Check for updates

FEBS Letters

FEBS PRESS science publishing by scientists

Insertion loop-mediated folding propagation governs efficient maturation of hyperthermophilic Tk-subtilisin at high temperatures

Ryo Uehara[1,2,3], Nanako Dan[1], Hiroshi Amesaka[4], Takuya Yoshizawa[1], Yuichi Koga[5], Shigenori Kanaya[5], Kazufumi Takano[4], Hiroyoshi Matsumura[1,2] and Shun-ichi Tanaka[1,2,4]

1 Department of Biotechnology, College of Life Sciences, Ritsumeikan University, Shiga, Japan
2 Ritsumeikan Global Innovation Research Organization, Ritsumeikan University, Shiga, Japan
3 Division of Cancer Cell Regulation, Aichi Cancer Center Research Institute, Nagoya, Japan
4 Department of Biomolecular Chemistry, Kyoto Prefectural University, Japan
5 Department of Material and Life Science, Graduate School of Engineering, Osaka University, Japan

Correspondence
R. Uehara, Division of Cancer Cell Regulation, Aichi Cancer Center Research Institute, 1-1 Kanokoden, Chikusa-ku, Nagoya 464-8681, Japan
Tel: +81 52 762 6111
E-mail: r.uehara@aichi-cc.jp
S.-i. Tanaka, Department of Biomolecular Chemistry, Kyoto Prefectural University, 1-5 Shimogamohangi-cho, Sakyo-ku, Kyoto 606-8522, Japan
Tel: +81 75 703 5659
E-mail: stanaka1@kpu.ac.jp

(Received 15 September 2020, revised 21 November 2020, accepted 23 November 2020, available online 19 December 2020)

doi:10.1002/1873-3468.14028

Edited by Peter Brzezinski

The serine protease Tk-subtilisin from the hyperthermophilic archaeon *Thermococcus kodakarensis* possesses three insertion loops (IS1-IS3) on its surface, as compared to its mesophilic counterparts. Although IS1 and IS2 are required for maturation of Tk-subtilisin at high temperatures, the role of IS3 remains unknown. Here, CD spectroscopy revealed that IS3 deletion arrested Tk-subtilisin folding at an intermediate state, in which the central nucleus was formed, but the subsequent folding propagation into terminal subdomains did not occur. Alanine substitution of the aspartate residue in IS3 disturbed the intraloop hydrogen-bonding network, as evidenced by crystallographic analysis, resulting in compromised folding at high temperatures. Taking into account the high conservation of IS3 across hyperthermophilic homologues, we propose that the presence of IS3 is important for folding of hyperthermophilic subtilisins in high-temperature environments.

Keywords: archaea; folding; hyperthermophile; maturation; subtilase

Tk-subtilisin is an extracellular serine protease from a hyperthermophilic archaeon *Thermococcus kodakaraensis* [1]. Like mesophilic subtilisins, Tk-subtilisin is synthesized as a precursor consisting of a signal sequence (Met[-24]-Ala[-1]) and a propeptide (TKpro, Gly1-Leu69) attached to the N terminus of the catalytic subtilisin domain (TKS, Gly70-Gly398). The inactive pro-form (proTKS, Gly1-Gly398) is secreted with the assistance of the signal sequence. Then the pro-form matures into an active-form (TKS) through folding, autoprocessing, and TKpro degradation [2–6]. The maturation process of Tk-subtilisin is similar to that of mesophilic subtilisins but different as Ca^{2+} ions are required for TKS folding, folding is completed before

Abbreviations
proSA, proTKS with Ser324 → Ala mutation; proSA/D356A, proSA with Asp356 → Ala mutation; proSA/L349A, proSA with Leu349 → Ala mutation; proSA/P350A, proSA with Pro350 → Ala mutation; proTKS, a precursor form of Tk-subtilisin (Gly1-Gly398); proTKS/D356A, proTKS with D356 → Ala mutation; proTKS/P350A, proTKS with Pro350 → Ala mutation; proTKS/ΔIS3, proTKS with deletion of IS3 (Gly346-Ser358); Suc-AAPF-*p*NA, *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide; TKpro, an N-terminal propeptide of Tk-subtilisin (Gly1-Leu69); TKS, a subtilisin domain of Tk-subtilisin (Gly70-Gly398); Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl] glycine.

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

autoprocessing, and a high temperature (≥ 80 °C) is necessary for TKpro degradation. Notably, the folding machinery of Tk-subtilisin is attributed to the unique insertion sequences in the subtilisin domain.

Tk-subtilisin harbors three insertion sequences, IS1 (Gly70-Pro82), IS2 (Pro207-Asp226), and IS3 (Gly346-Ser358) in the subtilisin domain unlike mesophilic subtilisins (Fig. 1). IS1 and IS2, but not IS3, are responsible for Ca^{2+} ion binding to Tk-subtilisin [7,8]. IS1 forms a long linker extending from the TKS N terminus to the TKpro C terminus and allows the formation of the Ca1 site before autoprocessing, resulting in a fully folded un-autoprocessed form [9,10]. The highly conserved Ca1 site formation is the final step in the folding of many subtilisin-like proteases (subtilases). Mesophilic subtilisins lacking IS1 are partially unfolded and thermally unstable until autoprocessing due to the absence of Ca^{2+} ions [11–13]. Tk-subtilisin bypasses such an unstable intermediate to enhance the folding efficiency at high temperatures. IS2 forms a Ca^{2+}-binding loop, to which four Ca^{2+} ions (Ca2-Ca5) bind and promote TKS folding by stabilizing the central αβα-substructure [14]. TKpro also acts as a chaperone through interactions with the same substructure [15,16] as commonly seen in mesophilic subtilisins [17]. This unique cooperative chaperone system, including the Ca^{2+}-binding loop and TKpro, ensures robust TKS folding at high temperature. Both IS1 and IS2 are important for efficient folding and thus maturation of Tk-subtilisin at high temperatures where the host organism grows.

