Structural and functional studies of subtilisin-like serine protease, Tk-SP, from *Thermococcus kodakaraensis* with β-jelly roll domain

TITA FOOPHOW

Biotechnology English Course
Division of Advanced Science and Biotechnology
Graduate School of Engineering
Osaka University, Japan

2010

Structural and functional studies of subtilisin-like serine protease, Tk-SP, from *Thermococcus kodakaraensis* with β-jelly roll domain

TITA FOOPHOW

Biotechnology English Course
Division of Advanced Science and Biotechnology
Graduate School of Engineering
Osaka University, Japan

2010

TABLE OF CONTENTS

ABBREVIATIONS

Following abbreviations are used in this thesis:

Tk-SP:	an active mature form (Val^{114}-Val^{539}) of subtilisin-like serine protease from *Thermococcus kodakaraensis*
Pro-Tk-SP:	Tk-SP in a pro-form (Ala^{1}-Gly^{640})
ProN:	N-terminal propeptide of Tk-SP (Ala^{1}-Ala^{113})
ProC:	C-terminal propeptide of Tk-SP (Asp^{540}-Gly^{640})
Pro-Tk-S359A/C:	Pro-Tk-SP derivatives with the single Ser^{359}→Ala (Ser^{359}→Cys) mutation
Tk-S359A/C:	Tk-SP derivatives with the single Ser^{359}→Ala (Ser^{359}→Cys) mutation
ProN-Tk-S359A:	Pro-Tk-S359A derivatives lacking C-propeptide
ProC-Tk-S359A/C:	Pro-Tk-S359A/C derivatives lacking N-propeptide
Tk-S359A/CΔJ:	Tk-S359A/C derivatives lacking C-terminal β-jelly roll domain (Ala^{422}-Val^{539})
Tk-$S359A^{EDTA}$:	Tk-S359A treated with EDTA
Tk-$S359A\Delta J^{EDTA}$:	Tk-S359AΔJ treated with EDTA
ProN:ProC-Tk-S359C:	Complex between ProN and ProC-Tk-S359C
Suç-AAPF-*p*NA:	*N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide
MALDI-TOF:	matrix-assisted laser desorption ionization reflectron-type time-of-flight
GdnHCl:	guanidine hydrochloride
TCA:	trichloroacetic acid
CBB:	Coomassie brilliant blue.
CD:	circular dichroism
EDTA:	ethylenediaminetetraacetic acid
PEG:	polyethyleneglycol
PDB:	Protein data bank

CHAPTER 1

General Introduction

1.1. Serine proteases

Serine proteases are the large family of enzymes, which hydrolyzes the peptide bond of substrate via a nucleophilic Ser in the active site. They are widely distributed in various organisms including bacteria, archaea, and eukaryotes (Siezen and Leunissen, 1997). These enzymes have many physiological functions, ranging from generalized protein digestion to more specific regulated processes such as the activation of zymogens, blood clotting, and the immune system (Neurath, 1984). Barrett and Rawlings have devised a classification scheme based on statistically significant similarities in sequence and structure of all known proteolytic enzymes and terms this database MEROPS (Rawlings *et al.*, 2006). This classification system divides proteases into clans based on catalytic mechanism and families on the basis of common ancestry. The two largest clans are known; the trypsin and subtilisin clans, clan PA and clan SB, respectively. These two clans are distinguished by a highly similar arrangement of catalytic His, Asp, and Ser in radically different β/β (trypsin) and α/β (subtilisin) protein scaffolds.

The main player in the catalytic mechanism in the trypsin and subtilisin clan enzymes is the three amino acids: His, Asp, and Ser, which are called a catalytic triad. The triad is located in the active site of the enzyme, where catalysis occurs, and is preserved in all serine protease enzymes. Each amino acid in the triad performs a specific task in catalytic process. The Ser has an -OH group that is able to act as a nucleophile, attacking the carbonyl carbon of the scissile peptide bond of the substrate. A pair of electrons on the His nitrogen has the ability to accept the hydrogen from the Ser -OH group, thus coordinating the attack of the peptide bond. The carboxyl group on the Asp in turn hydrogen bonds with the His, making the pair of electrons mentioned above much more electronegative. The catalytic mechanism of serine proteases is summarized in Fig. 1.1 (Voet *et al.*, 2006).

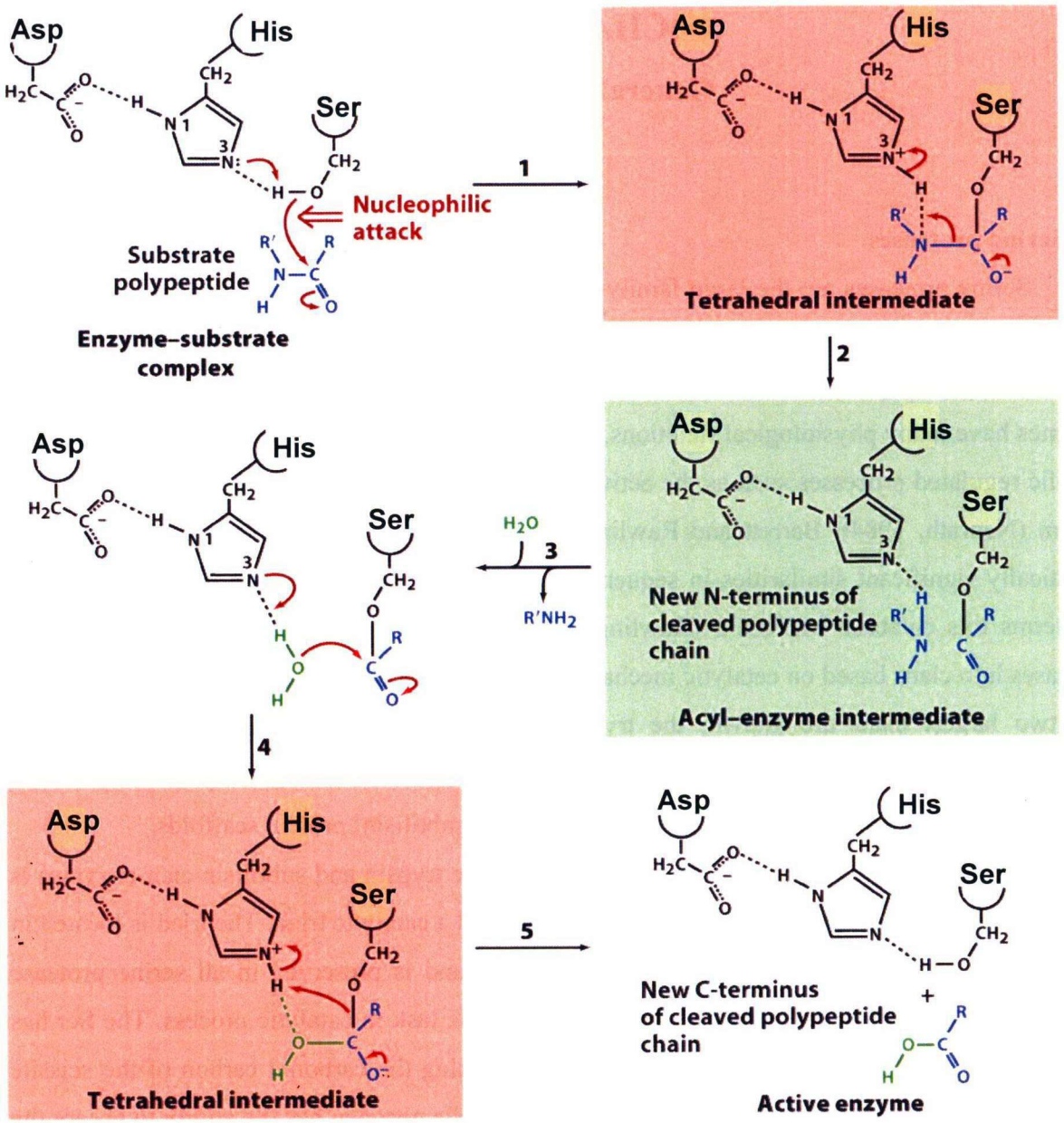

Figure 1.1. The catalytic mechanism of the serine protease. The reaction involves (1) the nucleophilic attack of the active site Ser on the carbonyl carbon atom of the scissile peptide bond to form the tetrahedral intermediate; (2) the decomposition of the tetrahedral intermediate to the acyl-enzyme intermediate through general acid catalysis by the active site Asp-polarized His, followed by loss of the amine product and its replacement by a water molecule; (3) the reversal of Step 2 to form a second tetrahedral intermediate; (4) the reversal of Step 1 to yield the reaction's carboxyl product and the active enzyme.

The catalytic mechanism of serine proteases based on considerable chemical and structural data is described in more detail below:

1. After protease has bound a substrate, Ser nucleophilically attacks the scissile peptide's carbonyl group to form the tetrahedral intermediate, which resembles the reaction's transition state (covalent catalysis). X-ray studies indicate that Ser is ideally positioned to carry out this nucleophilic attack (i.e., catalysis also occurs by proximity and orientation effects). This nucleophilic attack involves transfer of a proton to the imidazole ring of His, thereby forming an imidazolium ion (general base catalysis). This process is aided by the polarizing effect of the unsolvated carboxylate ion of Asp, which is hydrogen bonded to His (electrostatic catalysis). The tetrahedral intermediate has a well-defined, although transient, existence. We shall see that much of protease's catalytic power derives from its preferential binding of the transition state leading to this intermediate (transition state binding catalysis).

2. The tetrahedral intermediate decomposes to the acyl-enzyme intermediate under the driving force of proton donation from N3 of His (general acid catalysis). The amine leaving group ($R'NH_2$, the new N-terminal portion of the cleaved polypeptide chain) is released from the enzyme and replaced by water from the solvent. The acyl-enzyme intermediate is highly susceptible to hydrolytic cleavage.

3 and 4. The acyl-enzyme intermediate is deacylated by what is essentially the reversal of the previous steps followed by the release of the resulting carboxylate product (the new C-terminal portion of the cleaved polypeptide chain), thereby regenerating the active enzyme. In this process, water is the attacking nucleophile and Ser is the leaving group.

1.2. Subtilisin-like proteases (subtilases)

According to the MEROPS classification, subtilases constitute the S8 family within the SB clan of serine proteases. Among the subtilases, the subtilisin family (EC 3.4.21.108), which is represented by subtilisin E from *Bacillus subtilis* (Stahl and Ferrari, 1984), subtilisin BPN′ from *B. amyloliquefaciens* (Well *et al.*, 1983), and subtilisin Carlsberg from *B. licheniformis* (Jacobs *et al.*, 1985), has been the most extensively studied for structures and functions. The biochemical properties and crystal structures of these subtilisins have been determined to understand substrate specificity, enzymatic rate enhancements, and activity-stability-structure relationships (Wright *et*

al., 1969; Syed *et al.*, 1993; Jain *et al.*, 1998; Perona and Craik, 1995; Takagi and Takahashi, 2003). Because subtilisins are commercially valuable enzymes, attempts to improve their activity and stability with protein engineering technology have also been made extensively (Wells and Estell, 1988; Smith and Gottesman, 1989; Bryan, 2000).

1.2.1. Subtilisin: Secretion and maturation

The mechanism by which subtilisins are folded and matured has been the most extensively studied for bacterial subtilisins. According to these studies, subtilisins are first secreted into an external medium in a pro form (prosubtilisin) with the assistance of signal peptide, and activated upon autoprocessing and degradation of propeptide (Shinde and Inouye, 1996). Degradation of propeptide is necessary for the mature domain to function as an active enzyme, because the propeptide continues to bind tightly to the mature domain after autoprocessing and thereby inhibits its activity (Li *et al.*, 1995; Huang *et al.*, 1997; Yabuta *et al.*, 2001). It has been proposed that propeptides of subtilisins function not only as inhibitors of their cognate mature domains but also as intramolecular chaperones that facilitate folding of the mature domains after secretion (Eder *et al.*, 1993; Shinde *et al.*, 1997; Subbian *et al.*, 2005). The mature domains alone are not an active form but an inactive form with a molten globule-like structure in the absence of propeptides (Eder *et al.*, 1993; Shinde and Inouye, 1995). Dependence on the propeptide for maturation of its cognate mature domain has been reported not only for other members of the subtilisin family (Baier *et al.*, 1996; Marie-Claire *et al.*, 2001; Basak and Lazure, 2003) but also for other proteases (Silen and Agard, 1989; Smith and Gottesman, 1989; Winther and Sorensen, 1991; O'Donohue and Beaumont, 1996; Marie-Claire *et al.*, 1999; Nirasawa *et al.*, 1999). The proposed model for the activation pathway for subtilisin of bacterial origin is summarized in Fig. 1.2.

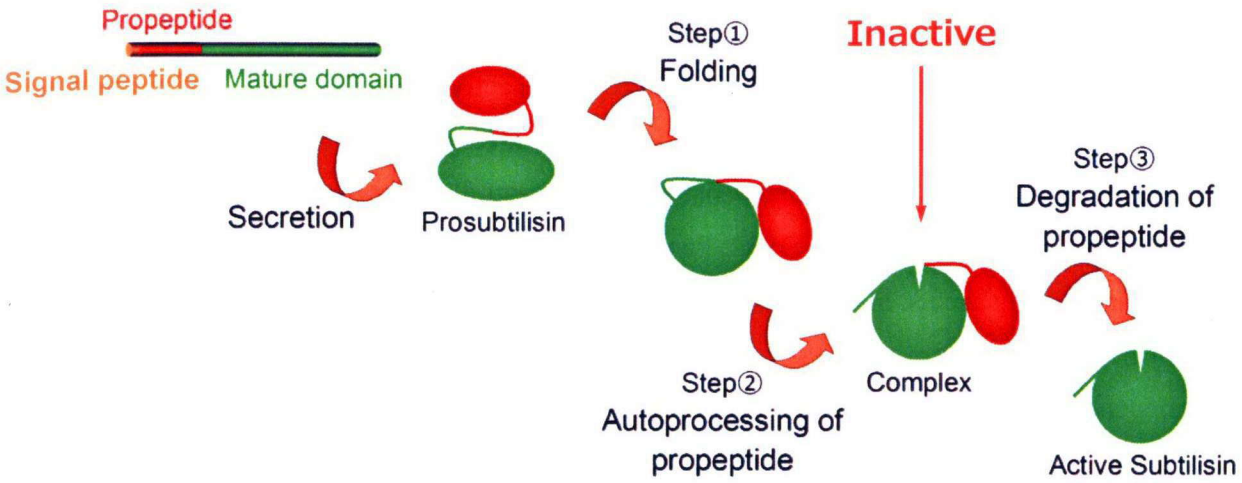

Figure 1.2. Maturation mechanism proposed for bacterial subtilisins. The signal peptide, propeptide, and mature domain are colored orange, red, and green, respectively.

1.2.2. Subtilisin: Structure and function

The crystal structures of bacterial subtilisins are available not only for the active mature form (Bode *et al.*, 1987; Pantoliano *et al.*, 1988) but also for the inactive complex between the propeptide and mature domain (Jain *et al.*, 1998; Gallagher *et al.*, 1995). Comparison of these structures indicates that the folding process of the mature domain is almost fully completed prior to degradation of the propeptide. The structures of the active mature domain fold into a conserved structure characterized by layers of the central α-helix, parallel β-sheet, and outer helices; this fold is named the subtilisin-like fold (Murzin *et al.*, 1995). The crystal structure of subtilisin BPN′ (PDB code 2ST1) is shown in Fig. 1.3.

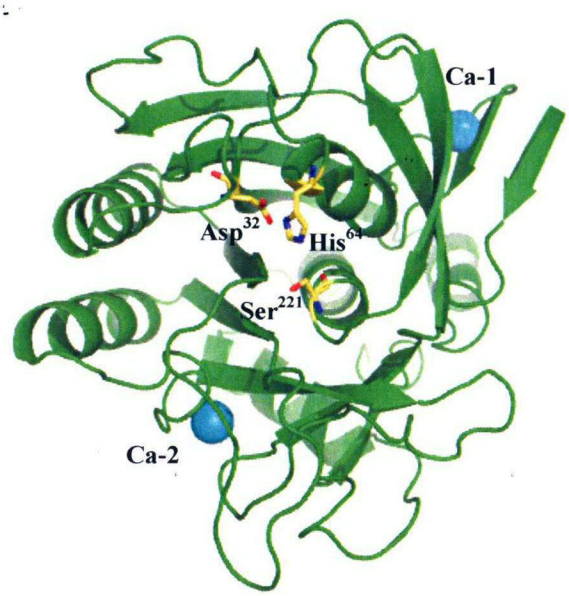

Figure 1.3. Overall structure of subtilisin BPN′. The catalytic triad residues (Asp^{32}, His^{64}, and Ser^{221}) are idicated by *yellow stick models*, in which oxygen and nitrogen atoms are *colored red* and *blue*, respectively. Two Ca^{2+} ions (Ca-1 and Ca-2) are shown in *cyan spheres*.

The subtilisin structures also indicate that bacterial subtilisins have two Ca^{2+} binding sites, Ca-1 site and Ca-2 site, both of which are located far from the active site. Ca-1 and Ca-2 represent high-affinity and low-affinity binding sites of the Ca^{2+} ions, respectively. The enzyme is greatly destabilized with respect to both thermal denaturation and autodegradation upon removal of Ca-1 (Voordouw *et al.*, 1976; Bryan *et al.*, 1992; Pantoliano *et al.*, 1989). Prosubtilisin E is folded and autoprocessed in the absence of Ca^{2+} (Yabuta *et al.*, 2002). Deletion of the loop forming Ca-1 of subtilisin BPN′, followed by directed evolution and selection for increased stability, results in a Ca^{2+}-independent subtilisin mutant with native-like activity (Bryan *et al.*, 1992; Gallagher *et al.*, 1993; Strausberg *et al.*, 1995). These results suggest that Ca^{2+} is not required for activity or folding of subtilisin but is required for stability. However, the number and location of the Ca^{2+} binding sites vary greatly for other members of the subtilase family. For example, thermitase has three Ca^{2+} binding sites (Gros *et al.*, 1988), proteinase K (Betzel *et al.*, 1988) and Ak.1 protease (Smith *et al.*, 1999) have four sites, sphericase has five sites (Almog *et al.*, 2003), and Tk-subtilisin has seven sites (Tanaka *et al.*, 2007b), all of which are located far from the active site, but do not necessarily correlate to Ca-1 or Ca-2 of bacterial subtilisins. In addition, various subtilases, such as sphericase (Almog *et al.*, 2003), cell envelope protease (Exterkate and Alting, 1999), and psychrophilic subtilisin (Davail *et al.*, 1994), have been reported to exhibit Ca^{2+}-dependent activity. These results suggest that the role of the Ca^{2+} ions varies for different subtilases.

1.2.3. Subtilisin: Application

Proteases have been widely used for industrial purposes, mainly for laundry detergents and dishwashing detergents. They have been used in household detergents for more than 40 years and are considered as a basic ingredient and are present in around 75% of the detergents world-wide (Schäfer *et al.*, 2007). Proteases are hydrolases that catalyse the breakdown of proteins and thus enhance the cleaning of proteinaceous stain and typical stains including blood, milk, egg, grass, and sauces. All proteases currently used in detergents belong to the class of serine proteases originating from the *Bacillus* species, such as *B. licheniformis*, *B. amyloliquefaciens*, *B. alkaophilis*, *B. lentus*, *B. halodurans*, and *B. clausii*. To be well suited for use in a typical detergent, some subtilisins have been modified by protein engineering to impove their stability and activity at broad temperature and pH ranges and in the presence of other

detergent ingredients. Subtilisin variants used in detergents are summarized in Table 1.1 (Maurer, 2004).

Table 1.1. Subtilisin variants used in detergents

Trade mark	Producer	Origin	WT/PE[c]	Production strain	Synonym
Alcalase®	Novozymes	*B. licheniformis*	WT	*B. licheniformis*	Subtilisin Carlsberg
FNA[a]	Genencor	*B. amyloliquefaciens*	PE	*B. subtilis*	
Savinase®	Novozymes	*B. clausii*	WT	*B. clausii*	Subtilisin 309
Purafect™	Genencor	*B. lentus*	WT	*B. subtilis*	
KAP[b]	Kao	*B. alkalophilus*	WT	*B. alkalophilus*	
Everlase™	Novozymes	*B. clausii*	PE	*B. clausii*	
Purafect OxP™	Genencor	*B. lentus*	PE	*B. subtilis*	
FN4[a]	Genencor	*B. lentus*	PE	*B. subtilis*	
BLAP S[b]	Henkel	*B. lentus*	PE	*B. licheniformis*	
BLAP X[b]	Henkel	*B. lentus*	PE	*B. licheniformis*	
Esperase®	Novozymes	*B. halodurans*	WT	*B. halodurans*	Subtilisin 147
Kannase™	Novozymes	*B. clausii*	PE	*B. clausii*	
Properase™	Genencor	*B. alkalophilus PB92*	PE	*B. alkaliphilus*	

[a]Exclusive molecules for specific customer.
[b]Exclusive molecules for captive use. The names of captive use products are often based on technical terms or acronyms.
[c]PE, protein engineered; WT, wild type.

1.3. Three subtilisin homologues from *Thermococcus kodakaraensis*

In recent years, there has been an increasing interest in hyperthermophilic and thermophilic microorganisms because they are considered as a source of industrially useful enzymes for biotechnological applications. Enzymes from heat-adapted microorganisms with highly specific activities at high temperatures have potential uses not only in various industrial fields, such as food processing, detergent business, and waste disposal business, but also in basic research field. These enzymes are usually highly thermostable and active at elevated temperatures. Therefore, they are good models to investigate the stabilization mechanism of proteins from thermophiles.

Thermococcus kodakaraensis is a hyperthermophilic archaea isolated from a solfatara at vent of Kodakara Island, Kagoshima, Japan. The growth temperature of the strain ranged from 65 to 95°C, and the optimal temperature was 90°C (Atomi *et al.*, 2004). This strain is expected to be a promising source of commercially valuable enzymes, because KOD polymerase for PCR (Toyobo) and hyper stable thiol protease (Morikawa *et al.*, 1994) have been isolated from this strain. The genome of *T. kodakaraensis* contains three genes encoding subtilisin homologues,

Tk-0076 (accession no. **YP 182489**), Tk-1675 (Prepro-Tk-subtilisin) (accession no. **BAB60701**), and Tk-1689 (Prepro-Tk-SP) (accession no. **YP 184102**), (Fukui *et al.*, 2005) as shown in Fig. 1.4. The primary structures of these proteins are schematically shown in Fig. 1.5 in comparison with that of subtilisin E.

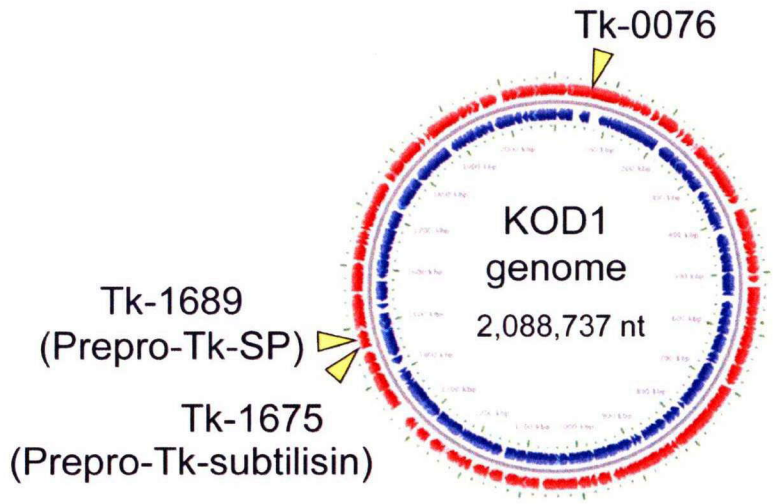

Figure 1.4. *Thermococcus kodakaraensis* genome. The locations of three subtilisin homologues are indicated.

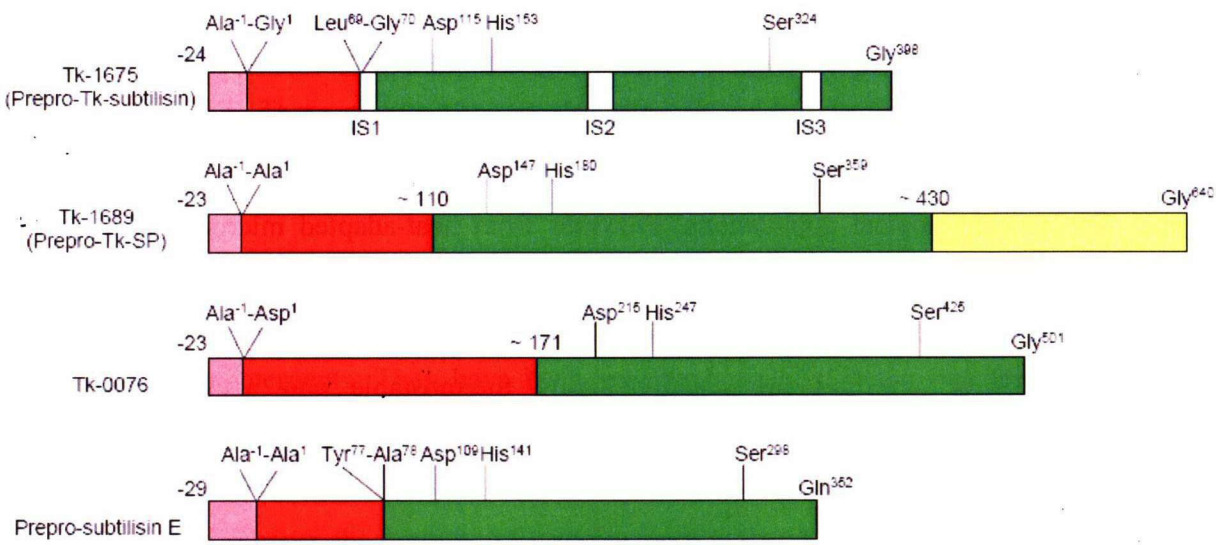

Figure 1.5. Schematic representation of the primary structures of Tk-1675 (Prepro-Tk-subtilisin), Tk-1689 (Prepro-Tk-SP), Tk-0076, and Prepro-subtilisin E. Pink box represents a signal peptide, red box represents a propeptide, green box represents a mature domain, yellow box represents C-terminal extension, and open box represents an insertion sequence (IS). The locations of the active site residues are shown. The accession numbers of these sequences are BAB60701 for Prepro-Tk-subtilisin, YP184102 for Prepro-Tk-SP, YP182489 for Tk-0076, and AAA22742 for Prepro-subtilisin E.

