Crystal Structure of Unautoprocessed Precursor of Subtilisin from a Hyperthermophilic Archaeon

**EVIDENCE FOR Ca\(^{2+}\)-INDUCED FOLDING**

Shun-ichi Tanaka, Kenji Saito, Hyongi Chon, Hiroyoshi Matsumura, Yuichi Koga, Kazufumi Takano, and Shigenori Kanaya

From the Departments of Material and Life Science and Applied Chemistry, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan, and CREST (Sosho Project), JST, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

The crystal structure of an active site mutant of pro-Tk-subtilisin (pro-S324A) from the hyperthermophilic archaeon *Thermococcus kodakaraensis* was determined at 2.3 Å resolution. The overall structure of this protein is similar to those of bacterial subtilisin-propeptide complexes, except that the peptide bond linking the propeptide and mature domain contacts with the active site, and the mature domain contains six Ca\(^{2+}\) binding sites. The Ca-1 site is conserved in bacterial subtilisins but is formed prior to autoprocessing, unlike the corresponding sites of bacterial subtilisins. All other Ca\(^{2+}\)-binding sites are unique in the pro-S324A structure and are located at the surface loops. Four of them apparently contribute to the stability of the central alpha substructure of the mature domain. The CD spectra, 1-anilino-8-naphthalenesulfonic acid fluorescence spectra, and sensitivities to chymotryptic digestion of this protein indicate that the conformation of pro-S324A is changed from an unstable molten globule-like structure to a stable native one upon Ca\(^{2+}\) binding. Another active site mutant, pro-S324C, was shown to be autoprocessed to form a propeptide-mature domain complex in the presence of Ca\(^{2+}\). The CD spectra of this protein indicate that the structure of pro-S324C is changed upon Ca\(^{2+}\) binding like pro-S324A but is not seriously changed upon subsequent autoprocessing. These results suggest that the maturation process of Tk-subtilisin is different from that of bacterial subtilisins in terms of the requirement of Ca\(^{2+}\) for folding of the mature domain and completion of the folding process prior to autoprocessing.

Subtilisin is a serine protease with broad substrate specificity. It is a member of the subtilase family (1) and is widely distributed in various organisms, including bacteria and archaea. These subtilisins are synthesized in a precursor form, in which a signal peptide and a propeptide are attached to the N terminus of the mature domain (1, 2). The mechanism by which subtilisins are folded and matured has been most extensively studied for bacterial subtilisins, such as subtilisin E, subtilisin BPN’, and subtilisin Carlsberg. 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 (3). 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 (4–6). 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 (7–9). The mature domains alone are not an active form but an inactive form with a molten globule-like structure in the absence of propeptides (7, 10). Dependence on the propeptide for maturation of its cognate mature domain has been reported not only for other members of the subtilase family (11–13) but also for other proteases (14–19).

The crystal structures of bacterial subtilisins are available not only for the active mature form (20, 21) but also for the inactive complex between the propeptide and mature domain (22, 23). Comparison of these structures indicates that the folding process of the mature domain is almost fully completed prior to degradation of the propeptide. These structures also indicate that bacterial subtilisins have two Ca\(^{2+}\) binding sites, site 1 (site A) and site 2 (site B), both of which are located far from the active site. Sites 1 and 2 have high and low binding affinity, respectively. The enzyme is greatly destabilized with respect to both thermal denaturation and autodegradation upon removal of Ca\(^{2+}\) from site 1 (24–26). Prosubtilisin E is folded and autoprocessed in the absence of Ca\(^{2+}\) (27). Deletion of the loop forming site 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 (25, 28, 29). 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 (30), proteinase K (31), and Ak1 protease (32) have four sites, and sphericase has five sites (33), which are located far from the active site but do not necessarily correlate to site 1 or site 2 of bacterial subtilisins. In addition,
Various subtilases, such as sphericase (33), cell envelope protease (34), and psychrophilic subtilisin (35), have been reported to exhibit Ca\(^{2+}\)-dependent activity. These results suggest that the role of the Ca\(^{2+}\) ions varies for different subtilases.

Tk-subtilisin\(^2\) from the hyperthermophilic archaean Thermococcus kodakaraensis consists of signal peptide (Met\(^{24}\)–Ala\(^{-1}\)), propeptide (Gly\(^{3}–\)Leu\(^{69}\)), and mature domain (Gly\(^{70}–\)Gly\(^{398}\)) (2). The mature domain of Tk-subtilisin contains three major insertion sequences as compared with those of bacterial subtilisins. Other than that, the sequence is similar to those of bacterial subtilisins both in size and amino acid sequence (amino acid sequence identities of 45–46%). Like bacterial subtilisins, Tk-subtilisin is matured from pro-Tk-subtilisin upon autoprocessing and degradation of the propeptide. In this process, the propeptide is autoprocessed at first to produce an inactive complex between the propeptide and mature domain. Then the propeptide, which is a potent inhibitor and, at the same time, a good substrate of the mature domain, is degraded by the mature domain to produce active enzyme. However, unlike bacterial subtilisins, Tk-subtilisin requires Ca\(^{2+}\) for activity and is therefore not matured at all even at high temperatures in the absence of Ca\(^{2+}\). In addition, the propeptide of pro-Tk-subtilisin is not required for folding of the mature domain, although it is involved in folding of the mature domain in an auxiliary manner. These results suggest that the maturation process of Tk-subtilisin is different from that proposed for bacteria.

To understand this unique maturation process, it is necessary to determine the crystal structure of pro-Tk-subtilisin in a Ca\(^{2+}\)-bound form. However, pro-Tk-subtilisin is not fully stable in the presence of Ca\(^{2+}\) even at 4 °C, especially when its concentration is high. In this condition, pro-Tk-subtilisin is gradually converted to an active mature form, which is finally self-degraded.