IS3 forms a hairpin loop following an α8 helix, containing Ser324, one of the catalytic residues (Fig. 1). Although IS3 is not responsible for Ca^{2+} ion binding, it is essential for the molecular architecture of Tk-subtilisin. Several studies have demonstrated the importance of the loops not only for connecting secondary structure elements but also for protein folding [18–23]. The loop formation allows the unfolded polypeptide chain to explore energetically favorable interactions during the early stage of the folding process. In addition, the amino acid sequence of IS3 is highly conserved among a wide range of hyperthermophilic subtilases (Fig. S1). These findings led us to hypothesize that IS3 also has a pivotal role in the folding and, thus, maturation of Tk-subtilisin at high temperatures, which is yet to be investigated. In this study, we constructed a series of mutants with deletion or single amino acid substitutions in IS3 and characterized the mutant proteins using biochemistry and crystallography. Based on the results of our experiments, we discuss the underlying mechanism of IS3-mediated Tk-subtilisin folding at extremely high temperatures.

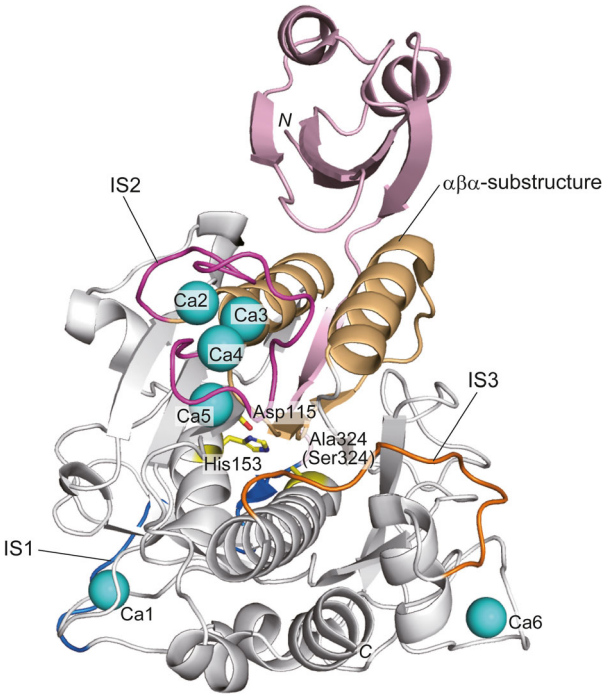

Fig. 1. Tertiary structure of a precursor form of Tk-subtilisin (PDB ID: 2E1P). The propeptide region is coloured light pink. In the subtilisin domain, three insertion sequences IS1, IS2, and IS3 are shown in blue, magenta, and orange, respectively, and the central αβα-substructure is shown in wheat. Other regions are coloured gray. The Ca^{2+} ions (Ca1–Ca6) are shown as cyan spheres. Two active-site residues (Asp115 and His153) and Ala324, which is substituted for the catalytic serine residue, are represented as yellow sticks, in which the oxygen and nitrogen atoms are coloured red and blue, respectively. N and C represent the N and C termini, respectively.

Materials and methods

Plasmid construction

We previously generated pET25b vectors for overproduction of proTKS and proSA [7]. These plasmids were used as a template for mutagenesis. The pET25b derivatives for overproduction of proTKS/ΔIS3, proTKS/P350A, proTKS/D356A, proSA/L349A, proSA/P350A, and proSA/D356A were constructed using a KOD mutagenesis kit (Toyobo, Kyoto, Japan). The PCR primers were designed such that IS3 (Gly346-Ser358) was removed from proTKS for proTKS/ΔIS3 and Leu349, Pro350, and Asp356 were substituted with alanine for other mutants. All DNA oligomers summarized in Table S1 were synthesized by Hokkaido System Science (Sapporo, Japan). PCR was performed with a GeneAmp PCR system 2400 (Applied Biosystems, Tokyo, Japan). The DNA sequences were confirmed by a Prism 310 DNA sequencer (Applied Biosystems).

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Protein preparation

Escherichia coli BL21-CodonPlus (DE3) was used as a host strain for overproduction of all proteins. Proteins were overproduced as inclusion bodies, solubilized with urea, and purified in a denatured form as described previously [2]. Briefly, an insoluble fraction of cell lysates was collected by centrifuge after sonication, dissolved in 20 mM Tris/HCl pH 9.0 containing 8 M urea and 5 mM EDTA and loaded onto Hitrap Q column (GE Healthcare, Little Chalfont, Buckinghamshire, England). Proteins were eluted by 0 → 0.3 M NaCl linear gradient. proTKS and its derivatives were dialyzed against 20 mM Tris/HCl (pH 7.0) and stored in a molten globular Ca^{2+}-free form at 4°C until use. To obtain active enzymes, the proteins were denatured in the presence of 6 M GdnHCl and in turn refolded by 100-fold dilution with 50 mM CAPS-NaOH (pH 9.5) containing 10 mM $CaCl_2$ and 1 mM DTT. For complete maturation, refolded proteins were incubated on ice for 10 min and then at 80 °C for 30 min to degrade TKpro. Degradation of TKpro was confirmed by 15% Tricine SDS/PAGE (SDS/PAGE using Tricine buffer) [24]. For refolding at high temperature, denatured proteins were added to refolding buffer pre-incubated at 80 °C and further incubated for 60 min. For proSA and its derivatives, denatured proteins were dialyzed against 20 mM Tri-HCl (pH 7.0) for a Ca^{2+}-free form or 20 mM Tris/HCl (pH 7.0) containing 10 mM $CaCl_2$ and 1 mM DTT for a Ca^{2+} bound form. For crystallization, the Ca^{2+} bound form of proSA/D356A was further purified by chymotrypic digestion, heat treatment and size exclusion chromatography as described for proSA [7]. The purity of the protein was analyzed by a 12% SDS/PAGE and 15% Tricine SDS/PAGE, followed by staining with Coomassie Brilliant Blue (CBB). The protein concentration was determined by measuring UV absorption at 280 nm (A280) using a cell with an optical path length of 10 mm. A280 values for 0.1% solution (1.0 $mg{\cdot}mL^{-1}$) are 1.24 for precursors (proTKS, proSA, and their derivatives), which contain 16 tyrosine and five tryptophan residues, and 1.47 for mature Tk-subtilisin, which contains 15 tyrosine and five tryptophan residues. These values were calculated by using an absorption coefficient of 1526 $M^{-1}{\cdot}cm^{-1}$ for tyrosine and 5225 $M^{-1}{\cdot}cm^{-1}$ for tryptophan at 280 nm [25].