Of these subtilisin homologues from *T. kodakaraensis*, Tk-subtilisin has been extensively studied structurally and functionally. Tk-subtilisin (Gly^{70}-Gly^{398}) is matured from Pro-Tk-subtilisin (Gly^{1}-Gly^{398}) upon autoprocessing and degradation of Tk-propeptide (Gly^{1}-Leu^{69}) (Pulido *et al.*, 2006). Pro-Tk-subtilisin is overproduced in *E. coli* in inclusion bodies, purified as a denatured form in the presence of 8 M urea, and refolded by removing urea. Tk-subtilisin is a highly thermostable serine protease with a half-life of 50 min at 100°C (Kannan *et al.*, 2001; Pulido *et al.*, 2006). Tk-propeptide functions not only as an intramolecular chaperone (Tanaka *et al.*, 2007a, 2008) but also as a strong inhibitor (Pulido *et al.*, 2006, 2007a, 2007b) of Tk-subtilisin. However, the maturation process of Tk-subtilisin is different from those of other subtilisins and subtilisin-like serine proteases in terms of the requirement of Ca^{2+} for folding of the mature domain. Tk-subtilisin requires Ca^{2+} for folding and assumes a molten globule-like structure in the absence of Ca^{2+} even in the presence of Tk-propeptide (Tanaka *et al.*, 2007a). In contrast, bacterial subtilisins and their homologues do not require Ca^{2+} for folding, but require it for stability (Voordouw *et al.*, 1976; Pantoliano *et al.*, 1989; Gros *et al.*, 1991; Bryan *et al.*, 1992; Smith *et al.*, 1999).

The crystal structure of the active-site mutant of Pro-Tk-subtilisin, Pro-S324A, which represents the unautoprocessed form (Tanaka *et al.*, 2007a), as well as those of the autoprocessed and mature forms (Tanaka *et al.*, 2007b), have been determined. The structure of the unautoprocessed form is basically identical to that of the autoprocessed form, except that the scissile peptide bond contacts with the active site, indicating that the folding process of Tk-subtilisin is almost fully completed prior to autoprocessing. The structure of the autoprocessed from is shown in Fig. 1.6. This structure is similar to those of the subtilisin E:propeptide (Jain *et al.*, 1998) and subtilisin BPN′:propeptide (Gallagher *et al.*, 1995) complexes. However, the structure of Tk-subtilisin is different from those of bacterial subtilisins in the number of the Ca^{2+}-binding sites. Tk-subtilisin has seven Ca^{2+}-binding sites, while bacterial subtilisins have only two. The Ca-1 site of Tk-subtilisin is conserved as the high affinity Ca^{2+}-binding site (Ca-1) in the structures of bacterial subtilisins, while other six sites are unique for Tk-subtilisin.

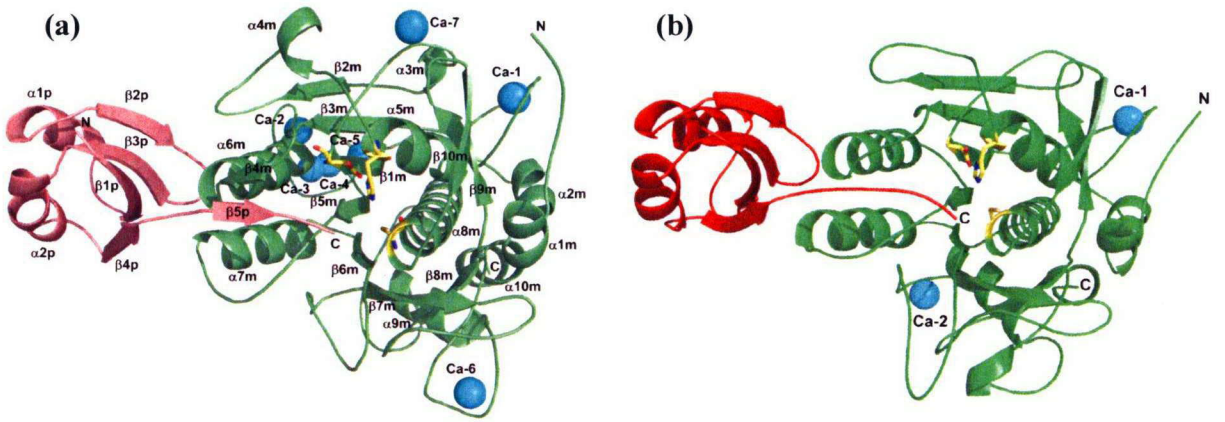

Figure 1.6. Three-dementional structures of the autoprocessed form of (a) Tk-subtilisin and (b) subtilisin E. The propeptide domain and mature domain are *colored pink* and *lime green* for Tk-subtilisin and *red* and *green* for subtilisin E, respectively. Two active-site residues (Asp and His) and Cys, which is substituted for the active-site serine residue, are indicated by *yellow stick models* in which the oxygen and nitrogen atoms are *colored red* and *blue*, respectively. Calcium ions are shown in *cyan spheres*. N and C represent the N and C-termini, respectively.

Tk-subtilisin has three insertion sequences as compared to bacterial subtilisins (Fig. 1.5). According to the crystal structure of Tk-subtilisin, one of the insertion sequences, IS2, assumes a long Ca^{2+}-binding loop (Gly^{206}-Glu^{229}) between α6m-helix and β5m-strand (Fig. 1.6). This loop is rich in aspartic acid residues and contains four Ca^{2+}-binding sites (Ca-2-Ca-5). Mutational and structural studies indicate that this Ca^{2+}-binding loop is required for folding of Tk-subtilisin (Takeuchi *et al.*, 2009). This loop is probably unfolded in the absence of Ca^{2+}, due to extensive negative charge repulsions among aspartic acid residues. However, in the presence of Ca^{2+}, these aspartic acid residues come close with one another to form Ca^{2+}-binding sites and directly coordinate with the Ca^{2+} ions at these sites. As a result, the Ca^{2+}-binding loop is folded into a correct structure, which permits proper arrangement of α6m-helix and β5m-strand and thereby folding of a central αβα substructure. α6m-helix, β5m-strand, and α7m-helix form a central αβα substructure, and formation of this αβα substructure has been proposed to be crucial for folding of bacterial subtilisins (Gallagher *et al.*, 1995; Bryan *et al.*, 1995). Insertion sequences 1 and 3 are required for stabilization of Tk-subtilisin, because they provide the Ca-1 and Ca-6 sites, respectively, and these sites contribute to the stabilization of Tk-subtilisin (Y. Takeuchi, unpublished results).

As to the propeptide-catalyzed folding, the C-terminal extended region of Tk-propeptide is not critical for folding of Tk-subtilisin, but accelerate it by binding to a folding intermediate of Tk-subtilisin with a native-like structure at its binding site, while Glu^{201}-mediated interactions are critical for initiation of propeptide-catalyzed folding of Tk-subtilisin (Tanaka *et al.*, 2009).

In contrast, the structures, functions, and maturation mechanisms of Tk-0076 and Tk-1689 (Prepro-Tk-SP) have not been analyzed at all.

1.4. Objective of the study

Tk-1689 (Prepro-Tk-SP) without a putative N-terminal signal sequence consists of 640 amino acid residues and shows the amino acid sequence identity of 29 % to Pro-Tk-subtilisin. All three amino acid residues that form a catalytic triad of Tk-subtilisin are fully conserved in the Tk-1689 sequence as Asp^{147}, His^{180}, and Ser^{359}. However, Tk-1689 has a long C-terminal extension as compared to Prepro-Tk-subtilisin and Prepro-subtilisin E (Fig. 1.5). In addition, none of the Ca^{2+}-binding sites of Tk-subtilisin, including the Ca^{2+}-binding loop, is conserved in the Tk-1689 sequence, suggesting that Tk-1689 has unique Ca^{2+}-binding sites. Then, the questions arise whether the C-terminal extension of Tk-1689 is folded into a unique structure, Tk-1689 exhibits unique enzymatic properties, and Tk-1689 is folded, matured, and stabilized with unique mechanisms. To answer these question, we decided to overproduce Tk-1689 in *E. coli*, purify, and structurally and biochemically characterize the protein. We focus on the structural and functional studies of Tk-1689 in this study, because attempts to overproduce Tk-0076 in functional form have so far been unsuccessful (T. Foophow, unpublished results).

The thesis is divided into five Chapters. In Chapter 1, definition, classification, and catalytic mechanism of serine proteases, structures, functions, and maturation process of bacterial subtilisins, biotechnological applications of subtilisins, and structure, function, and maturation mechanism of a subtilisin homologue, Tk-subtilisin, from *Thermococcus kodakaraensis* are summarized. In Chapter 2, the gene encoding the Tk-1689 derivative without a putative N-terminal signal sequence, termed Pro-Tk-SP (Ala^{1}-Gly^{640}), was overexpressed in *E. coli*. The recombinant protein was purified, however, as an active 44 kDa protease, termed Tk-SP, which lacks the N-terminal 113 and C-terminal 101 amino acid residues. The activity and stability of Tk-SP were analyzed and compared with those of Tk-subtilisin. Based on these

results, we discuss autoprocessing of Pro-Tk-SP and the role of propeptide and Ca^{2+} ion. In Chapter 3, the crystal structure of the active-site mutant of the Pro-Tk-SP derivative lacking a C-propeptide (ProN-Tk-S359A) was determined at 2.0 Å. The overall structure of this protein is similar to those of an autoprocessed form of Pro-Tk-subtilisin (Tk-S324A:propeptide complex) and bacterial subtilisin:propeptide complex, except that ProN-Tk-S359A contains a β-jelly roll domain at the C-terminus and does not contain any Ca^{2+} ion in the subtilisin domain. The role of the β-jelly roll domain was analyzed by using the Tk-SP derivatives (Tk-S359A/C and Tk-S359A/CΔJ). In Chapter 4, intermediate form of maturation process of Tk-SP was identified. The maturation process of Tk-SP is not similar to that of Tk-subtilisin because Tk-SP has C-propeptide and does not require Ca^{2+} for folding. The maturation process of Tk-SP was analyzed using Pro-Tk-SP derivative with the $Ser^{359} \rightarrow Cys$ mutation, Pro-Tk-S359C, because this mutation greatly reduces the enzymatic activity of Tk-SP by ~2000 fold and the maturation process of Tk-SP can be detected by SDS-PAGE. Lastly, in Chapter 5, the research works performed in this study are briefly summarized and future prospects on the structural and functional studies of Tk-SP and its possible biotechnological applications are described.

CHAPTER 2

Maturation of Tk-SP from its precursor form with N- and C-terminal propeptides

2.1. Introduction

Thermococcus kodakaraensis is a hyperthermophilic archaeon, which grows most optimally at 90°C (Atomi *et al.*, 2004). Its genome contains three genes encoding subtilisin-like serine proteases, Tk-0076 (accession no. **YP 182489**), Tk-subtilisin (accession no. **BAB60701**), and Tk-1689 (accession no. **YP 184102**) (Fukui *et al.*, 2005). The structure and function of Tk-subtilisin have been extensively studied as described in Chapter 1 (Kannan *et al.*, 2001; Pulido *et al.*, 2006, 2007a, 2007b; Tanaka *et al.*, 2006, 2007a, 2007b, 2008, 2009; Takeuchi *et al.*, 2009). However, other two archeal subtilisin-like proteases, Tk-1689 and Tk-0076, have not been characterized at all and therefore it remains to be determined whether they are activated by the mechanism similar to that proposed for Tk-subtilisin and exhibit similar enzymatic properties as Tk-subtilisin. Because Tk-0076 could not be overproduced in a functional form (T. Foophow, unpublished results), we decided to overproduce, purify, and biochemically characterize Tk-1689 in this study.

Tk-1689 is composed of 663 amino acid residues and contains a putative N-terminal signal sequence. In this report, we overexpressed the gene encoding the Tk-1689 derivative lacking this signal sequence, termed Pro-Tk-SP (Ala^{1}-Gly^{640}), in *E. coli*. We showed that the recombinant protein purified from the *E. coli* cells, termed Tk-SP, exhibits protease activity, but lacks the N-terminal 113 and C-terminal 101 amino acid residues of Pro-Tk-SP. Like Tk-subtilisin, Tk-SP is resistant to heat, denaturants, and detergents. However, unlike Tk-subtilisin, Tk-SP does not require Ca^{2+} for folding and is therefore resistant to EDTA as well.

2.2. Materials and Methods

2.2.1. Plasmid construction

The gene encoding Pro-Tk-SP (Met + Ala^{1}-Gly^{640}) was amplified by PCR with a combination of forward (5'-GGCCTTTATCATATGGCCCCCCAGAAG-3') and reverse (5'-GGCCTTGGATCCTCACCCGTAGTAAAC-3') primers, in which the *Nde*I and *Bam*HI sites are underlined, respectively. The genomic DNA of *T. kodakaraensis* (Kannan *et al.*, 2001) was used as a template. The resultant DNA fragment was digested with *Nde*I and *Bam*HI, and ligated into the *Nde*I-*Bam*HI sites of pET25b (Novagen) to generate pET25b-pro-Tk-sp. This plasmid was used as a template to construct the plasmid pET25b-pro-Tk-S359A for overproduction of the active site mutant Pro-Tk-S359A by PCR using the QuikChange II Site-directed mutagenesis kit (Stratagene). The mutagenic primers were designed such that the codon for Ser^{359} (TCG) is changed to GCG for Ala. All DNA oligomers for PCR were synthesized by Hokkaido System Science. PCR was performed in 25 cycles using a thermal cycler (Gene Amp PCR System 2400; Applied Biosystems) and KOD DNA polymerase (Toyobo). The DNA sequences of the genes encoding Pro-Tk-SP and Pro-Tk-S359A were confirmed by ABI Prism 310 DNA sequencer (Applied Biosystems).

2.2.2. Overproduction and purification

For gene expression, the *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) were transformed with the pET25b derivatives. These transformants were grown at 37°C in NZCYM medium (Novagen) containing 50 μg/ml ampicillin and 35 μg/ml chloramphenicol. When the optical density (O.D.) at 600 nm of the culture reached ~0.8, 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to the culture medium and cultivation was continued for an addition 4 h. Cells were then harvested by centrifugation at 6000*g* for 10 min at 4°C.

All protein purification procedures were performed at 4°C. For purification of Tk-SP, cells were suspended in 20 mM Tris-HCl (pH 9.0), disrupted by sonication and centrifuged at 30000*g* for 30 min to remove insoluble materials. The protein was precipitated by adding ammonium sulfate to the resultant supernatant to 30% saturation. The pellet was collected by centrifugation at 30000*g* for 30 min and dissolved in 20 mM Tris-HCl (pH 7.5). The resultant solution was dialyzed against 20 mM Tris-HCl (pH 7.5) and applied to a column (5 ml) of

HitrapQ (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column by linearly increasing the NaCl concentration from 0 to 1.0 M in the same buffer (20 column volumes in total). The fractions containing the protein were pooled, dialyzed against 20 mM Tris-HCl (pH 7.5), concentrated appropriately using the Centricon (Millipore) ultrafiltration system, incubated at 80°C for 2 h and used for further characterization.

For purification of Pro-Tk-S359A, sonication lysis of the cells, ammonium sulfate precipitation and the Hitrap Q column chromatography were carried out as described above for purification of Tk-SP. After this column chromatography, the fractions containing the protein were pooled, concentrated using the Centricon ultrafiltration system and loaded onto a Sephacryl S-200HR column (GE Healthcare) equilibrated with 20 mM Tris-HCl (pH 7.5) containing 50 mM NaCl. The fractions containing the protein were collected and dialyzed against 20 mM Tris-HCl (pH 7.5).

Tk-subtilisin was overproduced in *E. coli* and purified in a pro form (Pro-Tk-subtilisin), and matured from Pro-Tk-subtilisin as described previously (Pulido *et al.*, 2006).

The production level of the protein in the *E. coli* cells and the purity of the protein were analyzed by Tricine (*N*-[2-hydroxy-1,1-bis (hydroxymethyl)ethyl]glycine)-SDS-PAGE (SDS-PAGE using Tricine buffer) (Schägger, 2006) using a 15% polyacrylamide gel, followed by staining with Coomassie brilliant blue (CBB). For these analyses, the sample was prepared by washing the *E. coli* cells with 10% (v/v) trichloroacetic acid (TCA), suspending them in SDS sample buffer [50 mM Tris-HCl (pH 6.8), 0.1 M dithiothreitol, 2% SDS, 10% (v/v) glycerol, and 0.005% (w/v) bromophenol blue], neutralizing the pH of the resultant solution (whole cell extract), and boiling it for 5 min, or by precipitating the protein by 10% (v/v) TCA, dissolving the pellet in SDS sample buffer, neutralizing the pH of the resultant solution, and boiling it for 5 min.

The N-terminal amino acid sequence of the protein was determined by a Procise automated sequencer model 491 (Applied Biosystems). The protein concentration was determined from the UV absorption at 280 nm by the use of an $E^{0.1\%}$ value of 1.80 for Pro-Tk-S359A, 1.83 for Tk-SP, and 1.47 for Tk-subtilisin. These values were calculated by using ε = 1526 M^{-1} cm^{-1} for Tyr and 5225 M^{-1} cm^{-1} for Trp at 280 nm (Goodwin and Morton, 1946).

2.2.3. Molecular mass

The molecular mass of the protein was determined by a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometer (Autoflex or Ultraflex, Bruker Daltonik GmbH). Mass calibration was performed using protein calibration standard II (Bruker Daltonik GmbH). Raw data were analyzed by the program Findpept World Wide Server (Bruker Daltonik GmbH). The molecular mass of the protein was also estimated by gel filtration column chromatography using TSK-GEL G2000SWXL (Tosoh). Bovine serum albumin (67 kDa), ovalbumin (44 kDa), chymotrypsinogen A (25 kDa), and RNase A (14 kDa) were used as standard proteins.

2.2.4. Activity staining of gel (gel assay)

SDS-PAGE was carried out using a 12% polyacrylamide gel containing 0.1% (w/v) gelatin. Samples were precipitated with TCA and boiled for 5 min in SDS sample buffer before they were loaded onto the gel. After electrophoresis, the gel was washed by 2.5% (v/v) Triton X-100 at room temperature for 1 h, incubated in 50 mM Tris-HCl (pH 9.0) at 80°C for 2 h and stained with CBB. Protease bands were visualized as clear zones due to the hydrolysis of gelatin.

2.2.5. Enzymatic activity

The enzymatic activity was determined by using *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-*p*NA) (Sigma) as a peptide substrate. The reaction mixture (100 μl) contained 50 mM Tris-HCl (pH 7.5) for Tk-SP or 50 mM Tris-HCl (pH 8.0) for Tk-subtilisin and 2 mM Suc-AAPF-*p*NA. The amount of *p*-nitroaniline released from the substrate was determined from the absorption at 410 nm with an absorption coefficient of 8900 M^{-1} cm^{-1} by automatic UV spectrophotometer (model DU640, Beckman Coulter Inc.). One unit of enzymatic activity was defined as the amount of the enzyme that produced 1 μmol of *p*-nitroaniline per minute. The specific activity was defined as the enzymatic activity per milligram of protein. For kinetic analyses, the concentration of Suc-AAPF-*p*NA was varied from 0.01 to 2 mM for Tk-SP and from 0.5 to 5 mM for Tk-subtilisin. The enzymatic reaction followed the Michaelis-Menten kinetics, and the K_m and k_{cat} values were determined from the Lineweaver-Burk plots.

The enzymatic activity was also determined by using azocasein (Sigma) as a protein substrate at various temperatures. The reaction mixture (300 μl) contained 50 mM Tris-HCl (pH 7.5) for Tk-SP or 50 mM Tris-HCl (pH 8.0) for Tk-subtilisin and 2% (w/v) azocasein. The enzymatic reaction was initiated by adding an appropriate amount of the enzyme and terminated by adding 200 μl of 15% (v/v) TCA. The reaction time was usually 20 min. After centrifugation at 15000*g* for 15 min, an aliquot of the supernatant (160 μl) was withdrawn, mixed with 40 μl of 2 M NaOH and measured for absorption at 440 nm (A_{440}). One unit of enzymatic activity was defined as the amount of the enzyme that increased the A_{440} value of the assay reaction mixture by 0.1 in 1 min.

2.2.6. Cleavage of oxidized insulin chain B

Oxidized insulin chain B (Sigma) was digested by Tk-SP or Tk-subtilisin with an enzyme/substrate ratio of 1:10 (by weight) in 50 mM Tris-HCl (pH 7.5) or 50 mM CAPS-NaOH (pH 9.5), respectively, at 80°C for 30 min. Then, 0.1% (v/v) trifluoroacetic acid (TFA) was added to inactivate the enzyme and the sample was loaded onto Vivapure C18 micro spin column for concentration, purification, and desalting of peptides. The peptides were identified by determining their molecular masses by MALDI-TOF mass spectrometer as mentioned above, except that peptide calibration standard II (Bruker) was used.

2.2.7. Circular dichroism spectra

The far-UV (200-260 nm) circular dichroism (CD) spectra were measured at 25°C on a J-725 automatic spectropolarimeter (Japan Spectroscopic). The protein was dissolved in 20 mM Tris-HCl (pH 7.5). The protein concentration was 0.1 mg/ml and a cell with an optical path length of 2 mm was used. The mean residue ellipticity, θ, which has units of deg cm^2 $dmol^{-1}$, was calculated by using an average amino acid molecular weight of 110.

2.3. Results and Discussion

2.3.1. Amino acid sequence

In this study, Tk-1689 is designated as Prepro-Tk-SP, on the assumption that it consists of a signal sequence, a propeptide and a serine protease (Tk-SP) domain. The amino acid sequence of Prepro-Tk-SP is compared with those of a subtilisin-like serine protease from *Pyrococcus furiosus*, which is termed Prepro-Pf-SP in this study, prepro-aqualysin-I from *Thermus aquaticus* YT-1, prepro-subtilisin E from *Bacillus subtilis*, Prepro-Tk-subtilisin and Tk-0076 in Fig. 2.1. Pf-SP has not been biochemically characterized at all, like Tk-SP. Tk-SP in a putative pro form (Pro-Tk-SP, Ala^{1}-Gly^{640}) shows the amino acid sequence identities of 84% to putative Pro-Pf-SP (Ala^{1}-Gly^{631}), 31% to pro-aqualysin-I (Val^{1}-Pro^{499}), 29% to pro-subtilisin E (Ala^{1}-Gln^{352}), 29% to Pro-Tk-subtilisin (Gly^{1}-Gly^{398}), and 28% to a putative pro form of Tk-0076 (Asp^{1}-Gly^{501}). All three amino acid residues that form a catalytic triad of subtilisin-like serine proteases are fully conserved in the Prepro-Tk-SP sequence as Asp^{147}, His^{180}, and Ser^{359}. The Asn that is required to form an oxyanion hole is also conserved as Asn^{280}.

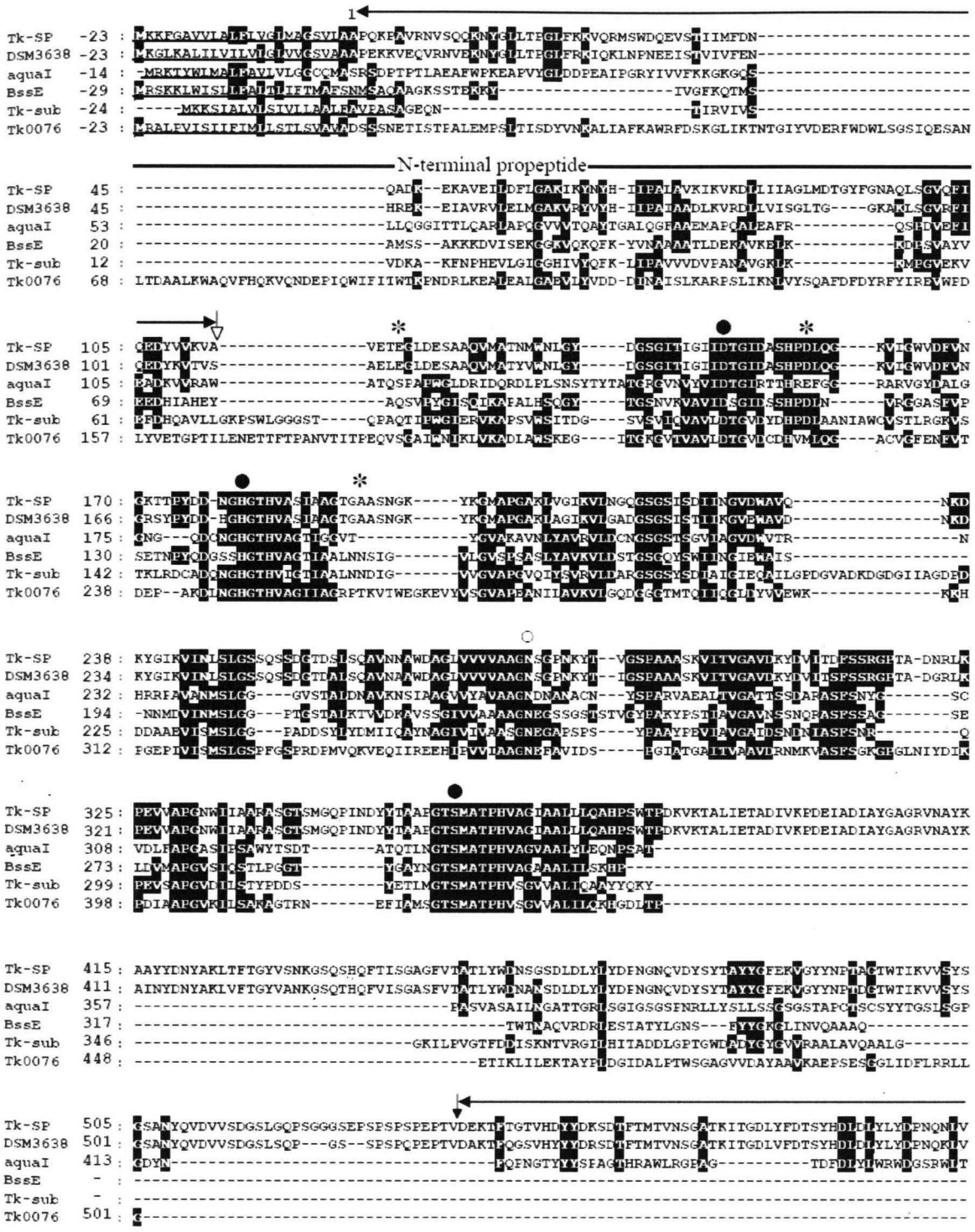
1
N-terminal propeptide
C-terminal propeptide

Figure 2.1. Alignment of the amino acid sequences of subtilisin-like serine proteases. The amino acid sequence of Prepro-Tk-SP (Tk-SP) is compared with those of Prepro-Pf-SP (DSM3638), prepro-aqualysin-I (aqual), prepro-subtilisin E (BssE), Prepro-Tk-subtilisin (Tk-sub) and Tk-0076. Gaps are denoted by dashes. The amino acid residues, which are conserved in at least three different proteins, are highlighted in black. The amino acid residues that form a catalytic triad and the Asn residue that forms an oxyanion hole are denoted by solid and open circles, respectively. The amino acid residues with the side chains, which coordinate with the Ca^{2+} ion at Ca-1 site of subtilisin E or Tk-subtilisin, are indicated by asterisks. The signal sequences are underlined. They are experimentally determined for aqual (Terada *et al.*, 1990) and BssE (Wong and Doi, 1986) or are estimated for other proteins using the SignalP V2.0 World Web Server. The open arrow indicates the position of the peptide bond, at which the N-terminal propeptides of Pro-Tk-SP, pro-aqualysin-I, pro-subtilisin E and Pro-Tk-subtilisin are autoprocessed. The filled arrow indicates the position of the peptide bond, at which the C-terminal propeptide of Pro-Tk-SP is autoprocessed. The C-terminal propeptide of pro-aqualysin-I is autoprocessed at Gly^{394}-Ser^{395}. The ranges of the N- and C-terminal propeptides of Tk-SP are indicated above the sequences. The numbers represent the positions of the amino acid residues starting from the N terminus of the protein in a pro form. The accession numbers are YP184102 for Prepro-Tk-SP, NP579399 for Prepro-Pf-SP (DSM3638), P08594 for prepro-aqualysin-I (aqual), AAA22742 for prepro-subtilisin E (BssE), BAB60701 for Prepro-Tk-subtilisin (Tk-sub) and YP182489 for Tk-0076.