We have recently succeeded in crystallizing the active site mutant of pro-Tk-subtilisin, pro-S324A, in a Ca\(^{2+}\)-bound form and performed preliminary x-ray diffraction studies (36). In this report, we determined the crystal structure of this protein, which represents the unautoprocessed precursor structure of subtilisin, at 2.3 Å resolution. We also showed that the conformation of pro-S324A is considerably changed from a molten globule-like structure to a native one upon Ca\(^{2+}\) binding. Furthermore, we constructed another active site mutant pro-S324C and characterized it biochemically. Based on these results, we discuss a unique maturation process of Tk-subtilisin.

**EXPERIMENTAL PROCEDURES**

**Construction of Plasmids**—The gene encoding pro-S324C was amplified by PCR using overlap extension method (37). The mutagenic primers were designed such that the codon for Ser\(^{324}\) (AGC) is changed to TGC for Cys. The pET25b derivative for overproduction of pro-S324A (36) was used as a template. The pET25b derivatives for overproduction of pro-S324C and mat-S324A were constructed as previously described for pro-Tk-subtilisin and mat-Tk-subtilisin, respectively (2). All DNA oligomers for PCR were synthesized by Hokkaido System Science (Sapporo, Japan). PCR was performed in 25 cycles using a thermal cycler (Gene Amp PCR System 2400; PerkinElmer Life Sciences) and KOD DNA polymerase (Toyobo). The DNA sequences of the genes encoding pro-S324C and mat-S324A were confirmed by ABI Prism 310 DNA sequencer (PerkinElmer Life Sciences).

**Overproduction and Purification of the Mutant Proteins**—Overproduction of pro-S324A, pro-S324C, and mat-S324A in Escherichia coli BL21-codonPlus(DE3) (Stratagene) was carried out as described previously for pro-Tk-subtilisin (38). The cells were collected by centrifugation, suspended in 20 mM Tris-HCl (pH 9.0), disrupted by sonication on ice, and centrifuged at 30,000 × g for 30 min at 4 °C. The pellet was dissolved in 20 mM Tris-HCl (pH 9.0) containing 8 M urea and 5 mM EDTA and applied to a HiTrap Q HP column (GE Healthcare) equilibrated with the same buffer. The protein was eluted from the column at NaCl concentration of ~0.1 M for pro-S324A and pro-S324C and ~0.2 M for mat-S324A by linearly increasing the NaCl concentration from 0 to 0.3 M. The fractions containing the protein were collected and refolded as described below. For refolding of pro-S324A, pro-S324C, and mat-S324A in a Ca\(^{2+}\)-free form, these proteins (25 mM) were dialyzed against 20 mM Tris-HCl (pH 7.0) at 4 °C. For refolding of pro-S324A and pro-S324C in a Ca\(^{2+}\)-bound form, these proteins (25 mM) were dialyzed against 20 mM Tris-HCl (pH 7.0) containing 10 mM CaCl\(_2\) and 1 mM DTT at 4 °C for 5 days, incubated at 80 °C for 30 min, and centrifuged at 30,000 × g for 30 min at 4 °C. The resultant supernatant was loaded onto a Sephacryl S-200HR column (GE Healthcare) equilibrated with 20 mM Tris-HCl (pH 7.0) containing 10 mM CaCl\(_2\) and 50 mM NaCl. The fractions containing the protein were collected and dialyzed against 20 mM Tris-HCl (pH 7.0) containing 10 mM CaCl\(_2\).

Pro-Tk-subtilisin, mat-Tk-subtilisin, and Tk-propeptide were overproduced and purified as described previously (2). mat-Tk-subtilisin\(^4\), which represents a Ca\(^{2+}\)-bound active form of mat-Tk-subtilisin, was prepared by incubating pro-Tk-subtilisin at 80 °C for 90 min in the presence of Ca\(^{2+}\) as described previously (2) and used without further purification.

The protein concentration was determined from UV absorption using a cell with an optical path length of 1 cm and \(A_{260}^\text{value}\) for 0.1% (1 mg/ml) solution of 1.25 for pro-Tk-subtilisin, pro-S324A, and pro-S324C; 1.47 for mat-Tk-subtilisin and mat-S324A; and 0.21 for Tk-propeptide. These values were calculated by using absorption coefficients of 1.526 M\(^{-1}\) cm\(^{-1}\) for tyrosine and 5.225 M\(^{-1}\) cm\(^{-1}\) for tryptophan at 280 nm (39). The purity of the protein was confirmed by SDS-PAGE (40), followed by staining with Coomassie Brilliant Blue (CBB). For N-terminal amino acid sequencing, the protein visualized with CBB staining was transferred to a nitrocellulose membrane (blotting). The N-terminal amino acid sequence was determined using a Procise 491 automated protein sequencer (PerkinElmer Life Sciences).

\(^2\)The abbreviations used are: Tk-subtilisin, subtilisin from Thermococcus kodakaraensis; pro-Tk-subtilisin, Tk-subtilisin in a pro-form (Gly\(^{3}–\)Gly\(^{69}\)); propeptide, propeptide of Tk-subtilisin (Gly\(^{3}–\)Leu\(^{69}\)); mat-Tk-subtilisin, Tk-subtilisin in a mature form (Gly\(^{70}–\)Gly\(^{398}\)); mat-Tk-subtilisin\(^*\), a Ca\(^{2+}\)-bound active form of mat-Tk-subtilisin; pro-S324A (S324C), pro-Tk-subtilisin with the Ser\(^{224}\) → Ala (Ser\(^{224}\) → Cys) mutation; mat-S324A(S324C), mat-Tk-subtilisin with the Ser\(^{224}\) → Ala (Ser\(^{224}\) → Cys) mutation; mat-S324C\(^*\), mat-Tk-subtilisin\(^*\) with the Ser\(^{224}\) → Cys mutation; CBB, Coomassie Brilliant Blue; DTT, dithiothreitol; ANS, 1-anilino-8-naphthalenesulfonic acid.
Crystal Structure of Pro-Tk-subtilisin

