St*ThrRS-2* crystal as a function of wavelength.

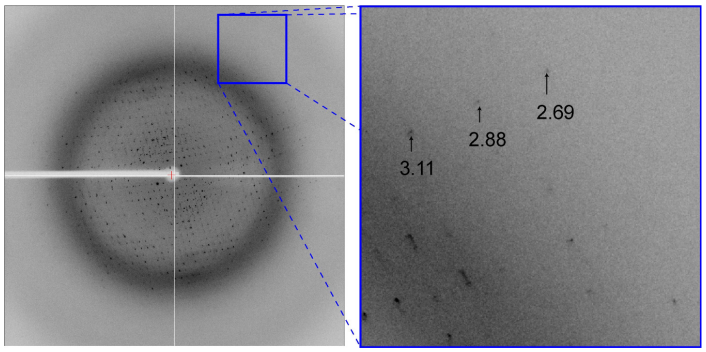


Figure 7-8-b-2 An X-ray diffraction image of *St*ThrRS-2* crystal, taken with an oscillation angle of 1° (left) and its zoomed up image in a box (right). Arrows indicate diffraction spots at the resolutions shown with values.

Table 7-8-b Statistics of intensity data of the *St*ThrRS-2* crystal and the crystallographic data at four wavelengths

	peak	edge	high remote	low remote
Wavelength (Å)	0.97893	0.97907	0.96384	0.98295
Resolution (Å)	50-2.9 (3.0-2.9)	50-2.9 (3.0-2.9)	50-2.9 (3.0-2.9)	50-2.9 (3.0-2.9)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell				
a (Å)	60.5	60.5	60.5	60.6
b (Å)	67.9	67.8	67.9	67.9
c (Å)	134.2	133.9	133.8	134.0
Observed reflections	83055	83857	84524	73735
Unique reflections	22735	22803	23010	22915
Completeness (%)	99.0 (100)	99.0 (100)	99.0 (100)	98.1 (99.9)
Redundancy	3.7 (3.8)	3.7 (3.8)	3.7 (3.5)	3.3 (3.4)
I/σ	17.7 (4.7)	18.6 (5.0)	17.4 (3.6)	14.2 (3.4)
R_{merge} (%)	11.2 (26.7)	11.2 (26.7)	11.5 (29.0)	14.7 (36.4)

$R_{merge} = 100 \times \sum_{hkl}\sum_i|I_i(hkl)-I(hkl)|/\sum_{hkl}\sum_i I(hkl)$

7-8-c.Phase angle determination of *St*ThrRS-2*

The space group was kept to be $P2_12_12_1$ during the X-ray experiments. The average cell constants are a = 60.5, b = 67.9 and c = 140.0 Å. An acceptable Z value (the number of protein molecule in the asymmetric unit) was estimated to be 2, the V_M value to be 2.32 Å^3/Da and the solvent content to be 47.1%. Four Se atoms will be found for two protein (subunit) in the asymmetric unit, because *St*ThrRS-2* contains two selenomethionine residues.

The MAD method using the four to two data sets and the SAD (single-wavelength anomalous dispersion) method of the peak data set were tried by the program *SHARP*.

Using the data sets measured at four wavelengths with resolution range of 42.75 - 4.03 Å, 8 Se atoms were found in the asymmetric unit, and then the solution showed that the correlation coefficient between F_o and F_c was 0.214. After the heavy atom position refinement, a total of eleven selenium atoms were found and two atoms with occupancies greater than 50% finally.

Using the peak, edge and remoteH data sets with resolution range of 42.75 - 5.0 Å, nine Se atoms were found in the asymmetric unit, and then the solution showed that the correlation coefficient between F_o and F_c was 0.217. After the heavy atom position refinement, a total of eleven selenium atoms were found and three atoms with occupancies greater than 50% finally.

Using the peak and edge data sets with resolution range of 42.75 - 5.0 Å, eight Se atoms were found in the asymmetric unit, and then the solution showed that the correlation coefficient between F_o and F_c was 0.109. After the heavy atom position refinement, a total of eleven selenium atoms were found and five atoms with occupancies greater than 50% finally.

A best result, chosen from three trials, has the maximum number of Se atoms with high occupancies. After heavy atom position refinement by maximum likelihood method, the final figure of merit was 0.386 for acentric reflections and the final phasing power was 0.651. After solvent flattering technique with solvent content fixed to be 37.0%, the final correlation coefficient in $|E^2|$ was 0.705 and the final R-factor between the structure factor amplitudes from the modified map and the observed data was 33.6%. However it is difficult to assign peptide chains on the density modified map. Therefore, the SAD method using the peak data set was also performed.

Using the peak data set with resolution range of 37.40 - 6.0 Å, nine Se atoms were found in the asymmetric unit, and then the solution showed that the correlation coefficient between F_o and F_c was 0.357. After the heavy atom position refinements, a total of eleven selenium atoms were found and three atoms with occupancies greater

than 50% finally. After heavy atom position refinement by maximum likelihood method, the final figure was merit of 0.166 for acentric reflections and the final phasing power was 0.519. After solvent flattering technique with solvent content fixed to be 49.6%, the final correlation coefficient in $|E^2|$ was 0.693 and the final *R*-factor between the structure factor amplitudes from the modified map and the observed data was 32.1%. However it is difficult to assign peptide chains on the density modified map.

It is difficult to estimate phase information using the MAD and SAD method. One reason is considered that the R_{marge} value is more than 10% (see Table 7-8-b). We are crystallizing *St*ThrRS-2* using the other conditions to get data resulting in the higher resolution and lower R_{marge} value.

7-9. Reference

1. E. Gilbert, D. Marc, P. Olivier, G. Jean & M. Dino (1990) “Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs” *Nature*, **347**, 203-206.
3. E. Krissinel & K. Henrick (2007) “Inference of macromolecular assemblies from crystalline state” *J. Mol. Biol*. **372**, 774-797

< Chapter 8; Discussion of ThrRS structures from crenarchaeota >

8-1. Structual comparesion of *Ap*ThrRS-1 with ThrRS from *E.coli*

In the case of *E. coli*, ThrRS is composed of the three domains; editing, catalytic and anticodon-binding [1]. Threonine is attached to the cognate $tRNA^{Thr}$ through a reactive threonyl-adenylate intermediate in the catalytic domain. A misacylated product, $Ser\text{-}tRNA^{Thr}$, is also synthesized, but it is hydrolyzed in the editing domain. In *Aeropyrum pernix* and *Sulfolobus tokodaii*, the genomes contain the two genes of ThrRS-1 and ThrRS-2. As described in the chapter 5, the sequence of ThrRS-1 contains the catalytic and anticodon-binding domains. The sequence of ThrRS-2 contains the editing and anticodon-binding domains. It is considered that ThrRS-1 and ThrRS-2 are complementary to each other in functions, one is only for catalysis and the other is only for editing. The crystal structure of *Ap*-ThrRS-1, which is one of two gene products, has been successfully determined at 2.3 Å resolution. This study has revealed that *Ap*ThrRS-1 is a dimeric enzyme composed of the two identical subunits, each containing the two domains, catalytic and anticodon-binding.

In general, the two subunits of ThrRS form a dimer as a member of the class II aaRS. Superimposition of the dimer of *Ap*ThrRS-1 on that of *Ec*ThrRS is shown in Figure 8-1-1. The dimer formations are similar to each other, but different in the peripheral regions. In *Ec*ThrRS, the dimer formation interface occurs mainly between the catalytic domains, and the editing domains are attached to the opposite side of the interface. The catalytic domains of *Ap*ThrRS-1 are also interacted to each other to form a dimer, like those of *Ec*ThrRS. However, *Ap*ThrRS-1 lacks the editing domains.

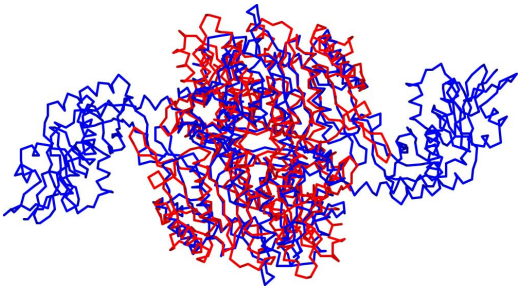


Figure 8-1-1. A superimposition of Cα chains between the dimers of *Ap*ThrRS-1 (red) and of *Ec*ThrRS (blue).