Enzymatic activities

The protease activity was determined at 80 °C using azocasein (Sigma) or at 20 °C using *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-*p*NA) (Sigma) as described previously [2]. Briefly, 10 μL of protein solution was mixed with 90 μL of substrate solution containing 2% (w/v) azocasein and incubated at 80 °C for 20 min. The reaction buffer was 50 mM CAPS-NaOH (pH 9.5) and 1 mM $CaCl_2$. Enzyme reaction was terminated by addition of 15% TCA. The mixture was cooled on ice for 15 min and centrifuged. An aliquot of the supernatant was withdrawn and mixed with 2 M NaOH and measured for absorption at 440 nm (A440). One unit of activity was defined as an amount of enzyme that increases the A440 value of the reaction mixture by 0.1 in 1 min. A change in activities during maturation was measured using 2 mM Suc-AAPF-*p*NA. Enzyme reaction was performed at 20 °C for 20 min to prevent further maturation. The reaction buffer was 50 mM Tris/HCl (pH 8.0) and 5 mM $CaCl_2$. After reaction, the amount of *p*-nitroaniline released from substrate was determined by measuring an absorption at 410 nm with an absorption coefficient of 8900 $M^{-1}{\cdot}cm^{-1}$. The specific activity was defined as an activity per milligram of protein.

CD spectra measurement

The far-UV CD spectra of protein were measured on a J-725 spectropolarimeter (Japan Spectroscopic Co., Tokyo, Japan) at 20 °C. Protein was dissolved in 20 mM Tris/HCl (pH 7.0) containing 10 mM $CaCl_2$ for a Ca^{2+} bound form and 20 mM Tris/HCl (pH 7.0) for a Ca^{2+}-free form. The protein concentration and optical path length were 0.1 $mg{\cdot}mL^{-1}$ and 2 mm. The mean residue ellipticity [θ], which has the units of deg $cm^2{\cdot}dmol^{-1}$, was calculated using an average amino acid molecular mass of 110 Da.

For refolding experiments, proteins were denatured by incubation in 20 mM Tris/HCl (pH 7.0) containing 6 M GdnHCl at room temperature for 1 h. Refolding reaction was initiated by a rapid dilution of the denatured proteins with 20-fold volume of refolding buffer (50 mM Tris/HCl (pH 7.0) containing 10 mM $CaCl_2$ and 1 mM DTT) pre-incubated at 30, 40 or 50 °C. The CD value at 222 nm was monitored for 10 min.

For thermal denaturation analysis, denaturation curves of proSA derivatives were obtained by monitoring a change in CD values at 222 nm as previously described [10]. Protein was dissolved in 10 mM Tris/HCl (pH 7.5) containing 6 M GdnHCl. The protein concentration and optical path length were 0.1 $mg{\cdot}mL^{-1}$ and 2 mm. The temperature was linearly increased from 25 to 100 °C at a rate of approximately 1 $°C{\cdot}min^{-1}$. The fraction unfolded (F_u) was determined by normalization of the molar ellipticity (θ) according to the following equation; $F_u = ([\theta]_{obs} - [\theta]_u)/([\theta]_n - [\theta]_u)$, where $[\theta]_{obs}$ is the observed molar ellipticity at the given temperature, $[\theta]_u$ is the molar ellipticity of fully unfolded proteins, and $[\theta]_n$ is the molar ellipticity of the native proteins. The midpoints of the transition of these thermal denaturation curves, $T_{1/2}$, were calculated from the resultant normalized curves.

Limited proteolysis

The denatured proteins of proSA and proTKS/ΔIS3 in the presence of 6 M urea were refolded by dialysis against

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

20 mM Tris/HCl (pH 7.0) containing 10 mM $CaCl_2$ and 1 mM DTT at 4 °C overnight. Proteins were then dialyzed against 20 mM Tris/HCl (pH 7.0) containing 10 mM $CaCl_2$ to remove DTT. The refolded proteins were treated with chymotrypsin at 30 °C for 1 h at the range of chymotrypsin : subtilisin ratios from 1 : 10 to 1 : 10 000 (w/w). The proteolytic reaction was terminated by addition of 10% TCA. The precipitated proteins were collected by centrifugation, washed with 70% Acetone, and analyzed by 15% SDS/PAGE.

Molecular mass analysis

The molecular mass of chymotrypsin-digested fragments was determined by a matrix-assisted laser desorption ionization reflectron-type time-of-flight (MALDI-TOF) mass spectrometer (Autoflex, Bruker Daltonik GmbH, Bremen, Germany). The TCA precipitated protein was dissolved in deionized water, mixed with 3,5-dimethoxy-4-hydroxycinnamic acid matrix and spotted onto the MALDI target plate for ionization. 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).

Crystallization and X-ray diffraction data collection

proSA/D356A was dialyzed against 10 mM Tris/HCl (pH 7.0), concentrated up to 10 $mg{\cdot}mL^{-1}$ using ultrafiltration system (Amicon Ultra, Millipore Co.) and crystallized under the same condition as proSA [7]. The single crystals suitable for X-ray diffraction data collection appeared after 1 week. X-ray diffraction data sets were collected at a wavelength of 0.9 Å without additional cryoprotectant at the BL44XU stations in SPring-8 (Hyogo, Japan). The data sets were indexed, integrated and scaled using the HKL2000 suite [26]. The structure was solved by the molecular replacement method using MOLREP [27] in the CCP4 program suite [28]. The crystal structure of proSA (PDB code: 2E1P) was used as a starting model. Model building and structure refinement were done using Refmac [29] and Coot [30]. Progress in the structure refinement was evaluated by the free R factor and by inspection of stereochemical parameters calculated by PROCHECK [31]. The statistics for data collection and refinement are summarized in Table S1. The figures were prepared using PYMOL (http://www.pymol.org).

Protein Data Bank accession number

The coordinates and structure factors for proSA/D356A have been deposited in the Protein Data Bank as entry 3WIV.