Enzymatic Activity—The enzymatic activity was determined by using N-succinyl-AAFP-p-nitroanilide (Sigma) as a substrate at 80 °C. The reaction mixture (100 µl) contained 50 mM CAPS-NaOH (pH 9.5), 5 mM CaCl₂, and 2 mM N-succinyl-AAFP-p-nitroanilide. The amount of p-nitroaniline released from the substrate was determined from the absorption at 410 nm with an absorption coefficient of 8,900 M⁻¹ cm⁻¹ by automatic UV spectrophotometer (Beckman model DU640). Each unit of enzymatic activity was defined as the amount of the enzyme that produced 1 µmol of p-nitroaniline/min. The specific activity was defined as the enzymatic activity/mg of protein.

Circular Dichroism Spectroscopy—The far-UV (200–260 nm) and near-UV (250–320 nm) CD spectra of the protein were measured by a J-725 automatic spectropolarimeter (Japan Spectroscopic Co.) at 20 °C. The buffer was 20 mM Tris-HCl (pH 7.0) containing 10 mM CaCl₂ for protein in a Ca²⁺-bound form and 20 mM Tris-HCl (pH 7.0) for protein in a Ca²⁺-free form. 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 the unit of degrees cm² dmol⁻¹, was calculated by using an average amino acid molecular weight of 110.

1-Anilino-8-naphthalenesulfonic Acid (ANS) Fluorescence Spectroscopy—Binding of ANS (Sigma) to the protein was analyzed by measuring the fluorescence of ANS at 20 °C. The protein (1 µM) and ANS (50 µM) were dissolved in 20 mM Tris-HCl (pH 7.0) either in the presence or absence of 10 mM CaCl₂. The protein was incubated at 20 °C for 4 h prior to the measurement of ANS fluorescence. The excitation wavelength was 380 nm, and the emission was monitored from 400 to 600 nm. The spectrum obtained in the absence of the protein was used as blank.

Analysis of Folding Efficiency—mat-Tk-subtilisin or mat-S324A (30 µM) was first unfolded by incubating it for 1 h at 25 °C in 20 mM Tris-HCl (pH 7.0) containing 6 M guanidine HCl. The protein was refolded by 100-fold dilution with 50 mM Tris-HCl (pH 7.0) containing 1 mM DTT, either in the presence or absence of 10 mM CaCl₂ and either in the presence or absence of 0.3 µM Tk-propeptide at 4 °C. Then refolded mat-Tk-subtilisin was incubated at 80 °C for 30 min, and refolded mat-S324A was digested with chymotrypsin at 30 °C for 1 h at an enzyme/substrate ratio of 1:100 (w/w) after incubation at 4 °C for 1 day. Finally, 112 µl of trichloroacetic acid was added to 1 ml of the reaction mixture to precipitate the proteins. The resultant precipitates were washed with 70% acetone and analyzed by 15% SDS-PAGE.

Crystallization of Pro-S324A in Ca²⁺-bound Form—Pro-S324A purified in a Ca²⁺-bound form was dialyzed against 10 mM Tris-HCl (pH 7.0), concentrated to 10 mg/ml using an ultrafiltration system Centricron (Millipore Corp.), and crystallized using 4 M sodium formate as described previously (36). X-ray diffraction studies on the resultant crystals indicated that these belonged to the space group I222 (unit cell parameters a = 92.69, b = 121.78, c = 77.53 Å) and contained single protein molecules per asymmetric unit.
The catalytic serine residue (Fig. 2) binds to the N termini of these two four-stranded antiparallel Ca\(^{2+}\) and colored red active site serine residue, are indicated by respectively. For the subtilisin E-propeptide complex, the entire structure, including two Ca\(^{2+}\) and orange, respectively. Two active site residues (Asp\(^{115}\) and His\(^{153}\)) and Ala\(^{324}\), which is substituted for the major insertion sequences (Gly\(^{70}\)-Pro\(^{82}\), Pro\(^{207}\)-Asp\(^{226}\), and Gly\(^{346}\)-Ser\(^{358}\)) are colored blue, magenta, and orange, respectively. Two active site residues (Asp\(^{115}\) and His\(^{153}\)) and Ala\(^{324}\), which is substituted for the catalytic serine residue (Fig. 2A). Leu\(^{69}\) and Gly\(^{70}\) occupy subsites S1 and S1’ of the mature domain, respectively. Leu\(^{69}\) oxygen forms a hydrogen bond with Ala\(^{324}\) nitrogen. The \(\beta_{5p}\)-strand of the propeptide domain interacts with \(\beta_{4m}\)-strand of the mature domain to form an antiparallel \(\beta\)-sheet in the vicinity of the active site. A similar structure has been reported for the linker peptide of prokumamolisin, which is a member of a different subtilase family (46).

The mature domain consists of a seven-stranded parallel \(\beta\)-sheet (\(\beta_{1m}\)-\(\beta_{3m}\) and \(\beta_{5m}\)-\(\beta_{8m}\), \(\beta_{4m}\)-strand, an anti-parallel \(\beta\)-sheet (\(\beta_{9m}\) and \(\beta_{10m}\)), and 10 \(\alpha\)-helices (\(\alpha_{1m}\)-\(\alpha_{10m}\)). Of them, the seven-stranded parallel \(\beta\)-sheet and two \(\alpha\)-helices (\(\alpha_{5m}\) and \(\alpha_{8m}\)) form a core structure. As in all members of the subtilase family, three active site residues, Asp\(^{115}\), His\(^{153}\), and Ser\(^{324}\), form a catalytic triad in Tk-subtilisin. In the present structure, Ser\(^{324}\) is replaced by Ala. Two hydrogen bonds are formed between His\(^{153}\) N\(^\beta\) and Asp\(^{115}\) O\(^\phi\) and between His\(^{153}\) N\(^\beta\) and Asp\(^{115}\) O\(^\phi\). These two hydrogen bonds are also observed in the subtilisin E-propeptide complex structure (22). The pro-S324A structure has one disulfide bond between Cys\(^{132}\) and Cys\(^{147}\), which are the only cysteine residues in the pro-S324A sequence. A disulfide bond is formed at the same position in subtilisin S39 and subtilisin S41 from psychrophilic Bacillus (35) and spherocase from mesophilic Bacillus (33) but is not formed at any position in bacterial subtilisins. It remains to be determined whether this disulfide bond contributes to the high stability of Tk-subtilisin.