2.3.2. Overproduction and purification

In order to identify an active mature form of Tk-SP, the Pro-Tk-SP gene was expressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells transformed with plasmid pET25b-pro-Tk-sp. Comparison of the proteins extracted from the cells with IPTG induction (Fig. 2.2a, lane 2) with those from the cells without IPTG induction (Fig. 2.2a, lane 1) indicates that the production level of Pro-Tk-SP is too low to be detected as a band on SDS-PAGE, followed by CBB staining. However, activity staining of the gel (gel assay) indicates that Pro-Tk-SP accumulates in the cells in three forms with molecular masses of 65, 55, and 44 kDa, mainly in a form with molecular mass of 55 kDa (Fig. 2.2b, lane 2). The 65 kDa protein may represent Pro-Tk-SP, because this size is comparable to that calculated from the amino acid sequence (68633). The 55 and 44 kDa proteins may represent the Pro-Tk-SP derivatives with N- and/or C-terminal truncations.

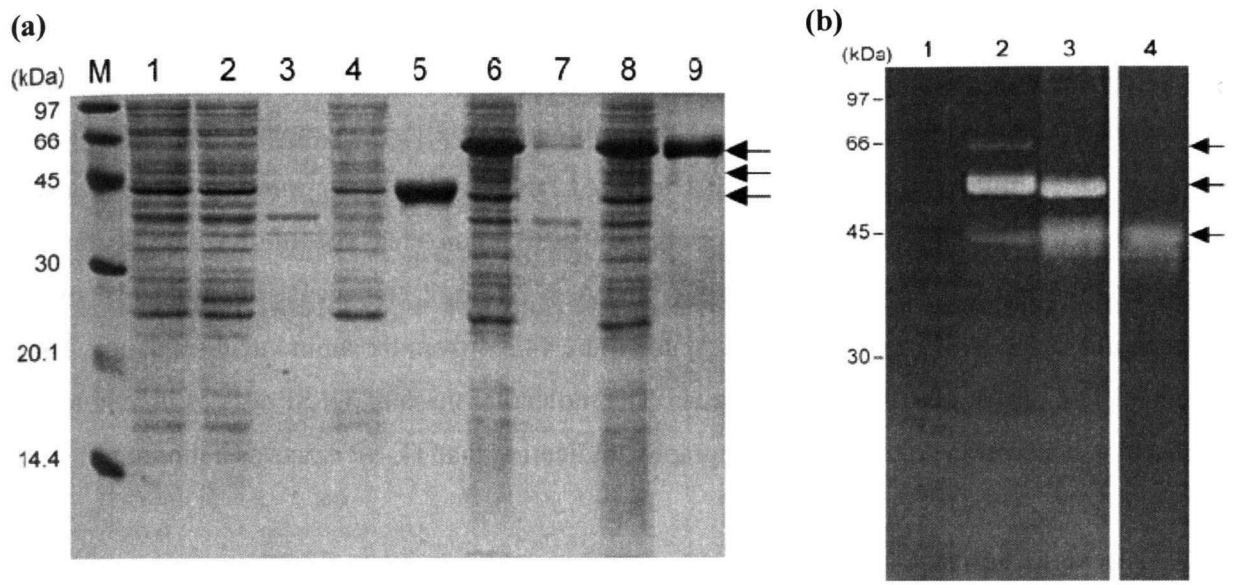

Figure 2.2. SDS-PAGE analyses for overproduction and purification of proteins. (**a**) *E. coli* BL21-CodonPlus(DE3)-RIL transformants with pET25b-pro-Tk-sp (lanes 1-4) or pET25b-pro-Tk-S359A (lanes 6-8) were subjected to electrophoresis on a 15% polyacrylamide gel in the presence of SDS. After electrophoresis, the gel was stained with CBB. Lane M, low-molecular-weight marker kit (GE Healthcare); lane 1, whole cell extract without IPTG induction; lanes 2 and 6, whole cell extract with IPTG induction; lanes 3 and 7, insoluble fractions after sonication lysis of the cells; lanes 4 and 8, soluble fractions after sonication lysis of the cells; lane 5, purified protein of Tk-SP (4 μg); lane 9, purified protein of Pro-Tk-S359A (4 μg). (**b**) Activity staining of gel. The same samples loaded onto lanes 1, 2, and 4 of the gel in A were loaded onto lanes 1, 2, and 3, respectively, of a 12% polyacrylamide gel containing 0.1% gelatin and SDS. Lane 4, purified protein of Tk-SP (0.1 μg). After electrophoresis, the gel was washed with 2.5% (v/y) Triton X-100 at room temperature for 1 h, incubated in 50mM Tris-HCl (pH 9.0) at 80°C for 2 h and stained with CBB. Protease bands were visualized as clear zones due to the hydrolysis of gelatin. For both A and B, the arrows indicate the positions of the 65, 55, and 44 kDa proteins from the top to the bottom. Numbers along the gel represent the molecular masses of individual standard proteins.

Tk-SP was purified from the cells, in which the Pro-Tk-SP gene was overexpressed. When the solution fraction obtained after sonication lysis was analysed by SDS-PAGE, none of the three forms mentioned above was detected as a distinctive band by CBB staining (Fig. 2.2a, lane 4). However, the 55 kDa and 44 kDa proteins were detected as two major bands by gel assay (Fig. 2.2b, lane 3), indicating that these forms accumulate in the cells in a soluble form. Finally, the 44 kDa protein was purified to give a single band on SDS-PAGE by both CBB staining (Fig. 2.2a, lane 5) and gel assay (Fig. 2.2b, lane 4). These results indicate that Pro-Tk-SP is processed into the 44 kDa protein through the 55 kDa protein during purification procedures.

The amount of the 44 kDa protein purified from 250 ml culture was ~0.2 mg. The N-terminal amino acid sequence of this protein was determined to be Val-Glu-Thr-Glu, indicating that the N-terminal amino acid residue of this protein is Val^{114}. The molecular mass of this protein was determined to be 44187 ± 202 by MALDI-TOF mass spectrometer. These results suggest that this protein is processed from Pro-Tk-SP upon cleavage of the peptide bonds between Ala^{113} and Val^{114} and between Val^{539} and Asp^{540} (Fig. 2.1), because the molecular mass of the peptide from Val^{114} to Val^{539} is calculated to be 44204 from the amino acid sequence. This protein will be designated as Tk-SP hereafter. The molecular mass of Tk-SP was estimated to be 42 kDa by gel filtration column chromatography, indicating that Tk-SP exists as a monomer.

2.3.3. Enzymatic activity

To analyze the pH dependence of Tk-SP, its activity was determined at various pH ranging from 4.5 to 11.5 and 20°C using Suc-AAPF-*p*NA as a substrate. Tk-SP did not clearly show an optimum pH for activity but exhibited high activity at a wide pH range between 7.0 and 11.5 (Fig. 2.3). Tk-subtilisin also exhibited high activity at a wide pH range between 8.0 and 11.5 (Fig. 2.3).

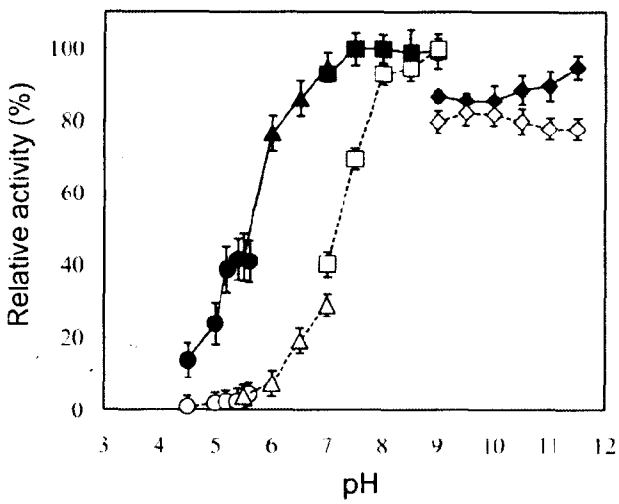

Figure 2.3. pH dependence of enzymatic activity. The enzymatic activities of Tk-SP (solid symbols) and Tk-subtilisin (open symbols) relative to those of Tk-SP at pH 7.5 and Tk-subtilisin at pH 9, respectively, are plotted as a function of pH. 100% enzymatic activities of Tk-SP and Tk-subtilisin are 2.3 and 38 U/mg, respectively. The buffers (50 mM) are sodium acetate (circle), 2-(*N*-morpholino)ethanesulfonic acid (MES) (triangle), Tris-HCl (square) and *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS) (diamond). The enzymatic activity was determined at 20°C by using Suc-AAPF-*p*NA as a substrate. Each experiment was carried out three times and the average values are shown together with the error bars.

To analyze the temperature dependence of Tk-SP, its activity was determined at various temperatures ranging from 20 to 100°C using azocasein as a substrate. For comparative purpose, the temperature dependence of Tk-subtilisin was also analyzed. Tk-SP and Tk-subtilisin exhibited the highest activities at 100°C and 90°C, respectively, although the difference between the activities at 90 and 100°C is not so significant for both enzymes (Fig. 2.4). The maximum specific activity of Tk-SP (510 ± 50 U/mg) was 5-fold lower than that of Tk-subtilisin (3100 ± 240 U/mg).

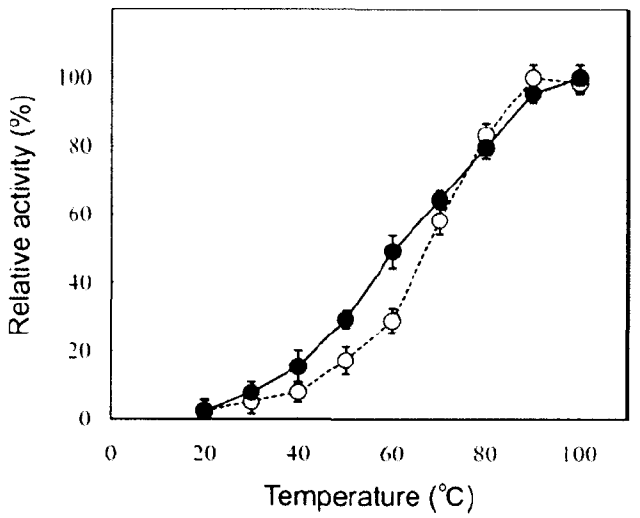

Figure 2.4. Temperature dependence of enzymatic activity. The enzymatic activities of Tk-SP (solid circle) and Tk-subtilisin (open circle) relative to those of Tk-SP at 100°C and Tk-subtilisin at 90°C, respectively, are plotted as a function of temperature. 100% enzymatic activities of Tk-SP and Tk-subtilisin are 510 and 3100 U/mg, respectively. The enzymatic activity was determined by using azocasin as a substrate. Each experiment was carried out three times and the average values are shown together with the error bars.

The Ca^{2+} dependence of Tk-SP was analyzed by measuring the activity at 80°C using azocasein as a substrate in the presence or absence of 10 mM $CaCl_2$, or in the presence or absence of 10 mM EDTA. The Tk-SP activity was not seriously changed regardless of whether 10 mM $CaCl_2$ or 10 mM EDTA was added. In contrast, Tk-subtilisin exhibited little activity at 80°C in the presence of 10 mM EDTA.

The kinetic parameters of Tk-SP and Tk-subtilisin were determined using Suc-AAPF-*p*NA as a substrate at 20 and 80°C (Table 2.1). It is noted that the K_m value of 7.9 mM of Tk-subtilisin at 80°C represents the apparent value because the solubility of the substrate at pH 8.0 is

too low to increase its concentration beyond 5 mM. Both of the K_m and k_{cat} values of Tk-SP were considerably lower than those of Tk-subtilisin at either temperature. These results indicate that Tk-SP is characterized by the high substrate binding affinity and low turnover number when compared with Tk-subtilisin. Nevertheless, the k_{cat}/K_m value of Tk-SP is rather similar to that of Tk-subtilisin at either 20°C or 80°C.

Table 2.1 Kinetic parameters of Tk-SP and Tk-subtilisin[a]

Enzyme	Temperature (°C)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m($mM^{-1}s^{-1}$)
Tk-SP	20	0.11 ± 0.015	1.6 ± 0.15	15 ± 1.0
	80	0.42 ± 0.02	25 ± 1.3	62 ± 2.4
Tk-subtilisin	20	2.4 ± 0.12	14 ± 0.67	6.0 ± 0.28
	80	7.9 ± 0.40	440 ± 22	56 ± 2.1

[a]The enzymatic activity was determined at the temperatures indicated in 50 mM Tris-HCl (pH 7.5) for Tk-SP or 50 mM Tris-HCl (pH 8.0) for Tk-subtilisin using Suc-AAPF-*p*NA as a substrate. The substrate concentration was varied from 0.01 to 2 mM for Tk-SP and from 0.5 to 5 mM for Tk-subtilisin. Each experiment was carried out three times and the average values are shown together with the errors.

The cleavage site specificity of Tk-SP was analyzed using oxidized insulin chain B as a substrate and compared with that of Tk-subtilisin (Fig. 2.5). Oxidized insulin chain B is cleaved by Tk-SP at multiple sites, preferably at the C-termini of the hydrophobic residues, such as Tyr, Phe, Leu, Val, and Ala. This peptide is also cleaved by Tk-subtilisin at the C-termini of the hydrophobic residues, but at more specific sites.

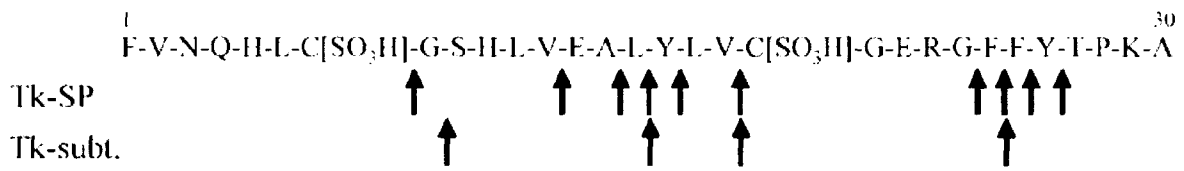

Figure 2.5. Cleavage-site specificity. Cleavage sites of oxidized insulin chain B by Tk-SP and Tk-subtilisin (Tk-subt) are indicated by arrows.

2.3.4. Stability

To analyze the stability of Tk-SP against irreversible heat inactivation, Tk-SP (0.05 mg/ml) was incubated in 50 mM Tris-HCl (pH 7.5) at 80°C, 90°C, and 100°C. With appropriate intervals, an aliquot of the solution was withdrawn and analyzed for residual activity at 80°C using azocasein as a substrate. As shown in Fig. 2.6, Tk-SP was stable at 80°C for at least 3 h. It lost activity at 90°C and 100°C with half-lives of >3 h and 100 min, respectively. For comparative purposes, the stability of Tk-subtilisin against heat inactivation was analyzed in 50 mM Tris-HCl (pH 8.0) at 100°C. Tk-subtilisin lost its activity at 100°C with a half-life of 50 min. Thus, Tk-SP is a highly thermostable enzyme and is slightly more stable than Tk-subtilisin.

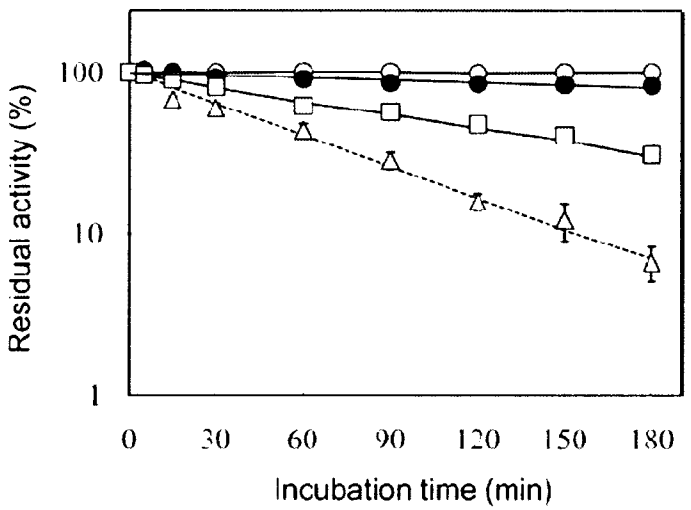

Figure 2.6. Thermal stability against heat inactivation. Semilog plots of the residual activity versus incubation time are shown. Tk-SP was incubated at 80°C (open circle), 90°C (solid circle) or 100°C (open square) in 50 mM Tris-HCl (pH 7.5). For comparative purpose, Tk-subtilisin was incubated at 100°C (open triangle) in 50 mM Tris-HCl (pH 8.0). An aliquot of the solution was withdrawn with appropriate intervals and the residual activity was determined at 80°C by using azocasein as a substrate. The lines were obtained by linear regression of the data. 100% enzymatic activities of Tk-SP and Tk-subtilisin are 35 and 420 U/mg, respectively. Each experiment was carried out three times and the average values are shown together with the error bars.

To analyze the effect of pH on the stability of Tk-SP, Tk-SP (0.05 mg/ml) was incubated at various pH ranging from 1 to 13 for 24 h at 30°C, and analyzed for residual activity at 20°C and pH 7.5 using Suc-AAPF-*p*NA as a substrate. As shown in Fig. 2.7, Tk-SP is fully stable between pH 7 and 11 whereas it is not fully stable at ≤pH 6 and ≥pH 12. It loses >85% of its activity at ≤pH 3 and pH 13. In contrast, Tk-subtilisin is fully stable between pH 2 and 12, indicating that Tk-SP is less stable than Tk-subtilisin at acidic pH, especially at pH 2 and 3.

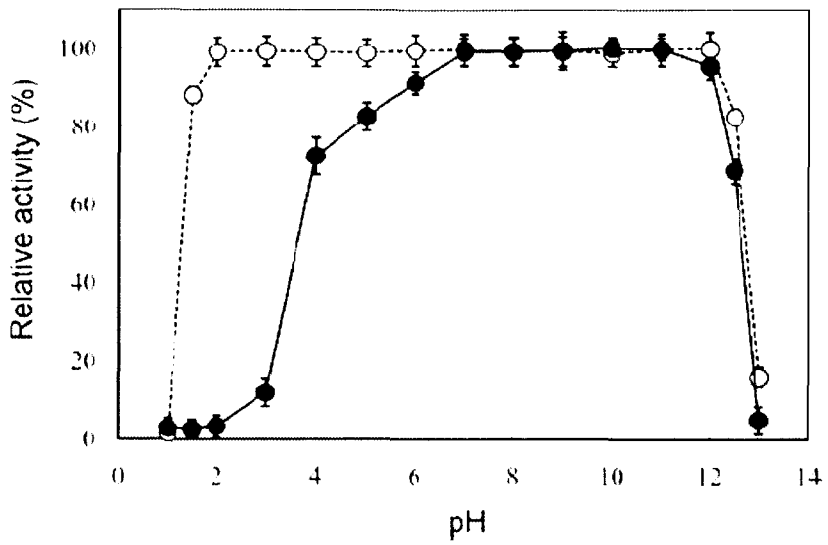

Figure 2.7. Effect of pH on protein stability. Proteins were incubated at various pH at 30°C for 24 h and analyzed for residual activities at 20°C by using Suc-AAPF-*p*NA as a substrate. The enzymatic activities of Tk-SP (solid circle) and Tk-subtilisin (open circle) relative to those of Tk-SP at pH 7 and Tk-subtilisin at pH 10, respectively, are plotted as a function of pH. 100% enzymatic activities of Tk-SP and Tk-subtilisin are 2.3 and 38 U/mg, respectively. The buffers (50 mM) are KCl-HCl for pH 1.0 and 1.5, Glycine-HCl for pH 2 and 3, sodium acetate for pH 4 and 5, MES for pH 6, Tris-HCl for pH 7-9, Glycine-NaOH for pH 10 and 11, and KCl-NaOH for pH 12-13. Each experiment was carried out three times and the average values are shown together with the error bars.

To analyze the effect of the EDTA treatment on the stability of Tk-SP, Tk-SP (0.05 mg/ml) was incubated at 55°C, 80°C and 90°C in 20 mM Tris-HCl (pH 8.0) containing 10 mM EDTA. With appropriate intervals, an aliquot was withdrawn and the residual activity was determined at 20°C using Suc-AAPF-*p*NA as a substrate. For comparative purposes, Tk-subtilisin was also incubated at 55°C in 20 mM Tris-HCl (pH 8.0) containing 10 mM EDTA. As shown in Fig. 2.8, Tk-subtilisin rapidly lost activity at 55°C, whereas Tk-SP was stable at this temperature for at least 1 h. However, Tk-SP lost activity at 80°C and 90°C with half-lives of ~3 h and 10 min, respectively, indicating that Tk-SP with the EDTA treatment is less stable than that without this treatment. These results suggest that Tk-SP requires Ca^{2+} ion(s) for maximal stability and these Ca^{2+} ion(s) bind to the protein too tightly to be removed in the absence of EDTA.

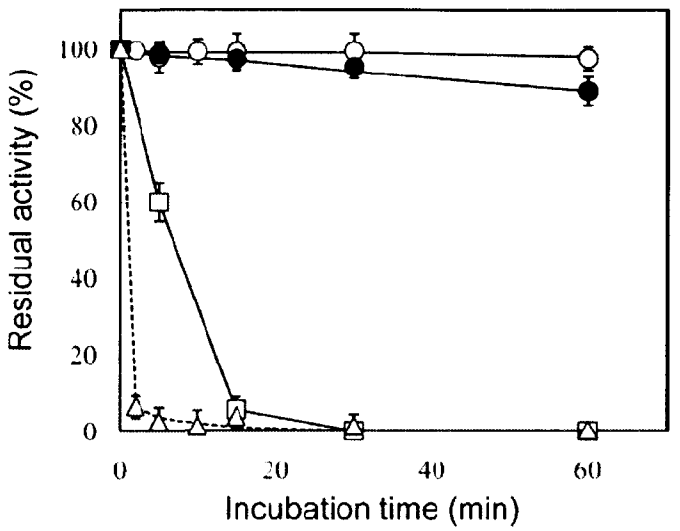

Figure 2.8. Effect of EDTA on protein stability. Tk-SP was incubated at 55°C (open circle), 80°C (solid circle) or 90°C (open square) in 50 mM Tris-HCl (pH 7.5) containing 10 mM EDTA. For comparative purpose, Tk-subtilisin was incubated at 55°C (open triangle) in 50 mM Tris-HCl (pH 8.0) containing 10 mM EDTA. An aliquot of the solution was withdrawn with appropriate intervals and the residual activity was determined at 20°C by using Suc-AAPF-*p*NA as a substrate. 100% enzymatic activities of Tk-SP and Tk-subtilisin are 2.3 and 38 U/mg, respectively. Each experiment was carried out three times and the average values are shown together with the error bars.

To analyze the effects of detergents and denaturants on the stability of Tk-SP, Tk-SP (0.05 mg/ml) was incubated in 20 mM Tris-HCl (pH 8.0) containing 10% (v/v) Tween-20, 10% (v/v) Triton X-100, 2 or 5% (w/v) SDS, 4, 6, or 8 M urea, or 2 or 4 M GdnHCl for 1 h at 55°C and analyzed for residual activity at 20°C using Suc-AAPF-*p*NA as a substrate. The results are summarized in Table 2.2. Tk-SP retained at least 87% activity upon incubation with anyone of 10% Tween-20, 10% Triton X-100, 5% SDS, and 8 M urea, indicating that the protein is stable at 55°C even in the presence of these detergents and denaturants. However, Tk-SP is unstable in the presence of ≥2 M GdnHCl, and lost 35% and 99% of its activity upon incubation with 2 and 4 M GdnHCl, respectively, for 1 h at 55°C. In contrast, Tk-subtilisin retained all of its activity at 55°C in the presence of anyone of 10% Tween-20, 10% Triton X-100, 5% SDS, 8 M urea, and 6 M GdnHCl (Table 2.2).

Table 2.2 Effect of the treatment with a detergent or denaturant on the stability of Tk-SP and Tk-subtilisin[a]

Chemical substance	Concentration	Residual activity (%)	
		Tk-SP	Tk-subtilisin
None		100	100
Tween-20	10% (v/v)	104 ± 5.9	102 ± 3.7
Triton X-100	10% (v/v)	105 ± 2.5	103 ± 4.7
SDS	2% (w/v)	97 ± 1.8	100 ± 3.2
	5% (w/v)	91 ± 5.8	101 ± 1.9
Urea	4 M	93 ± 4.7	102 ± 1.5
	6 M	90 ± 4.9	100 ± 1.2
	8 M	87 ± 4.5	102 ± 3.2
GdnHCl	2 M	65 ± 1.2	104 ± 4.4
	4 M	<1	102 ± 1.8

[a] Tk-SP and Tk-subtilisin were incubated at 55°C for 60 min in 20 mM Tris-HCl (pH 7.5) and Tris-HCl (pH 8.0), respectively, containing a detergent or denaturant at the concentration indicated and their residual activities were determined at 20°C using Suc-AAPF-*p*NA as a substrate. 100% enzymatic activities of Tk-SP and Tk-subtilisin are 2.3 and 38 U/mg, respectively. Each experiment was carried out three times and the average values are shown together with the errors.

2.3.5. Overproduction and purification of Pro-Tk-S359A

To examine whether the N-terminal (Ala^{1}-Ala^{113}) and C-terminal (Asp^{540}-Gly^{640}) propeptides of Pro-Tk-SP are autoprocessed by its mature domain or processed by *E. coli* proteases, the active-site mutant Pro-Tk-S359A was constructed. Upon overproduction, Pro-Tk-S359A accumulated in the cells in a soluble form (Fig. 2.2a, lane 8) and purified to give a single band on SDS-PAGE (Fig. 2.2a, lane 9). The amount of the protein purified from 250 ml culture was 0.8 mg. The N-terminal amino acid sequence of this protein was determined to be Ala-Pro-Gln, indicating that the N-terminal methionine residue was post-translationally removed. The molecular mass of Pro-Tk-S359A was estimated to be 65 kDa from SDS-PAGE. These results indicate that Tk-SP is not processed from Pro-Tk-SP by *E. coli* proteases but is autoprocessed.