Results and Discussion

IS3-deleted mutant reveals the importance of IS3 in TKS folding

Three insertion sequences (IS1-IS3) in the subtilisin domain of Tk-subtilisin form short surface loops in the tertiary structure [7,8]. We previously showed that IS1 and IS2 allow Tk-subtilisin to bind Ca^{2+} ions, which facilitate the folding and maturation at high temperatures [10,13,14]. IS3, in contrast, is not responsible for Ca^{2+} ion binding, yet this insertion loop seems necessary for the molecular architecture. To directly address the role of IS3, we analyzed the maturation of the proTKS mutant lacking IS3 (proTKS/ΔIS3). When proTKS and proTKS/ΔIS3 were refolded in the presence of Ca^{2+} ions, proTKS/ΔIS3 did not undergo autoprocessing and TKpro degradation and remained inactive until the end of incubation, whereas proTKS completed these processes within 10 min (Fig. 2A,B). Because the C-terminal peptide of TKpro (Ala66-Leu69) lies in the active site of TKS, autoprocessing between Leu69 and Gly70 occurs immediately upon proTKS folding [8]. We expected refolded proTKS/ΔIS3 to assume a molten globule state where the active site is not fully formed. Then, we investigated the secondary structure of proTKS/ΔIS3 using far-UV CD spectra (Fig. 2C). The spectrum of refolded proTKS/ΔIS3 in the presence of 10 mM Ca^{2+} ions, which is far beyond the concentration of Ca^{2+} ions (1 mM) sufficient for complete folding of Tk-subtilisin [15,16], was distinct from that of refolded proSA in the same condition, while the spectra of both proteins were similar in the absence of Ca^{2+} ions. proSA, in which a catalytic Ser324 is replaced with alanine, assumes the un-autoprocessed native structure of proTKS in the presence of Ca^{2+} ions, while the Ca^{2+}-free protein has a molten globule state [7]. Therefore, proTKS/ΔIS3 assumed an intermediate structure, which was intrinsically formed in the transition from the Ca^{2+}-free molten globule state to the native structure, indicating that proTKS/ΔIS3 could not complete folding.

To further obtain insights into the intermediate structure of proTKS/ΔIS3 formed in the presence of Ca^{2+} ions, we performed limited proteolysis and sequencing of proTKS/ΔIS3. Consistent with the results of previous study [7], we found that the native structure of proSA was highly tolerant to chymotryptic digestion (Fig. 2D, lanes 1 and 2). In contrast, proTKS/ΔIS3 in the intermediate state was vulnerable, and it was degraded into two major fragments, large and small fragments (Fragm L and Fragm S,

