BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

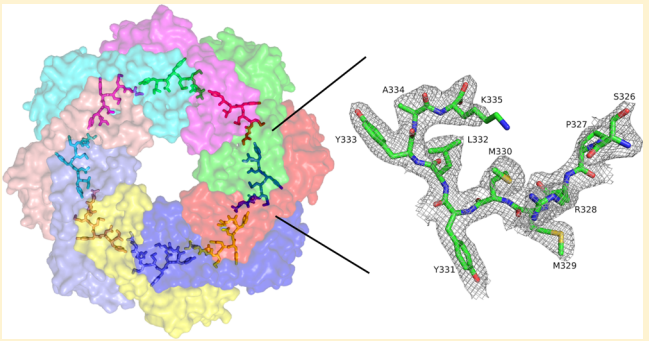
Structure of Signal Peptide Peptidase A with C-Termini Bound in the Active Sites: Insights into Specificity, Self-Processing, and Regulation

Sung-Eun Nam and Mark Paetzel*

Department of Molecular Biology and Biochemistry, Simon Fraser University, South Science Building, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6

Supporting Information

ABSTRACT: Bacterial signal peptide peptidase A (SppA) is a membrane-bound enzyme that utilizes a serine/lysine catalytic dyad mechanism to cleave remnant signal peptides within the cellular membrane. *Bacillus subtilis* SppA ($SppA_{BS}$) oligomerizes into a homo-octameric dome-shaped complex with eight active sites, located at the interface between each protomer. In this study, we show that $SppA_{BS}$ self-processes its own C-termini. We have determined the crystal structure of a proteolytically stable fragment of $SppA_{BS}$K199A that has its C-terminal peptide bound in each of the eight active sites, creating a perfect circle of peptides. Substrate specificity pockets S1, S3, and S2′ are identified and accommodate C-terminal residues Tyr331, Met329, and Tyr333, respectively. Tyr331 at the P1 position is conserved among most *Bacillus* species. The structure reveals that the C-terminus binds within the substrate-binding grooves in an antiparallel β-sheet fashion. We show, by C-terminal truncations, that the C-terminus is not essential for oligomeric assembly. Kinetic analysis shows that a synthetic peptide corresponding to the C-terminus of $SppA_{BS}$ competes with a fluorometric peptide substrate for the $SppA_{BS}$ active site. A model is proposed for how the C-termini of SppA may function in the regulation of this membrane-bound self-compartmentalized protease.



In bacteria, proteins destined for locations other than the cytosolic space are synthesized in a precursor form (preprotein) with a signal peptide preceding what will be the N-terminus of the mature protein.[1] The preprotein is transported to and translocated across the cytosolic membrane via the Sec machinery.[2] Upon translocation, the signal peptide is cleaved off by signal peptidase.[3]

Signal peptide peptidase A (SppA) functions to clear the membrane of remnant signal peptides. In 1984, Ichihara et al. observed degradation of signal peptides upon addition of a membrane extract containing *Escherichia coli* SppA.[4] In *Bacillus subtilis*, the level of formation of mature secretory proteins was decreased in an SppA-deleted strain, suggesting that SppA is involved in the processing of the signal peptide from the precursor.[5] In *E. coli*, four SppA protomers or, in the case of *B. subtilis*, eight SppA protomers come together to form a dome-shaped structure with a small opening at the top and a wider opening at the base.[6,7] The wider opening of the dome sits on the extracellular side of the cytosolic membrane.[8] The protomer of *E. coli* SppA is a tandem repeat of an α/β-fold domain with a high degree of structural homology but a moderate degree of sequence homology between the N-terminal and C-terminal domains. The *B. subtilis* SppA protomer is half the size of the *E. coli* SppA protomer and is most similar to the C-terminal domain of the *E. coli* SppA protomer. Both *E. coli* SppA ($SppA_{EC}$) and *B. subtilis* SppA ($SppA_{BS}$) utilize a serine as the nucleophile and a lysine as the general base.[6−8] $SppA_{EC}$ contains four complete active sites, each located at the interface of the two domains of the protomer. $SppA_{BS}$ has eight active sites, located at the interface of the protomers.[6,7]

In this work, we show that $SppA_{BS}$ self-processes its own C-termini. We have determined a crystal structure of an active site mutant of $SppA_{BS}$ with its C-termini bound within its active sites. Each of the eight active sites accommodates one C-terminal peptide, creating a continuous circular density inside the binding groove of the octameric SppA. The C-terminal peptides show antiparallel β-sheet interactions with the substrate-binding grooves. The C-terminal peptide residues Tyr331, Tyr333, and Met329 are accommodated by the S1, S2′, and S3 substrate specificity pockets of $SppA_{BS}$, respectively (nomenclature of Schechter and Berger).[9] We show by C-terminal truncation analysis that the C-terminus is not essential for $SppA_{BS}$ oligomer assembly. In addition, we show, via a fluorogenic peptide cleavage assay, that a synthetic peptide corresponding to the bound C-terminal peptide can inhibit $SppA_{BS}$ in a competitive manner. We propose a model for how the C-termini of SppA may compete for the active sites and function in regulation.