2.3.6. CD spectra of Pro-Tk-S359A

The far-UV CD spectrum of Pro-Tk-S359A measured at 25°C and pH 7.5, which represents that of the protein in a native form, gives a broad trough with a single minimum value of -11400 at 218 nm (Fig. 2.9). This spectrum is greatly changed in the presence of 6 M GdnHCl, such that the depth of the trough is greatly reduced, indicating that Pro-Tk-SP is denatured in this condition (Fig. 2.9). However, the spectrum of Pro-Tk-S359A, which was first denatured by GdnHCl and then renatured by removing GdnHCl by dialysis in the absence of Ca^{2+}, is indistinguishable from that of the protein in a native form, indicating that Pro-Tk-S359A is refolded into a native structure even in the absence of Ca^{2+} (Fig. 2.9). To examine whether Pro-Tk-S359A is denatured at acidic pH, the far-UV CD spectrum of this protein was measured at pH 2.0 and 25°C. This spectrum is significantly different from that of the protein in a native form (Fig. 2.9). This result, as well as the result that Tk-SP almost fully loses activity upon incubation at pH 2.0 (Fig. 2.7), indicates that Pro-Tk-S359A is irreversible denatured at pH 2.0. In gel assay, Tk-SP is fully denatured prior to SDS-PAGE by the TCA treatment, followed by boiling for 5 min in the presence of SDS. Nevertheless, Tk-SP exhibits activity probably because it is not refolded from an acid-denatured state, but is at least partially refolded from a fully unfolded state. In fact, Tk-SP lost activity after TCA precipitation and boiling of the pellet in 1% SDS, but exhibited 6.3% of its activity upon subsequent extensive dialysis against 50 mM Tris-HCl (pH 9.0) containing 2.5% Triton X-100.

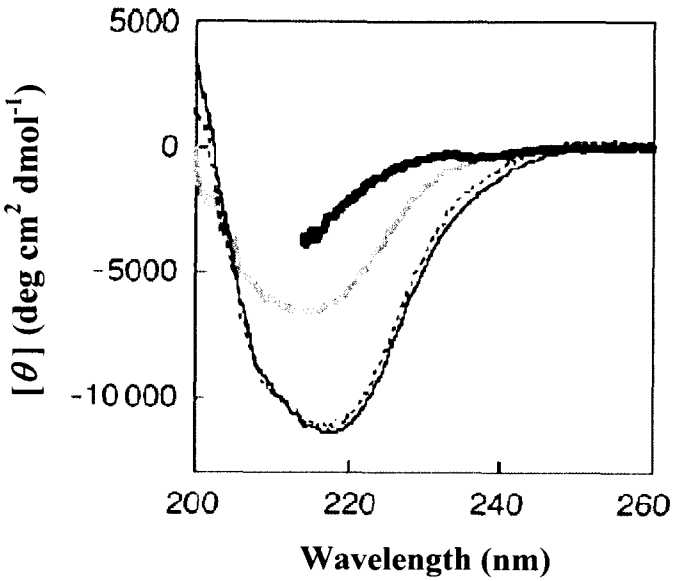

Figure 2.9. CD spectra. The far-UV CD spectra of Pro-Tk-S359A measured in 20 mM Tris-HCl (pH 7.5) (thin solid line), 20 mM Glycine-HCl (pH 2.0) (thick gray line) and 20 mM Tris-HCl (pH 7.5) containing 6 M GdnHCl (thick solid line) are shown. The far-UV CD spectrum of the refolded protein of Pro-Tk-S359A measured in 20 mM Tris-HCl (pH 7.5) (broken line) is also shown. For refolding, Pro-Tk-SP was first unfolded in the presence of 6 M GdnHCl and then refolded by removing GdnHCl in the absence of Ca^{2+}. These spectra were measured at 25°C. The protein concentration was 0.1 mg/ml. The mean residue ellipticity, $[\theta]$, with units of deg cm^2 $dmol^{-1}$, was calculated using an average amino acid molecular weight of 110.

2.3.7. Autoprocessing of Pro-Tk-SP

Pro-Tk-SP is characterized by the presence of a long C-terminal extension when compared with bacterial prosubtilisins and Pro-Tk-subtilisin (Fig. 2.1). This C-terminal extension is also present in Pro-Pf-SP and pro-aqualysin-I. The autoprocessing of N-terminal (Val^{1}-Trp^{113}) and C-terminal (Ser^{395}-Pro^{499}) propeptides to produce mature aqualysin-I (Ala^{114}-Gly^{394}) has been reported (Terada *et al.*, 1990). In this study, we showed that Tk-SP is also matured from Pro-Tk-SP upon autoprocessing of N- and C-terminal propeptides.

The N-terminal amino acid residue of Tk-SP was determined to be Val^{114}, suggesting that the peptide bond between Ala^{113} and Val^{114} of Pro-Tk-SP or its derivative is cleaved by its mature domain to release an N-teminal propeptide. Similar autoprocessing has been reported for Pro-Tk-subtilisin (Pulido *et al.*, 2006), pro-aqualysin-I (Terada *et al.*, 1990) and pro-subtilisin E (Jain *et al.*, 1998). The C-terminal residue of Tk-SP was not determined but estimated to be Val^{539} from its molecular mass and cleavage-site specificity. The molecular mass of 44187 with

an error of 202 allows a potential cleavage anywhere between Pro^{537} and Glu^{541}. The amino acid sequence between these residues is -Pro^{537}-Thr^{538}-Val^{539}-Asp^{540}-Glu^{541}-. The molecular mass of the protein calculated from the amino acid sequence is 44004 for Val^{114}-Pro^{537} and 44448 for Val^{114}-Glu^{541}. However, among these residues, Val^{539} is the only residue with bulky hydrophobic side chain. Because Tk-SP preferably cleaves the peptide bonds located at the C-termini of hydrophobic residues (Fig. 2.5), the peptide bond between Val^{539} and Asp^{540} of Pro-Tk-SP or its derivative might be autoprocessed by its mature domain to release a C-teminal propeptide.

Identification of the 55 kDa protein as an intermediate form of the autoprocessing reaction of Pro-Tk-SP indicates that the N- and C-terminal propeptides are autoprocessed in a stepwise manner. However, it remains to be determined whether the N- or C-terminal propeptide is autoprocessed first, because these propeptides are similar in size (12686 for Ala^{1}-Ala^{113} and 11778 for Asp^{540}-Gly^{640}).

2.3.8. Role of propeptides

Subtilisin-like serine proteases can be classified into three groups based on the difference in the location of propeptides. Members of the first group, which are represented by bacterial subtilisins, contain propeptides at the N-termini. These propeptides have been reported to function not only as an intramolecular chaperon but also as a potent inhibitor of their cognate mature domains (Zhu *et al.*, 1989; Eder and Fersht, 1995; Li *et al.*, 1995; Yabuta *et al.*, 2001). Members of the second group, which are represented by extracellular proteases from Gram-negative bacteria, such as IgA protease from *Neisseria gonorrhoeae* (Pohlner *et al.*, 1987) and serine protease SSP from *Serratia marcescens* (Yanagida *et al.*, 1986), contain propeptides at the C-termini. These propeptides have been reported to be required for translocation of the protease across the outer membrane. Members of the third group, which are represented by aqualysin-I, contain propeptides at both the N- and C-termini. It has been reported that the N-terminal propeptide of aqualysin-I functions similarly to the propeptides of bacterial subtilisins (Marie-Claire *et al.*, 2001), whereas its C-terminal propeptide directs the protein to the outer membrane and traps it there (Kim *et al.*, 1997).

Tk-SP is a member of the third group. However, the observation that Tk-SP without its N- and C-terminal propeptides can be refolded and exhibits activity in gel assay (Fig. 2.2b) indicates that these propeptides are not required for folding of Tk-SP. It remains to be

determined whether these propeptides inhibit Tk-SP activity, because these propeptides are autoprocessed and degraded by Tk-SP during purification procedures. The observation that the 65 and 55 kDa proteins exhibit activity in gel assay (Fig. 2.2b) does not necessarily indicate that these propeptide do not inhibit Tk-SP activity, because these propeptides may be autoprocessed and degraded during activity staining of gel. Further mutational and structural studies will be required to understand the role of the propeptides of Tk-SP.

2.3.9. Role of Ca^{2+}

Subtilisin-like serine proteases usually contain several Ca^{2+}-binding sites, which vary in number from two to seven for different proteases (Betzel *et al.*, 1988; Gros *et al.*, 1991; Bryan *et al.*, 1992; Smith *et al.*, 1999; Almog *et al.*, 2003; Barrette-Ng *et al.*, 2003; Tanaka *et al.*, 2007b). Binding of the Ca^{2+} ions to these sites has been reported to be required to greatly stabilize proteins (Voordouw *et al.*, 1976; Pantoliano *et al.*, 1989; Gros *et al.*, 1991; Bryan *et al.*, 1992; Smith *et al.*, 1999), to make the conformation of the active site functional (Bajorath *et al.*, 1989), or to induce folding of the entire molecule (Tanaka *et al.*, 2007a). The following observations suggest that Tk-SP contain Ca^{2+}-binding sites that are important for stability but not required for activity or folding: (1) Tk-SP exhibits enzymatic activity in gel assay even in the absence of Ca^{2+}; (2) EDTA treatment destabilizes Tk-SP without seriously affecting its enzymatic activity; (3) CD measurements indicate that Pro-Tk-S359A can be almost fully refolded even in the absence of Ca^{2+}. Of the two Ca^{2+}-binding sites of bacterial subtilisins, Ca-1 site is conserved in Tk-subtilisin (Tanaka *et al.*, 2007a). This site is formed by the side chains of Gln^{79}, Asp^{118} and Asn^{154}, and the main chain carbonyl oxygen atoms of Leu^{152}, Ile^{156} and Val^{158} in subtilisin E. Of these residues, only Asp^{118} is conserved as Asp^{156} in Tk-SP (Fig. 2.1), suggesting that this site is not conserved in Tk-SP. Other Ca^{2+}-binding sites of bacterial subtilisins and Tk-subtilisin are also not conserved in Tk-SP. Tk-SP may have unique Ca^{2+}-binding site(s).

2.3.10. Practical usefulness of Tk-SP

Subtilisin-like serine proteases have been widely used for industrial purposes, mainly for laundry detergents and for food processing, silk refining and leather processing (Schäfer *et al.*, 2007). To be well suited for these uses, enzymes should have high activity and stability at broad temperature and pH ranges and should also be compatible with various detergents and

denaturants (Kumar and Takagi, 1999; Maurer, 2004). Thermostable subtilisin-like serine proteases from thermophiles and hyperthermophiles have received increased attention because of their inherent stability at high pH and temperature values, and in the presence of various detergents and denaturing agents, which enable their use in processes that limit conventional enzymes (Choi *et al.*, 1999; Gödde *et al.*, 2005; Pulido *et al.*, 2006). In this study, Tk-SP was shown to be highly stable even at 100°C and pH 12, and in the presence of 5%SDS, 8 M urea and 10 mM EDTA. In these harsh conditions, protein-derived environmental pollutants and disease-causing protein aggregates, which are hardly degraded by conventional enzymes at mild conditions, become susceptible to proteolytic degradation. Therefore, it is highly expected that Tk-SP will find new application in the industrial practice.

2.4. Summary

The genome of the hyperthermophilic archaeon *T. kodakaraensis* contains three genes encoding subtilisin-like serine proteases, Tk-1689, Tk-0076, and Tk-subtilisin. Of them, the structure and function of Tk-subtilisin have been extensively studied. To examine whether Tk-1689 is matured to an active form and functions as a hyperthermostable protease as is Tk-subtilisin, the gene encoding the Tk-1689 derivative without a putative N-terminal signal sequence, termed Pro-Tk-SP, was overexpressed in *E. coli*. Pro-Tk-SP is composed of 640 amino acid residues and its molecular mass is 68.6 kDa. The recombinant protein was purified, however, as an active 44-kDa protease, termed Tk-SP, which lacks the N-terminal 113 and C-terminal 101 amino acid residues. This result suggests that Pro-Tk-SP consists of a N-terminal propeptide (Ala^{1}-Ala^{113}), a mature domain (Tk-SP, Val^{114}-Val^{539}), and a C-terminal propeptide (Asp^{540}-Gly^{640}). Like Tk-subtilisin, Tk-SP showed a broad substrate specificity and was highly thermostable. Its optimum temperature for activity was around 100°C and its half-life at 100°C was 100 min. It was fully resistant to treatment with 5% SDS, 8 M urea, or 10% Triton X-100. However, unlike Tk-subtilisin and bacterial subtilisins, Tk-SP requires neither Ca^{2+} nor propeptide for folding. As a result, Tk-SP was fully active even in the presence of 10 mM EDTA. Thus, Tk-SP has a great advantage over other proteases in high resistance to heat, denaturants, detergents, and chelating agents and therefore has great potential for application in biotechnology fields.

CHAPTER 3

Crystal structure of Tk-SP: requirement of a C-terminal β-jelly roll domain for hyperstability

3.1. Introduction

In Chapter 2, Tk-SP (Vla^{114}-Vla^{539}) has been biochemically characterized. It is matured from Pro-Tk-SP (Ala^{1}-Gly^{640}) upon autoprocessing and degradation of N-propeptide (Ala^{1}-Ala^{113}) and C-propeptide (Asp^{540}-Gly^{640}). Like Tk-subtilisin, Tk-SP is a highly thermostable enzyme with optimum temperature of 100°C and half-life of 100 min at 100°C. It requires Ca^{2+} for maximal stability. However, unlike Tk-subtilisin and bacterial subtilisins, Tk-SP requires neither Ca^{2+} nor propeptides for folding. Comparison of the amino acid sequence of Tk-SP with those of Tk-subtilisin and bacterial subtilisins indicates that none of the Ca^{2+}-binding sites are conserved and Tk-SP has a long C-terminal extension with 100-150 residues. Without this extension, Tk-SP shows the amino acid sequence identities of 41% to subtilisin E and 31% to Tk-subtilisin. However, it remains to be determined whether Tk-SP has unique Ca^{2+}-binding sites and its C-terminal extended region is folded into a unique structure.

In this study, we determined the crystal structure of the Pro-Tk-SP derivative lacking C-propeptide. We showed that the overall structure of this protein is similar to those of Pro-Tk-subtilisin (Tanaka *et al.*, 2007a) and bacterial subtilisin:propeptide complexes (Gallagher *et al.*, 1995; Jain *et al.*, 1998), except that it contains a β-jelly roll domain at the C-terminus, does not contain any Ca^{2+} ion in the subtilisin domain, and has a long N-terminal extension. The β-jelly roll domain contains two Ca^{2+} ions. We constructed the Tk-SP derivatives (Tk-S359AΔJ and Tk-S359CΔJ) lacking β-jelly roll domain and analyzed their biochemical properties. We showed that Tk-S359CΔJ is enzymatically active. We also showed that Tk-S359AΔJ is greatly destabilized as compared to Tk-S359A. Based on these results, we propose that the β-jelly roll domain of Tk-SP is required for its adaptation to high temperature environment.

3.2. Materials and Methods

3.2.1. Plasmid construction

The pET25b derivatives for overproduction of ProN-Tk-S359A, Tk-S359A, ProC-Tk-S359A, and Tk-S359AΔJ were constructed by polymerase chain reaction (PCR). The pET25b derivative for overproduction of Pro-Tk-S359A (Chapter 2, Section 2.2.2) was used as a template. The sequences of the PCR primers are

5'-GGCCTTTATCATATGGCCCCCCAGAAG-3' for primer 1

5'-GGCCTTATCATATGGTTGAGACCGAGG-3' for primer 2

5'-GGCCTGGATCCTCAATAGTTGTCGTAGTAG-3 for primer 3

5'-GCCTTGGATCCTCAAACGGTCGGCTCTG-3' for primer 4

5'-GGCCTTGGATCCTCACCCGTAGTAAAC-3' for primer 5

where the *Nde*I (primers 1 and 2) and *Bam*HI (primers 3-5) sites are underlined. Primers 1 and 2 are designed such that the ATG codon is attached to the 5'-termini of the genes encoding the proteins mentioned above. Primers 3-5 are designed such that the TGA codon is attached to the 3'-termini of the genes encoding the proteins mentioned above. Primers 1 and 4, primers 2 and 4, primers 2 and 5, and primers 2 and 3 were used to amplify the genes encoding ProN-Tk-S359A, Tk-S359A, ProC-Tk-S359A, and Tk-S359AΔJ, respectively. The resultant DNA fragments were digested with *Nde*I and *Bam*HI, and ligated into the *Nde*I-*Bam*HI sites of pET25b.

The pET25b derivative for overproduction of Pro-Tk-S359C was constructed by PCR using the QuikChange II site-directed mutagenesis kit (Stratagene). The pET25b derivative for overproduction of Pro-Tk-SP (Chapter 2, Section 2.2.2) was used as a template. The mutagenic primers were designed such that the codon for Ser^{359} (TCG) is changed to TGC for Cys. The pET25b derivatives for overproduction of Tk-S359C and Tk-S359CΔJ were constructed by PCR. The pET25b derivative for overproduction of Pro-Tk-S359C was used as a template. Primers 2 and 4 were used to amplify the gene encoding Tk-S359C and primers 2 and 3 were used to amplify the gene encoding Tk-S359CΔJ. The resultant DNA fragments were digested with *Nde*I and *Bam*HI and ligated into the *Nde*I-*Bam*HI sites of pET25b.

All DNA oligomers for PCR were synthesized by Hokkaido System Science. PCR was performed in 25 cycles using a thermal cycler (Gene Amp PCR System 2400; Applied Biosystems) and KOD DNA polymerase (Toyobo). The DNA sequences of the genes encoding

all proteins mentioned above were confirmed by ABI Prism 310 DNA sequencer (Applied Biosystems).

3.2.2. Overproduction and purification

Pro-Tk-S359A was overproduced and purified as described previously (Chapter 2, Section 2.2.2). Overproductions of ProN-Tk-S359A, Tk-S359A/C, ProC-Tk-S359A, Tk-S359AΔJ, Tk-S359C, and Tk-S359CΔJ in *E. coli* BL21-CodonPlus(DE3) (Stratagene) were carried out as described previously for Pro-Tk-SP (Chapter 2, Section 2.2.2). For purification of ProN-Tk-S359A, Tk-S359A, Tk-S359C, and Tk-S359CΔJ, the cells were collected by centrifugation, suspended in 20 mM Tris-HCl (pH 9.0), disrupted by sonication on ice, and centrifuged at 30000*g* for 30 min at 4°C. The supernatant was collected and applied to a HitrapQ HP column (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column by linearly increasing the NaCl concentration from 0 to 1.0 M in the same buffer. The fractions containing the protein were pooled, dialyzed against 20 mM Tris-HCl (pH 7.5), and applied to a Hitrap DEAE FF column (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column by linearly increasing the NaCl concentration from 0 to 1.0 M in the same buffer. The fractions containing the protein were pooled, concentrated using ultrafiltration system Centricon (Millipore), and loaded onto a Sephacryl S-200HR column (GE Healthcare) equilibrated with 20 mM Tris-HCl (pH 7.5) containing 10 mM $CaCl_2$ and 50 mM NaCl. The fractions containing the protein were collected and dialyzed against 20 mM Tris-HCl (pH 7.5). Tk-S359AΔJ was purified equally, except that one purification step was added prior to the last gel filtration chromatography. For this step, the fractions obtained after Hitrap DEAE FF column chromatography were dialyzed against 20 mM Tris-HCl (pH 7.5) containing 0.3 M $(NH_4)_2SO_4$ and loaded onto a Hitrap phenyl HP column (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column by linearly decreasing the $(NH_4)_2SO_4$ concentration from 0.3 to 0 M in the same buffer.

The protein concentration was determined from UV absorption at 280 nm using the absorbance of a 0.1% (1.0 mg/ml) solution at 280 nm of 1.80 for Pro-Tk-S359A, 1.66 for ProN-Tk-S359A, 1.84 for Tk-S359A and Tk-S359C, and 1.65 for Tk-S359AΔJ and Tk-S359CΔJ. These values were calculated using absorption coefficients of 1526 $M^{-1}cm^{-1}$ for Tyr and 5225

$M^{-1}cm^{-1}$ for Trp at 280 nm (Googwin and Morton, 1946). The purity of the protein was confirmed by SDS-PAGE (Laemmli, 1970), followed by staining with Coomassie brilliant blue (CBB).

3.2.3. Activity staining of gel (gel assay)

The hydrolysis of gelatin by *E. coli* BL21-CodonPlus(DE3)-RIL transformants with the pET25b derivatives for overproduction of Tk-S359C and Tk-S359CΔJ were analyzed as described in Chapter 2 (Section 2.2.4).

3.2.4. Enzymatic activity

The enzymatic activity of Tk-S359C and Tk-S359CΔJ was determined by using azocasein (Sigma) as a protein substrate as described in Chapter 2 (Section 2.2.5), except that the reaction mixture was performed in the presence of 10 mM $CaCl_2$.

3.2.5. Circular dichroism spectra

The far-UV (200-260 nm) and near-UV (250-320 nm) CD spectra of the protein were measured at 25°C on a J-725 automatic spectropolarimeter (Japan Spectroscopic). The proteins were dissolved in 20 mM Tris-HCl (pH 7.5) or the same buffer containing 10 mM $CaCl_2$. The protein concentration and optical path length were 0.15 mg/ml and 2 mm for far-UV CD spectra and 1.0 mg/ml and 1 cm for near-UV CD spectra, respectively. The mean residue ellipticity, θ, which has units of deg cm^2 $dmol^{-1}$, was calculated by using an average amino acid molecular weight of 110.

3.2.6. Thermal denaturation

Thermal denaturation curves of the proteins were obtained by plotting the change in CD values at 222 nm against increasing temperature. The protein was dissolved in 50 mM Tris-HCl (pH 7.5) containing 1 M GdnHCl or the same buffer containing 10 mM $CaCl_2$. The protein concentration and optical path length were 0.2 mg/ml and 2 mm, respectively. The linear rate for temperature increase was approximately 1.0°C/min. The thermal denaturation processes of Tk-S359AΔJ, Tk-S359AΔJ^{EDTA}, and Tk-S359A^{EDTA} were reversible, while that of Tk-S359A was irreversible under these conditions. The thermal denaturation curves were normalized, assuming a linear temperature dependence of the baselines of native and denatured states. The midpoints of

the transition of these thermal denaturation curves, T_m, were calculated from the resultant normalized curves on the basis of a least-squares analysis.

3.2.7. Protein crystallization

The protein was dialyzed against 5 mM Tris-HCl (pH 7.0), concentrated to 7 mg/ml using an ultrafiltration system Centricon (Millipore). The crystallization conditions were initially screened using crystallization kits from Hampton Research (Crystal Screens I and II) and Emerald Biostructures (Wizard I and II). The conditions were surveyed using the sitting-drop, vapour-diffusion method at 20°C. Drops were prepared by mixing 1 µl each of the protein solution and reservoir solution and were vapour-equilibrated against 100 µl of reservoir solution.

Crystals of ProN-Tk-S359A appeared after a few weeks using Crystal Screen II solution No. 22 [0.1 M 4-morpholineethanesulfonic acid (pH 6.5) and 12% w/v polyethylene glycol (PEG) 20000]. However, the crystals were clusters and were not suitable for X-ray diffraction analysis. The crystallization conditions were further optimized by alteration of pH, protein concentration, or crystallant (PEG). Diffraction-quality crystals grew after 2 weeks using 0.1 M 4-morpholineethanesulfonic acid (pH 6.5) and 5 % (w/v) PEG 1500 at 20°C. One of the single crystals separated from them was used for data collection. X-ray diffraction studies on the resultant crystals indicated that these belonged to the space group $P2_12_12_1$ (unit cell parameters a = 63.01 Å, b = 124.00 Å, c = 141.04 Å) and contained 2 protein molecules per asymmetric unit.

3.2.8. X-ray diffraction data collection and structure determination

X-ray diffraction data were collected at a wavelength of 1.0 Å with the beam line BL44XU station at SPring-8, Japan. Diffraction data were indexed, integrated, and scaled using the HKL2000 program suite (Otwinowski and Minor, 1997). The crystal structure was solved by the molecular replacement method using MOLREP (Vagin and Teplyakov, 1997) in the CCP4 program suite (collaborative computational project, 1994). The 2.1-Å structure of propeptide:subtilisin E complex [Protein Data Bank (PDB) code 1SCJ] was used as a starting model. Automated model extension was done using ARP/wARP (Langer *et al.*, 2008) to obtain the atomic coordinates of β-jelly roll domain. Refinement of the structure was carried out with the programs Coot (Emsley and Cowtan, 2004) and REFMAC (Murshudov *et al.*, 1997) in the CCP4 program suite. After molecular replacement, two Ca^{2+} ions coordinated by five or six

ligands were observed, refining with full occupancy in the electron density ($2F_o$-F_c). Progress in the structure refinement was evaluated at each stage by the free R-factor and by inspection of stereochemical parameters calculated by the program PROCHECK (Laskowski *et al.*, 1993). The Ramachandran plot produced by PROCHECK showed that all residues in the structure are in the favored and allowed regions. RMSD calculation was done using the secondary-structure matching function (Krissinel and Henrick, 2004) of the LSQKAB program from the CCP4 program suite. The statistics for data collection and refinement are presented in Table 3.1. The figures were prepared by PyMol (available on the World Wide Web at www.pymol.org).

3.3. Results and Discussion

3.3.1. Protein preparation and crystallization

Tk-SP has been purified from *E. coli* cells, in which the Pro-Tk-SP gene is overexpressed (Chapter 2, Section 2.3.2). In these cells, Pro-Tk-SP accumulates in three forms, intact, partially autoprocessed, and fully autoprocessed forms, because the maturation process of this protein is initiated immediately when it is synthesized. None of these proteins are detected as a band by SDS-PAGE, followed by CBB staining, probably because the production level of Pro-Tk-SP in the cells is low. In contrast, the active site mutant of Pro-Tk-SP, Pro-Tk-S359A, accumulates in the cells in a soluble form as the most abundant protein (Chapter 2, Section 2.3.2). Because a large amount of the protein is necessary for crystallographic studies and the protein of interest is not always crystallized, we constructed a series of the Pro-Tk-S359A derivatives lacking C-propeptide (ProN-Tk-S359A), N-propeptide (ProC-Tk-S359A), and N- and C-propeptides (Tk-S359A). The primary structures of these proteins are schematically shown in Fig. 3.1. Upon overproduction, these proteins, except for ProC-Tk-S359A, accumulated in the cells in a soluble form. They were purified to give a single band on SDS-PAGE (Fig. 3.2). The amount of the protein purified from 1 l culture was approximately 10 mg for these proteins. The reason why ProC-Tk-S359A is not overproduced in *E. coli* remains to be understood. It is unlikely that N-propeptide is required for folding of the mature domain and therefore ProC-Tk-S359A is not correctly folded and rapidly degraded by *E. coli* proteases, because Tk-S359A without N- and C-propeptides was overproduced in *E. coli* in a soluble form.

The molecular masses of Pro-Tk-S359A, ProN-Tk-S359A, and Tk-S359A estimated from gel filtration chromatography were 63, 53, and 42 kDa, respectively, suggesting that they exist as a monomer in solution. The far- and near-UV CD spectra of ProN-Tk-S359A and Tk-S359A are similar to those of Pro-Tk-S359A (Fig. 3.3), suggesting that truncations of N- and C-propeptides do not seriously affect the conformation of the mature domain of Pro-Tk-S359A. Pro-Tk-S359A, ProN-Tk-S359A and Tk-S359A were individually screened for crystallization conditions using the kits available from commercial sources. However, only ProN-Tk-S359A gave crystals suitable for X-ray crystallographic analyses.