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

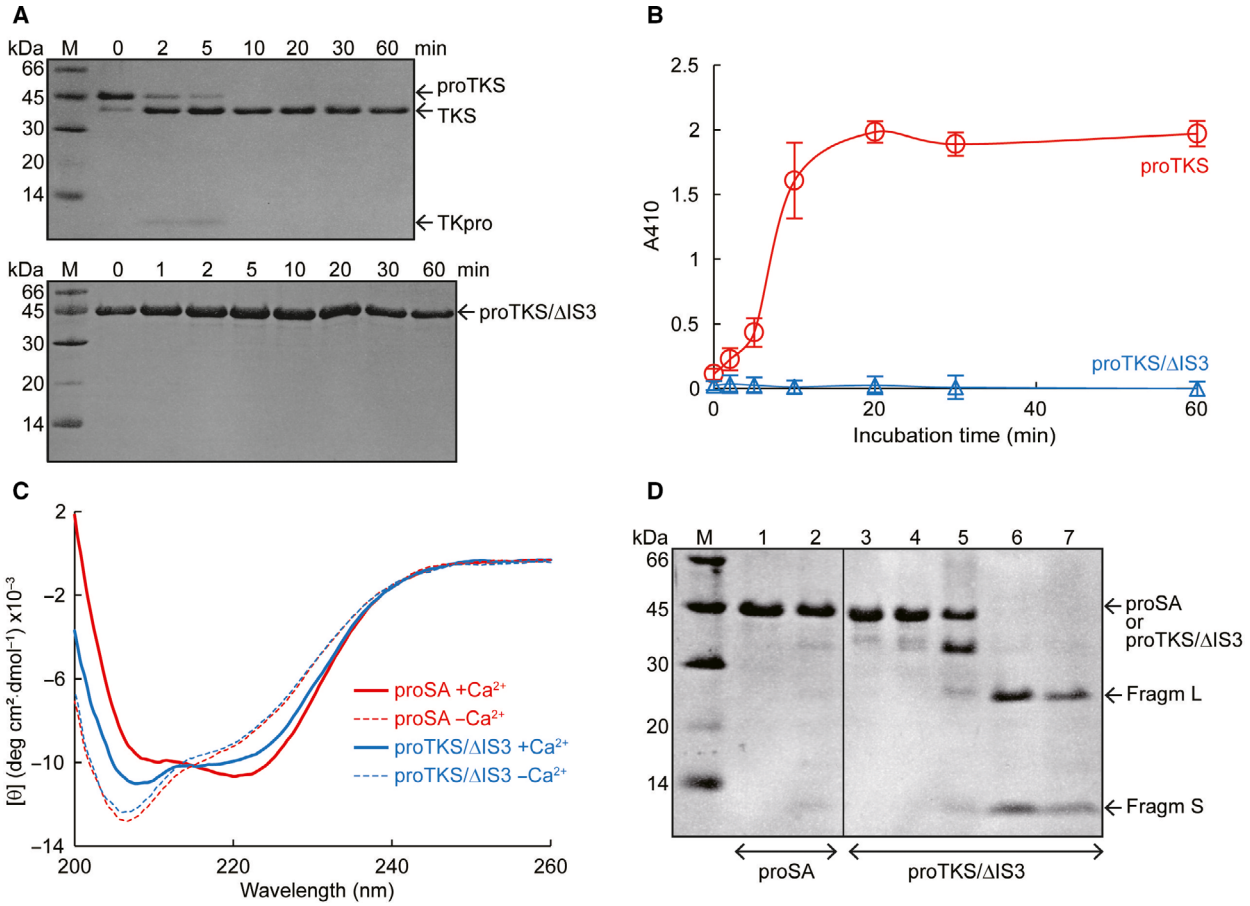

Fig. 2. Folding deficiency of proTKS/ΔIS3 into the native structure. (A, B) Maturation of proTKS and proTKS/ΔIS3. Denatured proteins were refolded on ice for 10 min and then incubated at 80 °C for up to 60 min. At the time indicated at the top of the gel, (A) proteins were analyzed by 15% Tricine SDS/PAGE (upper panel in A, proTKS; lower panel in A, proTKS/ΔIS3) and (B) the activities of proTKS (open red circle) and proTKS/ΔIS3 (open blue triangle) were measured using synthetic peptide substrate Suc-AAPF-*p*NA. In B, the data represent the mean of three independent experiments. The error bars indicate the standard deviation (sd) ($n = 3$). (C) The far-UV CD spectra of proSA refolded in the absence of Ca^{2+} ions (broken red line) and in the presence of Ca^{2+} ions (thick red line), and proTKS/ΔIS3 refolded in the absence of Ca^{2+} ions (broken blue line) and in the presence of Ca^{2+} ions (thick blue line). (D) Limited proteolysis by chymotrypsin. The Ca^{2+} bound forms of proSA (lanes 1, 2) and proTKS/ΔIS3 (lane 3–7) were treated with chymotrypsin at 30 °C for 1 h. The chymotrypsin : subtilisin ratios were 1 : 10 000 (lane 4), 1 : 1000 (lane 5), 1 : 100 (lane 6) and 1 : 10 (lane 2, 7). Lane 1 and 3 contained undigested proteins. Lane M was a low molecular weight marker. The image is composite of different loading positions on the same gel, and the stitch is indicated by a vertical line.

respectively), that were tolerant to further proteolysis at chymotrypsin : subtilisin ratios of 1 : 100 and 1 : 10 (Fig. 2D, lanes 6 and 7). This suggested that the intermediate form of proTKS/ΔIS3 included a folded core structure along with the unstructured portions that could be susceptible to proteolysis. It should be noted that the Ca^{2+}-free forms of proSA and proTKS/ΔIS3 were highly susceptible to chymotryptic digestion, and Fragm L did not appear in any chymotrypsin : subtilisin ratio (Fig. S2). These results emphasized that the intermediate structure formation was dependent on the presence of Ca^{2+} ions. The N-terminal four residues of Fragm L were determined to be S_{100}ITD by peptide sequencing. Using MALDI-TOF mass spectrometry, we found that the molecular size of Fragm L was 28 566.8 daltons, which was nearly identical with the theoretical molecular mass of the Ser100-Gly377 fragment (28 570.73 daltons). We did not obtain the peak corresponding to Fragm S in the MALDI-TOF mass analyses. We presumed that Fragm S would include the N-terminal TKpro (Gly1-Leu69), which maintains the native structure without TKS and is highly resistant to proteolysis [3], because Fragm S appeared by proteolysis of proTKS/ΔIS3

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

both in the presence and absence of Ca^{2+} ions (Fig. 2D and Fig. S2). Consequently, Fragm L was devoid of terminal subdomains but included the central αβα-substructure as shown in Fig. 1. Several studies have proposed that subtilisin folding initially occurs in the αβα-substructure with the assistance of an intramolecular chaperone and subsequently propagates into the N and C terminus [32,33]. ProTKS/ΔIS3 could fold the central αβα-substructure that was induced by Ca^{2+} ion binding to IS2. However, the mutant failed to propagate the folding, leading to the trapped intermediate structure. These results, in turn, suggested that the IS3 loop was a key structural element for efficient Tk-subtilisin folding.

Identification of a critical residue primarily involved in IS3-mediated Tk-subtilisin folding

In order to identify the residue(s) within IS3 responsible for the IS3-mediated Tk-subtilisin folding, we mutated Leu349, Pro350, and Asp356, which are highly conserved among hyperthermophilic subtilases (Fig. S1). We substituted these residues in proSA with alanine producing proSA/L349A, proSA/P350A, and proSA/D356A, and kinetically analyzed their folding using CD spectroscopy as described previously [10]. All three mutant proteins folded as fast as proSA at 30 °C (Fig. 3A, top panels). However, when the temperature was increased to 40 and 50 °C, proSA/P350A and proSA/D356A exhibited slower refolding than proSA and proSA/L349A (Fig. 3A, middle and bottom panels). Notably, while proSA, proSA/L349A, and proSA/P350A exhibited rapid refolding with the increasing temperature, proSA/D356A showed a reduced refolding rate. Consequently, the folding defect in proSA/D356A increased in a temperature-dependent manner. ProSA/P350A also exhibited slightly slower refolding than proSA, but its refolding rate increased depending on temperature, similar to that of proSA and proSA/L349A. Next, we examined the thermal denaturation of these proteins in the presence of 6 M GdnHCl, where the T_M value of proSA was 77.1 ± 0.3 °C (Fig. 3B). The T_M values of the mutant proteins were not markedly different from that of proSA (Fig. 3B), indicating that the mutation of any of the three conserved residues did not affect the stability of proSA. In addition, the far-UV CD spectra revealed that these mutations on IS3 did not significantly alter the native structure of proSA (Fig. S3). These results suggested that the structural vulnerability was not the cause of the compromised folding ability of proSA/P350A and proSA/D356A at higher temperatures.

To explore the impact of the reduced folding rate caused by P350A and D356A mutations on Tk-subtilisin maturation, we tested whether proTKS, harboring P350A and D356A, undergoes maturation in two distinct folding conditions. First, we investigated the maturation of these proteins under moderate folding conditions. Proteins were refolded on ice for 10 min and subsequently incubated at 80 °C for 60 min. The prolonged incubation at 80 °C was designed to ensure complete degradation of TKpro, which is required to release active TKS [2–6]. At this temperature, proTKS and its derivatives were converted to mature proteins at a similar efficiency within 60 min, and each matured protein exhibited comparable enzymatic activities (Fig. 3C). These results indicated that D356A and P350A did not severely impair the catalytic activity or the thermal stability of Tk-subtilisin. Next, we added the denatured proteins directly to the refolding buffer pre-incubated at 80 °C and continued incubation for 60 min. In this extreme folding condition, even proTKS produced a modest fraction of mature proteins (Fig. 3C). The enzymatic activities of the protein folded under the extreme folding condition relative to that under the moderate folding condition (E/M ratio) was 15%, 16%, and 1.6% for proTKS, proTKS/P350A, and proTKS/D356A, respectively. Remarkably, proTKS/D356A was poorly activated under the extreme folding condition.