Comparison of the amino acid sequence of pro-Tk-subtilisin with those of bacterial subtilisins based on the pro-S324A structure indicates that the pro-Tk-subtilisin sequence contains three major insertion sequences. They are Gly\(^{70}\)-Pro\(^{82}\), Pro\(^{207}\)-Asp\(^{226}\), and Gly\(^{346}\)-Ser\(^{358}\). These insertion sequences form a loop at the protein surface, and two of them (the first and second one) are responsible for the formation of Ca\(^{2+}\)-binding sites (Fig. 1A). The overall structure of pro-S324A in a Ca\(^{2+}\)-bound form is compared with that of subtilisin E-propeptide complex (22) in Fig. 1B. The former and latter structures represent those before and after cleavage of the propeptide, respectively. Comparison of these structures indicates that the main chain fold of pro-S324A is similar to that of the subtilisin E-propeptide complex, except that the N terminus of the mature domain moves away from the active site in the subtilisin E-propeptide complex structure, and the structure of the mature domain of pro-S324A contains three additional loops formed by the insertion sequences. These two structures could be superimposed with root mean square deviation values of 1.8
Å for 65 Ca atoms from Thr^5 to Leu^69 in the propeptide domain and 0.6 Å for 264 Ca atoms from Ala^83 to Gln^394, except Ser^105, Val^108, Ala^127-Asn^128, Gly^138-Arg^145, Pro^207-Asp^226, Gln^343-Lys^347, Thr^353-Asn^360, and Thr^376-Trp^378, in the mature domain. However, this structure was considerably different from that of the subtilisin E-propeptide complex in the positions of Ca^2+ binding sites as described below.

Ca^2+ Binding Sites—Six Ca^2+ binding sites were identified in the mature domain of pro-S324A. This number, which was the highest among those reported for other subtilases to date, was consistent with that determined by atomic absorption spectrometry (36). Only one of them (the Ca-1 site) was conserved in the structures of bacterial subtilisins, whereas others were unique in the pro-S324A structure. Four of these unique Ca^2+ binding sites were located at the surface loop, which was formed by the second insertion sequence of Tk-subtilisin. The structural details for these Ca^2+ binding sites are presented in Table 2.

The 2F_o - F_c map of pro-S324A around the Ca-1 site is shown in Fig. 2B as a representative. This site correlates to the high affinity site (site 1 or site A) of bacterial subtilisins. In this site, the Ca^2+ ion is heptacoordinated with six amino acid residues (Fig. 3A). All of these residues, except Ile^168, which is replaced by Thr in one bacterial subtilisin sequence, are fully conserved in various bacterial subtilisin sequences. However, it has been proposed for bacterial subtilisins that one of these residues, Gln^2, cannot make a direct coordination with the Ca^2+ ion in an unautoprocessed precursor form, because of the large distance between the high affinity site and active site (22, 47). This means that the high affinity site of bacterial subtilisins is formed only when the peptide bond between the propeptide and mature domain is cleaved and the N-terminal region of the mature domain moves away from the active site. Hence, the unautoprocessed form of bacterial subtilisins is much less stable than the autoprocessed complex (6, 27). In contrast, the corresponding residue of Tk-subtilisin, Gln^84, can coordinate with Ca-1 in an unautoprocessed precursor form. This is clearly due to the presence of the loop formed by the insertion sequence (Gly^70-Pro^83), which also exists as an N-terminal extension of the mature domain after autoprocessing. These results suggest that, unlike bacterial subtilisins, the conformation of pro-Tk-subtilisin is not seriously changed upon autoprocessing.

The other five Ca^2+ binding sites are unique to Tk-subtilisin and do not correlate to any site so far identified in various subtilases. Of them, four sites (Ca-2 to Ca-5) are located at a long external loop between α6m-helix and β5m-strand (Gly^206-Glu^229) (Fig. 1). This Ca^2+ binding loop is rich in Asp and is mostly formed by a unique insertion sequence of

FIGURE 2. A stereo view of electron density around the active site (A) and Ca-1 site (B) of pro-S324A. In A, the 2F_o - F_c map contoured at the 1.5σ level is shown. In B, the 2F_o - F_c maps contoured at the 2.0σ and 5.0σ levels are shown in blue and magenta, respectively. The active site residues, the nitrogen and oxygen atoms, and the Ca^2+ ions are indicated as in Fig. 1.
The Ca-6 site is located at a surface loop between the α9m- and α10m-helices. In this site, the Ca\(^{2+}\) ion is hexacoordinated with five residues and one water molecule (Fig. 3D). Bacterial subtilisins contain a similar surface loop at the same position. However, the Ca\(^{2+}\) ion cannot bind to this loop. Insertion of the three amino acid residues (Thr\(^{376}\)–Trp\(^{378}\)) widens this loop, and the integrations of two aspartic acid residues (Asp\(^{372}\) and Asp\(^{379}\)) into this loop, which provide ligands for Ca\(^{2+}\) binding, may be responsible for the formation of the Ca-6 site in pro-Tk-subtilisin. Binding of the Ca\(^{2+}\) ion to this site may be responsible for tight packing of the C-terminal α9m- and α10m-helices. Because the α10m-helix interacts with the N-terminal loop region between the α1m- and α2m-helices, this site may be important for both thermal and proteolytic stability of Tk-subtilisin.