In the case of *Ec*ThrRS, the contacting area of the dimerization interface is estimated to be 1926.9 $Å^2$ by *PISA* server¶ [2]. It is 10.3% of the total surface area

¶This program was developed for the prediction of probable biological unit from protein structures obtained by means of protein crystallography. The approach is based on the consistent enumeration of all possible assemblies in a crystal with subsequent analysis for chemical stability. The analysis is based on the evaluation of free energy of complex dissociation, which includes the free energy of binding and the entropy change term.

(18618.0 Å^2), when the editing domain is truncated. In the case of *Ap*ThrRS-1, the corresponding value is 18.5% (3856.7 Å^2) of the total surface area (20848.7 Å^2), so that the interface area of *Ap*ThrRS-1 is larger than that of *Ec*ThrRS, as shown in Figure 8-1-2. This is consistent to the fact that ThrRS-1 from *Aeropyrum pernix* (an aerobic hyper-thermophilic crenarchaeon) is optimally grown at 95 ºC [3], and it is more stable than *Ec*ThrRS.

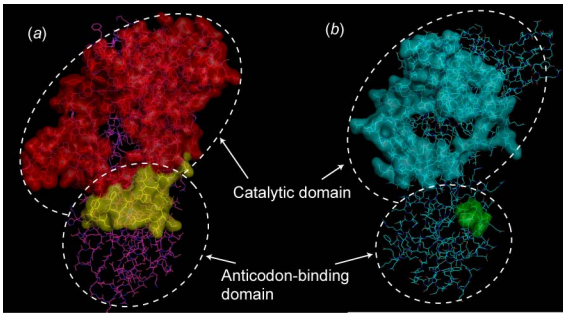


Figure 8-1-2. Dimer interfaces of *Ap*ThrRS-1 (*a*) and *Ec*ThrRS (*b*). Red and cyan surfaces show the residues of the interface in the catalytic domains and yellow and green in the anticodon-binding domains, respectively.

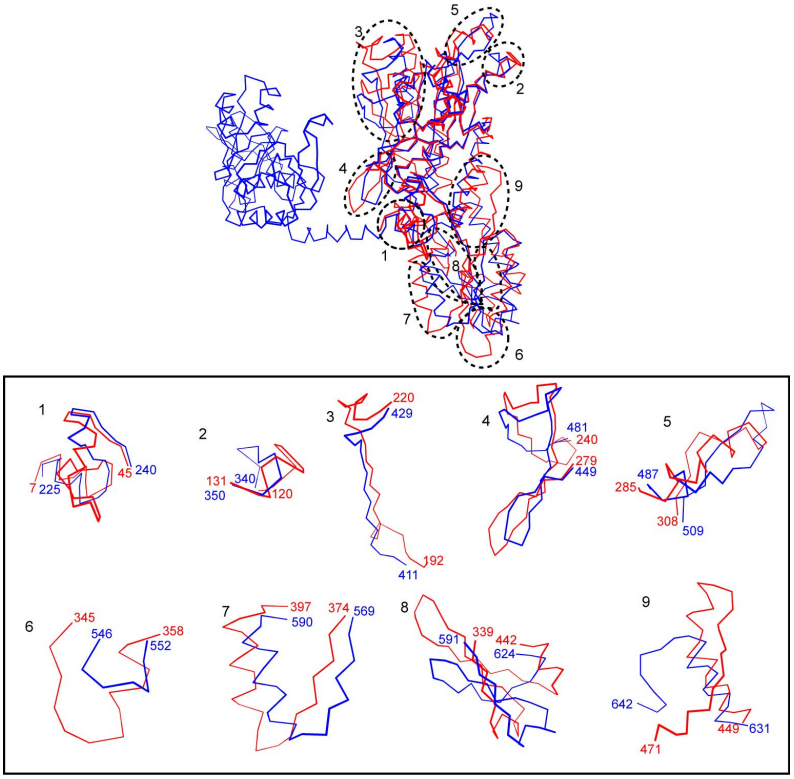


Figure 8-1-3. A superimposition of Cα chains between the subunits of *Ap*ThrRS-1 (red) and of *Ec*ThrRS (blue), is drawn in wire models (upper). The bottom figures show their partial structures. The colored numbers indicate the residue numbers.

Figure 8-1-3 shows nine regions in which the main chain conformations are largely different between *Ap*ThrRS-1 and *Ec*ThrRS. These differences occur at several inserted

residues in the *Ap*ThrRS-1 sequence. In total, the sequence length of 471 residues is longer than that of *Ec*ThrRS (400 residues) without the editing domain. The regions 1~5 are in the peripheral parts of the catalytic domain. In the region 1 at the N-terminus of *Ap*ThrRS-1, there is an additional loop in *Ap*ThrRS-1. In the regions 2~4, the three loops are longer in *Ap*ThrRS-1. In the region 5, an α-helix is added in *Ap*ThrRS-1, while the corresponding part in *Ec*ThrRS contains a shorter loop. The regions 6~9 in the anticodon-binding domains contain long loops and additional helices in *Ap*ThrRS-1. In the region 9 at the C-terminus, the main chain conformation is largely different. These longer loops of *Ap*ThrRS-1 increase the area of the dimerization interface, as shown in Figure 8-1-2. Although the main chain of *Ap*ThrRS-1 adopts a different conformation locally, the catalytic site and the anticodon-binding site are highly conserved as described below.

The motifs 1, 2 and 3, which are the typical features of the class II aaRSs, are conserved completely in both *Ec*ThrRS and *Ap*ThrRS-1. These motifs contain the characteristic sequences: gϕxxϕxxPϕϕ in the motif 1; (F/Y/H)Rx(E/D) and (R/H)xxxFxxx(D/E) in the motif 2; and λxϕgϕgeRϕϕϕϕ in the motif 3 (x stands for any residues, λ for small amino acids P, G, S, T; ϕ for hydrophobic residues F, Y, W, I, L, V, M, A) [4]. Figure 8-1-4 (*a*) shows sequence alignments of the three motifs between *Ec*ThrRS and *Ap*ThrRS-1. In the motif 1 sequences, the proline and the hydrophobic residues are conserved to a loop and an α helix (Figure 8-1-4 (*b*)). In the motif 2, HRxE and RxxxFxxxD are also conserved to form the two β strands and loops, which also conserved in *Ec*ThrRS and *Ap*ThrRS-1. The motif 3 is involved in formation of an α helix and loops, and ϕgϕgϕ and λxϕgϕgϕeRϕϕϕϕ are high sequence homology with both of *Ap*ThrRS-1 and *Ec*ThrRS. These features suggest that *Ap*ThrRS-1 belongs to the class II aaRSs.

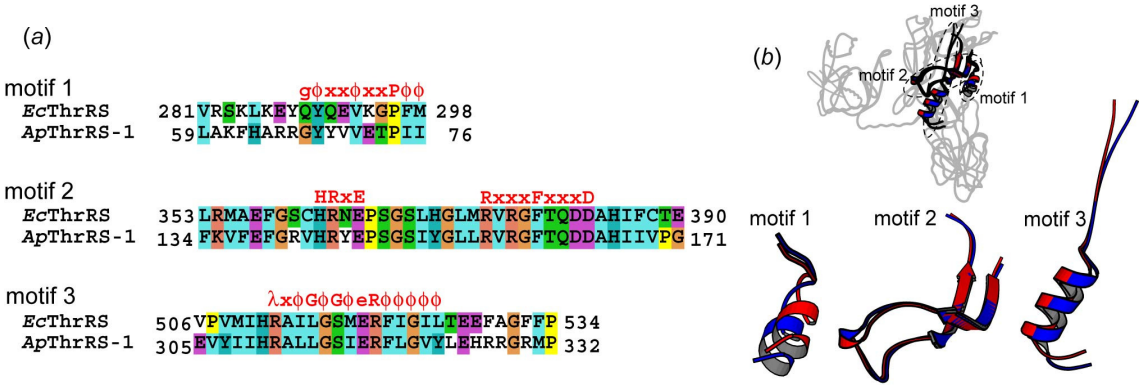


Figure 8-1-4. Sequence alignments of three motifs of *Ec*ThrRS and *Ap*ThrRS-1 (*a*). Red characters indicate the characteristic sequences described in [3]. Numbers indicate the terminal residue number of each motif. Right figure (*b*) shows a superimposition of the three motif structures of *Ap*ThrRS-1 (red) onto those of *Ec*ThrRS (blue) in ribbon models.