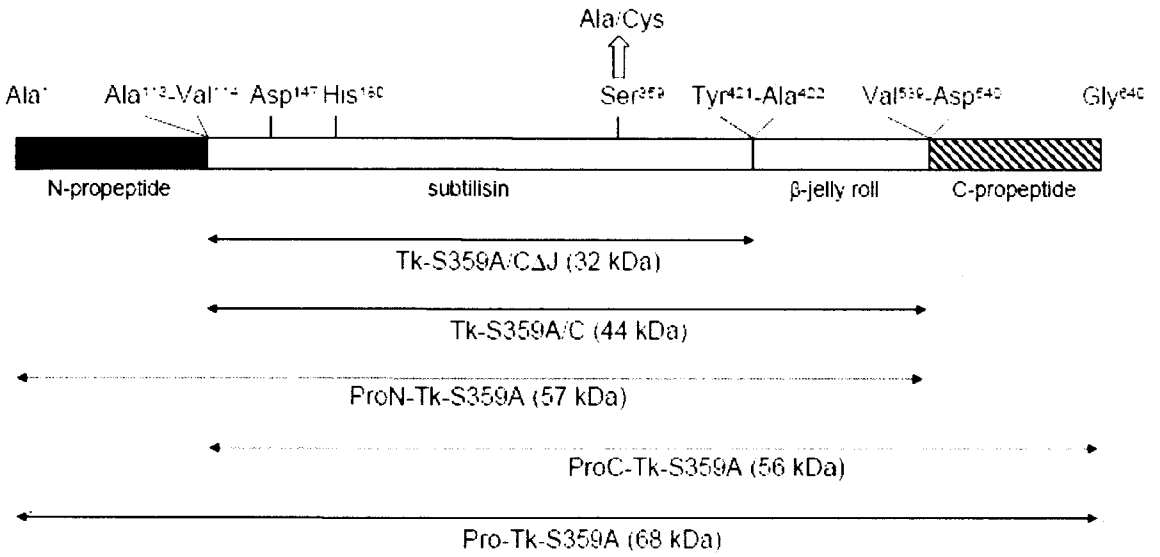

Figure 3.1. Schematic representation of the primary structure of Pro-Tk-SP. The dark box represents the N-propeptide domain, the open box represents the subtilisin domain, the gray box represents the β-jelly domain, and the hatched box represents the C-propeptide domain. The locations of the active-site residues, the mutated residue, and the regions of the active-site mutants analyzed in this study are shown. The N- and C-terminal residues of each domain are also shown. The accession number of this sequence is YP184102.

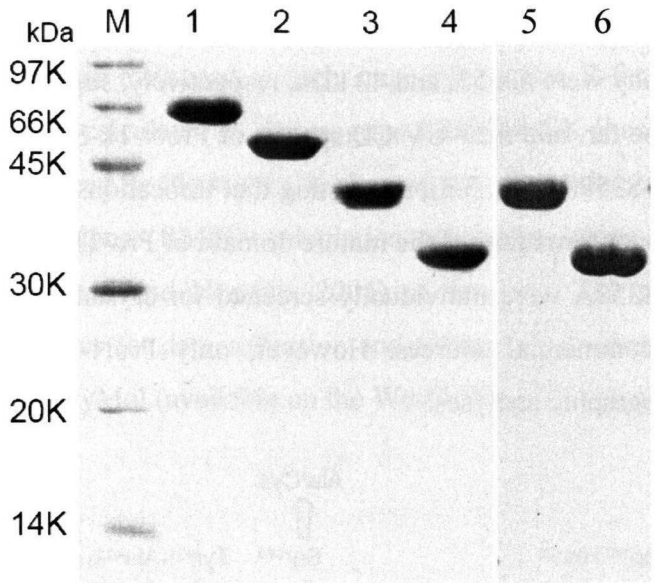

Figure 3.2. SDS-PAGE of the Pro-Tk-S359A and Tk-S359C derivatives. Samples were subjected to electrophoresis on a 15% polyacrylamide gel in the presence of SDS. After electrophoresis, the gel was stained with CBB. Lane M, low-molecular-weight marker kit (GE Healthcare); lane 1, Pro-Tk-S359A (4 μg); lane 2, ProN-Tk-S359A (4 μg); lane 3, Tk-S359A (4 μg); lane 4 Tk-S359AΔJ (4 μg); lane 5, Tk-S359C (6 μg); lane 6, Tk-S359CΔJ (6 μg). The molecular mass of each standard protein is indicated beside the gel.

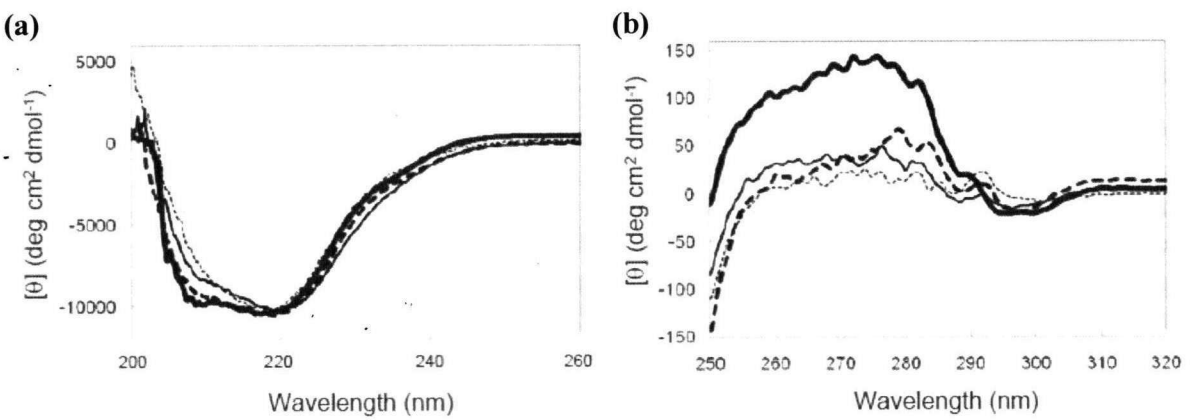

Figure 3.3. CD-spectra. The far-UV (a) and near-UV (b) CD spectra of Pro-Tk-S359A (thin broken line), ProN-Tk-S359A (thick broken line), Tk-S359A (thin solid line), and Tk-S359AΔJ (thick solid line) measured in 20 mM Tris-HCl (pH 7.5) are shown. These spectra were measured at 25°C as described under Material and Method (Section 3.2.5).

3.3.2. Overall three-dimensional structure of ProN-Tk-S359A

The crystal structure of ProN-Tk-S359A was determined at 2.0 Å resolution. The asymmetric unit of the crystal structure consists of two molecules (molecules A and B). Molecule A contains 505 of 539 residues, with Ala^{1}-Gln^{3}, Gly^{91}-Phe^{93}, Ala^{113}-Ala^{123}, and Ser^{523}-Val^{539} missing, and two Ca^{2+} ions. Molecule B contains 507 of 539 residues, with Ala^{1}-Gln^{3}, Thr^{90}-Asn^{95}, Glu^{117}-Ser^{122}, and Ser^{523}-Val^{539} missing, and two Ca^{2+} ions. The structures of these two molecules are virtually identical with each other with a root-mean-square deviation (RMSD) value of 0.7 Å for 502 C^{α} atoms from Lys^{4}-Pro^{522}, except for Thr^{90}-Asn^{95} and Ala^{113}-Ala^{123}. We used the structure of molecule B in this study. The overall structure of this molecule is shown in Fig. 3.4a. The ProN-Tk-S359A structure consists of the N-propeptide domain (Lys^{4}-Ala^{113}), the subtilisin-like α/β domain (subtilisin domain, Val^{114}-Tyr^{421}), and the β-jelly roll domain (Ala^{422}-Pro^{522}). Thus, the mature domain (Tk-SP) of Pro-Tk-SP consists of the subtilisin and β-jelly roll domains.

Table 3.1. Data collection and refinement statistics

Data collection	
Wavelength (Å)	1.0
Temperature (K)	100
Resolution (Å)	50.0–2.00
Highest-resolution shell (Å)	2.07–2.00
Total no. of measured reflections	814,987
Unique reflections	74,932
Data completeness (%)	99.0 (95.6)
R_{merge} (%)[a]	11.3 (36.5)
Average $I/\sigma(I)$	13.1 (2.6)
Refinement statistics	
R_{work} (%)	18.1
R_{free} (%)[b]	24.0
No. of molecules	2
No. of protein atoms	7580
No. of water molecules	611
No. of calcium ions	4
RMSD from ideal	
Bond lengths (Å)	0.012
Bond angles (°)	1.2
Average *B*-factor (Å^2)	
Protein atoms	21.7
Water atoms	30.6
Calcium ions	23.7
Ramachandran plot (%)	
Most favored regions	87.6
Additional allowed regions	12.4
Generously allowed regions	0

[a] $R_{merge} = \sum |I_{hkl} - \langle I_{hkl} \rangle| / \sum I_{hkl}$, where I_{hkl} is the intensity measurement for reflection with indices *hkl* and $\langle I_{hkl} \rangle$ is the mean intensity for multiply recorded reflections.

[b] R_{free} was calculated using 5% of the total reflections chosen randomly and omitted from refinement.

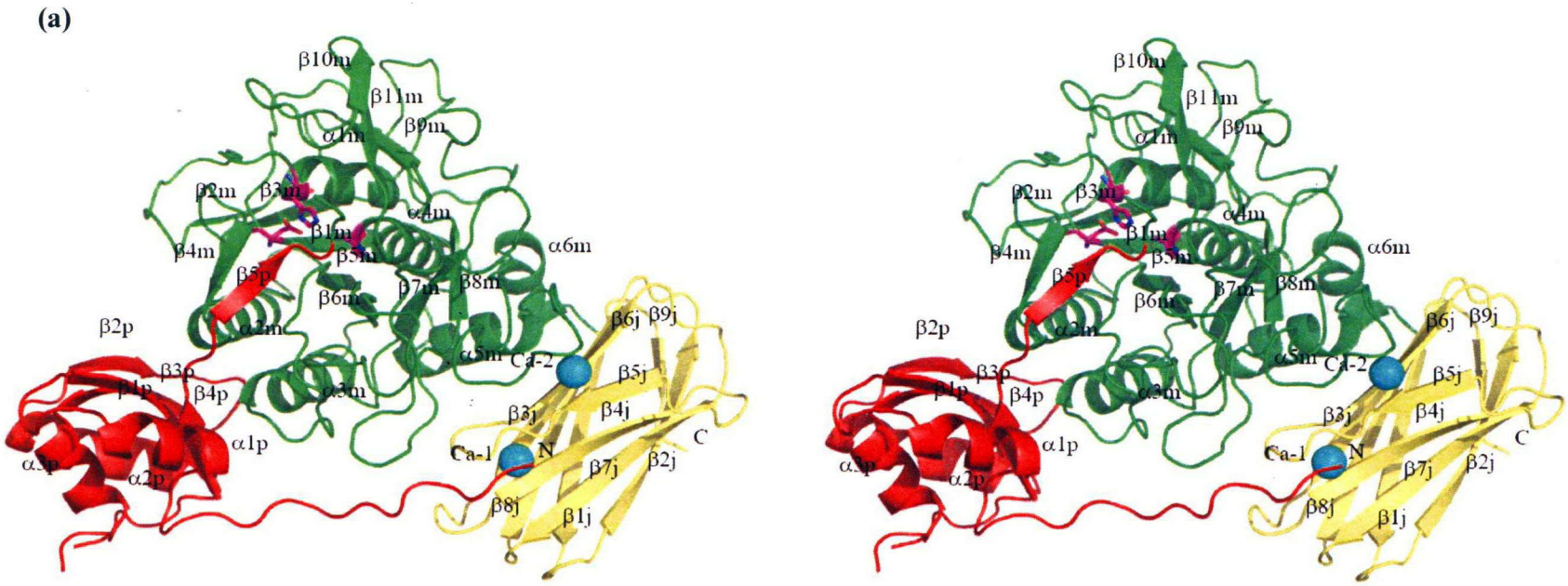

Figure 3.4. Three-dimensional structure of ProN-Tk-S359A. (a) Stereo view of the main-chain folding of ProN-Tk-S359A. The N-propeptide domain (Lys^{4}-Ala^{113}), subtilisin domain (Val^{114}-Tyr^{421}), and C-terminal β-jelly roll domain (Ala^{422}-Pro^{522}) are *colored red, green*, and *yellow*, respectively. Two active-site residues (Asp^{147} and His^{180}) and Ala^{359}, which is substituted for the active-site serine residue, are indicated by *magenta stick models*, in which the oxygen and nitrogen atoms are *colored red* and *blue*, respectively. Two Ca^{2+} ions (Ca-1 and Ca-2) are shown in *cyan spheres*. N and C, N and C-terminus, respectively. (b) Stereo view of electron density around the active site of ProN-Tk-S359A. The $2F_o$ - F_c map contoured at the 2.0 σ level is shown. The active-site residues are indicated by *yellow stick models*, and the oxygen and nitrogen atoms are *colored red* and *blue*, respectively. (c) Superposition of the ProN-Tk-S359A structure on the structure of the active-site mutant of Pro-Tk-subtilisin (Pro-S324A) (PDB code 2E1P). For the ProN-Tk-S359A structure, the active site residues, the oxygen and nitrogen atoms, and the Ca^{2+} ions are indicated as shown in Fig. 3.4a. For the Pro-S324A structure, the entire structure, including six Ca^{2+} ions (Ca-1 to Ca-6), is *colored gray*. Two active-site residues (Asp^{115} and His^{153}) and Ala^{324}, which is substituted for the active-site serine residue, are indicated by *stick models*. (d) Superposition of the ProN-Tk-S359A structure on the structure of Kp-43 (PDB code 1WMD). The ProN-Tk-S359A structure is indicated as shown in Fig. 3.4c. For the Kp-43 structure, the entire structure, including three Ca^{2+} ions (Ca-1 to Ca-3), is *colored gray*. Three active-site residues (Asp^{30}, His^{68}, and Ser^{255}) are indicated by *stick models*. (e) Superposition of the structure of the β-jelly roll domain of ProN-Tk-S359A on that of the C-terminal domain of Kp-43. These structures are indicated as shown in Fig. 3.4d. (f) The structure of the Ca-1 site. The Ca^{2+} ion and water molecule are shown in *cyan* and *red spheres*, respectively. The residues, which are located in the β-jelly roll domain and coordinated with the Ca^{2+} ion, are indicated by *yellow stick models*, in which the oxygen and nitrogen atoms are *colored red* and *blue*, respectively. (g) The structure of the Ca-2 site. The Ca^{2+} ion, water molecule, and the residues coordinated with the Ca^{2+} ion are indicated as shown in Fig. 3.4f. The residues located in the subtilisin domain are indicated by *green stick models*.

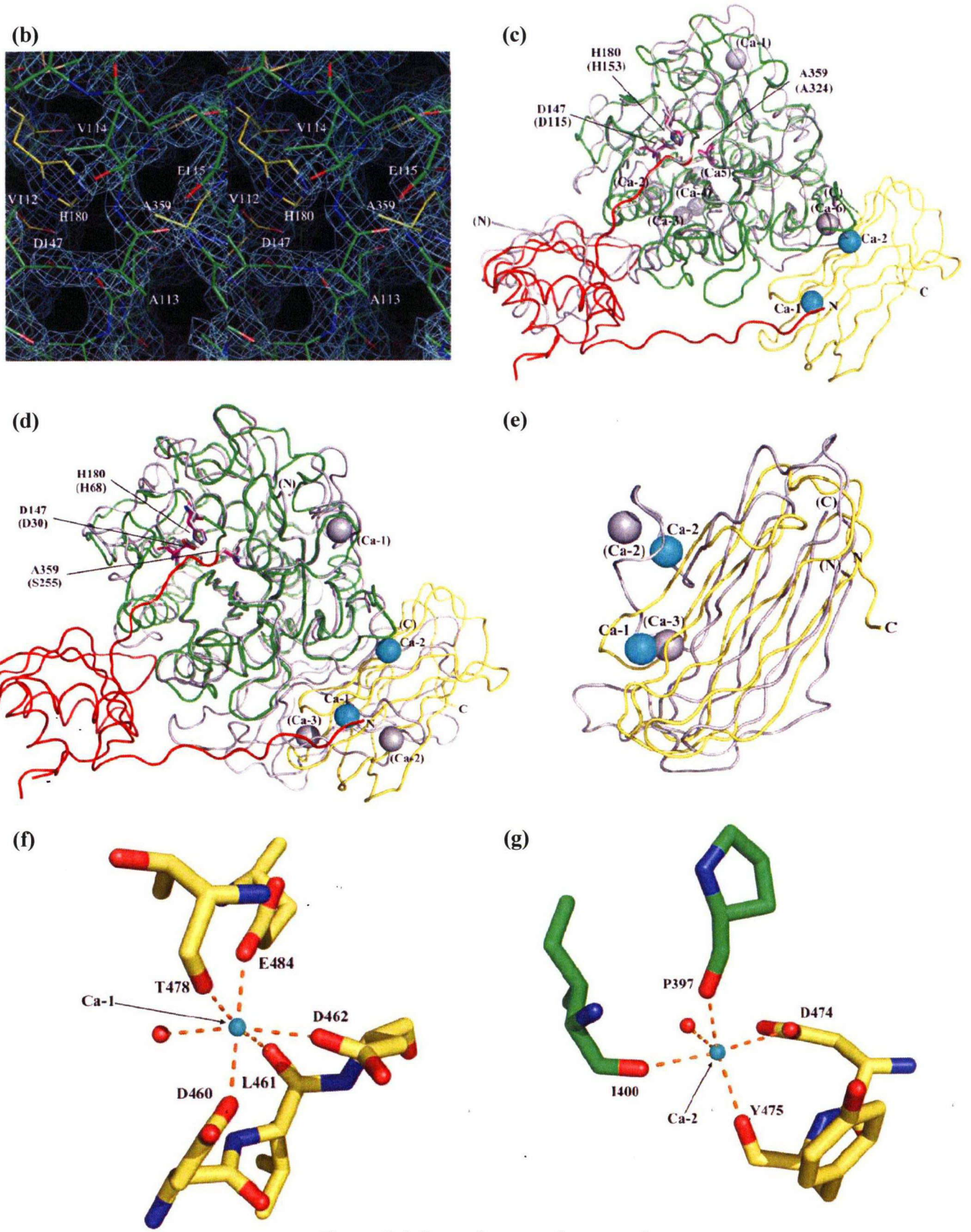

Figure 3.4 (legend on previous page)

The N-propeptide domain folds into a compact structure with a four-stranded antiparallel β-sheet (β1p-β4p) and three α-helices (α1p-α3p). This β-sheet packs tightly to the two nearly parallel α-helices (α2m and α3m) located at the surface of the subtilisin domain. Gln^{105} and Asp^{107} of the N-propeptide domain also bind to the N-termini of these two α-helices to form helix caps. In addition, the linker peptide between the N-propeptide and subtilisin domains runs through the active-site cleft in a substrate/product-like manner, with the Ala^{113}-Val^{114} peptide bond (cleavage site) docked close to Ala^{359}, which is substituted for the catalytic serine residue (Fig. 3.4b). Ala^{113} and Val^{114} occupy subsites S1 and S1′ of the subtilisin domain, respectively. Ala^{113} oxygen forms a hydrogen bond with Ala^{359} nitrogen. The β5p-strand of the N-propeptide domain interacts with the β4m-strand of the subtilisin domain to form an antiparallel β-sheet in the vicinity of the active site. The subtilisin domain consists of a seven-stranded parallel β-sheet (β1m-β3m and β5m-β8m), β4m-strand, an antiparallel β-sheet (β9m-β11m), and six α-helices (α1m-α6m). Of them, the seven-stranded parallel β-sheet and two α-helices (α1m and α4m) form a core structure of this domain. As in all members of the subtilase family, three active-site residues, Asp^{147}, His^{180}, and Ser^{359}, form a catalytic triad. In the present structure, Ser^{359} is replaced by Ala. Two hydrogen bonds are formed between His^{180} $N^{\delta 1}$ and Asp^{147} $O^{\delta 1}$ and between His^{180} $N^{\delta 1}$ and Asp^{147} $O^{\delta 2}$. These two hydrogen bonds are also observed in Pro-Tk-subtilisin (Tanaka *et al.*, 2007a). The β-jelly roll domain consists of 2 four-stranded antiparallel β-sheets (β1j- β8j- β3j- β6j and β2j- β7j- β4j- β5j) and β9j-strand. These two sheets form a core structure of this domain and are oppositely oriented to each other within a β-barrel like structure, in which the main chain is rotated once in every two β-strands (Fig. 3.5).

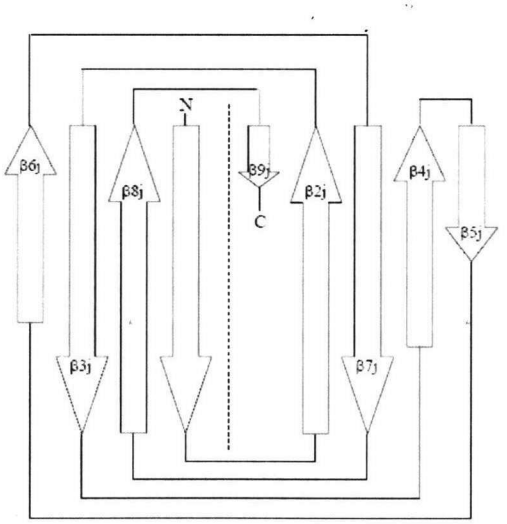

Figure 3.5. Topological analysis of the β-jelly roll domain of Tk-SP. Topology diagrams with β-strands are present as arrows. The dotted line represents a divider between two discontinuous β-sheets.

Comparison of the overall structure of ProN-Tk-S359A with that of Pro-Tk-subtilisin indicates that the main-chain fold of ProN-Tk-S359A is similar to that of Pro-Tk-subtilisin, except that ProN-Tk-S359A has a long extension and an extra α-helix (α1p) at the N-terminal region, lacks four α-helices (α1m-α4m for Pro-Tk-subtilisin) and Ca^{2+} ions in the subtilisin domain, and contains the C-terminal β-jelly roll domain (Fig. 3.4c). The RMSD values between these two structures are 1.5 Å for the N-propeptide domain for 63 C^{α} atoms from Glu^{34}-Ala^{113} and 1.5 Å for the subtilisin domain for 253 C^{α} atoms from Ala^{123}-Tyr^{421}.

The C-terminal β-jelly roll domain is not present in bacterial subtilisins as well. However, subtilisin-like alkaline serine protease, Kp-43, from *Bacillus* sp. (Nonaka *et al.*, 2004), kexin-like proteases (Holyoak *et al.*, 2003; Henrich *et al.*, 2003; Holyoak *et al.*, 2004; Kobayashi *et al.*, 2009), and tomato subtilase 3 (Ottmann *et al.*, 2009) contain a C-terminal jelly roll-like structure with eight β-strands. Of them, Kp-43 shows the highest amino acid sequence identity to Tk-SP for both the subtilisin domain (32%) and β-jelly roll domain (16%). Therefore, the overall structure of ProN-Tk-S359A is compared with that of Kp-43 in Fig. 3.4d. When the ProN-Tk-S359A structure is superimposed onto the Kp-43 structure, such that the structures of their subtilisin domains are superimposed to each other with the minimum RMSD value of 1.4 Å for 265 C^{α} atoms from Tyr^{136}-Ala^{416}, the structures of their β-jelly roll domains are considerably different with each other. However, without the subtilisin domains, the β-jelly roll domains of these two proteins can be superimposed to each other with the RMSD value of 1.4 Å for 90 C^{α} atoms from Lys^{423}-Ser^{515} (Fig. 3.4e). These results indicate that the β-jelly roll domain of Tk-SP is topologically similar to that of Kp-43, but its arrangement relative to the subtilisin domain is significantly different from that of Kp-43.

3.3.3. Ca^{2+}-binding sites

Two Ca^{2+}-binding sites are identified in the β-jelly roll domain of ProN-Tk-S359A, based on the $2F_o$ - F_c maps of ProN-Tk-S359A around these Ca^{2+}-binding sites contoured at 2.0 σ and 5.0 σ levels (data not shown) and coordination. The Ca-1 site is located in the β-jelly roll domain, whereas the Ca-2 site is located at the interface between the subtilisin and β-jelly roll domains. These Ca^{2+}-binding sites are relatively well conserved in a β-jelly roll-like domain of Kp-43 (Nonaka *et al.*, 2004) and kexin-like proteases (Holyoak *et al.*, 2003; Henrich *et al.*, 2003;

Kobayashi *et al.*, 2009) (Fig. 3.4e), although the site corresponding to the Ca-2 site of ProN-Tk-S359A is not located at the interface between the subtilisin and β-jelly roll domains of Kp-43 and kexin-like proteases (Fig. 3.4d). The Ca^{2+} ions are hexacoordinated with three side-chain oxygen atoms of Asp^{460} ($O^{\delta1}$), Asp^{462} ($O^{\delta1}$), and Glu^{484} ($O^{\delta2}$), two main chain carbonyl oxygen atoms of Leu^{461} and Thr^{478}, and one water molecule in the Ca-1 site (Fig. 3.4f), and pentacoordinated with one side chain oxygen atom of Asp^{474} ($O^{\delta1}$), three main-chain carbonyl oxygen atoms of Pro^{397}, Ile^{400}, and Tyr^{475}, and one water molecule in the Ca-2 site (Fig. 3.4g).

3.3.4. Stability of Tk-S359A derivative without β-jelly roll domain

To examine whether β-jelly roll domain is important for folding and stability of Tk-S359A, we constructed the Tk-S359A derivative lacking this domain (Tk-S359AΔJ) (Fig. 3.1). Tk-S359AΔJ was overproduced in *E. coli* cells in a soluble form, and was purified to give a single band on SDS-PAGE (Fig. 3.2, lane 4). The production level and purification yield of this protein were nearly identical to those of Tk-S359A. The far-UV and near-UV CD spectra of Tk-S359AΔJ are compared with those of Tk-S359A in Fig. 3.3. The far-UV CD spectrum of Tk-S359AΔJ is similar to that of Tk-S359A, suggesting that the removal of β-jelly roll domain does not seriously affect the protein conformation. Likewise, the near-UV CD spectrum of Tk-S359AΔJ is similar to that of Tk-S359A in shape, although these spectra greatly differ in CD values around 260-280 nm. The CD values of Tk-S359AΔJ are significantly higher than those of Tk-S359A around this region, probably because the aromatic residues located at the interface between the subtilisin and β-jelly roll domains are exposed to the solvent and thereby give the strong signals upon removal of the β-jelly roll domain. These results suggest that the β-jelly roll domain is not required for folding of Tk-SP.

Thermal denaturation of Tk-S359AΔJ and Tk-S359A was analyzed by monitoring the change in CD values at 222 nm as the temperature was increased. Thermal denaturation of Tk-S359AΔJ was reversible, while that of Tk-S359A was irreversible at any condition examined. The thermal denaturation curve of Tk-S359A was reproducible, unless the protein concentration, pH, and the rate of temperature increase were seriously changed. The thermal denaturation curves of Tk-S359AΔJ and Tk-S359A measured in 50 mM Tris-HCl (pH 7.5) containing 1 M GdnHCl and 10 mM $CaCl_2$ are shown in Fig. 3.6. The midpoints of the transition of these

thermal denaturation curves, T_m, are summarized in Table 3.2. The T_m value of Tk-S359AΔJ is lower than that of Tk-S359A by 29.4°C in T_m, indicating that the β-jelly roll domain contributes to the stabilization of Tk-S359A by 29.4°C in T_m.