TKS degrades the unfolded and misfolded proteins during maturation [2]. Even proTKS underwent autodegradation under the extreme folding condition, yielding only 15% of active Tk-subtilisin obtained in the moderate folding condition. More severely, proTKS/D356A failed to yield approximately 98% of active proteins due to inefficient folding at a high temperature (Fig. 3A). ProTKS/P350A yielded active subtilisin at a comparable amount to proTKS, although P350A slightly reduced the refolding rate (Fig. 3A). Given that proSA/P350A refolding was significantly tolerant to the increasing temperature, the reduced refolding rate by P350A might be dispensable for the maturation of proTKS even under the extreme folding condition. These results indicated that Asp356, which contributes to IS3-mediated folding, was required for the enhanced maturation efficiency of proTKS at high temperatures.

Asp356 enables efficient folding through IS3 at high temperatures

Our mutational analysis revealed that Asp356, but not Leu349 and P350, was a critical residue responsible for the IS3-mediated Tk-subtilisin folding. Its significance

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

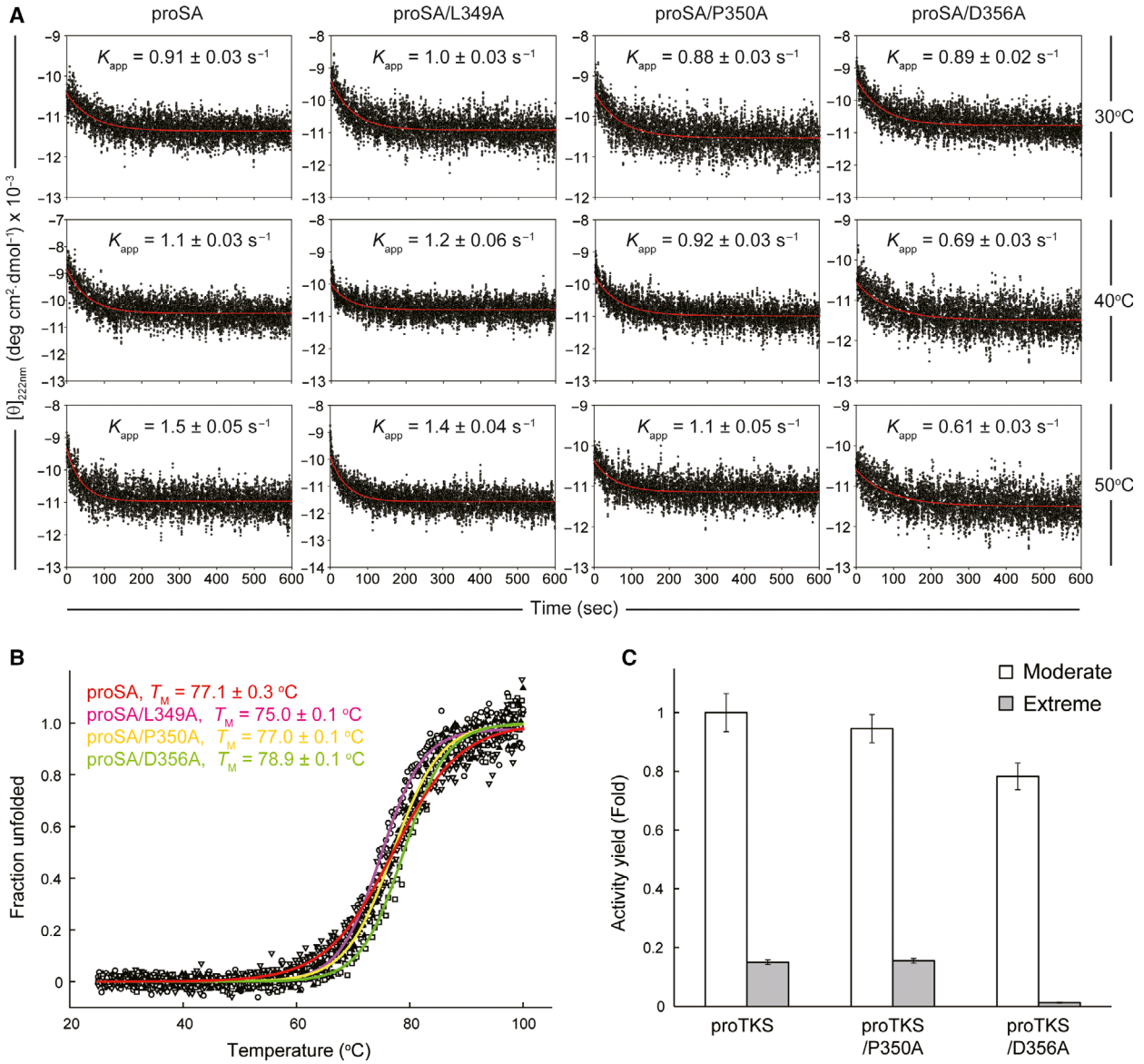

Fig. 3. Effect of mutation of the highly conserved residues in IS3 on folding, stability, and maturation of Tk-subtilisin. (A) Refolding traces of proSA and its derivatives at various temperatures. The trace is fit to a single exponential function as represented by red line. (B) Thermal stabilities of proSA derivatives. The thermal denaturation of proSA (open inverted triangle), proSA/L349A (open circle), proSA/P350A (closed triangle), and proSA/D356A (open square) was measured by monitoring the change in CD values at 222 nm in the presence of 6 M GdnHCl. The denaturation curves and T_M values are shown in the graph. (C) Maturation of proTKS and its derivatives in the moderate and extreme refolding condition. Denatured proteins were refolded in the refolding buffer pre-incubated on ice or at 80 °C for the moderate or extreme condition, respectively. Proteolytic activities toward azocasein were measured after incubation for 60 min. The error bars indicate the standard deviation (sd) (n = 3).

on Tk-subtilisin maturation became more apparent when folding occurred at higher temperatures. In the tertiary structure of proSA, Asp356 was located in the turn structure of the IS3 loop flanked with α8 and α9 helices (Fig. 4). The side chain of Asp356 served as a center of intra-loop hydrogen-bonding interactions: Asp356 $O^{\delta1}$-Thr353 N, Asp356 $O^{\delta1}$-Thr353 O^{γ}, Asp356 $O^{\delta2}$-Ile357 N, Asp356 $O^{\delta2}$-Ser358 N, and Asp356 $O^{\delta2}$-Ser358 O^{γ}.