8-2. Catalytic site of *Ap*ThrRS-1

The structure described above suggests that *Ap*ThrRS-1 catalyzes the first reaction to produce the Thr-tRNA^{Thr}. To examine the plausibility of the catalytic site of *Ap*ThrRS-1, the catalytic domain is compared to that of *Ec*ThrRS (PDBID 1EVK [5]). Figure 8-2-1 (*a*) and (*b*) shows their superimpositions including threonine found in *Ec*ThrRS. The zinc finger formation is absolutely conserved. The zinc atom is tetrahedrally coordinated by Cys112, His166 and His310. At the forth position, the hydroxyl group and the amino group of the threonine can be bound so that the zinc atom forms a square pyramidal pentacoordinated state. In *Ec*ThrRS, the bound threonine is trapped by hydrogen bonds between the threonine carboxyl group and the side chains of Arg363, Glu381 and Gln484, and between the threonine amino group and the hydroxyl group of Tyr482 (see Figure 8-2-1 (*a*)). These interacting residues are conserved in *Ap*ThrRS-1. They are Arg144, Glu162, Glu383 and Tyr257, respectively (Figure 8-2-1 (*b*)). The hydroxyl group of threonine forms a hydrogen bond with the side chain of Asp383 in *Ec*ThrRS. This interaction and the zinc atom coordination are essential for recognition of the threonine. These specific features are also conserved in *Ap*ThrRS-1. The methyl group of threonine is placed in a hydrophobic environment by Thr482 and Ala513 in *Ec*ThrRS. This situation is formed by Thr280 and Ala312 in *Ap*ThrRS-1. From these common features, it is concluded that threonine can be bound in the catalytic site of *Ap*ThrRS-1.

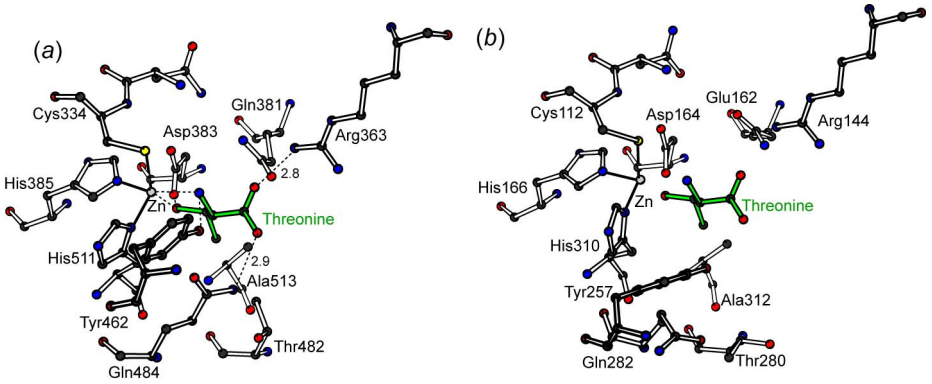


Figure 8-2-1. The catalytic site of *Ec*ThrRS (PDBID 1EVK) (*a*) and the catalytic site of *Ap*ThrRS-1 (*b*). Black, blue, red and yellow spheres show carbon, nitrogen, oxygen and sulfur atoms, respectively. The bound threonine molecule in (*b*) is generated by superimposition of the two catalytic sites, and colored in green.

In addition, to examine a possibility of ATP binding in the catalytic site of *Ap*ThrRS-1, the two structures of *Ap*ThrRS-1 and *Ec*ThrRS (PDBID 1EVL [5]) are superimposed, because 1EVL contains Thr-AMS (5'-O-(N-(L-threonyl)-sulfamoyl)adenosine), as an analog of Thr-AMP. Figure 8-2-2 (*a*)

and (*b*) shows the superimposed catalytic sites of *Ec*ThrRS and *Ap*ThrRS-1 with Thr-AMS. In *Ec*ThrRS, the adenine ring is surrounded by the side chains of Val376, Pro379 and Arg520, which belong to the motifs 2 and 3, respectively (see Figure 8-1-4). The corresponding residues in *Ap*ThrRS-1 are Val157, Phe160 and Arg319. The side chain of Gln479 in *Ec*ThrRS interacts to the 3'-hydroxy group of the ribose. The corresponding residue in *Ap*ThrRS-1 is Gln277. Although it points away in *Ap*ThrRS-1, it can be back to interact to the ribose when ATP is practically bound. These structural consistencies also support that this domain of *Ap*ThrRS-1 is catalytic.

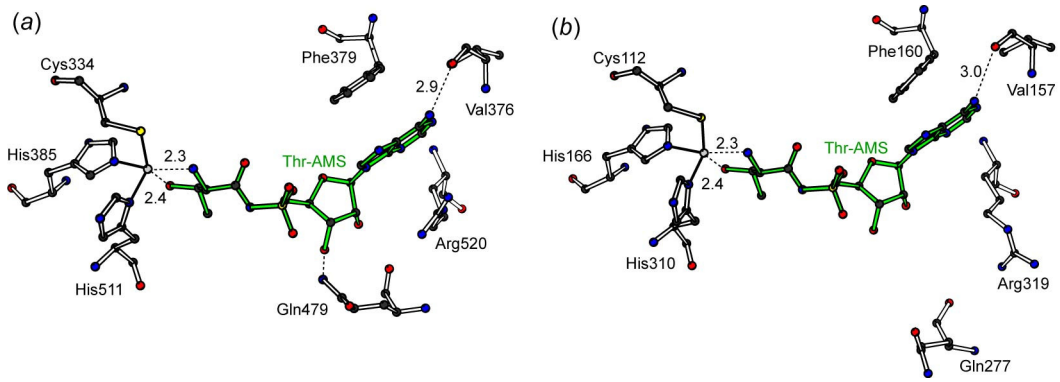


Figure 8-2-2. The catalytic site of *Ec*ThrRS (PDBID 1EVL) (*a*) and the catalytic site of *Ap*ThrRS-1 (*b*). Black, blue, red and yellow spheres show carbon, nitrogen, oxygen and sulfur atoms, respectively. The bound Thr-AMS molecule in (*b*) is generated by superimposition of the two catalytic sites and are colored in green.

Furthermore, to examine a possibility to recognize the CCA end of $tRNA^{Thr}$, the catalytic site of *Ap*ThrRS-1 is compared with that of *Ec*ThrRS (PDBID 1QF6 [4]), because 1QF6 contains $tRNA^{Thr}$. Figure 8-2-3 (*a*) and (*b*) shows the catalytic site of *Ap*ThrRS-1 superimposed on that of *Ec*ThrRS with the bound $tRNA^{Thr}$. In *Ec*ThrRS, the terminal adenine (A76 residue) of $tRNA^{Thr}$ is intercalated between Tyr313 and Arg363, and the N6 atom forms a hydrogen bond with the carbonyl group of Ala316. These three residues correspond to Tyr91, Arg144 and Asn94 in *Ap*ThrRS-1, respectively. In Figure 8-2-3 (*a*), the oxygen atoms of A76 ($tRNA^{Thr}$) are involved in a hydrogen-bonded network; the O2′ atom with His309 and Tyr462 and the O3′ atom with Gln484. These residues also correspond to His87, Tyr257, Gln282 and Tyr91 in *Ap*ThrRS-1, respectively. These structural features in the catalytic sites support that $tRNA^{Thr}$ can be bound in the catalytic site of *Ap*ThrRS-1.