**Ca\(^{2+}\)-induced Folding**—The crystallographic studies indicate that the structure of pro-S324A in a Ca\(^{2+}\)-bound form represents a native structure of this protein. To examine whether the conformation of pro-Tk-subtilisin is changed upon Ca\(^{2+}\) binding, the far-UV and near-UV CD spectra of pro-S324A in a Ca\(^{2+}\)-free and Ca\(^{2+}\)-bound form were measured. The far-UV and near-UV CD spectra of pro-S324A in a Ca\(^{2+}\)-free form were nearly identical to those of pro-Tk-subtilisin in a Ca\(^{2+}\)-free form (data not shown), suggesting that the conformation of pro-Tk-subtilisin is not seriously changed by the mutation. However, the far-UV CD spectrum of pro-S324A in a Ca\(^{2+}\)-bound form was considerably different from that of pro-S324A in a Ca\(^{2+}\)-free form (Fig. 4A). The far-UV CD spectrum of pro-S324A in a Ca\(^{2+}\)-free form gave a broad trough with a single minimum around at 208 nm, whereas that of pro-S324A in a Ca\(^{2+}\)-bound form gave a trough with a double minimum around at 210 and 222 nm. Likewise, the near-UV CD spectrum of pro-S324A in a Ca\(^{2+}\)-bound form was considerably different from that of pro-S324A in a Ca\(^{2+}\)-free form (Fig. 4B). These results suggest that the conformation of pro-S324A is greatly changed upon Ca\(^{2+}\) binding, such that the helical content of the protein increases and the environment of the aromatic residues is altered. Pro-S324A changes its conformation in a reversible manner, because the far-UV CD spectrum of pro-S324A in a Ca\(^{2+}\)-free form was changed to that in a Ca\(^{2+}\)-bound form in the presence of 10 mM CaCl\(_2\), and the far-UV CD spectrum of pro-S324A in a Ca\(^{2+}\)-bound form was changed to that in a Ca\(^{2+}\)-free form upon treatment with 50 mM EDTA.

**ANS Fluorescence Spectroscopy**—To obtain more information on the Ca\(^{2+}\)-induced conformational change of pro-Tk-subtilisin, pro-S324A was analyzed for its binding ability to ANS. ANS is known to bind more effectively to the protein in a partially folded state than to that in a fully folded or unfolded state (48). As shown in Fig. 5, ANS fluorescence of pro-S324A in a Ca\(^{2+}\)-bound form was greatly reduced as compared with that of pro-S324A in a Ca\(^{2+}\)-free form. These results, as well as the CD spectra of pro-S324A, strongly suggest that this protein is folded into a molten globule-like structure in the absence of Ca\(^{2+}\) but is folded into a native structure in the presence of Ca\(^{2+}\).

**Protein Stability**—To examine whether Ca\(^{2+}\) is required not only for folding but also for stabilization of pro-Tk-subtilisin, the susceptibility of pro-S324A to chymotryptic digestion was

---

**TABLE 2**

| Binding site | Residue | Distance (Å) | B factor (Å\(^2\)) |
|-------------|---------|-------------|-------------------|
| Ca-1        | Glu\(^{34}\) O\(^{2+}\) | 2.5          | 22.6              |
|             | Asp\(^{24}\) O\(^{2+}\) | 2.5          | 20.5              |
|             | Asp\(^{34}\) O\(^{2+}\) | 2.6          | 23.1              |
|             | Leu\(^{14}\) O         | 2.3          | 23.1              |
|             | Asn\(^{14}\) O\(^{2+}\) | 2.5          | 26.3              |
|             | Ile\(^{14}\) O         | 2.3          | 26.5              |
|             | Val\(^{34}\) O         | 2.2          | 21.7              |
| Ca-2        | Leu\(^{55}\) O         | 2.4          | 21.4              |
|             | Asp\(^{55}\) O\(^{2+}\) | 2.4          | 27.6              |
|             | Val\(^{55}\) O         | 2.1          | 31.0              |
|             | Asp\(^{65}\) O         | 2.2          | 24.5              |
|             | Water 261 O             | 2.4          | 21.8              |
|             | Water 269 O             | 2.1          | 18.8              |
| Ca-3        | Asp\(^{10}\) O\(^{2+}\) | 2.4          | 32.6              |
|             | Asp\(^{14}\) O\(^{2+}\) | 2.4          | 39.4              |
|             | Asp\(^{216}\) O\(^{2+}\) | 2.4         | 28.7              |
|             | Ile\(^{34}\) O         | 2.2          | 28.3              |
|             | Asp\(^{222}\) O         | 2.3          | 23.9              |
|             | Asp\(^{222}\) O         | 2.2          | 25.0              |
| Ca-4        | Asp\(^{10}\) O\(^{2+}\) | 2.4          | 39.3              |
|             | Asp\(^{14}\) O\(^{2+}\) | 2.6          | 30.4              |
|             | Asp\(^{222}\) O         | 2.6          | 23.9              |
|             | Asp\(^{222}\) O         | 2.4          | 24.7              |
|             | Asp\(^{224}\) O         | 2.4          | 30.2              |
|             | Water 52 O              | 2.7          | 21.7              |
| Ca-5        | Val\(^{10}\) O\(^{2+}\) | 2.3          | 22.3              |
|             | Glu\(^{10}\) O\(^{2+}\) | 2.2          | 20.1              |
|             | Ala\(^{17}\) O          | 2.3          | 24.5              |
|             | Glu\(^{24}\) O\(^{2+}\) | 2.7          | 28.1              |
|             | Glu\(^{24}\) O\(^{2+}\) | 2.2          | 29.9              |
|             | Water 1 O               | 2.6          | 25.0              |
|             | Water 86 O              | 2.4          | 21.3              |
| Ca-6        | Asp\(^{372}\) O\(^{2+}\) | 2.2          | 32.1              |
|             | Leu\(^{372}\) O         | 2.3          | 30.5              |
|             | Pro\(^{372}\) O         | 2.4          | 37.7              |
|             | Gly\(^{372}\) O         | 2.3          | 32.4              |
|             | Asp\(^{379}\) O         | 2.5          | 31.2              |
|             | Water 17 O              | 2.5          | 24.3              |