In protein folding, the turn structure acts as a nucleation site for structural formation that propagates from the corner residues toward the flanking secondary structure elements, or as a flexible connector that is passively formed as a consequence of structural

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

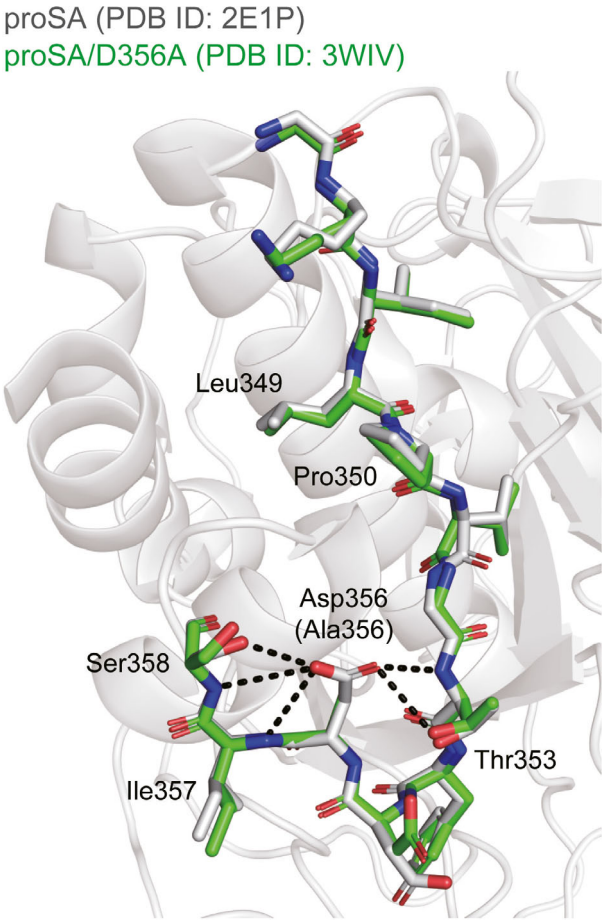

Fig. 4. The structures of IS3. The residues of IS3 in proSA/D356A (PDB ID: 3WIV) and proSA (PDB ID: 2E1P) are shown as green sticks and gray sticks, respectively. The entire structure of proSA is shown. Hydrogen-bonding network mediated by Asp356 is shown as dashed lines.

consolidation from other regions [34]. Formation of the turn structure, primarily mediated by the backbone hydrogen bonds in a transition state, is critical for the folding of several proteins [35–37]. Similarly, the turn structure in the IS3 loop may have a role as a mechanical propagator, which is initially formed through Asp356-mediated interactions in the intermediate state and prompts terminal substructure folding, such as α9 helix. Despite its importance for Tk-subtilisin folding, loss of the hydrogen bonds in proSA/D356A did not reduce the thermal stability of the native structure (Fig. 3B). This finding was consistent with the X-ray crystal structure of proSA/D356A, which greatly resembled proSA (Fig. S4). The IS3 loop structure of proSA/D356A was nearly identical to that of proSA, despite the complete loss of intraloop interactions (Fig. 4). Hence, the intra-loop interactions mediated by Asp356 were required to form the turn structure during folding but not to maintain it upon completion of folding. BLAST search showed that Asp356 was conserved in all hyperthermophilic subtilases harboring IS3 and its hydrogen-bond partner residues, such as Thr353 and Ser358, which were occasionally substituted with small polar residues that acted as a hydrogen donor (Fig. S1). As observed for proTKS in extreme folding conditions (Fig. 3D), Asp356-assisted folding might be required for these subtilases to become active in hyperthermophilic environments.

In general, folding at high temperatures is a challenge for proteins. Tk-subtilisin is equipped with dual chaperone machinery to efficiently achieve folding in harsh environments [14,15]. We previously reported that IS1 and IS2 are required for the efficient maturation of proTKS at high temperatures where the host organism grows [7,14]. In this study, we found that IS3 further enhances the folding rate and, thereby, the maturation efficiency of proTKS by 10-fold. Hence, all three insertions in TKS (IS1-IS3) are eventually required for maturation at a high temperature, although the underlying mechanisms are different [10,13,14]. Such insertions can also be found in other groups of hyperthermophilic subtilases [38]. Pyrolysin is one of the well-characterized examples. This protein has large insertions including calcium-binding motifs, which are important for protein stability and maturation [39]. Given that both Tk-subtilisin and pyrolysin work at extremely high temperatures, acquiring insertions would be a common strategy among hyperthermophilic subtilases to adapt to their native environments.

Acknowledgements

This work was supported by Grant-in-Aids for Scientific Research 19K23733 (to Ryo Uehara), 18K05445 (to Shun-ichi Tanaka), Scientific Research on Innovative Areas 'Integrated Bio-metal Science' 20H05516 (to Shun-ichi Tanaka), and The Hattori Hokokai Foundation 19-003 (to Ryo Uehara).

Author contributions

RU, SK, and ST conceived and supervised the study; RU, TY, SK, KT, HM, and ST designed experiments; RU, DA, and HA performed experiments; RU, KY, and ST analyzed data; RU and ST wrote and edited the manuscript; All authors read and approved the final manuscript.

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

References

1 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**, 2445–2452.