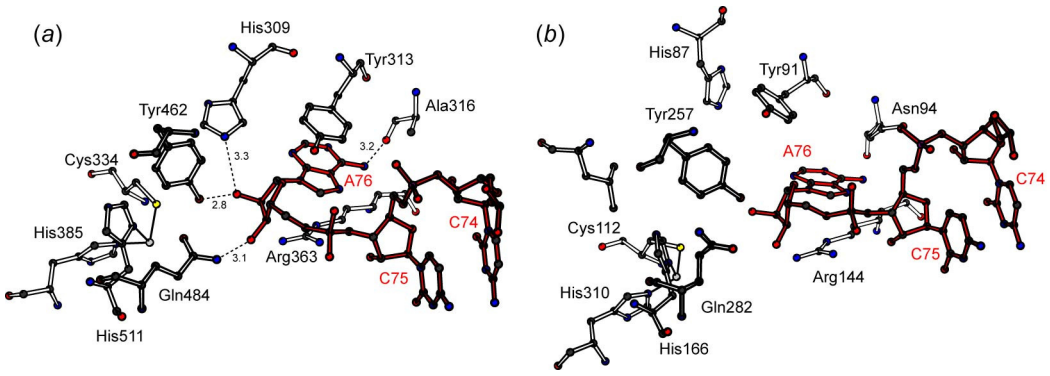


Figure 8-2-3. The catalytic site of *Ec*ThrRS (PDBID 1QF6) (*a*) and the catalytic site of *Ap*ThrRS-1 (*b*). Black, blue, red and yellow spheres show carbon, nitrogen, oxygen and sulfur atoms, respectively. The bound $tRNA^{Thr}$ and AMP molecules in (*b*) are generated by superimposition of the two catalytic sites, and are colored in red and green, respectively.

8-3. Anticodon-binding site of *Ap*ThrRS-1

A superimposition of the anticodon-binging site of *Ap*ThrRS-1 onto that of *Ec*ThrRS (PDBID 1QF6) [1] is shown in Figure 8-3 (*b*). Glu600 and Arg609 in *Ec*ThrRS, which interact with the anticodon G35 and U36 residues, correspond to Glu404 and Arg416 in *Ap*ThrRS-1, respectively. Although their side chains are pointed away in *Ap*ThrRS-1, they can interact to the anticodon bases when $tRNA^{Thr}$ is practically bound in this domain of *Ap*ThrRS-1. In *Ec*ThrRS, the wobble base C34, which is not strictly recognized, interacts with Gln575. The corresponding residue of *Ap*ThrRS-1 is Thr382, the Oγ atom of which may interact with C34 of $tRNA^{Thr}$. These features suggest that the anticodon-binding domain of *Ap*ThrRS-1 is able to recognize the anticodon of $tRNA^{Thr}$.

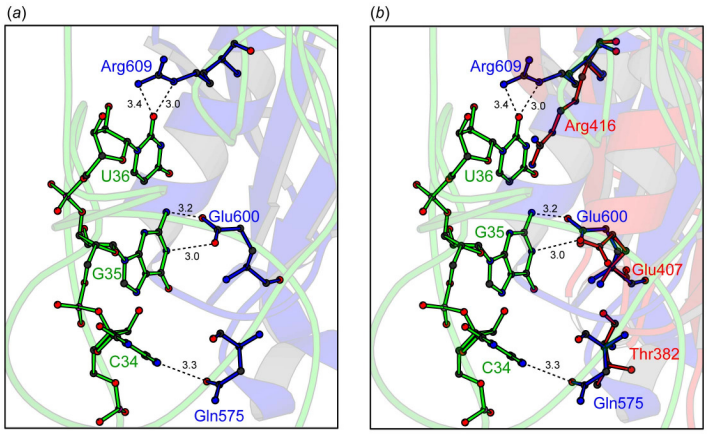


Figure 8-3. The anticodon-binging site of *Ec*ThrRS (blue) (*a*). $tRNA^{Thr}$ bound in *Ec*ThrRS is colored in green. The anticodon binding site of *Ap*ThrRS-1 superimposed onto that of *Ec*ThrRS (red) is shown in (*b*). Black, blue, red and pink spheres show carbon, nitrogen, oxygen and phosphor atoms, respectively.

8-4. Structure of *Ap*ThrRS-2

The structure of *Ap*ThrRS-2 has not yet been determined. However, a possibility of the editing and anticodon-binding functions is examined. The sequence of *Ap*ThrRS-2 contains the two domains; editing and anticodon-binding (referred to chapter 5). The sequence of the editing domain is similar to that of *Pa*ThrRS with the identity of 37%. The two editing domains of *Pa*ThrRS form a dimer [6]. The structure of the editing domain of *Ap*ThrRS-2 might be similar to that of *Pa*ThrRS which is different from that of *Ec*ThrRS (see Figure 5-1-3 in chapter5). The crystal structure of the editing domain of *Pa*ThrRS in complex with seryl-3′-aminoadenosine, an analog of Ser-AMP, is reported [6] (PDBID 2HL0). The proposed catalytic residues His83 and Lys121 and the interacting residues with the adenine base, Val45, Ala94 and Phe117 in *Pa*ThrRS are shown in Figure 8-4-1 (*a*). The substrate may be hydrolyzed by the water molecule bound to the side chain of Lys121 [6]. These residues surrounding the substrate are conserved in *Ap*ThrRS-2, as shown in Figure 8-4-1 (*b*). These features suggest that the editing domain of *Ap*ThrRS-2 hydrolyzes the misacylated Ser-tRNA^{Thr} in a way similar to *Pa*ThrRS.

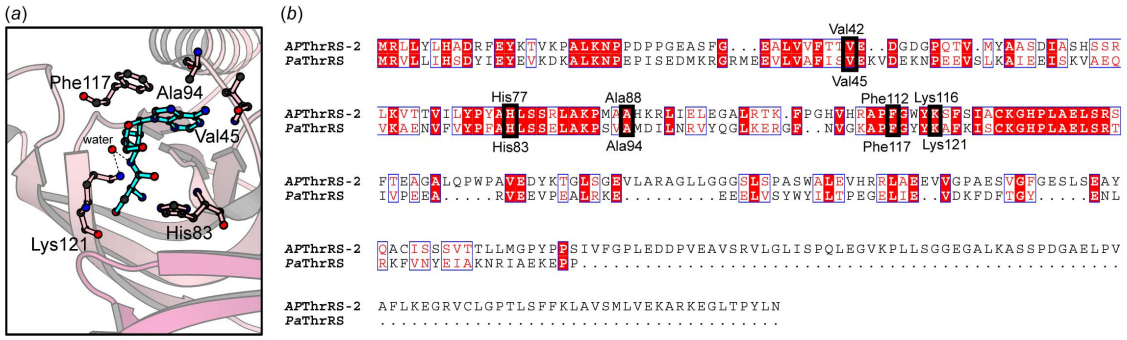


Figure 8-4-1. The substrate binding site of the truncated editing domain of *Pa*ThrRS (PDBID 2HL0) (*a*). An analog, seryl-3'-aminoadenosine, is colored in cyan. Black, blue and red spheres show carbon, nitrogen and oxygen atoms, respectively. (*b*) shown a sequence alignment of the editing domains of *Ap*ThrRS-2 and *Pa*ThrRS using the program *Kalign* and colored by the program *ESPript*. The alignment parameters are 11.0% for the gap opening penalty which is a gap for opening and 0.85% for the gap extending penalty which is for extending a gap by one residue.

In *Ec*ThrRS, Gln575, Glu600 and Arg609 form hydrogen bonds with the anticodon residues of tRNA^{Thr} (see Figure 8-3). Glu600 and Arg609 in *Ec*ThrRS correspond to Glu376 and Arg385 in *Ap*ThrRS-2 by a sequence alignment, as shown in Figure 8-4-2. Glu575 in *Ec*ThrRS interacting with the 34th residue of tRNA^{Thr} anticodon corresponds to Gly351 in *Ap*ThrRS-2. In *Aeropyrum pernix* genome, four genes of tRNA^{Thr} are encoded [3]. Their anticodons are assigned UGU, UGU, CGU, and GGU. Three bases, U, C and G are available as the first anticodon residue (34th) of tRNA^{Thr} in *Aeropyrum*

pernix. Therefore, it is considered that the amino acid residue interacting with the 34th residue of $tRNA^{Thr}$ is not required strict recognition. These features of the anticodon-binding domain of *Ap*ThrRS-2 are acceptable for recognition of $tRNA^{Thr}$ anticodon.