Tk-subtilisin. The Ca\(^{2+}\) ions are hexacoordinated with four residues and two water molecules in the Ca-2 site, six residues in the Ca-3 site, and four residues and one water molecule in the Ca-4 site and heptacoordinated with four residues and two water molecules in the Ca-5 site (Fig. 3B). The Ca\(^{2+}\) ions bound to the Ca-3 and Ca-4 sites are located very close to each other, at a distance of 4.2 Å, and share the three aspartic acid residues (Asp\(^{216}\), Asp\(^{216}\), and Asp\(^{222}\)) for coordination.

The Ca\(^{2+}\)-binding loop apparently contributes to the stabilization of an αβα substructure, which consists of α6m-helix, β5m-strand, and α7m-helix, because this loop interacts with the α7m-helix (Fig. 3C). In addition, this Ca\(^{2+}\)-binding loop probably stabilizes the central seven-stranded parallel β-sheet and thereby the core structure, because the Ca-5 site is located between the β5m and β1m-strands, both of which are constituents of the central seven-stranded parallel β-sheet. Bacterial subtilisins also have αβα substructures (Fig. 1B). However, these structures greatly differ from that of Tk-subtilisin, because they lack a Ca\(^{2+}\)-binding loop. As a result, they do not contain any Ca\(^{2+}\) binding site. It has been proposed that the stabilization of the αβα substructure by the propeptide domain is crucial for subtilisin folding (47). However, in the case of Tk-subtilisin, the αβα substructure is stabilized by the Ca\(^{2+}\) ions instead of the propeptide. This may be the reason why the mature domain of Tk-subtilisin requires Ca\(^{2+}\) instead of the propeptide for folding.
analyzed. When this protein was incubated with chymotrypsin at 30°C for 1 h in 20 mM Tris-Cl (pH 7.0) at an enzyme/substrate ratio of 1:10–1:10,000 (w/w), pro-S324A in a Ca²⁺-free form was almost fully degraded even at an enzyme/substrate ratio of 1:1000, whereas pro-S324A in a Ca²⁺-bound form was not degraded at all even at an enzyme/substrate ratio of 1:10 (Fig. 6). These results suggest that pro-S324A is greatly stabilized upon Ca²⁺ binding. It has been reported that Ca²⁺ is required for stability of bacterial subtilisins (24). However, this stabilization effect has only been analyzed for the mature form. The conformation of the unautoprocessed precursor of bacterial subtilisins has been reported to be significantly different from that of the autoprocessed complex, because the unautoprocessed precursor of bacterial subtilisins exists in a molten globule-like state with a high content of solvent-accessible hydrophobic surface area and is a proteolytically unstable conformer (10). Thus, the maturation process of pro-Tk-subtilisin is different from that of bacterial prosubtilisins, because Ca²⁺, instead of the propeptide, induces folding of the mature domain into a highly stable functional structure.

Role of Propeptide—When the mature domain of Tk-subtilisin (mat-Tk-subtilisin) was refolded by dilution with the buffer containing Ca²⁺ in the presence or absence of Tk-propeptide, followed by the incubation at 80°C for 30 min, the refolding yield was ~1 and 100% in the absence and presence of an equal molar concentration of Tk-propeptide, respectively (Fig. 7, lanes 2 and 3). At the lower Tk-propeptide concentrations, this yield increased as the concentration of Tk-propeptide increased (data not shown). The specific activity of the refolded protein was comparable with that of mat-Tk-subtilisin* regardless of the refolding yield. These results indicate that mat-Tk-subtilisin is converted to mat-Tk-subtilisin* with a native structure.

FIGURE 3. Ca²⁺ binding sites. The structures of the Ca-1 (A), Ca-2 to Ca-5 (B), and Ca-6 (D) sites are shown. The Ca²⁺ ions and water molecules are represented by spheres. The residues that bind to the Ca²⁺ ions are labeled. In C, a stereo view of the central αβα substructure consisting of the α6m-helix, β5m-strand, and α7m-helix and the β1m-strand are shown in a ribbon drawing. One of the active site residues, Asp₁¹⁵, is indicated by a stick model.
protein with a nonnative structure was removed by degradation with chymotrypsin instead of mat-Tk-subtilisin. The chymotrypsin digestion was carried out in a condition in which pro-S324A in a Ca$^{2+}$-bound form (Fig. 6) and mat-Tk-subtilisin (Fig. 7, lane 5) are not significantly degraded, whereas pro-S324A in a Ca$^{2+}$-free form is fully degraded (Fig. 6). In this condition, mat-S324A refolded in the absence of Ca$^{2+}$ was fully degraded by chymotrypsin, regardless of whether it was refolded in the presence of Tk-propeptide (Fig. 7, lanes 7 and 9). This result indicates that mat-S324A is not refolded into a native structure in the absence of Ca$^{2+}$. In contrast, mat-S324A was refolded into a chymotrypsin-resistant form, which represents a Ca$^{2+}$-bound form with a native structure, with a yield of 20% in the presence of Ca$^{2+}$, even in the absence of Tk-propeptide (Fig. 7, lane 4). This yield increased to nearly 100% when an equal molar concentration of Tk-propeptide was added in trans (Fig. 7, lane 10). These results suggest that the propeptide of pro-Tk-subtilisin is required to increase the refolding yield of the mature domain not only by preventing degradation of a preactivated form of the protein but also by assisting the folding of the mature domain as an intramolecular chaperon.