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

3 Pulido MA, Koga Y, Takano K and Kanaya S (2007) Directed evolution of Tk-subtilisin from a hyperthermophilic archaeon: identification of a single amino acid substitution responsible for low-temperature adaptation. *Protein Eng Des Sel* **20**, 143–153.

4 Pulido MA, Tanaka S-I, Sringiew C, You D-J, Matsumura H, Koga Y, Takano K and Kanaya S (2007) 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.

5 Uehara R, Ueda Y, You D-J, Koga Y and Kanaya S (2013) Accelerated maturation of Tk-subtilisin by a Leu→Promutation at the C-terminus of the propeptide, which reduces the binding of the propeptide to Tk-subtilisin. *FEBS J* **280**, 994–1006.

6 Yuzaki K, Sanda Y, You D-J, Uehara R, Koga Y and Kanaya S (2013) Increase in activation rate of Pro-Tk-subtilisin by a single nonpolar-to-polar amino acid substitution at the hydrophobic core of the propeptide domain. *Protein Sci* **22**, 1711–1721.

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

8 Tanaka S-I, Matsumura H, Koga Y, Takano K and Kanaya S (2007) 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.

9 Uehara R, Takeuchi Y, Tanaka S-I, Takano K, Koga Y and Kanaya S (2012) Requirement of Ca^{2+} ions for the hyperthermostability of Tk-subtilisin from *Thermococcus kodakarensis*. *Biochemistry* **51**, 5369–5378.

10 Uehara R, Tanaka S-I, Takano K, Koga Y and Kanaya SJE (2012) Requirement of insertion sequence IS1 for thermal adaptation of Pro-Tk-subtilisin from hyperthermophilic archaeon. *Extremophiles* **16**, 841–851.

11 Jain SC, Shinde U, Li Y, Inouye M and Berman HM (1998) The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 å resolution. *J Mol Biol* **284**, 137–144.

12 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.

13 Uehara R, Angkawidjaja C, Koga Y and Kanaya S (2013) Formation of the high-affinity calcium binding site in pro-subtilisin E with the insertion sequence IS1 of Pro-Tk-subtilisin. *Biochemistry* **52**, 9080–9088.

14 Takeuchi Y, Tanaka S-I, 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 archaeon. *Biochemistry* **48**, 10637–10643.

15 Tanaka S-I, 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.

16 Tanaka S-I, 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.

17 Chen Y-J and Inouye M (2008) The intramolecular chaperone-mediated protein folding. *Curr Opin Struct Biol* **18**, 765–770.

18 Fierz B, Satzger H, Root C, Gilch P, Zinth W and Kiefhaber T (2007) Loop formation in unfolded polypeptide chains on the picoseconds to microseconds time scale. *Proc Natl Acad Sci USA* **104**, 2163–2168.

19 Krieger F, Möglich A and Kiefhaber T (2005) Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. *J Am Chem Soc* **127**, 3346–3352.

20 Berezovsky IN, Grosber AY and Trifonov EN (2000) Closed loops of nearly standard size: common basic element of protein structure. *FEBS Lett* **466**, 283–286.

21 Viguera AR and Serrano L (1997) Loop length, intramolecular diffusion and protein folding. *Nat Struct Biol* **4**, 939–946.

22 Nagi AD, Anderson KS and Regan L (1999) Using loop length variants to dissect the folding pathway of a four-helix-bundle protein. *J Mol Biol* **286**, 257–265.

23 Orevi T, Rahamim G, Hazan G, Amir D and Haas E (2013) The loop hypothesis: contribution of early formed specific non-local interactions to the determination of protein folding pathways. *Biophys Rev* **5**, 85–98.

24 Schägger H (2006) Tricine–SDS-PAGE. *Nat Protoc.* **1**, 16.

25 Goodwin TW and Morton RA (1946) The spectrophotometric determination of tyrosine and tryptophan in proteins. *Biochem J* **40**, 628–632.

26 Otwinowski Z and Minor W (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode in

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Methods in Enzymology (pp. 307–326). San Diego: Academic Press.
27 Vagin A and Teplyakov A (1997) MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* **30**, 1022–1025.
28 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D* **67**, 235–242.
29 Murshudov GN, Vagin AA and Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53**, 240–255.
30 Emsley P and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132.
31 Laskowski RA, MacArthur MW, Moss DS and Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**, 283–291.
32 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.
33 Kim SG, Chen YJ, Falzon L, Baum J and Inouye M (2020) Mimicking cotranslational folding of prosubtilisin E *in vitro*. *J Biochem* **167**, 473–482.
34 Marcelino AM and Gierasch LM (2008) Roles of beta-turns in protein folding: from peptide models to protein engineering. *Biopolymers* **89**, 380–391.
35 Nauli S, Kuhlman B and Baker D (2001) Computer-based redesign of a protein folding pathway. *Nat Struct Biol* **8**, 602–605.
36 Deechongkit S, Nguyen H, Powers ET, Dawson PE, Gruebele M and Kelly JW (2004) Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature* **430**, 101–105.
37 Deechongkit S, Nguyen H, Jager M, Powers ET, Gruebele M and Kelly JW (2006) β-sheet folding mechanisms from perturbation energetics. *Curr Opin Struct Biol* **16**, 94–101.
38 Siezen RJ and Leunissen JA (1997) Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci* **6**, 501–523.
39 Gao X, Zeng J, Yi H, Zhang F, Tang B and Tang XF (2017) Four inserts within the catalytic domain confer extra stability and activity to hyperthermostable pyrolysin from *Pyrococcus furiosus*. *Appl Environ Microbiol* **83**, e03228-16.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Multiple sequence alignment of Pre-Pro-subtilisins across species.

Fig. S2. Limited proteolysis of the Ca^{2+}-free forms of proSA and proTKS/ΔIS3 by chymotrypsin.

Fig. S3. Far-UV CD spectra of proSA derivatives.

Fig. S4. Crystal structure of proSA/D356A.

Table S1. Primer pairs used in this study.

Table S2. Data collection and refinement statistics.

18733468, 2021, 4, Downloaded from https://febs.onlinelibrary.wiley.com/doi/10.1002/1873-3468.14028 by CochraneChina, Wiley Online Library on [02/03/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License