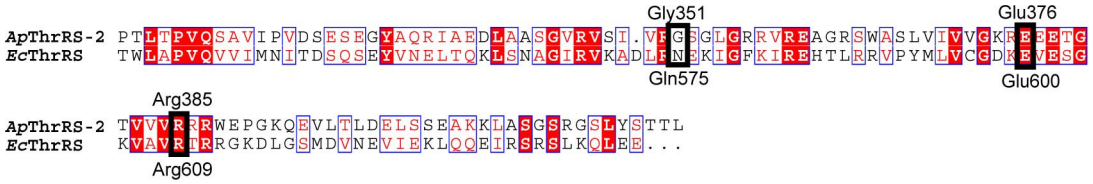


Figure 8-4-2. A sequence alignment of the anticodon-binding domains of *Ap*ThrRS-2 and *Ec*ThrRS using the program *Kalign* and colored by the program *ESPript*. The alignment parameters are 11.0% for the gap opening penalty which is a gap for opening and 0.85% for the gap extending penalty which is for extending a gap by one residue.

A model structure of *Ap*ThrRS-2 is shown in Figure 8-4-3. The editing domains may form a dimer through a β-sheet as like the editing domain of *Pa*ThrRS. The inserted sequences with 110 residues between the editing domain and anticodon-binding domain (see Figure 5-2-2-2 in the chaper 5) may be liker domain as shown in broken ellipses. The anticodon-binding domains may be folded on the opposite side of the dimer formation interface between the editing domains

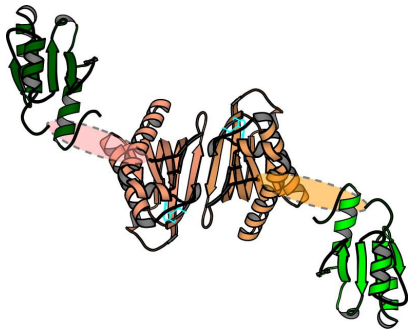


Figure 8-4-3. A model of *Ap*ThrRS-2 based on the structures of the editing domain of *Pa*ThrRS (PDBID 2HL0) colored in orange and the anticodon-binding domains of *Ec*ThrRS (PDBID 1QF6) colored in green. The substrate models bound to the editing domains are colored in cyan.

8-5. Structure of *St*ThrRS-1

*St*ThrRS-1 and *St*ThrRS-1* are crystallized successfully, but the crystal structure is not yet determined. Here a possibility that *St*ThrRS-1 has the catalytic and anticodon-binding functions is described. Figure 8-5-1 shows a sequence alignment of the catalytic domains of *St*ThrRS-1 and *Ec*ThrRS. The Cys334, His385 and His511 residues essential for the zinc finger in *Ec*ThrRS correspond to Cys226, His277 and His405 in *St*ThrRS-1. This suggests that these residues form the zinc finger. The conservative motifs 1, 2 and 3 as the class II aaRS are found in the catalytic domain of

St-ThrRS-1, as indicated in green boxes of Figure 8-5-1. These support that the catalytic domain of *St*-ThrRS-1 may catalyze the threonylation of $tRNA^{Thr}$.

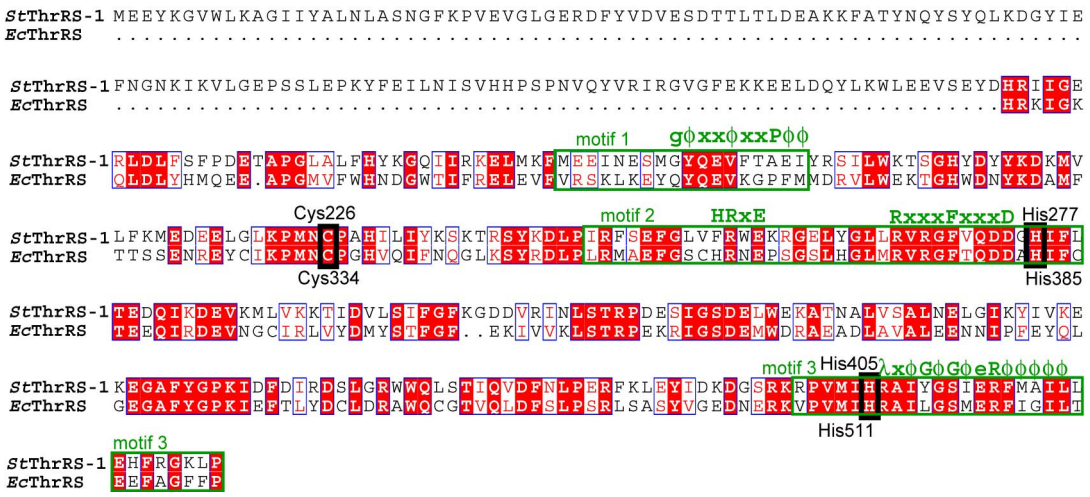


Figure 8-5-1. A sequence alignment of the catalytic domains of *St*ThrRS-1 and *Ec*ThrRS using the program *Kalign* and colored by the program *ESPript*. The alignment parameters are 11.0% for the gap opening penalty which is a gap for opening and 0.85% for the gap extending penalty which is for extending a gap by one residue.

In *Ec*ThrRS, the Gln575, Glu600 and Arg609 form hydrogen bonds with the anticodon of $tRNA^{Thr}$ (see Figure 8-3). Glu600 and Arg609 correspond to Glu494 and Arg505 in *St*ThrRS-1, as shown in Figure 8-5-2. In *Sulfolobus tokodaii* genome, the anticodons of three $tRNA^{Thr}$ genes [6] are UGU, CGU, and GGU. The anticodons of their second and third residues are strictly fixed with G and U. The first residue is acceptable to one of three bases, U, C and G. Therefore, the recognition is not strict at the first anticodon base. These support that the anticodon-binding domain of *St*ThrRS-1 recognize the anticodon of $tRNA^{Thr}$.

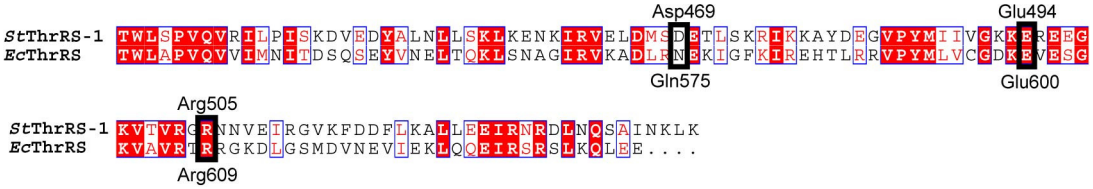


Figure 8-5-2. A sequence alignment of the anticodon-binding domains of *St*ThrRS-1 and *Ec*ThrRS using the program *Kalign* and colored by the program *ESPript*. The alignment parameters are 10.0% for the gap opening penalty which is a gap for opening and 0.2% for the gap extending penalty which is for extending a gap by one residue.

*St*ThrRS-1 has an additional sequence at the N-terminus containing about 134 residues (see Figure 5-2-2-1). The function of the sequence is unclear because it has no sequence homology with any other sequences found on the FASTA server except for the

N-terminal sequences of *Sf*ThrRS-1. The crystal structure of *St*ThrRS-1 will provide an insight of the detail structure and the function.

8-6. Structure of *St*ThrRS-2

X-Ray diffraction data were collected at 2.9 Å resolution, but the crystal structure is not yet determined. A sequence alignment of *St*ThrRS-2 with *Pa*ThrRS is shown in Figure 8-6-1. The editing domain sequence of *St*ThrRS-2 is quite different from that of *Pa*ThrRS with the identity of 15%. The residues of *Pa*ThrRS, which interact with an analog, seryl-3'-aminoadenosine (see Figure 8-4-1 (*a*)) cannot be corresponded any residues in *St*ThrRS-2 by sequence alignment. The sequence of *St*ThrRS-2 is also different from that of *Ec*ThrRS with identity of 23%. However, the only 13 amino acids in *St*ThrRS-2 correspond to the interfacing residues of a dimer formation of *Pa*ThrRS (see Figure 5-2-2-2 in the chapter 5). Therefore the sequences might occupy the center of the most essential part of the interface as shown in Figure 8-6-3.