Conformation of Autoprocessed Protein—It has been reported for a bacterial subtilisin that the mutation of the active site serine residue to Cys greatly reduces the enzymatic activity, such that the protein is autoprocessed but the propeptide is not further degraded by the mature domain (49). It has also been reported that the conformation of prosubtilisin is significantly changed upon autoprocessing (10). To examine whether this mutation blocks the maturation process of pro-Tk-subtilisin prior to the degradation of the propeptide and the conformation of pro-Tk-subtilisin is changed upon autoprocessing, the mutant protein pro-S324C was constructed.

Upon overproduction, pro-S324C accumulated in an insoluble form in the _E. coli_ cells, solubilized by 8 M urea, purified, and refolded by removing urea in the absence or presence of Ca$^{2+}$, like pro-S324A. The production level of pro-S324C was nearly identical to that of the wild-type protein, and 10 mg of the purified protein was obtained from 1 liter of culture. Pro-S324C refolded in the presence of Ca$^{2+}$ was eluted from the gel filtration column as a single peak at the position where pro-S324A was eluted. Nevertheless, it gave two bands on SDS-PAGE with the molecular masses of 35 and 10 kDa (Fig. 8, lane 2). These proteins were identified as mat-S324C and Tk-propeptide, respectively, by N-terminal amino acid sequence determination. In contrast, pro-S324C refolded in the absence of Ca$^{2+}$ gave a single band on SDS-PAGE at the position where pro-S324A is migrated (Fig. 8, lane 1). Neither pro-S324C refolded in the presence of Ca$^{2+}$ nor that refolded in the absence of Ca$^{2+}$ exhibited enzymatic activity. These results indicate that pro-S324C is autoprocessed into the propeptide and mature domain, but the propeptide continues to bind tightly to the

Upon overproduction, pro-S324A accumulated in an insoluble form in the _E. coli_ cells, solubilized by 8 M urea, purified, and refolded by removing urea in the absence or presence of Ca$^{2+}$, like pro-S324A. The production level of pro-S324C was nearly identical to that of the wild-type protein, and 10 mg of the purified protein was obtained from 1 liter of culture. Pro-S324C refolded in the presence of Ca$^{2+}$ was eluted from the gel filtration column as a single peak at the position where pro-S324A was eluted. Nevertheless, it gave two bands on SDS-PAGE with the molecular masses of 35 and 10 kDa (Fig. 8, lane 2). These proteins were identified as mat-S324C and Tk-propeptide, respectively, by N-terminal amino acid sequence determination. In contrast, pro-S324C refolded in the absence of Ca$^{2+}$ gave a single band on SDS-PAGE at the position where pro-S324A is migrated (Fig. 8, lane 1). Neither pro-S324C refolded in the presence of Ca$^{2+}$ nor that refolded in the absence of Ca$^{2+}$ exhibited enzymatic activity. These results indicate that pro-S324C is autoprocessed into the propeptide and mature domain, but the propeptide continues to bind tightly to the

FIGURE 4. CD spectra. Shown are far-UV (left) and near-UV (right) CD spectra of pro-S324A in a Ca$^{2+}$-free form (thin line), pro-S324A in a Ca$^{2+}$-bound form (thick line), and the mat-S324C-propeptide complex (broken line). Spectra were measured at 20 °C as described under “Experimental Procedures.”

FIGURE 5. ANS fluorescence spectra. The ANS fluorescence spectra of pro-S324A in a Ca$^{2+}$-free form (thin line), pro-S324A in a Ca$^{2+}$-bound form (thick line), and the mat-S324C-propeptide complex (broken line) are shown. These spectra were measured at 20 °C as described under “Experimental Procedures.”
Crystal Structure of Pro-Tk-subtilisin

FIGURE 6. Susceptibility of pro-S324A to chymotryptic digestion. Pro-S324A in a Ca$^{2+}$-free form (lanes 1–4) and pro-S324A in a Ca$^{2+}$-bound form (lanes 5 and 6) were digested with chymotrypsin at 30 °C for 1 h in 20 mm Tris-HCl (pH 7.0) at enzyme/substrate ratios (by weight) of 1:10,000 (lane 2), 1:1000 (lane 3), 1:100 (lane 4), and 1:10 (lane 6) and subjected to 15% SDS-PAGE. The protein was stained with CBB. M, low molecular weight marker kit (Amersham Biosciences); lane 1, undigested pro-S324A in a Ca$^{2+}$-free form; lane 2, undigested pro-S324A in a Ca$^{2+}$-bound form; lane 3, mat-Tk-subtilisin* digested with chymotrypsin as mentioned above; lane 4, mat-Tk-subtilisin* digested with chymotrypsin at 4 °C, digested at 80 °C for 30 min, and subjected to 15% SDS-PAGE. The protein was stained with CBB. M, low molecular weight marker kit (Amersham Biosciences); lane 1, undigested pro-S324A in a Ca$^{2+}$-free form; lane 5, undigested pro-S324A in a Ca$^{2+}$-bound form. The molecular mass of each standard protein is indicated beside the gel.