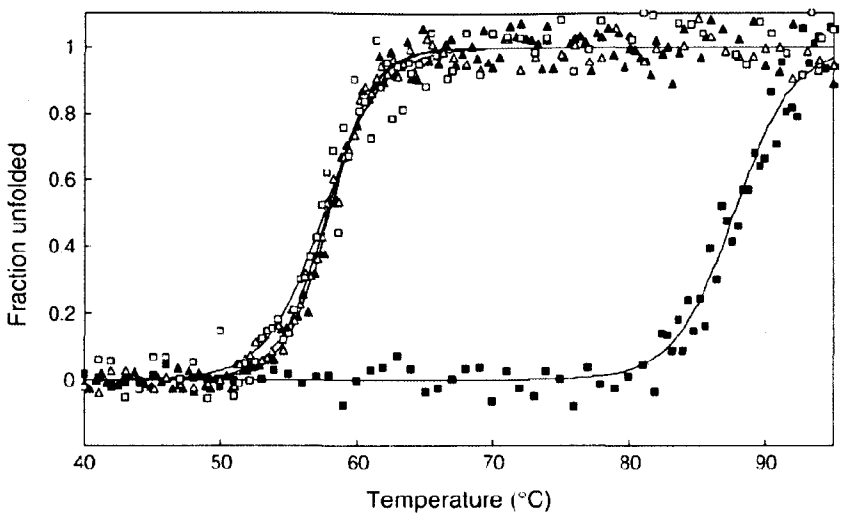

Figure 3.6. Thermal denaturation curves of Tk-S359A and Tk-S359AΔJ. The thermal denaturation curves of Tk-S359A (solid square) and Tk-S359AΔJ (solid triangle) measured in the presence of 10 mM $CaCl_2$ are shown in comparison with those of Tk-S359A^{EDTA} (open square) and Tk-S359AΔJ^{EDTA} (open triangle) measured in the absence of 10 mM $CaCl_2$. The curves were recorded by monitoring the change in CD values at 222 nm as described under Material and Methods (Section 3.2.5).

Table 3.2. Thermal stabilities of Tk-S359A and Tk-S359AΔJ

Protein[a]	[CaCl2] (mM)	T_m[b] (°C)	ΔT_m[c] (°C)
Tk-S359A	10	88.3 ± 0.86	-
Tk-S359A^{EDTA}	0	58.8 ± 0.93	-29.5
Tk-S359AΔJ	10	58.9 ± 1.3	-29.4
Tk-S359AΔJ^{EDTA}	0	58.4 ± 0.93	-29.9

[a] Tk-S359A^{EDTA} and Tk-S359AΔJ^{EDTA} were prepared by dissolving Tk-S359A and Tk-S359AΔJ in 50 mM Tris-HCl (pH 7.5) containing 10 mM EDTA and dialyzed against 50 mM Tris-HCl (pH 7.5).

[b] Thermal denaturation of the protein was analyzed in 50 mM Tris-HCl (pH 7.5) containing 1 M GdnHCl in the presence or absence of 10 mM $CaCl_2$ by monitoring the change in CD values at 222 nm. T_m represents the midpoint of the transition of the thermal denaturation curve. The experiment was carried out at least twice and the errors from the average values are indicated.

[c] $\Delta T_m = T_m$(examined) – 88.3 (the T_m value of Tk-S359A).

3.3.5. Protease activity of Tk-S359C derivative without β-jelly roll domain

To prove that the β-jelly roll domain is not required for folding of Tk-SP, it is necessary to show that the Tk-SP derivative lacking this domain, Tk-SPΔJ (residues 114-421), retains enzymatic activity. However, attempts to overproduce Tk-SP and Tk-SPΔJ in *E. coli* have so far been unsuccessful, probably due to their cytotoxicities to *E. coli* cells. These proteins are cytotoxic, possibly because they exhibit protease activity. Therefore, the Tk-SP derivative with the mutation of the active-site Ser to Cys (Tk-S359C) and the Tk-S359C derivative lacking the β-jelly roll domain (Tk-S359CΔJ) were constructed. It has been reported that the corresponding mutation greatly reduces the enzymatic activity of bacterial subtilisin (Li and Inouye, 1994) and Tk-subtilisin (Tanaka *et al.*, 2007a), but the resultant mutant proteins retain a weak enzymatic activity. Great reduction in enzymatic activities of Tk-SP and Tk-SPΔJ by the mutation should reduce their cytotoxicities and increase their accumulation levels in *E. coli* cells.

Overproduction of Tk-S359C and Tk-S359CΔJ in *E. coli* cells, sonication lysis of the cells and separation of the soluble and insoluble fractions were carried out as described for Pro-Tk-S359A (Chapter 2, Section 2.2.2). SDS-PAGE analyses indicated that Tk-S359C and Tk-S359CΔJ accumulate in *E. coli* cells in a soluble form upon overproduction (Fig. 3.7a, lanes 2 and 4). The molecular masses of Tk-S359C and Tk-S359CΔJ estimated from SDS-PAGE are 44 and 32 kDa, respectively, which are comparable to those calculated from their amino acid sequences (44223 for Tk-S359C and 31688 for Tk-S359CΔJ). The soluble fractions obtained after sonication lysis of the cells were subjected to gel assay at 40 and 80°C. Tk-S359C and Tk-S359CΔJ gave a band at the positions to which they migrated, even in the absence of Ca^{2+} (Fig. 3.7b), indicating that both proteins exhibit activity. The amounts of these proteins loaded on the gel were comparable with each other, as judged by the intensities of their bands visualized by CBB staining (Fig. 3.7a). The enzymatic activities of these proteins were also comparable with each other at 40°C, as judged by the intensities of their bands visualized by gel assay (Fig. 3.7b, lanes 2 and 3). In contrast, Tk-S359C exhibited much higher activity than Tk-S359CΔJ at 80°C. The enzymatic activity of Tk-S359CΔJ decreased at 80°C compared to that at 40°C, probably because Tk-S359CΔJ is not stable at 80°C. These results indicate that β-jelly roll domain is not required for folding of Tk-SP but is required for its hyperstability. It remains to be determined, however, whether this domain accelerates the folding rate of the protein. It is noted that several

additional weak bands are observed in gel assay (Fig. 3.7b). These bands may represent the self-degradation products of Tk-S359C.

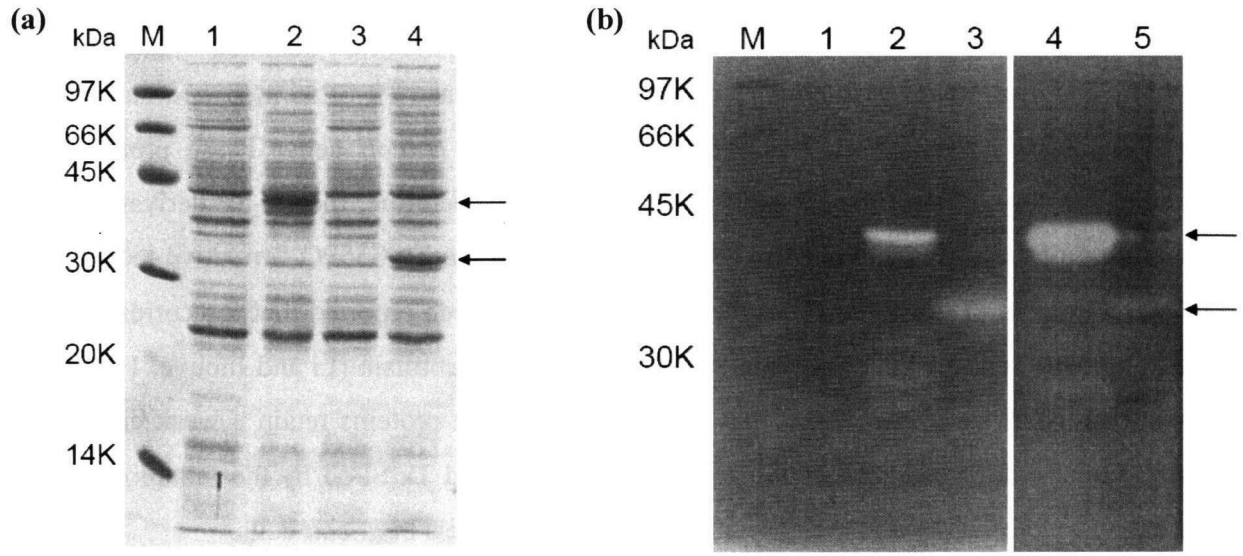

Figure 3.7. SDS-PAGE analyses for overproduction of Tk-S359C and Tk-S359CΔJ. (a) The *E. coli* BL21-CodonPlus(DE3)-RIL transformants with the pET25b derivatives for overproduction of Tk-S359C (lanes 1 and 2) and Tk-S359CΔJ (lanes 3 and 4) were subjected to 15% SDS-PAGE. After electrophoresis, the gel was stained with CBB. Lane M, low-molecular-weight marker kit (GE Healthcare); lanes 1 and 3, whole-cell extract without IPTG induction; lanes 2 and 4, soluble fraction obtained after sonication lysis of the cells with IPTG induction. (b) Activity staining of gel. The same samples loaded onto lanes 1, 2, and 4 of the gel in (a) were loaded onto lane 1, lanes 2 and 4, and lanes 3 and 5, respectively, of a 12% polyacrylamide gel containing 0.1% gelatin and SDS. After electrophoresis, the gel was washed with 2.5% (v/v) Triton X-100 at room temperature for 1 h, incubated in 50 mM Tris-HCl (pH 9.0) at 40°C (lanes 1-3) or 80°C (lanes 4 and 5) for 10 h, and stained with CBB. Protease bands were visualized as clear zones due to the hydrolysis of gelatin. Lane M, low-molecular-weight marker kit (GE Healthcare). The upper and lower arrows indicate the positions of Tk-S359C and Tk-S359CΔJ, respectively. The molecular mass of each standard protein is indicated beside the gel.

To determine the enzymatic activities of Tk-S359C and Tk-S359CΔJ, these proteins were purified as described for Tk-S359A and Tk-S359AΔJ to give a single band on SDS-PAGE (Fig. 3.2, lanes 5 and 6). The amount of the protein purified from 1 l culture was approximately 3 mg for these proteins. The molecular masses of Tk-S359C and Tk-S359CΔJ estimated from gel filtration chromatography are 43 and 30 kDa, respectively, which are comparable to those calculated from the amino acid sequences, suggesting that they exist as a monomer in solution.

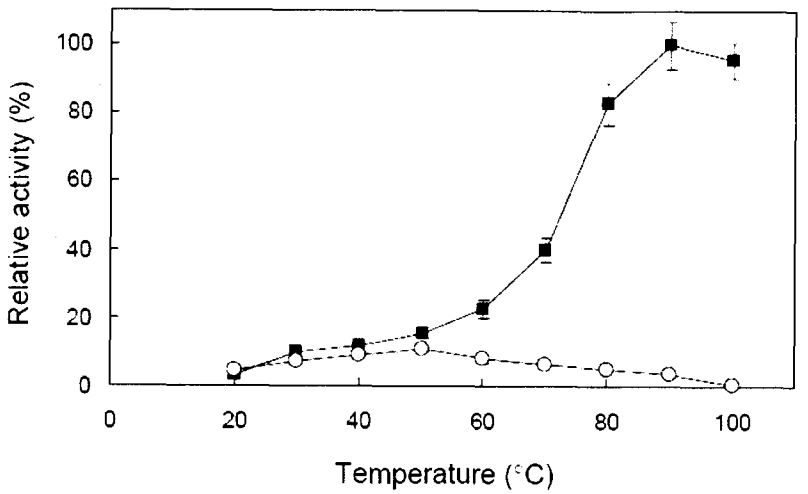

Figure 3.8. Temperature dependence of enzymatic activity. The enzymatic activities of Tk-S359C (solid square) and Tk-S359CΔJ (open circle) relative to that of Tk-S359C at 90°C are plotted as a function of temperature. 100% enzymatic activity of Tk-SP is 0.28 U/mg. The enzymatic activity was determined by using azocasin as a substrate. Each experiment was carried out twice and the average values are shown together with the error bars.

The enzymatic activities of Tk-S359C and Tk-S359CΔJ were determined at various temperatures ranging from 20 to 100°C using azocasien as a substrate (Fig. 3.8). The optimum temperature of Tk-S359C for activity was 90°C, which is comparable to that of Tk-SP (Chapter 2, Section 2.3.3). In contrast, Tk-S359CΔJ exhibited the highest activity at 50°C. The maximum specific activities of Tk-S359C and Tk-S359CΔJ were 0.28±0.020 and 0.030±0.002 units/mg, respectively. The specific activity of Tk-S359CΔJ was comparable to that of Tk-S359C at ≤40°C (0.025±0.016 for Tk-S359CΔJ and 0.032±0.013 for Tk-S359C at 40°C), but was considerably lower than that of Tk-S359C at ≥60°C. It did not increase beyond the maximum one at ≥60°C, probably because Tk-S359CΔJ is not fully stable and thermally denatured during assay at these temperatures. This result is consistent with the previous one that Tk-S359AΔJ is less stable than Tk-S359A by 30°C in T_m (Chapter 3, Section 3.3.4). Thus, the β-jelly roll domain contributes to the stabilization of Tk-SP. However, the result that the enzymatic activity of Tk-S359CΔJ is comparable to that of Tk-S359C at ≤40°C indicates that this domain is not important for activity.

3.3.6. Stability of Tk-S359A in Ca^{2+}-free form

It has previously been shown that the Ca^{2+} ions bind to Tk-SP too tightly to be removed upon extensive dialysis against the Ca^{2+}-free buffer but can be removed from the protein by the treatment with 10 mM EDTA (Chapter 2, Section 2.3.4). Tight binding of the Ca^{2+} ions to the protein is supported by the observation that the crystal structure of ProN-Tk-S359A contains two Ca^{2+} ions, because the protein is dialyzed against the Ca^{2+}-free buffer prior to crystallization. These results suggest that Pro-Tk-S359A and Tk-S359A are purified in a Ca^{2+}-bound form as well. Therefore, to analyze the role of these Ca^{2+} ions, Tk-S359A was dissolved in the buffer containing 10 mM EDTA and then dialyzed against the Ca^{2+}-free buffer. The resultant protein, termed Tk-S359A^{EDTA}, is expected to represent a Ca^{2+}-free form of Tk-S359A. As a reference, Tk-S359AΔJ^{EDTA} was also prepared. The far-UV and near-UV CD spectra of Tk-S359A^{EDTA} were nearly identical to those of Tk-S359A shown in Fig. 3.3 (data not shown), suggesting that the conformation of the β-jelly roll domain is not seriously changed upon removal of the Ca^{2+} ions. All of these spectra were not seriously changed in the presence of 10 mM $CaCl_2$ as well (data not shown). Because the β-jelly roll domain is approximately one-fourth of the entire Tk-S359A molecule in size, a possibility that the structural change of this domain upon removal of the Ca^{2+} ions is too small to be detected by CD spectra cannot be ruled out. However, the near-UV and far-UV CD spectra of Tk-S359A are significantly and slightly changed upon removal of β-jelly roll domain, respectively (Fig. 3.3). Therefore, it seems unlikely that β-jelly roll domain is unfolded upon removal of the Ca^{2+} ions.

Thermal denaturation of these proteins was analyzed by monitoring the change in CD values at 222 nm as the temperature was increased. Thermal denaturation of Tk-S359A^{EDTA} was reversible, while that of Tk-S359A was irreversible at any condition examined. The thermal denaturation curve of Tk-S359A was reproducible, unless the protein concentration, pH, and the rate of temperature increase were seriously changed. The thermal denaturation curves of Tk-S359A^{EDTA} and Tk-S359A measured in 50 mM Tris-HCl (pH 7.5) containing 1 M GdnHCl in the absence and presence of 10 mM $CaCl_2$, respectively, are shown in Fig. 3.6. The thermal denaturation curve of Tk-S359AΔJ^{EDTA} measured in the same buffer containing 1 M GdnHCl is also shown in Fig. 3.6. The midpoints of the transition of these thermal denaturation curves, T_m, are summarized in Table 3.2. The T_m value of Tk-S359A^{EDTA} is lower than that of Tk-S359A by 29.5°C, indicating that the Ca^{2+} ions contribute to the stabilization of Tk-S359A by 29.5°C in T_m.

In contrast, the T_m value of Tk-S359AΔJ^{EDTA} is nearly identical with that of Tk-S359AΔJ as expected, because Tk-S359AΔJ does not contain the Ca^{2+} ions. The difference in the T_m values between Tk-S359A and Tk-S359A^{EDTA} is nearly identical with that between Tk-S359A and Tk-S359AΔJ, indicating that the β-jelly roll domain does not contribute to the stabilization of Tk-SP in a Ca^{2+}-free form but contributes to it in a Ca^{2+}-bound form.

3.3.7. Role of the β-jelly roll domain

In this study, we showed for the first time that the β-jelly roll domain is not required for folding or activity but is required for hyperstabilization of the protein. This domain contains two Ca^{2+} ions and contributes to the stabilization of the protein only in a Ca^{2+}-bound form. It has been reported that the P-domains of kexin-like proteases with a β-jelly roll-like structure are essential for cleavage of the N-terminal pro-domain (Gluschankof and Fuller, 1994) and are important for activity, stability, and folding of the subtilisin domain (Zhou *et al.*, 1998 and Ueda *et al.*, 2003). However, the role of these P-domains has been analyzed by constructing the mutant proteins at the P-domains and comparing their expression and transportation in eukaryotic cells with those of the wild-type protein. None of these mutant proteins have been purified and biochemically characterized. Therefore, our results strongly suggest that these mutant proteins are not expressed and transported in cells normally, probably due to a great reduction in their stability.

Database searches indicate that none of the proteins, except for the C-terminal regions of alkaline serine proteases from *Pyrococcus furiosus* (accession no. NP579399), *Thermococcus gammatolerans*, (accession no. YP002958734) and *Thermococcus onnurineus* (accession no. YP002307742), show the amino acid sequence similarities to the β-jelly domain of Tk-SP. Alkaline serine proteases from *P. furiosus*, *T. gammatolerans* and *T. onnurineus*, all of which are from hyperthermophilic archaea, show the high amino acid sequence identities to Tk-SP both for the subtilisin domain (93%, 91%, and 81%, respectively) and the β-jelly roll domain (92%, 89%, and 84%, respectively). The Ca^{2+}-binding sites identified in the β-jelly roll domain of Tk-SP are almost fully conserved in these proteins. These results suggest that the β-jelly roll domain also contributes to the hyperstabilization of these proteins. These hyperthermophilic proteases probably adapt to high-temperature environment by attaching the β-jelly roll domain to their C-termini.

3.3.8. Role of N- and C-propeptides

In order to analyze the role of N- and C-propeptides of Tk-SP, we tried to overproduce the Pro-Tk-SP derivatives without N-propeptide (ProC-Tk-SP, residues 114-640), C-propeptide (ProN-Tk-SP, residues 1-539), and N- and C-propeptides (Tk-SP, residues 114-539) in *E. coli*. However, none of these proteins accumulate in the cells upon overproduction, such that they are detected as a band on SDS-PAGE by CBB staining or gel assay. These results suggest that all of these proteins exhibit cytotoxicities to *E. coli* cells. Because Pro-Tk-SP accumulates in *E. coli* cells upon overproduction, such that Pro-Tk-SP and its mature forms are detected as a band on SDS-PAGE by gel assay (Chapter 2, Section 2.3.2), N- and C-propeptides may be required to inhibit the activity of Tk-SP and thereby to prevent the damage of *E. coli* cells caused by the attack of the active protease molecules.

Tk-SP exhibits activity in gel assay (Chapter 2, Section 2.3.2), indicating that neither N- nor C-propeptide is required for folding of the mature domain of Pro-Tk-SP. However, a possibility that these propeptides accelerate folding of the mature domain cannot be ruled out. It has previously been shown that Tk-propeptide is not required for folding of Tk-subtilisin but is required to accelerate it (Tanaka *et al*, 2008, 2009). According to the crystal structure of Pro-Tk-subtilisin, the propeptide interacts with the mature domain mainly at the two regions. One is the C-terminal extended region, which interacts with the active-site cleft of the subtilisin domain. The other is the core region, which mainly interacts with two surface helices of the central αβα substructure of the subtilisin domain. The former interaction accelerates folding of the mature domain by binding to a folding intermediate of the mature domain with a native-like structure at its binding site. The latter interaction is critical for initiation of propeptide-catalyzed folding of the mature domain. Because both interactions are well conserved in the ProN-Tk-S359A structure, the N-propeptide domain of Pro-Tk-SP may also accelerate folding of its subtilisin domain.

3.3.9. Possible maturation process of Tk-SP

Tk-SP is matured from Pro-Tk-SP through an intermediate form with N- or C-propeptide. Binding of the C-terminal extended region of N-propeptide to the subtilisin domain of Tk-SP in a substrate-like manner indicates that this intermediate is Tk-SP with C-propeptide (ProC-Tk-SP), because the active site of the subtilisin domain of Pro-Tk-SP cannot attack the peptide bond

between Tk-SP and C-propeptide unless N-propeptide is autoprocessed and dissociated from the subtilisin domain. Therefore, Tk-SP is probably matured from Pro-Tk-SP with the following process: (1) folding of the mature domain; (2) autoprocessing of N-propeptide; (3) degradation of N-propeptide; (4) autoprocessing of C-propeptide. Bacterial subtilisins require propeptide for folding (Eder *et al.*, 1993; Shinde *et al.*, 2000; Falzon *et al.*, 2007; Eder and Fersht, 1995; Bryan, 2000; Fisher *et al.*, 2007) and Tk-subtilisin requires Ca^{2+} for it (Tanaka *et al.*, 2007a; Takeuchi *et al.*, 2009). In contrast, Tk-SP is folded into a functional structure without the assistance of propeptide and Ca^{2+}. α2m helix, β5m strand, and α3m helix form a central αβα substructure of Tk-SP. Stabilization of this substructure by the propeptide or Ca^{2+}-binding loop located between the first α helix and β strand of this αβα substructure has been reported to be crucial for folding of bacterial subtilisins (Bryan *et al.*, 1995) and Tk-subtilisin (Takeuchi *et al.*, 2009), respectively. Tk-SP contains a short loop (Asp^{237}-Gly^{240}) at the corresponding region of the Ca^{2+}-binding loop. This loop is not present in bacterial subtilisins and Kp-43 and is stabilized by hydrogen bonds. Therefore, this loop may function as a substitute of the Ca^{2+}-binding loop of Tk-subtilisin.

3.4. Summary

Tk-SP is a hyperthermostable subtilisin-like serine protease from *Thermococcus kodakaraensis* and is autoprocessed from its precursor (Pro-Tk-SP) with N- and C-propeptides. The crystal structure of the active-site mutant of Pro-Tk-SP lacking C-propeptide, ProN-Tk-S359A, was determined at 2.0 Å resolution. ProN-Tk-S359A consists of the N-propeptide, subtilisin, and β-jelly roll domains. Two Ca^{2+} ions bind to the β-jelly roll domain. The overall structure of ProN-Tk-S359A without the β-jelly roll domain is similar to that of bacterial propeptide:subtilisin complex, except that it does not contain Ca^{2+} ions. To analyze the role of the β-jelly roll domain of Tk-SP, we constructed a series of the active-site mutants of Tk-SP with (Tk-S359A/C) and without (Tk-S359A/CΔJ) β-jelly roll domain. Both Tk-S359C and Tk-S359CΔJ exhibited protease activities in gel assay, whereas Tk-S359CΔJ greatly reduced the activity at ≥60°C, but did not seriously affect it at ≤40°C, indicating that this domain is not important for folding and activity but is important for stability. In addition, the T_m value of Tk-S359AΔJ determined by far-UV CD spectroscopy in the presence of 10 mM $CaCl_2$ was lower

than that of Tk-S359A by 29.4°C. The T_m value of Tk-S359A was decreased by 29.5°C by the treatment with 10 mM EDTA, indicating that the β-jelly roll domain contributes to the stabilization of Tk-S359A only in a Ca^{2+}-bound form. Tk-SP highly resembles subtilisin-like serine proteases from *P. furiosus*, *T. gammatolerans*, and *T. onnurineus* in size and amino acid sequence. We propose that attachment of a β-jelly roll domain to the C-terminus is one of the strategies of the proteases from hyperthermophiles to adapt to high-temperature environment.

CHAPTER 4

Maturation process of Tk-SP: identification of intermediate form

4.1. Introduction

Bacterial subtilisins (Shinde and Inouye, 1996; Subbian *et al.*, 2005) and Tk-subtilisin (Tanaka *et al.*, 2007a, 2007b, 2009; Pulido *et al.*, 2006; Takeuchi *et al.*, 2009) are matured from their precursors (pro-subtilisin) by three steps: folding of the subtilisin domain, autoprocessing of propeptide, and degradation of propeptide, which proceed sequentially. By contrast, Tk-SP is matured from Pro-Tk-SP through an intermediate form with N- or C-propeptide because Tk-SP is composed of propeptides at both the N- and C-termini (Chapter 2, Section 2.3.2). However, it remains to be determined whether N- or C-propeptide is autoprocessed first, because these propeptides are similar in size (13 kDa for N-propeptide and 12 kDa for C-propeptide) and the amount of the intermediate form is too low to be identified by N-terminal amino acid sequencing or mass spectrometry.

We have previously shown that Tk-S359C, in which Ser^{359} of Tk-SP is replaced by Cys, and Tk-S359CΔJ, in which the β-jelly roll domain of Tk-S359C is removed, exhibit protease activities (Chapter 3, Section 3.3.5). These activities are much lower than those of Tk-SP and Tk-SPΔJ, because they do not exhibit cytotoxicity to *E. coli* cells, while Tk-SP and Tk-SPΔJ exhibit it (Chapter 3, Section 3.3.8). Therefore, it is expected that the Pro-Tk-SP derivative with the Ser^{359}→Cys mutation, Pro-Tk-S359C, is matured very slowly, such that an intermediate form can be identified. In fact, the corresponding mutations arrest maturation of bacterial subtilisin (Li and Inouye, 1994) and Tk-subtilisin (Tanaka *et al.*, 2007a) at the step, in which propeptide is autoprocessed and forms a complex with the mature domain. However, the effects of the Ser^{359}→Cys mutation on kinetic parameter of Tk-SP remain to be determined.

In this study, we overproduced, purified, and biochemically characterized Tk-S359C and Pro-Tk-S359C. Based on these results, we discuss the maturation process of Pro-Tk-SP.