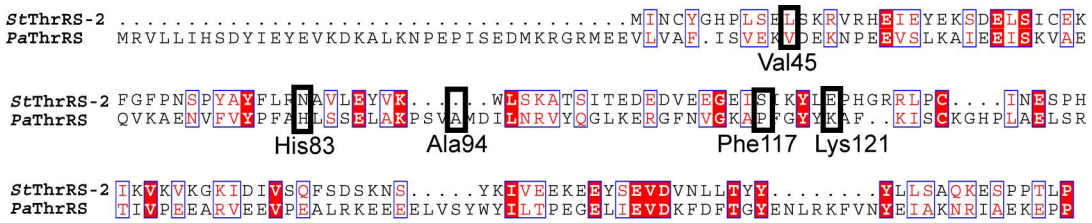


Figure 8-6-1. A sequence alignment of the editing domains of *St*ThrRS-2 and *Pa*ThrRS using the program *Kalign* and colored by the program *ESPript*. The alignment parameters are 11.0% for the gap opening penalty which is a gap for opening and 0.85% for the gap extending penalty which is for extending a gap by one residue.

In *Ec*ThrRS, Gln575, Glu600 and Arg609 form hydrogen bonds with the anticodon of $tRNA^{Thr}$ (see Figure 8-3). Glu600 and Arg609 in *Ec*ThrRS correspond to Glu212 and Arg223 in *St*ThrRS-2 by a sequence alignment, as shown in Figure 8-6-2. In *Sulfolobus tokodaii* genome, three $tRNA^{Thr}$ anticodons [6] are UGU, CGU and GGU. The second and third anticodon residues are strictly fixed to G and U bases, respectively, however, the first antcodon residue is acceptable to any of the three bases, U, C or G. Therefore, this site is not required for strict recognition. These features are consistent to that the anticodon-binding domain of *St*ThrRS-2 recognizes the anticodon of $tRNA^{Thr}$.

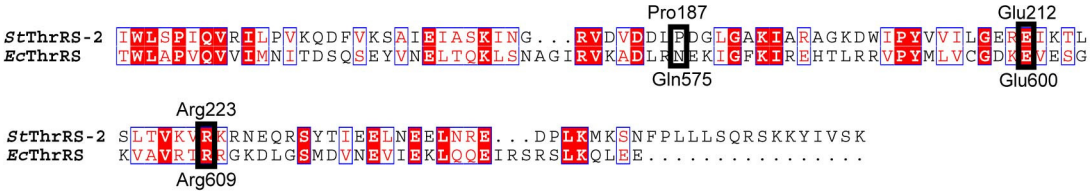


Figure 8-6-2. A sequence alignment of the anticodon-binding domains of *St*ThrRS-2 and *Ec*ThrRS using the program *Kalign* and colored by the program *ESPript*. The alignment parameters are 11.0% for the gap opening penalty which is a gap for opening and 0.85% for the gap extending penalty which is for extending a gap by one residue.

The model structure of *St*ThrRS-2 is shown in Figure 8-6-3. The editing domains may form a dimer through a β-sheet as like the editing domain of *Pa*ThrRS. However, the other foldings of the editing domain may be different from that of *Pa*ThrRS because the sequence identity is very low (15%). The inserted sequences with 110 residues between the editing domain and anticodon-binding domain (see Figure 5-2-2-2 in the chaper 5) may form a domain for editing as shown in broken ellipses. The anticodon-binding domains may be folded on the opposite side of the dimer formation interface between the editing domains.

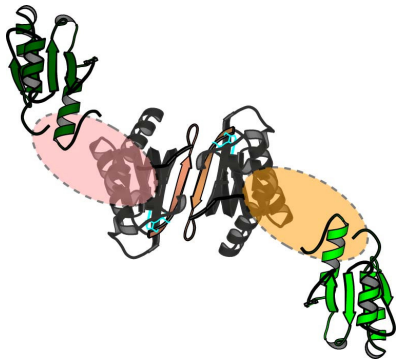


Figure 8-6-3. A model of *St*ThrRS-2 based on the structures of the editing domain of *Pa*ThrRS (PDBID 2HL0) colored in orange and the anticodon-binding domains of *Ec*ThrRS (PDBID 1QF6) colored in green. Black colored models indicate non-corresponding regions by a sequence alignment. The substrate models bound to the editing domains are colored in cyan.

8-7. Reference

1. R. Sankaranarayanan, A. C. Dock-Bregeon, P. Romby, J. Caillet, M. Springer, B. Rees, C. Ehresmann, B. Ehresmann & D. Moras (1999) "The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site" *Cell*. **97**, 371-381.
2. E. Krissinel & K. Henrick (2007) "Inference of macromolecular assemblies from crystalline state" *J. Mol. Biol*. **372**, 774-797.
3. Y. Kawarabayasi, Y. Hino, H. Horikawa, S. Yamazaki, Y. Haikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama, S. Fukui, Y. Nagai, K. Nishijima, H. Nakazawa, M. Takamiya, S. Masuda, T. Funahashi, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, H. Kikuchi (1999) "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1" *DNA Res*., **6**, 83-101, 145-152.
4. E. Gilbert, D. Marc, P. Olivier, G. Jean & M. Dino (1990) "Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs" *Nature*, **347**, 203-206.
5. R. Sankaranarayanan, A. C. Dock-Bregeon, B. Rees, M. Bovee, J. Caillet, P. Romby, C. S. Francklyn & D.Moras (2000) "Zinc ion mediated amino acid discrimination by threonyl-tRNA synthetase" *Nat Struct Biol*. **6**, 461-465.
6. Y. Kawarabayasi, Y. Hino, H. Horikawa, K. Jin-no, M. Takahashi, M. Sekine, S. Baba, A. Ankai, H. Kosugi, A. Hosoyama, S. Fukui, Y. Nagai, K. Nishijima, R. Otsuka, H. Nakazawa, M. Takamiya, Y. Kato, Yoshizawa, T. Tanaka, Y. Kudoh, J. Yamazaki, N. Kushida, A. Oguchi, K. Aoki, S. Masuda, M. Yanagii, M. Nishimura, A. Yamagishi, T. Oshima, H. Kikuchi (2001) "Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7" *DNA Res*., **8**, 123-140.

Part III

< Chapter 9; Conclusions >

Many proteins are involved in producing functional RNAs, such as rRNA, mRNA, tRNA etc. Although some of them were revealed the tertiary structures, their detailed mechanisms are not fully understood. In this study, two enzymes related to RNA processing, RNA 3′-terminal phosphate cyclase (Rtc) and threonyl-tRNA synthetases (ThrRS-1 and ThrRS-2) from crenarchaea, have been subjected to reveal the tertiary structures.

9-1. Part I

X-ray analyses RNA 3'-terminal phosphate cyclase from *Sulfolobus tokodaii*

9-1-a. Purpose of the study

In pre-tRNA splicing, an RNA fragment cleaved by RNase has a phosphate group at the 3′-terminus. This 3′-terminal phosphate is required to remove in prior to its ligation with the 5′-terminal phosphate of other RNA. Rtc catalyzes the ATP-dependent conversion of the 3′-phosphomonoester end to the 2′,3′-cyclic phosphodiester end. X-ray analyses of Rtc from *Sulfolobus tokodaii* (*St*-Rtc) and its several complexes with substrates have been performed in order to establish the reaction mechanism,

9-1-b. Structural features of *St*-Rtc

In the six different crystals, *St*-Rtc*, *St*-Rtc, *St*-Rtc*+Mn, *St*-Rtc*+ATP, *St*-Rtc+AMP and *St*-Rtc*-AMP crystals obtained under different conditions, a similar dimer formation commonly occurs by association of the two subunits, in which the interface is formed by a β-sheet extension though the hydrogen bonds between the two Ile223 residues related by a two-fold symmetry. Each subunit consists of a large domain and a protruded small domain. The βαβαββ motifs, repeated by three times, are folded into the large domain. According to a pseudo three-fold symmetry, three α-helices from each motif are packed to form the central core with the outer three α-helices. The core is surrounded by the twelve β-strands. The large domain is deformed from the regular three-fold symmetry by insertion of the small domain with the βαβββα motif into the third motif of the large domain.