FIGURE 7. Comparison of the refolding yield of mat-Tk-subtilisin and mat-S324A in the presence and absence of Tk-propeptide. For refolding of mat-Tk-subtilisin, the unfolded protein (30 μM) was diluted 100-fold with buffer A (50 mM Tris-HCl (pH 7.0) containing 1 mM DTT) (lane 1), Ca$^{2+}$ buffer (50 mM Tris-HCl (pH 7.0) containing 1 mM DTT and 10 mM CaCl$_2$) (lane 2), or Ca$^{2+}$/propeptide buffer (50 mM Tris-HCl (pH 7.0) containing 1 mM DTT, 10 mM CaCl$_2$, and Tk-propeptide (0.3 μM)) (lane 3) at 4 °C, incubated at 80 °C for 30 min, and subjected to 15% SDS-PAGE. For refolding of mat-S324A, the unfolded protein (30 μM) was diluted 100-fold with buffer A (lane 7), Ca$^{2+}$ buffer (lane 8), propeptide buffer (50 mM Tris-HCl (pH 7.0) containing 1 mM DTT and Tk-propeptide (0.3 μM)) (lane 9), or Ca$^{2+}$/propeptide buffer (lane 10) at 4 °C, digested with chymotrypsin at 30 °C for 1 h at an enzyme/substrate ratio of 1:100 (w/w), and subjected to 15% SDS-PAGE. Lane 4, undigested pro-Tk-subtilisin*; lane 5, mat-Tk-subtilisin* digested with chymotrypsin as mentioned above; lane 6, undigested mat-S324A. The protein was stained with CBB. M, low molecular weight marker kit. The molecular mass of each standard protein is indicated beside the gel. The upper and lower arrows indicate the positions of mat-Tk-subtilisin* and mat-S324A and Tk-propeptide, respectively.

FIGURE 8. SDS-PAGE of pro-S324A and pro-S324C refolded in the presence of Ca$^{2+}$. Pro-S324A (lane 1) and pro-S324C (lane 2) refolded in the presence of Ca$^{2+}$ were analyzed by 15% SDS-PAGE. The arrows indicate the positions of pro-S324A, mat-S324C*, and Tk-propeptide from the top to the bottom. The molecular mass of each standard protein is indicated beside the gel.

not sufficient for subsequent degradation of the propeptide, like the corresponding bacterial subtilisin mutant. The auto-processed form of pro-S324C will be designated as mat-S324C*-propeptide complex hereafter.

The far-UV and near-UV CD spectra of the mat-S324C*-propeptide complex were almost identical to those of pro-S324A in a Ca$^{2+}$-bound form (Fig. 4). Likewise, ANS fluorescence of the mat-S324C*-propeptide complex was nearly identical to that of pro-S324A in a Ca$^{2+}$-bound form (Fig. 5). The far-UV and near-UV CD spectra and ANS fluorescence of pro-S324C in a Ca$^{2+}$-free form were indistinguishable from those of pro-S324A in a Ca$^{2+}$-free form. Furthermore, pro-S324C in a Ca$^{2+}$-free form was highly sensitive to chymotryptic digestion, whereas the mat-S324C*-propeptide complex was highly resistant to it, like pro-S324A. These results indicate that the conformation of pro-S324C is considerably changed upon Ca$^{2+}$ binding but is not seriously changed upon subsequent autoprocessing. Thus, the folding process of pro-Tk-subtilisin is almost fully completed prior to autoprocessing.

Adaptation to High Temperatures—Comparison of the crystal structures of pro-S324A and the subtilisin E-propeptide complex indicates that the amino acid residues forming an interface between the propeptide and mature domain are relatively well conserved, whereas those forming a hydrophobic core of the propeptide are not. The number of the hydrophobic residues located at the core of the propeptide of Tk-subtilisin (11 residues) is much higher than that of subtilisin E (six), indicating that the propeptide of Tk-subtilisin is highly more stable.
than that of subtilisin E. In addition, as mentioned above, the mature domain of Tk-subtilisin is highly stabilized by Ca$$^{2+}$$ binding. The source organism of Tk-subtilisin, \textit{T. kodakaraensis}, grows most optimally at 90 °C (50). Tk-subtilisin probably adapts to a hyperthermic environment by increasing stability of both the propeptide and mature domain.

**CONCLUSION**

In this study, the crystal structure of pro-S324A, which represents an unautoprocessed precursor form of subtilisin, was determined at 2.3 Å resolution. In addition, the effects of Ca$$^{2+}$$ binding and autoprocessing on structure and stability of pro-Tk-subtilisin mutants were biochemically analyzed. These results show that the folding of unautoprocessed precursor of Tk-subtilisin is induced by Ca$$^{2+}$$ binding and is almost fully completed prior to autoprocessing. The mature domain of pro-S324A contains six Ca$$^{2+}$$ binding sites, all of which contribute to the stabilization of the protein. Bacterial subtilisins require a propeptide for folding, because their structure is unstable in the absence of propeptide. In contrast, Tk-subtilisin does not require propeptide for folding, probably because its structure is highly stabilized by Ca$$^{2+}$$ binding. These results suggest that the role of Ca$$^{2+}$$ and propeptide for maturation greatly varies for different subtilisases.

**REFERENCES**

1. Siezen, R. J., and Leunissen, J. A. (1997) \textit{Protein Sci.} \textbf{6}, 501–523
2. Pulido, M., Saito, K., Tanaka, S., Koga, Y., Morikawa, M., Takano, K., and Kanaya, S. (2006) \textit{J. Mol. Biol.} \textit{415}, 4145–4162
3. Shinde, U. P., and Inouye, M. (1996) \textit{Adv. Exp. Med. Biol.} \textbf{379}, 147–154
4. Li, Y., Hu, Z., Jordan, F., and Inouye, M. (1995) \textit{Biochemistry} \textbf{34}, 7205–7213
5. Almog, O., Gonzalez, A., Klein, D., Greenblatt, M. H., Braun, S., and Shoham, G. (2003) \textit{Protein Eng. Des. Sel.} \textbf{16}, 131–133

**Acknowledgments—**The synchrotron radiation experiments were performed at the BL41XU beamline of SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (Proposal 2005B1766). We thank Drs. T. Inoue and Y. Kai for support in X-ray crystallography and for helpful discussions. We thank M. Pulido for critical reading of the manuscript.

**Crystal Structure of Pro-Tk-subtilisin**