4.2. Materials and Methods

4.2.1. Overproduction and purification

The pET25b derivative for overproduction of Pro-Tk-S359C was previously constructed (Chapter 3, Section 3.2.1). This plasmid was used to transform *E. coli* BL21-CodonPlus(DE3) (Stratagene). Overproduction of Pro-Tk-S359C using resultant *E. coli* transformant and purification of this proteins was carried out as described previously for Pro-Tk-S359A (Chapter 2, Section 2.2.2), except that the HiTrap Q (GE Healthcare) column chromatography was performed at pH 9.0 in the presence of 5 mM EDTA, the protein was eluted from this column by linearly increasing the NaCl concentration from 0 to 2 M. The solution containing the purified protein was dialyzed against 20 mM Tris-HCl (pH 7.5) and used for further biochemical characterization. Tk-SP and Tk-S359C were overproduced in *E. coli* and purified as described preciously (Chapter 2, Section 2.2.2 and Chapter 3, Section 3.2.2, respectively).

The N-terminal amino acid sequence of the protein was determined by a Procise automated sequencer model 491 (Applied Biosystems). The protein concentration was determined from UV absorption at 280 nm using the absorbance of a 0.1% (1.0 mg/ml) solution at 280 nm of 1.80 for Pro-Tk-S359C and 1.84 for Tk-SP and Tk-S359C. These values were calculated by using absorption coefficients of 1526 $M^{-1}cm^{-1}$ for Tyr and 5225 $M^{-1}cm^{-1}$ for Trp at 280 nm (Goodwin and Morton, 1946). The purity of the protein was confirmed by SDS-PAGE using a 15% polyacrylamide gel (Laemmli, 1970), followed by staining with Coomassie brilliant blue (CBB).

4.2.2. Molecular mass

The molecular mass of the protein was determined by a matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometer (Autoflex or Ultraflex, Bruker Daltonik GmbH) and gel filtration column chromatography using TSK-GEL G2000SWXL (Tosoh) as described in Chapter 2 (Section 2.2.3).

4.2.3. Circular dichroism spectra

The far-UV (200-260 nm) and near-UV (250-320 nm) CD spectra were measured at 25°C on a J-725 automatic spectropolarimeter (Japan Spectroscopic) with protein concentration and optical path length as described in Chapter 3 (Section 3.2.4).

4.2.4. Heat treatment of Pro-Tk-S359C

Pro-Tk-S359C was incubated at 80°C in 20 mM Tris-HCl (pH 7.5) or the same buffer containing 10 mM $CaCl_2$. With appropriate intervals, the protein was precipitated by 10% (v/v) trichloroacetic acid, washed with 70% acetone, and analyzed for autoprocessing by SDS-PAGE using a 15% polyacrylamide gel (Laemmli, 1970), followed by staining with CBB.

4.2.5. Enzymatic activity

The enzymatic activities of Tk-SP and Tk-S359C were determined using *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-*p*NA) (Sigma) as a peptide substrate and azocasein (Sigma) as a protein substrate as described in Chapter 2 (Section 2.2.5), except that the reaction mixture was performed in the presence of 10 mM $CaCl_2$.

4.3. Results and Discussion

4.3.1. Enzymatic activity of Tk-S359C

The specific activities of Tk-S359C were determined at 40 and 80°C using azocasien as a substrate and the kinetic parameters of Tk-S359C determined at 40°C using Suc-AAPF-*p*NA as a substrate are compared with those of Tk-SP in Table 4.1. The specific activity of Tk-S359C is lower than that of Tk-SP by 2500 folds at 40°C and 1800 folds at 80°C. Likewise, the k_{cat} value of Tk-S359C is lower than that of Tk-SP by 2200 folds at 40°C. While, the K_m value of Tk-S359C is comparable to that of Tk-SP. These results indicate that the Ser^{359}→Cys mutation does not inactivate Tk-SP, but greatly reduces its activity by ~2000 folds without seriously affecting its binding affinity to the substrate. The corresponding mutations probably reduce the enzymatic activities of bacterial subtilisins and Tk-subtilisin to a similar extent. It has been reported, however, that the subtilisin BPN′ derivative, in which the active-site serine residue (Ser^{221}) is

chemically converted to Cys, does not exhibit protease activity but exhibits esterase activity (Philipp and Bender, 1983). The conformation of the active site of subtilisin BPN′ may be altered by chemical treatment, so that it is unfavorable for protease activity.

Table 4.1. Kinetic parameters of Tk-S359C and Tk-SP[a]

Enzyme	Temp. (°C)	V_{max} (units/mg)	K_m (mM)	k_{cat} (m^{-1})	k_{cat}/K_m (m^{-1}/mM)
Tk-S359C	40	0.032 ± 0.013	0.11 ± 0.001	0.16 ± 0.01	1.5 ± 0.042
	80	0.23 ± 0.017	-	-	-
Tk-SP	40	79 ± 9.0	0.15 ± 0.02	350 ± 18	2340 ± 200
	80	410 ± 25	-	-	-

[a] The enzymatic activity was determined at the temperatures indicated in 50 mM Tris-HCl (pH 7.5) containing 10 mM $CaCl_2$. For determination of the V_{max} value, azocasein (2%) was used as a substrate. For determination of the K_m and k_{cat} values, Suc-AAPF-*p*NA was used as a substrate. The concentration of this substrate was varied from 0.01 to 2 mM. Each experiment was carried out three times and the average values are shown together with the errors.
-, not determined.

4.3.2. Preparation of Pro-Tk-S359C

When Pro-Tk-SP is overproduced in *E. coli*, it accumulates in cells in three forms with molecular masses of 65, 55, and 44 kDa, mainly in a form with molecular mass of 55 kDa (Chapter 2, Section 2.3.2). The 65, 55, and 44 kDa proteins represent Pro-Tk-SP, Tk-SP with N- or C-propeptide, and Tk-SP, respectively. However, the 55 kDa protein remains to be identified, because the production level of this protein is too low to be detected by SDS-PAGE, followed by CBB staining. This protein can only be detected by activity staining of the gel (gel assay) following SDS-PAGE. Production level of this protein is quite low, because the maturation process of Pro-Tk-SP is initiated when it is synthesized in the cells and the resultant mature form, Tk-SP, is cytotoxic to *E. coli* cells due to its protease activity. Therefore, the Pro-Tk-SP derivative with the Ser^{359}→Cys mutation, Pro-Tk-S359C, was overproduced and purified. Because this mutation greatly reduces the enzymatic activity of Tk-SP as mentioned above, Pro-Tk-S359C would not exhibit cytotoxicity to *E. coli* cells and therefore would be purified in an amount sufficient to identify an intermediate form of the maturation process.

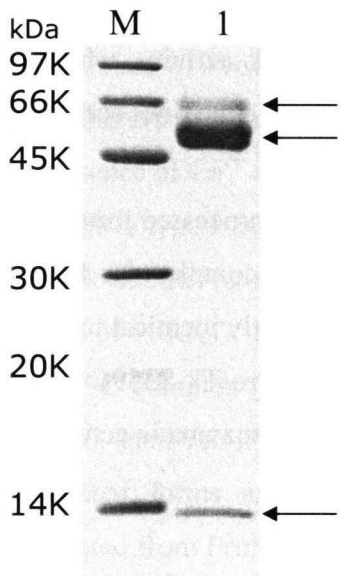

Figure 4.1. SDS-PAGE of Pro-Tk-S359C. Sample was subjected to electrophoresis on a 15% polyacrylamide gel in the presence of SDS. After electrophoresis, the gel was stained with CBB. Lane M, low-molecular-weight marker kit (GE Healthcare); lane 1, Pro-Tk-S359C (6 μg). The arrows indicate the positions of Pro-Tk-S359C, ProC-Tk-S359C, and ProN from the top to the bottom. The molecular mass of each standard protein is indicated beside the gel.

Upon overproduction, Pro-Tk-S359C accumulated in cells in a soluble form (data not shown). The production level of this protein and the amount of the protein purified from 1 l culture (3 mg) were comparable to those of Pro-Tk-S359A. Pro-Tk-S359C was eluted from the gel filtration column as a single peak. The molecular mass of this protein was estimated to be 63 kDa, which is comparable to that calculated from the amino acid sequence (68637), suggesting that it exists as a monomer in solution. Nevertheless, Pro-Tk-S359C gave three bands on SDS-PAGE with the molecular masses of 66, 55, and 13 kDa, although the amount of the 66 kDa protein estimated from the intensity of the band visualized with CBB staining was lower than that of the 55 kDa protein by >10 folds (Fig. 4.1). The N-terminal amino acid sequences of the 55 and 13 kDa proteins were determined to be VETE and PQKP, respectively, indicating that the N-terminal residues of the 55 and 13 kDa proteins are Val^{114} and Pro^{2}, respectively. The molecular masses of the 66, 55, and 13 kDa proteins were determined to be 68475 ± 255, 55852 ± 222, and 12631 ± 47 Da, respectively, by MALDI-TOF mass spectrometry, which are comparable to those of Pro-Tk-S359C (68637 Da), Tk-S359C with C-propeptide (ProC-Tk-S359C) (55968 Da), and N-propeptide (ProN) (12687 Da) calculated from the amino acid sequences. These results indicate that Pro-Tk-S359C was purified mostly in an autoprocessed form, in which ProN and ProC-Tk-S359C form a stable complex (ProN:ProC-Tk-S359C). It is noted that 55 kDa protein, instead of 66 kDa protein, accumulated in cells as a major protein upon induction for overproduction of Pro-Tk-S359C (data not shown). This result suggests that

Pro-Tk-S359C is autoprocessed into the ProN:ProC-Tk-S359C complex immediately when it is synthesized in *E. coli* cells. ProC-Tk-S359C probably exhibits very weak activity, which is comparable to that of Tk-S359C and is sufficient to promote autoprocessing but is not sufficient for subsequent degradation of ProN.

Pro-Tk-S359C purified as a mixture of the unautoprocessed and autoprocessed forms, but mostly in the autoprocessed form, was used as the ProN:ProC-Tk-S359C complex for further characterization. The far- and near-UV CD spectra of this complex were nearly identical to those of Pro-Tk-S359A (data not shown), indicating that the conformation of Pro-Tk-S359C is not seriously changed by autoprocessing of ProN. This complex exhibited little enzymatic activity at any temperature examined, indicating that this complex is inactive.

4.3.3. Maturation of ProN:ProC-Tk-S359C complex

Identification of the ProN:ProC-Tk-S359C complex as an intermediate form of the maturation process of Pro-Tk-S359C indicates that ProN is autoprocessed first. To examine whether the ProN:ProC-Tk-S359C complex is matured to Tk-S359C at elevated temperatures, this complex was incubated at 80°C in the presence or absence of Ca^{2+} in 20 mM Tris-HCl (pH 7.5) and analyzed by SDS-PAGE. When this complex was incubated in the absence of Ca^{2+}, the 44 kDa protein was generated (Fig. 4.3, lanes 2-4). This protein, as well as the 55 and 13 kDa proteins, did not give a clear band on SDS-PAGE. However, these proteins were identified as a single species of Tk-S359C (44392 kDa), ProC-Tk-S359C (55680 kDa), or ProN (12577 kDa) by MALDI-TOF mass spectrometry. As the incubation time increases, the amounts of Pro-Tk-S359C and ProC-Tk-S359C decrease, while that of Tk-S359C increases, indicating that ProC-Tk-S359C is gradually autoprocessed into Tk-S359C upon incubation at 80°C. Roughly 50% of ProC-Tk-S359C was autoprocessed into Tk-S359C upon incubation at 80°C for 3 h. Nevertheless, ProC was not detected by MALDI-TOF mass spectrometry. Instead, two peptides with molecular masses of 6360 and 4021 were detected, in addition to Pro-Tk-S359C, ProC-Tk-S359C, Tk-S359C, and ProN. This result suggests that ProC is degraded by the mature domain of ProC-Tk-S359C immediately when it is released from ProC-Tk-S359C upon autoprpcessing. Because the sum of the molecular masses of two peptides are smaller than the molecular mass of ProC calculated from the amino acid sequence (11780), ProC is probably degraded at multiple sites. In contrast, when the ProN:ProC-Tk-S359C complex was incubated in the presence of Ca^{2+},

it was not autoprocessed into Tk-S359C at all (Fig. 4.3, lane 5). These results suggest that the ProC domain of ProC-Tk-S359C does not assume a stable conformation in the absence of Ca^{2+} and is therefore highly susceptible to autoproessing, while it assumes a stable conformation in the presence of Ca^{2+} and is therefore highly resistant to autoprocessing.

It is noted that the amount of ProN is not seriously changed during autoprocessing of ProC-Tk-S359C into Tk-S359C at 80°C (Fig. 4.3). This result suggests that the enzymatic activity of ProC-Tk-S359C is too weak to degrade ProN. As mentioned above, the ProN:ProC-Tk-S359C complex is inactive. Nevertheless, ProC-Tk-S359C is autoprocessed into Tk-S359C. At 80°C, the ProN:ProC-Tk-S359C complex may exist in equilibrium between associated and dissociated forms and ProC-Tk-S359C may be autoporcessed into Tk-S359C when it is dissociated from ProN.

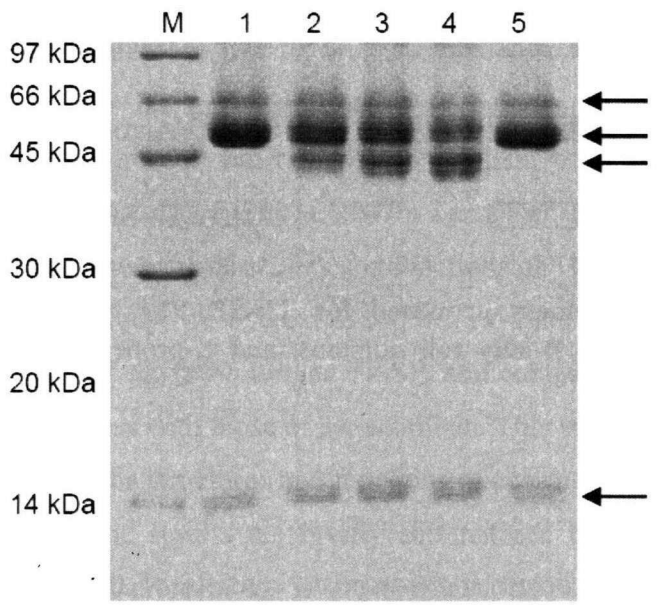

Figure 4.3. SDS-PAGE analyses for maturation of Pro-Tk-S359C. Pro-Tk-S359C (6 μg), was incubated in 1 ml of 20 mM Tris-HCl (pH 7.5) at 80°C for 0 h (lane 1), 1 h (lane 2), 2 h (lane 3), and 3 h (lane 4) or the same buffer containing 10 mM $CaCl_2$ at 80°C for 5 h (lane 5), and the protein was precipitated by the addition of 112 μl of trichloroacetic acid (100%, wt/vol) and subjected to 15% SDS-PAGE. The protein was stained with CBB. M, low-molecular-weight marker kit (GE Healthcare). The arrows indicate Pro-Tk-S359C, ProC-Tk-S359C, Tk-S359C, and ProN from the top to the bottom. The molecular mass of each standard protein is indicated beside the gel.

4.3.4. Proposed maturation process of Tk-SP

According to the maturation process of Tk-SP using Pro-Tk-S359C (Chapter 4, Section 4.3.2 and 4.3.3), the first step of maturation process is folding of the mature domain. Unlike, Tk-subtilisin and bacterial subtilisins, Tk-SP does not require propeptide and Ca^{2+} for folding of the mature domain. The second and third steps are autoprocessing and degradation of N-propeptide, respectively because N-propeptide of Tk-SP functions as an inhibitor of its mature domain. The last step is autodegradation of C-propeptide and the active Tk-SP is released. The proposed model for maturation process of Tk-SP is summarized in Fig. 4.4.

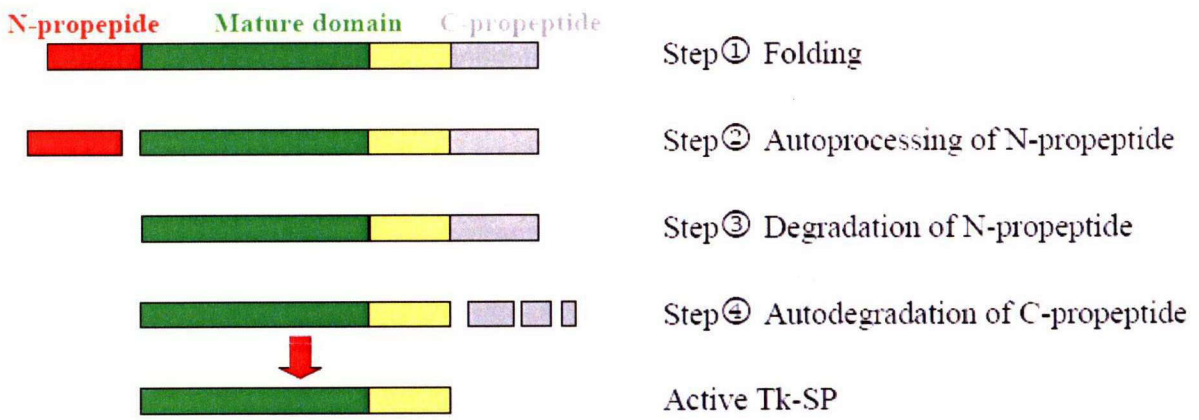

Figure 4.4. Maturation process proposed for Tk-SP. The N-propeptide, mature domain containing the subtilisin and β-jelly roll domains, and C-propeptide are colored red, green, yellow, and gray, respectively.

4.4. Summary

Tk-SP from *Thermococcus kodakaraensis* consists of the subtilisin and β-jelly roll domains and is matured from Pro-Tk-SP by removal of N- and C-propeptides. Mutation of the catalytic residue, Ser359, to Cys greatly reduced the k_{cat} value of Tk-SP by 2000 folds without seriously affecting the K_m value. As a result, this mutation arrested maturation of Pro-Tk-SP at the step in which only N-propeptide is autoprocessed, indicating that N-propeptide is autoprocessed first.

CHAPTER 5

General Conclusion

Subtilisin-like serine proteases are industrially valuable enzymes because they have diverse application in a wide variety of industries, such as detergent business, food processing, pharmaceutical development, and leather processing. Therefore, researchers still continue their efforts to identify novel subtilisin-like serine proteases with desirable functions. Subtilisin-like serine proteases from hyperthermophilic organisms have great potential for biotechnological application, because they are expected to be highly stable and to exhibit enzymatic activities even in very harsh conditions. To use them for industrial purposes, detailed information on their structures and functions is necessary. However, our knowledge on the structures and functions of subtilisin-like serine proteases from hyperthermophilic organisms is still limited.

The genome of the hyperthermophilic archaeon *Thermococcus kodakaraensis* contains three genes encoding subtilisin-like serine proteases, Prepro-Tk-subtilisin, Tk-1689 (Prepro-Tk-SP), and Tk-0076. The structure and function of Tk-subtilisin have been extensively studied (Kannan *et al.*, 2001; Pulido *et al.*, 2006, 2007a, 2007b; Tanaka *et al.*, 2007a, 2007b, 2008, 2009; Takeuchi *et al.*, 2009). In contrast, the structures and functions of Tk-1689 and Tk-0076 have not been analyzed at all. In this study, we overexpressed the Tk-1689 gene in *E. coli*, purified and characterized a mature form of Tk-1689, termed Tk-SP, and compared its enzymatic properties, structure, and maturation process with those of Tk-subtilisin. This study is expected to contribute to increase our knowledge on the structures and functions of subtilisin-like serine proteases from hyperthermophilic archaea. The results are briefly summarized below. Future prospects and possible biotechnological applications of Tk-SP are also described.

First, to examine whether Tk-SP is matured from its precursor form, Pro-Tk-SP, by autoprocessing and degradation of N-propeptide as Tk-subtilisin and bacterial subtilisins are, the gene encoding Pro-Tk-SP was overexpressed in *E. coli*. Upon overproduction, the protein accumulated in cells in three forms with molecular masses of 65, 55, and 44 kDa, which represent an intact form, a partially autoprocessed form, and a fully autoprocessed form (Tk-SP), respectively. All of them accumulated in cells in a soluble form. However, the protein was purified as Tk-SP, indicating that the N-terminal 113 and C-terminal 101 amino acid residues are autoprocessed during overproduction and purification of the protein. This result suggests that

Pro-Tk-SP consists of an N-terminal propeptide (Ala^{1}-Ala^{113}), a mature domain (Tk-SP, Val^{114}-Val^{539}), and a C-terminal propeptide (Asp^{540}-Gly^{640}). Thus, Tk-SP is different from Tk-subtilisin and bacterial subtilisins in maturation mechanism, because not only N-propeptide but also C-propeptide is autoprocessed from Pro-Tk-SP for maturation of Tk-SP. Tk-SP exhibits protease activity in gel assay in the absence of Ca^{2+} following SDS-PAGE (Fig. 2.2b), suggesting that Tk-SP requires neither propeptide nor Ca^{2+} for folding. Because Tk-subtilisin and bacterial subtilisins require Ca^{2+} and propeptide for folding, respectively, Tk-SP is different from Tk-subtilisin and bacterial subtilisins in folding mechanism as well. As to the enzymatic properties, Tk-SP is highly active at alkaline pH even at 100°C and highly resistant to thermal and chemical denaturation as Tk-subtilisin is. However, both the k_{cat} and K_m values of Tk-SP determined for peptide substrate are lower than those of Tk-subtilisin by roughly 20 folds. As a result, the k_{cat}/K_m value of Tk-SP is comparable to that of Tk-subtilisin. Thus, Tk-SP is characterized by high substrate binding affinity and low turnover number as compared to Tk-subtilisin.

Secondly, to examine whether Tk-SP has a unique structure, the crystal structure of the active-site mutant of the Pro-Tk-SP derivative lacking C-propeptide, ProN-Tk-S359A, was determined. The overall structure of ProN-Tk-S359A is similar to that of the active-site mutant of Pro-Tk-subtilisin, Pro-S324A (Tanaka *et al.*, 2007a), except that it contains the β-jelly roll domain at the C-terminus of the subtilisin domain and does not contain Ca^{2+} ions in the subtilisin domain. Two Ca^{2+}-binding sites were identified in the β-jelly roll domain. Pro-Tk-SP has a C-terminal extension of ~200 residues as compared to Pro-Tk-subtilisin. The C-terminal half of this extension is autoprocessed as C-propeptide when Pro-Tk-SP is matured to Tk-SP. However, we showed that the N-terminal half of this extension assumes a β-jelly roll structure, which is located at the C-terminal region of Tk-SP as a separate domain. This structure is not present in Tk-subtilisin and bacterial subtilisins. The finding that the subtilisin domain of Tk-SP does not contain any Ca^{2+}-binding site is consistent with the fact that Tk-SP does not require Ca^{2+} for folding or activity (Chapter 2). Construction of the active-site mutant of the Tk-SP derivatives lacking β-jelly roll domain, Tk-S359AΔJ and Tk-S359CΔJ, followed by biochemical characterization, indicate that the β-jelly roll domain contributes to hyperstabilization of Tk-SP only in a Ca^{2+}-bound form. Thus, Tk-SP is different from Tk-subtilisin in stabilization mechanism, because Tk-subtilisin is stabilized by the Ca^{2+} ions bound to Ca-1 and Ca-6 sites, which are formed by unique insertion sequences. The β-jelly roll domain of Tk-SP shows high

amino acid sequence identities to the C-terminal regions of alkaline serine proteases from *Pyrococcus furiosus*, *Thermococcus gammatolerans*, and *Thermococcus onnurineus* (92%, 89%, and 84%, respectively), all of which are from hyperthermophilic archaea. The Ca^{2+}-binding sites identified in the β-jelly roll domain of Tk-SP are almost fully conserved in these proteins. These results suggest that the β-jelly roll domain also contributes to hyperstabilization of these proteins in a Ca^{2+}-bound form. We propose that attachment of β-jelly roll domain to the C-terminus is one of the strategies of the proteases from hyperthermophiles to adapt to high-temperature environment.

Finally, to examine whether N- or C-propeptide is autoprocessed first in the maturation process of Pro-Tk-SP to Tk-SP, the active-site mutant of Pro-Tk-SP, Pro-Tk-S359C, was constructed. The Tk-SP derivative with the Ser^{359}→Cys mutation, Tk-S359C, was also constructed to analyze the effect of this mutation on enzymatic activity of Tk-SP. The specific activity of Tk-S359C was lower than that of Tk-SP by 2000 folds, indicating that this mutation greatly reduces enzymatic activity of Tk-SP, but does not inactivate it. This result suggests that Pro-Tk-S359C is suitable to analyze maturation process of Pro-Tk-SP, because the enzymatic activity of its subtilisin domain is probably sufficient to initiate maturation process, but is not sufficient to complete it. Pro-Tk-S359C was purified mostly in an autoprocessed form, in which ProN and ProC-Tk-S359C form a stable complex, indicating that ProN is autoprocessed first and ProC-Tk-SP is the intermediate form of maturation process.

Tk-SP is characterized by the presence of propeptides at both N- and C-termini. These propeptides are also present in aqualysin-I. Propeptides of bacterial subtilisin and N-propeptide of aqualysin I function as an intramolecular chaperone, that facilitates folding of the mature domain, and acts as a strong inhibitor of their cognate mature domains (Zhu *et al.*, 1989; Eder and Fersht, 1995; Li *et al.*, 1995; Yabuta *et al.*, 2001; Marie-Claire *et al.*, 2001). As to the chaperone function, N- and C-propeptides are not required for folding of Tk-SP, because Tk-SP without its N- and C-propeptides can be refolded and exhibits activity in gel assay following SDS-PAGE (Fig. 2.2b). However, a possibility that these propeptides accelerate folding of the mature domain cannot be excluded. It has previously been shown that Tk-propeptide is not required for folding of Tk-subtilisin but is required to accelerate it (Tanaka *et al.*, 2008, 2009). As to the inhibitory function, N- and C-propeptides may be required to inhibit the activity of Tk-SP and thereby to prevent the damage of *E. coli* cells caused by the attack of the active protease

molecules. In fact, overproducing strains of Pro-Tk-SP derivatives, which lack either N- or C-propeptide or both propeptides, cannot be constructed, probably due to cytotoxicities of the mature domain to *E. coli* cells. The crystal structure of ProN-Tk-S359A indicates that the C-terminal extended region of N-propeptide runs through the active-site cleft in a substrate-like manner. It is likely that N-propeptide acts as an inhibitor, which inhibits the enzymatic activity of the mature domain. However, it remains to be determined whether these propeptides inhibit enzymatic activity of Tk-SP. Overproduction of N- and C-propeptides alone and biochemical characterization of these propeptides may facilitate understanding of the role of these propeptides. Database searches indicate that none of the proteins, except for the C-terminal regions of alkaline serine proteases from *P. furiosus*, *T. gammatolerans*, and *T. onnurineus*, show significant amino acid sequence similarities to C-propeptide of Tk-SP. The C-terminal regions of these alkaline serine proteases show amino acid sequence identities of 76-87% to C-propeptide of Tk-SP, suggesting that C-propeptide is also conserved in these proteins. However, the amino acid sequence of C-propeptide is not similar to that of any protein, for which the crystal structure is available. Therefore, it is necessary to determine the crystal structure of C-propeptide to understand its role. To determine this structure, it would be necessary to crystallize Pro-Tk-SP or its derivative containing C-propeptide. Construction of a variety of the Pro-Tk-SP derivatives, followed by screening for crystallization conditions, may facilitate formation of crystals suitable for X-ray crystallographic analyses and determination of the crystal structure of C-propeptide. Futhermore, as the result of Tk-SP structure, Tk-SP contains two Ca^{2+}-binding sites in the β-jelly roll domain. The Ca-1 site is located in the β-jelly roll domain, whereas the Ca-2 site is located at the interface between the subtilisin and β-jelly roll domains. Ca^{2+} ions are required for stabilization of Tk-SP because T_m value of Tk-S359A in a Ca^{2+}-bound form is higher than that of Tk-S359A in a Ca^{2+}-free form by 29.5°C in T_m. However, it remains to be determined whether only single Ca^{2+} ion may be required for stabilization of Tk-SP. Therefore, characterization of the mutant proteins with single Ca^{2+}-binding sites is necessary to understand the role of Ca^{2+} ions.