9-1-c. Reaction mechanism at the first step

Based on the structures in different states, the following reaction mechanism has

been proposed. When ATP enters the first catalytic site of *St*-Rtc, the adenine moiety is stacked between the side chains of Pro126 and His283, and the ATP orientation is fixed by hydrogen bonds between Asp250 and the N6 atom of the adenine, and between Asp286 and the ribose hydroxyl groups. The β- and γ-phosphate groups are bound to the nearby basic residues. The bound Mg^{2+} ion changes the conformation of the α-phosphate group, so that the lone-paired electron on the Nε atom of His307 will attack the α-phosphorus atom. After a phosphoamide bond is formed between the Nε atom and the phosphorus atom, a pyrophosphate ion is ejected. The covalently bounded *St*-Rtc-AMP intermediate is used in the second reaction.

9-1-d. Reaction mechanism at the second step

Nearby the first catalytic site, seven basic residues, Arg17, Arg36, Arg39, Lys41, Arg46, His48 and Lys124, are assembled suggesting the second catalytic site, in which an RNA carrying a phosphate group at its 3'-terminus is bound. An RNA binding model constructed reasonably allows us to speculate the reaction mechanism at the second step. His48 forms a hydrogen bond with the 2′-hydroxyl group of the bound RNA. The oxygen atom of the 3′-phosphate group is forced to attack the α-phosphorus atom bound to His307 of the *St*-Rtc-AMP intermediate. The P-N bond is cleaved to produce an intermediate linked by a pyrophosphate bond between the 3′-phosphate group of RNA and AMP. When His48 withdraws a proton from the 2′-hydroxyl group, the 2′-oxygen atom attacks the 3′-phosphorus atom to form a free AMP and the 2′,3′-cyclic phosphate group at the 3′ end of RNA.

9-1-e. Significance of the study and future prospects

In this study, a reaction mechanism at the first step of 2′,3′-cyclic phosphate formation has been proposed based on the crystal structures. For the second reaction, a reaction mechanism is postulated, but further experiments are required to be established. It is expected to be clarified in the near future.

9-2. Part II

X-ray analyses of Threonyl-tRNA synthetases from creanarchaea)

9-2-a. Purpose of the study

In general, every organism possesses only one gene for each aminoacyl-tRNA synthetase. However, some crenarchaea organisms have two genes in their genomes. Sequence alignment analyses of the two gene products (ThrRS-1 and ThrRS-2) suggests us that ThrRS-1 is composed of bacterial-type catalytic and anticodon-binding domains,

but no editing domain, while ThrRS-2 composed of archaea-type editing domain and anticodon-binding domain, but no catalytic domain. This study has been aimed to clarify their complementary functions and the reaction mechanisms by X-ray analyses of ThrRS-1 and ThrRS-2 from *Aeropyrum pernix* (*Ap*) and those from *Sulfolobus tokodaii* (*St*) have been subjected in order to prove the above speculations,

9-2-b. Structural features of *Ap*ThrRS-1

X-Ray analysis of *Ap*ThrRS-1 has been successfully performed. Two subunits of *Ap*ThrRS-1 form a dimer, and each subunit contains the three motifs, 1, 2 and 3, in the amino acid sequence, as the typical features of the class II aminoacyl-tRNA synthetase. In comparison of the structure with those of ThrRS from *E. coli* (*Ec*ThrRS), it is easy to identify that *Ap*ThrRS-1 is composed of the catalytic and the anticodon-binding domains, but there is no editing domain, as expected.

9-2-c. Catalytic site of *Ap*ThrRS-1

A superimposition of *Ap*ThrRS-1 onto *Ec*ThrRS in complex with several substrates has shown that the residues interacting to threonine, ATP and $tRNA^{Thr}$ in *Ec*ThrRS are highly conserved in *Ap*ThrRS-1. When threonine enters the catalytic site, it is trapped by hydrogen bonds between the side chains of Arg144, Glu162, Glu383 and Tyr257. The methyl group of threonine is accepted in a hydrophobic environment of Thr280 and Ala312. The adenine moiety of ATP is surrounded by the side chains of Val157, Phe160 and Arg319, and the 3'-hydroxy group of the ribose interacts with the side chain of Gln277. The bound threonine points the carboxyl oxygen atom to the α-phosphorus atom of the ATP. This geometrical situation makes the two substrates easy to form Thr-AMP intermediate. When $tRNA^{Thr}$ enters the catalytic site, the adenine moiety at the 3'-end of the bound tRNA intercalates between Tyr91 and Arg144, and the N6 atom of the A76 residue forms a hydrogen bond with the main-chain carbonyl oxygen atom of Asn94. This situation forces Thr-AMP to transfer Thr to the A76 end of the bound $tRNA^{Thr}$.

9-2-d. Anticodon recognition of *Ap*ThrRS-1

The superimposed *Ap*ThrRS-1 described above also suggests that the residues interacting to the anticodon of $tRNA^{Thr}$ in *Ec*ThrRS are highly conserved in *Ap*ThrRS-1. Glu404 and Arg416 in *Ap*ThrRS-1 interact with the anticodon G35 and U36 of $tRNA^{Thr}$. The wobble base C34 is also able to interact with the Oγ atom of Thr382. These consistencies suggest that the anticodon-binding domain of *Ap*ThrRS-1 recognizes the

anticodon of $tRNA^{Thr}$.

9-2-e. Possible structure of ThrRS-2

As described above, the structural feature of ThrRS-1 suggests the necessity of ThrRS-2. As the N-terminal sequence of ThrRS-2 is similar to that of the editing domain of archaea ThrRS, ThrRS-2 is expected to form a dimer between the editing domains, and it catalyzes de-aminoacylation of the misacylated serine moiety at the CCA terminus. In addition, ThrRS-2 contains an additional domain different from archaea ThrRS.

9-2-f. Significance of the study and feature prospects

In this study, by solving the crystal structure of *Ap*ThrRS-1, it has been proved that ThrRS-1 can catalyze threonylation as well as serinylaiton of $tRNA^{Thr}$ without editing the misacylated tRNAs. Further X-ray analyses will be continued to solve the structures of several complexes with threonine, ATP and $tRNA^{Thr}$. For the structure of ThrRS-2, a possible structure has been proposed, but the structural study will also be continued to reveal the true structure and functions. Based on those structures, it will be possible to discuss the evolutionary relationship of the two-gene organisms.

List of publication

(1) Shimizu S., Juan, E. C. M., Miyashita Y., Sato Y., Suzuki K., Yogiashi M., Tsunoda M., Dock-Bregeon A. C., Moras D., Sekiguchi T. and Takénaka A., "Crystallographic study on two types of putative threonyl-tRNA synthetases fromcrenarchaea *Aeropyrum pernix* and those from *Sulfolobus tokodaii*", *Acta Crystallogr.*, **F64**, 903-910 (2008)

(2) Juan, E. C. M., Hoque M. M., Shimizu S., Hossain M. T., Yamamoto T., Imamura S., Suzuki K., Tsunoda M., Amano H., Sekiguchi T. and Takénaka A., "Structures of *Arthrobacter globiformis* urate oxidase-ligand complexes", *Acta Crystallogr.*, **D64**, 815-822 (2008)

(3) Hoque M. M., Shimizu S., Hossain M. T., Yamamoto T., Imamura S., Suzuki K., Tsunoda M., Amano H., Sekiguchi T. and Takénaka A., "The structures of *Alcaligenes faecalis* D-3-hydroxybutyrate dehydrogenase before and after NAD^+ and acetate binding suggest a dynamical reaction mechanism as a member of the SDR family", *Acta Crystallogr.*, **D64**, 496-505 (2008)

(4) Sato Y., Maeda Y., Shimizu S., Hossain M. T., Ubukata S., Suzuki K., Sekiguchi T. and Takénaka A., "Crystal structure of nondiscriminating aspartyl-tRNA synthetase from crenarchaea *Sulfolobus tokodaii* reveals the recognition mechanism of two different tRNA anticodons", *Acta Crystallogr.*, **D63**, 1042-1047 (2007)

(5) Suzuki K., Sato Y., Maeda Y., Shimizu S., Hossain M. T., Ubukata S., Sekiguchi T. and Takénaka A., "Crystallization and preliminary X-ray crystallographic study of putative aspartyl-tRNA synthetase from crenarchaea *Sulfolobus tokodaii* strain 7", *Acta Crystallogr.*, **F63**, 608-612 (2007)

(6) Kondo J., Ciengshin T., Juan, E. C. M., Sato Y., Mitomi K., Shimizu S. and Takénaka A., "Crystal Structure of d(gcGXGAgc) with X=G: a Mutation at X is Possible to Occur in a Base-Intercalated Duplex for Multiplex Formation", *Nucleosides Nucleotides Nucleic Acids*, **25**, 693-704 (2006)

Papers presentation at international conferences

(1) Shimizu S., Masanori Ohki, Nami Ohkubo, Suzuki K., Tsunoda M., Sekiguchi T. and Takénaka A., "Crystal structures of RNA 3'-terminal phosphate cyclase and its complexes with Mg^{2+}+ATP, ATP or Mn^{2+}", *Nucleic Acids Symp Ser*. **52**, 221-222 (2008).

(2) Haraguchi T, Shimizu S, Ma X, Kurose T, Juan E. C. M, Ohkubo A, Sekine M, Shibata T, Millington C. L., Williams D. M., Takénaka A. "Crystal structures of DNA duplexes stabilized by bicyclic-C residues" *Nucleic Acids Symp Ser*. **52**, 127-128 (2008)

(3) Takénaka A., Shimizu S., Miyashita Y., Sato Y., Juan, E. C. M., Suzuki K., Tsunoda M., Dock-Bregeon A. C., Moras D. and Sekiguchi T. "Two threonyl-tRNA synthetases with complementary functions; Crystal structure of ThrRS-1 from *Aeropyrum pernix*", *Acta Cryst*., **A64**, C311 (2008)

(4) Shimizu S., Ohki M., Ohkubo N., Suzuki K., Tsunoda M., Sekiguchi T. and Takénaka A. "RNA splicing related proteins; Crystal structure of RNA 3'-terminal phosphate cyclase", *Acta Cryst*., **A64**, C307 (2008)

(5) Suzuki K., Shimizu S., Ohbayashi R., Sato Y., Takénaka A., Sekiguchi T. and Yang S. Y. "Crystal structure of α-carbonic anhydrase from *Chlamydomonas reinhardtii*" *Acta Cryst*., **A64**, C263 (2008)

(6) Haraguchi T., Shimizu S., Ma X., Kurose T., Juan E. C. M., Sekine M., Shibata T., Williams D. M. and A. Takénaka "X-ray analyses of DNA duplexes stabilized by bicyclic-C residues" *Acta Cryst*., **A64**, C298 (2008)

(7) Hoque M. M., Shimizu S., Hossain M. T., Yamamoto T., Imamura S., Suzuki K., Tsunoda M., Amano H., Sekiguchi T. and Takénaka A. "The substrate recognition and the catalytic reaction mechanisms of D-3-hydroxybutyrate dehydrogenase" *Acta Cryst*. **A64**, C252 (2008)

(8) Shimizu S., Miyashita Y., Sato Y., Juan E. C. M., Suzuki K., Tsunoda M., Sekiguchi T. and Takénaka A. "Crystal structure of threonyl-tRNA synthetase ThrRS-1 from *Aeropyrum pernix* ", 22nd tRNA workshop, (2007)

(9) Suzuki K., Shimizu S., Ohbayashi R., Sato Y., Takénaka A., Sekiguchi T. and Yang S. Y. "Crystal structure of α-carbonic anhydrase from *Chlamydomonas reinhardtii*" 24th European Crystallographic Meeting (ECM24), (2007)

(10) Shimizu S., Miyashita Y., Sato, Y., Juan E. C. M., Suzuki K., Sekiguchi T. and Takénaka A., "Crystallizations and preliminary structural determinations of two types of putative threonyl-tRNA synthetases from *Sulfolobus tokodaii* strain 7 and those from *Aeropyrum pernix* K1", 2006 International Conference on Aminoacyl-tRNA Synthetase (AARS2006), (2006)

(11) Miyashita Y., Shimizu S., Sato Y., Juan E. C. M., Suzuki K., Sekiguchi T. and Takénaka A. "Crystallographic study on two types of threonyl-tRNA synthetases from *Sulfolobus tokodaii*", Joint Conference of the Asian Crystallographic Association and the Crystallographic Society of Japan (AsCA '06/CrSJ), (2006)

(12) Sato Y., Mitomi K., Sugiyama H., Shimizu S., Juan E. C. M., Kondo J. and Takenaka A., "Multiplex formation of repetitive sequences d([G]nA)m by non-Watson-Crick pairings", AsCA '06/CrSJ (2006)

(13) Hoque M. M., Juan E. C. M., Shimizu S., Kondou H., Hossain M. T., Suzuki K., Imamura S., Yamamoto T., Sekiguchi T. and Takénaka A., "X-Ray analyses of 3-hydroxybutyrate dehydrogenase from *Alcaligenes faecalis*"; AsCA'06/CrSJ, (2006)

(14) Sato Y., Mitomi K., Shimizu S., Juan E.C.M., Kondo J. and Takénaka A., "Duplex with non-WC pairings; crystal structure of d(gcGAGGGAgc)", 23rd European Crystallographic Meeting (ECM23), (2006)

(15) Shimizu S., Kobuna T., Sunami T., Kondo J and Takénaka A. "Crystal structures of d(CGAA) and d(CGAAGC); parallel-stranded duplexes and their dimmer", ECM23, (2006)

(16) Shimizu S., Kobuna T., Sunami T. and Takénaka A., "Crystal structure of d(CGAAGC); association of two parallel-stranded duplexes through anti-parallel-stranded duplex formation", *Nucleic Acids Symp Ser*. **48**, 119-20. (2004)

Acknowledgment

I am indebted to the many people who have supported and guided me throughout my studies. I wishes first to express my sincere thanks to my advisor Professor Akio Takénaka, who introduced me to this field of investigation, imparted his knowledge and advice on my work, offered various opportunities to me and encouraged me during the whole period of research work and in preparing this dissertation. I am also extremely grateful and express my heartfelt gratitude to Professor Mitsuo Sekine, who is bearing all of my responsibilities after retirement of Professor Akio Takénaka from this institution.

I have benefited greatly from the help of Professor Takeshi Sekiguchi, Dr. Kaoru Suzuki and Dr. Masaru Tsunoda of Iwaki Meisei University who prepared *Ap*ThrRS-1and *Ap*ThrRS-2, and advised me kindly. I also thank the outstanding staff of the Photon Factory, for their round-the-clock assistance during the X-ray diffraction experiments.

I extend my gratitude to the members of the Takénaka Laboratory; Dr. Md. Tofazzal Hossain, Dr. Jiro Kondo, Mr. Wataru Adachi, Mr. Kosei Kimura, Mr. Kazuhiro Fujita, Dr. Yoshiteru Sato, Dr. Ella Czarina Magat Juan, Mr. Toshiyuki Sasaki, Mr. Isao Koike, Mr. Yohei Maeda, Mr. Junichi Wakabayashi, Dr. Md. Mominul Hoque, Mr. Kenta Mitomi, Miss. Tanashaya Ciengshin, Mr. Keita Kikuchi, Mr. Yu-ichiro Miyashita, Mr. Hiroshi Kondo, Mr. Xiao Ma, Mr. Masanori Ohki, Mr. Shinsuke Kashiwai, Mr. Taizo Kurose, Mr. Hiromu Sugiyama, Mr. Tsuyoshi Haraguchi, Mr. Christopher L. Millington, Dr. Nami Okubo and Dr. Tomohiko Toyoda. And I would like to thank the members of the Sekine and Seio Laboratory, who allowed me to join for their seminars and helped in my moving office.

I would like to thank Professor Thomas Simonson of Laboratoire de Biochimie, Department of Biology, Ecole Polytechnique in France, Professor Shigehisa Hirose and Professor Akio Takénaka who introduced me to the internship of Ribosome Project at the Simonson's Laboratory in summer of 2006 and provided me with the financial supports. I would also like to thank the members of the Simonson's Laboratory, especially Dr. Alexey Aleksandrov.

I also thanks to the COE and G-COE for providing me with financial supports during my Doctoral course studies in these three years. I could fully concentrate on my research. It goes without saying that this thesis could not have been written without the strong support of my father, mother and brother.