For biotechnological application, broad substrate specificity and high resistance to heat, denaturants, and detergents are great advantage for Tk-SP over other subtilisin-like serine proteases from mesophiles. Tk-subtilisin also shows broad substrate specificity and high resistance to heat, denaturants, and detergents. However, Tk-subtilisin is overproduced in *E. coli* in inclusion bodies. Therefore, it is necessary to solubilize the protein in the buffer containing 8

M urea, purified the protein in a urea-denatured form, and refolded the protein by removing urea. In contrast, Tk-SP is overproduced in *E. coli* in a soluble form, which is enzymatically active. Production in a soluble form is an advantage for Tk-SP over Tk-subtilisin, because complicated refolding procedure is not necessary for purification of Tk-SP. Moreover, Tk-SP is superior to Tk-subtilisin in high resistance to chelating agents. Tk-subtilisin requires Ca^{2+} for folding and therefore it is highly sensitive to EDTA. In contrast, Tk-SP does not require Ca^{2+} for folding and therefore it is fully active even in the presence of 10 mM EDTA (Chapter 2). Detergents used for laundry machines usually contain chelating agents. Therefore, proteases, which are sensitive to chelating agents, cannot be used as additives of detergents. However, low production level and low enzymatic activity are disadvantage for Tk-SP over Tk-subtilisin. The amount of Tk-SP purified from 1 liter culture (*E. coli*) is only 0.8 mg. The specific activity of Tk-SP is lower than that of Tk-subtilisin by 5 folds for protein substrate. These disadvantages limit the biotechnological application of Tk-SP. The production level of Tk-SP is quite low, probable because Tk-SP is matured and activated during overproduction and purification, and the resultant mature form is cytotoxic to *E. coli* cells due to its protease activity. Therefore, development of an overproduction system, in which Tk-SP is overproduced in inactive form and is activated when needed, is necessary to increase production level of Tk-SP. The cost for purification of Tk-SP would be greatly reduced if its production level increases. Likewise, construction of the Tk-SP derivatives with increased enzymatic activity at broad temperature and pH ranges will also be necessary to reduce the cost for purification of the enzyme. Mutational works on Tk-SP using directed evolution, which is the way for a great variety of subtilisin variants with better specificity and stability, will be useful to construct the Tk-SP derivatives with increased enzymatic activity. If these disadvantages for Tk-SP were overcome, application possibility of Tk-SP would greatly increase. Possible application of Tk-SP would be prevention of secondary infection by degrading disease-causing amyloid fibers, prevention of allergy by degrading allergens, and removal of environmental pollutants by degrading persistent protein substances, such as hair and feather.

REFERENCES

Almog, O., Gonzalez, A., Klein, D., Greenblatt, H. M., Braun, S., and Shoham, G. 2003. The 0.93Å crystal structure of sphericase: a calcium-loaded serine protease from *Bacillus sphaericus*. *J. Mol. Biol.* **332:** 1071-1082.

Atomi, H., Fukui, T., Kanai, T., Morikawa, M., and Imanaka, T. 2004. Description of *Thermococcus kodakaraensis* sp. nov., a well studied hyperthermophilic archaeon previously reported as *Pyrococcus* sp. KOD1. *Archaea* **1:** 263-267.

Baier, K., Nicklisch, S., Maldener, I., and Lockau, W. 1996. Evidence for propeptide-assisted folding of the calcium-dependent protease of the cyanobacterium Anabaena. *Eur. J. Biochem.* **241:** 750-755.

Bajorath, J., Raghunathan, S., Hinrichs, W., and Saenger, W. 1989. Long-range structural changes in proteinase K triggered by calcium ion removal. *Nature* **337:** 481-484.

Barrette-Ng, I. H., Ng, K. K., Cherney, M. M., Pearce, G., Ryan, C. A., and James, M. N. 2003. Structural basis of inhibition revealed by a 1:2 complex of the two-headed tomato inhibitor-II and subtilisin Carlsberg. *J. Biol. Chem.* **278:** 24062-24071.

Basak, A., and Lazure, C. 2003. Synthetic peptides derived from the prosegments of proprotein convertase 1/3 and furin are potent inhibitors of both enzymes. *Biochem. J.* **373:** 231-239.

Betzel, C., Pal, G. P., and Saenger, W. 1988. Three-dimensional structure of proteinase K at 0.15-nm resolution. *Eur. J. Biochem.* **178:** 155-171.

Bode, W., Papamokos, E., and Musil, D. 1987. The high-resolution X-ray crystal structure of the complex formed between subtilisin Carlsberg and eglin c, an elastase inhibitor from the leech *Hirudo medicinalis*. Structural analysis, subtilisin structure and interface geometry. *Eur. J. Biochem.* **166:** 673-692.

Bryan, P. N. 2000. Protein engineering of subtilisin. *Biochim. Biophys. Acta.* **1543**: 203-222.

Bryan, P., Alexander, P., Strausberg, S., Schwarz, F., Lan, W., Gilliland, G., and Gallagher, D. T. 1992. Energetics of folding subtilisin BPN'. *Biochemistry* **31:** 4937-4945.

Bryan, P., Wang, L., Hoskins, J., Ruvinov, S., Strausberg, S., Alexander, P., Almog, O., Gilliland, G, L., and Gallagher, T. 1995. Catalysis of a protein folding reaction: mechanistic implications of the 2.0 Å structure of the subtilisin-prodomain complex. *Biochemistry* **34:** 10310-10318.

Choi, I. G., Bang, W. G., Kim, S. H., and Yu, Y. G. 1999. Extremely thermostable serine-type protease from *Aquifex pyrophilus*. Molecular cloning, expression, and characterization. *J. Biol. Chem.* **274:** 881-888.

Collaborative Computational Project, Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallog. Sect. D: Biol. Crystallogr.* **50:** 760-763.

Davail, S., Feller, G., Narix, E., and Gerday, C. 1994. Cold adaptation of proteins. Purification, characterization, and sequence of the heat-labile subtilisin from the Antarctic psychrophile *Bacillus* TA41. *J. Biol. Chem.* **269:** 17448-17453.

Eder, J., and Fersht, A. R. 1995. Pro-sequences assisted protein folding. *Mol. Microbiol.* **16:** 609-614.

Eder, J., Rheinnecker, M., and Fersht, A. R. 1993. Folding of subtilisin BPN': role of the pro-sequence. *J. Mol. Biol.* **233:** 293-304.

Emsley, P., and Cowtan, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **60:** 2126-2132.

Exterkate, F. A., and Alting, A. C. 1999. Role of calcium in activity and stability of the *Lactococcus lactis* cell envelope proteinase. *Appl. Environ. Microbiol.* **65:** 1390-1396.

Falzon, L., Patel, S., Chen, Y. J., and Inouye, M. 2007. Autotomic behavior of the propeptide in propeptide-mediated folding of prosubtilisin E. *J. Mol. Biol.* **366:** 494-503.

Fisher, K. E., Ruan, B., Alexander, P. A., Wang, L., and Bryan, P. N. 2007. Mechanism of the kinetically-controlled folding reaction of subtilisin. *Biochemistry* **46:** 640-651.

Fukui, T., Atomi, H., Kanai, T., Matsumi, R., Fujiwara, S., and Imanaka, T. 2005. Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. *Genome Res.* **15:** 352-363.

Gallagher, T., Bryan, P., and Gilliland, G. L. 1993. Calcium-independent subtilisin by design. *Protein* **16:** 205-213.

Gallagher, T., Gilliland, G., Wang, L., and Bryan, P. 1995. The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure* **3:** 907-914.

Gluschankof, P., and Fuller, R. S. 1994. A C-terminal domain conserved in precursor processing proteases is required for intramolecular N-terminal maturation of pro-Kex2 proteasse. *EMBO J.* **13,** 2280-2288.

Gödde, C., Sahm, K., Brouns, S. J., Kluskens, L. D., van der Oost, J., de Vos, W. M., and Antranikian, G. 2005. Cloning and expression of islandisin, a new thermostable subtilisin from Fervidobacterium islandicum, in *Escherichia coli*. *Appl. Environ. Microbiol.* **71:** 3951-3958.

Goodwin, T. W., and Morton, R. A. 1946. The spectrophotometric determination of tyrosine and tryptophan in proteins. *Biochem. J.* **40:** 628-632.

Gros, P., Betzel, C., Dauter, Z., Wilson, K. S., and Hol, W. G. 1988. Molecular dynamics refinement of a thermitase-eglin-c complex at 1.98 Å resolution and comparison of two crystal forms that differ in calcium content. *J. Mol. Biol.* **210:** 347-367.

Gros, P., Kalk, K. H., and Hol, W. G. 1991. Calcium binding to thermitase. Crystallographic studies of thermitase at 0, 5, and 100 mM calcium. *J. Biol. Chem.* **266:** 2953-2961.

Henrich, S., Cameron, A., Bourenkov, G. P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W., and Than, M. E. 2003. The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. *Nat. Struct. Biol.* **10:** 520-526.

Holyoak, T., Wilson, M. A., Fenn, T. D., Kettner, C.A., Petsko, G. A., Fuller, R. S., and Ringe, D. 2003. 2.4 Å resolution crystal structure of the prototypical hormone-processing protease Kex2 in complex with an Ala-Lys-Arg boronic acid inhibitor. *Biochemistry* **42:** 6709-6718.

Holyoak, T., Kettner, C. A., Petsko, G. A., Fuller, R. S., and Ringe, D. 2004. Structure basis for differences in substrate selectivity in Kex2 and furin protein convertases. *Biochemistry* **43:** 2412-2421.

Huang, H. W., Chen, W. C., Wu, C. Y., Yu, H. C., Lin, W. Y., Chen, S. T., and Wang, K. T. 1997. Kinetic studies of the inhibitory effects of propeptides subtilisin BPN′ and Carlsberg to bacterial serine proteases. *Protein Eng.* **10:** 1227-1233.

Jacobs, M., Eliasson, M., Uhlen, M., and Flock, J. I. 1985. Cloning, sequencing, and expression of subtilisin Carlsberg from *Bacillus licheniformis*. *Nucleic Acids Res.* **13:** 8913-8926.

Jain, S. C., Shinde, U., Li, Y., Inouye, M., and Berman, H. M. 1998. The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 Å resolution. *J. Mol. Biol.* **284:** 137-144.

Kannan, Y., Koga, Y., Inoue, Y., Haruki, M., Takagi, M., Imanaka, T., Morikawa, M., and Kanaya, S. 2001. Active subtilisin-like protease from a hyperthermophilic archaeon in a form with a putative prosequence. *Appl. Environ. Microbiol.* **67:** 2445-2552.

Kim, D. W., Lee, Y. C., and Matsuzawa, H. 1997. Role of the COOH-terminal pro-sequence of aqualysin I (a heat-stable serine protease) in its extracellular secretion by *Thermus thermophilus*. *FEMS Lett.* **157:** 39-45.

Kobayashi, H., Utsunomiya, H., Yamanaka, H., Sei, Y., Katunuma, N., Okamoto, K., and Tsuge, H. 2009. Structural basis for the kexin-like serine protease from *Aeromonas sobria* as sepsis-causing factor. *J. Biol. Chem.* **284:** 27655-25663.

Krissinel, E., and Henrick, K. 2004. Secondary-structure matching (SSM), a new tool for fast protein alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60:** 2256-2268.

Kumar, C. G., and Takagi, H. 1999. Microbial alkaline proteases: from a bioindustrial viewpoint. *Biotechnol Adv.* **17:** 561-594.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227:** 680-685.

Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3:** 1171-1179.

Laskowski, R. A., MacSrthur, M. W., Moss, D. S., and Thornton, J. M. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26:** 283-291.

Li, Y., Hu, Z., Jordan, F., and Inouye, M. 1995. Functional analysis of the propeptide of subtilisin E as an intramolecular chaperone for protein folding. Refolding and inhibitory abilities of propeptide mutants. *J. Biol. Chem.* **270:** 25127-25132.

Li. Y., and Inouye, M. 1994. Autoprocessing of prothiolsubtilisin E in which active-site serine 221 is altered to cysteine. *J. Biol. Chem.* **269:** 4169-4174.

Marie-Claire, C., Ruffet, E., Beaumont, A., and Roques, B. P. 1999. The prosequence of thermolysin acts as an intramolecular chaperone when expressed in *trans* with the mature sequence in *Escherichia coli*. *J. Mol. Biol.* **285:** 1911-1915.

Marie-Claire, C., Yabuta, Y., Suefuji, K., Matsuzawa, H., and Shinde, U. 2001. Folding pathway mediated by an intramolecular chaperone: the structural and functional characterization of the aqualysin I propeptide. *J. Mol. Biol.* **305:** 151-165.

Maurer, K. H. 2004. Detergent proteases. *Curr. Opin. Biotechnol.* **15:** 330-334.

Morikawa, M., Izawa, Y., Rashid, N., Hoaki, T., and Imanaka, T. 1994. Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic Pyrococcus sp. *Appl. Environ. Microbiol.* **60:** 4559-4566.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **53:** 240-255

Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. 1995. SOCP: a structureal classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247:** 536-540.

Neurath, H. 1984. Evolution of proteolytic enzymes. *Science* **224:** 350-357.

Nirasawa, S., Nakajima, Y., Zhang, Z. Z., Yoshida, M., and Hayashi, K. 1999. Intramolecular chaperone and inhibitor activities of a propeptide from a bacterial zinc aminopeptidase. *Biochem. J.* **341:** 25-31.

Nonaka, T., Fujihashi, M., Kita, A., Saeki, K., Ito, S., Horikoshi, K., and Miki, K. 2004. The crystal structure of an oxidatively stable subtilisin-like alkaline serine protease, KP-43, with a C-terminal beta-barrel domain. *J. Biol. Chem.* **279:** 47344-47351.

O'Donohue, M. J., and Beaumont, A. 1996. The roles of the prosequence of thermolysin in enzyme inhibition and folding *in vitro*. *J. Biol. Chem.* **271:** 26477-26481.

Ottmann, C., Rose, R., Huttenlocher, F., Cedzich, A., Hauske, P., Kaiser, M., Huber, R., and Schaller, A. 2009. Structural basis for Ca^{2+}-independence and activation by homodimerization of tomato subtilase 3. *Proc. Natl. Acad. Sci. U. S. A.* **106:** 17223-17228.

Otwinowski, Z., and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276:** 307-326.

Pantoliano, M. W., Whitlow, M., Wood J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L., and Bryan, P. N. 1989. Large increases in general stability for subtilisin BPN' through incremental changes in the free energy of unfolding. *Biochemistry* **28:** 7205-7213.

Pantoliano, M. W., Whitlow, M., Wood J. F., Rollence, M. L., Finzel, B. C., Gilliland, G. L., Poulos, T. L., and Bryan, P. N. 1988. The engineering of binding affinity at metal ion binding sites for the stabilization of proteins: subtilisin as a test case. *Biochemistry* **27:** 8311-8317.

Perona, J., and Craik C. S. 1995. Structural basis of substrate specificity in the serine proteases. *Prot. Sci.* **4:** 337-360.

Philipp, M. and Bender, M. L. 1983. Kinetics of subtilisin and thiolsubtilisin. *Mol. Cell. Biochem.* **51:** 5-32.

Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T. F. 1987. Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature* **325:** 458-462.

Pulido, M. A., Saito, K., Tanaka, S., Koga, Y., Morikawa, M., Takano, K., and Kanaya, S. 2006. Ca^{2+}-dependent maturation of Tk-subtilisin from a hyperthermophilic archaeon: propeptide is a potent inhibitor of the mature domain but is not required for its folding. *Appl. Environ. Microbiol.* **72:** 4154-4162.

Pulido, M. A., Koga, Y., Takano, K., and Kanaya, S. 2007a. Directed evolution of Tk-subtilisin from a hyperthermophilic archaeon: identification of a single amino acid substitution in the propeptide region responsible for low-temperature adaptation. *Protein Eng. Des. Sel.* **20:** 143-153.

Pulido, M., Tanaka, S., Sringiew, C., You, D-J., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2007b. Requirement of left-handed glycine residue for high stability of the Tk-subtilisin propeptide as revealed by mutational and crystallographic analyses. *J. Mol. Biol.* **374:** 1359-1373.

Rawlings, N.D., Morton, F.R., and Barrett, A.J. 2006. MEROPS: the peptidase database. *Nucleic Acids Res.* **34:** 270-272.

Schäfer, T., Borchert, T. W., Nielsen, V. S., Skagerlind, P., Gibson, K., Wenger, K., Hatzack, F., Nilsson, L. D., Salmon, S., Pedersen, S., Heldt-Hansen, H. P., Poulsen, P. B., Lund, H., Oxenbøll, K. M., Wu, G. F., Pedersen, H. H., and Xu, H. 2007. Industrial enzymes. *Adv. Biochem. Eng. Biotechnol.* **105:** 59-131.

Schägger, H. 2006. Tricine-SDS-PAGE. *Nat. Protoc.* **1:** 16-22.

Shinde, U. P., and Inouye, M. 1995. Folding pathway mediated by an intramolecular chaperone: characterization of the structural changes in pro-subtilisin E coincident with autoprocessing. *J. Mol. Biol.* **252:** 25-30.

Shinde, U. P., and Inouye, M. 1996. propeptide-mediated folding in subtilisin: the intramolecular chaperone concept. *Adv. Exp. Med. Biol.* **379:** 147-154.

Shinde, U. P., and Inouye, M. 2000. Intramolecular chaperones: polypeptide extensions that modulate protein folding. *Semin. Cell Dev. Biol.* **11:** 35-44.

Shinde, U. P., Liu, J. J., and Inouye, M. 1997. Protein memory through altered folding mediated by intramolecular chaperones. *Nature* **389:** 520-522.

Siezen, R. J., and Leunissen, J. A. 1997. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci.* **6:** 501-523.

Silen, J. L., and Agard, D. A. 1989. The α-lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*. *Nature* **341:** 462-464.

Smith, C. A., Toogood, H. S., Baker, H. M., Daniel, R. M., and Baker, E. N. 1999. Calcium-mediated thermostability in the subtilisin superfamily: the crystal structure of *Bacillus* Ak.1 protease at 1.8 Å resolution. *J. Mol. Biol.* **294:** 1027-1040.

Smith, S. M., and Gottesman, M. M. 1989. Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli*. *J. Biol. Chem.* **264:** 20487-20495.

Stahl, M.L., and Ferrari, E. 1984. Replacement of the *Bacillus subtilis* subtilisin structural gene with an in vitro-derived deletion mutation. *J. Bacteriol.* **158:** 411-418.

Strausberg, S. L., Alexander, P. A., Gallagher, D. T., Gilliland, G. L., Barnett, B. L., and Bryan, P. N. 1995. Directed evolution of a subtilisin with calcium-independent stability. *Biotechnology* **13:** 669-673.

Subbian, E., Yabuta, Y., and Shinde, U. P. 2005. Folding pathway mediated by an intramolecular chaperone: Intrinsically unstructured propeptide modulates stochastic activation of subtilisin. *J. Mol. Biol.* **347:** 367-383.

Syed, R., Wu, Z. P., Hogle, J. M., and Hilvert, D. 1993. Crystal structure of selenosubtilisin at 2.0-Å resolution. *Biochemistry* **32:** 6157-6164.

Takagi, H., and Takahashi, M. 2003. A new approach for alteration of protease functions: pro-sequence engineering. *Appl. Microbiol. Biotechnol.* **63:** 1-9.

Takeuchi, Y., Tanaka, S., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2009. Requirement of a unique Ca^{2+}-binding loop for folding of Tk-subtilisin from a hyperthermophilic aechaeon. *Biochemistry* **48:** 10637-10643.

Tanaka, S., Saita, K., Chon, H., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2006. Crystallization and preliminary X-ray diffraction study of active-site mutant of pro-Tk-subtilisin from a hyperthermophilic archaeon. *Acta Crystallogr. Sect. F* **62:** 902-905.

Tanaka, S., Saito, K., Chon, H., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2007a. Crystal structure of unautoprocessed precursor of subtilisin from a hyperthermophilic archaeon: evidence for Ca^{2+}-induced folding. *J. Biol. Chem.* **282:** 8246-8255.

Tanaka, S., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2007b. Four new crystal structures of Tk-subtilisin in unautoprocessed, autoprocessed and mature forms: insight into structural changes during maturation. *J. Mol. Biol.* **372:** 1055-1069.

Tanaka, S., Takeuchi, Y., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2008. Crystal structure of Tk-subtilisin folded without propeptide: requirement of propeptide for acceleration of folding. *FEBS Lett.* **582:** 3875-3878.

Tanaka, S., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. 2009. Identification of the interactions critical for propeptide-catalyzed folding of Tk-subtilisin. *J. Mol. Biol.* **394:** 306-319.

Terada, I., Kwon, S. T., Miyata, Y., Matsuzawa, H., and Ohta, T. 1990. Unique precursor structure of an extracellular protease, aqualysin I, with NH_2- and COOH-terminal pro-sequences and its processing in Escherichia coli. *J. Biol. Chem.* **265:** 6576-6581.

Ueda, K., Lipkind, G. M., Zhou, A., Zhu, X., Kuznetsov, A., Philipson, L., Gardner, P., Zhang, C., and Steiner, D. F. 2003. Mutational analysis of predicted interactions between the catalytic and P domains of prohormone convertase 3 (PC3/PC1). *Proc. Natl. Acad. Sci. U. S. A.* **100:** 5622-5627.

Vagin, A., and Teplyakov, A. 1997. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30:** 1022-1025.

Voet, D., Voet, J. G., and Pratt, C. 2006. Fundamentals of biochemistry: Life at the Molecular level. 2nd edition. John Wiley anCo.

Voordouw, G., Milo, C. and Roche, R. S. 1976. Role of bound calcium ions in thermostable, proteolytic enzymes. Separation of intrinsic and calcium ion contributions to the kinetic thermal stability. *Biochemistry* **15:** 3716-3724.

Wells, J. A., and Estell, D. A. 1988. Subtilisin-an enzyme designed to be engineered. *Trends Biochem. Sci.* **13:** 291-297.

Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., and Chen, E. Y. 1983. Cloning, sequencing, and secretion of Bacillus amyloliquefaciens subtilisin in *Bacillus subtilis*. *Nucleic Acids Res.* **11:** 7911-7925.

Winther, J. R., and Sorensen, P. 1991. Propeptide of carboxypeptidase Y provides a chaperone-like function as well as inhibition of the enzymatic activity. *Natl. Acad. Sci. USA* **88:** 9330-9334.

Wong, S. L. and Doi, R. H. 1986. Determination of the signal peptidase cleavage site in the preprosubtilisin of *Bacillus subtilis*. *J. Biol. Chem.* **261:** 10176-10181.

Wright, C. S., Alden, R.A., and Kraut, J. 1969. Structure of subtilisin BPN′ at 2.5 angstrom resolution. *Nature* **221:** 235-242.

Yabuta, Y., Subbian, E., Takagi, H., Shinde, U., and Inouye, M. 2002. Folding pathway mediated by an intramolecular chaperone: dissecting conformational changes coincident with autoprocessing and the role of Ca^{2+} in subtilisin maturation. *J. Biochem.* **131:** 31-37.

Yabuta, Y., Takagi, H., Inouye, M., and Shinde, U. 2001.Folding pathway mediated by an intramolecular chaperone: propeptide release modulates activation precision of pro-subtilisin. *J. Biol. Chem.* **276:** 4427-44434.

Yanagida, N., Uozumi, T., and Beppu, T. 1986. *Specific excretion of Serratia marcescens* protease through the outer membrane of *Escherichia coli*. *J. Bacteriol.* **166:** 937-944.

Zhou, A., Martin, S., Lipkind, G., LaMendola, J., and Steiner, D. F. 1998. Regulatory roles of the P domain of the subtilisin-like prohormone convertases. *J. Biol. Chem.* **273:** 11107-11114.

Zhu, X. L., Ohta, Y., Jordan, F., and Inouye, M. 1989. Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. *Nature* **339:** 483-484.

LIST OF PUBLICATIONS

1. **Foophow, T.,** Tanaka, S., Koga, Y., Takano, K., and Kanaya, S. 2010. Subtilisin-like serine protease from hyperthermophilic archaeon *Thermococcus kodakaraensis* with N- and C-terminal propeptides. *Prot. Eng. Des. Sel.* 23(5): 347-355.

2. **Foophow, T.,** Tanaka, S., Angkawidjaja, C., Koga, Y., Takano, K., and Kanaya, S. 2010. Crystal structure of a subtilisin homologue, Tk-SP, from *Thermococcus kodakaraensis*: requirement of a C-terminal β-jelly roll domain for hyperstability. *J. Mol. Biol.* 400(4): 865-877.

PATENT

1. **Foophow, T.,** Kanaya, S., Takano, K., and Koga, Y. (July 31, 2008) NOVEL PROTEASE AND USE THEREOF. (JP 2008-197467)

ACKNOWLEDGMENTS

This thesis would not have been possible without the support of many people.

I would like to convey thanks to the Japanese Ministry of Education, Culture, Sports, Science and Technology for providing the financial support during my study.

I wish to express my gratitude to my academic supervisor, Prof. Dr. Shigenori Kanaya who was abundantly helpful and offered invaluable assistance, support, and guidance. I greatly learn and benefit from his valuable ideas, suggestions, and criticisms.

Deepest gratitude is also due to the members of the supervisory committee, Assoc. Prof. Dr. Kazufumi Takano and Assist. Prof. Dr. Yuichi Koga without whose knowledge and assistance this study would not have been successful.

I greatly appreciate the helpful suggestions and comments from Prof. Dr. Satoshi Harashima and Prof. Dr. Kazuhito Fujiyama on the preparation of this Thesis.

I would like to show my gratitude to Dr. Marian Pulido, Dr. Shun-ichi Takana, and Assist. Prof. Dr. Clement Angkawidjaja for guidance of laboratory techniques and the significant contribution and assistance in completing this research study.

Special thanks to Ms. Reiko Matsumoto, the secretary of Kanaya laboratory, for all the support in helping me about my life and educational process in Japan.

I am grateful to all the students in the Molecular Biotechnology laboratory especially the international students for friendships, consulting, encouragement, and all the wonderful memories during my master and doctor study.

I would like to thanks to all the Thai friends and students in Osaka for sharing in the fun and relax times together.

Lastly, I wishes to express my love and gratitude to my beloved families; my parents, sister, brother and especially my husband and closed friend, Pipat, for their understanding, encouragement, and endless love, through the duration of my studies.