Received: August 21, 2013
Revised: November 13, 2013
Published: November 14, 2013

EXPERIMENTAL PROCEDURES

Cloning and Mutagenesis. The protocol for cloning the $SppA_{BS}$ (UniProt entry O34525) construct that lacks residues 1−25 ($SppA_{BS}^{\Delta 1-25}$) or 2−54 ($SppA_{BS}^{\Delta 2-54}$) and contains an N-terminal six-histidine (His_{X6}) tag as well as Lys199 mutated to alanine has been previously described.[7]

The QuickChange site-directed mutagenesis procedure (Stratagene) was used to mutate the Tyr331 codon to an alanine codon (underlined) using oligonucleotides 5′-ccgagaatgatggctctctatgcgaag-3′ and 5′-cttcgcatagagagccatcattctcgg-3′.

C-Terminally truncated constructs were made using the $SppA_{BS}^{\Delta 1-25}$K199A plasmid as a template, and the following oligonucleotides were used to mutate amino acid residues Met307, Gly295, and Met329 to stop codons (underlined): $SppA_{BS}^{\Delta 1-25,\ \Delta 329-335}$K199A, 5′-ggcgcgaacaaatagtttaaaagtgaa-3′ and 5′-ttcacttttaaactatttgttcgcgcc-3′; $SppA_{BS}^{\Delta 1-25,\ \Delta 307-335}$K199A, 5′-gaggaaagcttctgattaggctcactg-3′ and 5′-cagtgagcctaatcagaagctttcctc-3′; $SppA_{BS}^{\Delta 1-25,\ \Delta 295-335}$K199A, 5′-ggttcgccgagatagatgtatctctat-3′ and 5′-atagagatacatctatctcggcgaacc-3′. The QuickChange site-directed mutagenesis procedure was utilized, and the sequences were confirmed by DNA sequencing (Genewiz).

Expression, Purification, and Limited Proteolysis. The $SppA_{BS}^{\Delta 1-25}$K199A construct was expressed in *E. coli* Tuner (DE3) cells, and 5 g of cells (wet weight) was lysed in 45 mL of buffer A [20 mM Tris-HCl (pH 8.0) and 150 mM NaCl] as described previously.[7] The protein was purified from the lysate as follows. The cell lysate was centrifuged at 28912*g* for 35 min, and the supernatant was applied to a Ni^{2+}-NTA affinity resin (2 mL) pre-equilibrated with buffer A. The supernatant was rocked gently with the resin at 4 °C overnight and then washed in the following order: with buffer A (30 mL), with 50 mM imidazole in buffer A (30 mL), with 75 mM imidazole in buffer A (20 mL), and finally with buffer A (40 mL) alone. Thermolysin (Sigma, 0.5 mg) in 25 mL of buffer A was then added to the resin and the mixture rocked gently at room temperature overnight. The flow through was collected and concentrated to 7 mg/mL using an Amicon ultracentrifugal filter device (Millipore) with a 100 kDa cutoff. The 500 μL sample was then applied to a Superdex-200 size-exclusion chromatography column, equilibrated with buffer A, on an Amersham ÄKTA fast performance liquid chromatography (FPLC) system at a flow rate of 0.5 mL/min. The extinction coefficient (17880 M^{-1} cm^{-1}), molecular mass (36182 Da), and pI (6.9) of $SppA_{BS}^{\Delta 1-25}$K199A, including a hexahistidine and thrombin cleavage site (MGSSHHHHHHSSGLVPRGSH), were calculated using ProtParam.[10]

$SppA_{BS}^{\Delta 1-25,\ \Delta 329-335}$K199A, $SppA_{BS}^{\Delta 1-25,\ \Delta 307-335}$K199A, $SppA_{BS}^{\Delta 1-25,\ \Delta 295-335}$K199A, and $SppA_{BS}^{\Delta 2-54}$ were overexpressed and purified using a procedure that was previously described[7] with the exception that the $SppA_{BS}^{\Delta 2-54}$ construct cell lysate supernatant was added to a buffer A pre-equilibrated Ni^{2+}-NTA affinity resin and gently rocked overnight at 4 °C before the chromatographic elution via an imidazole step gradient.

Electrospray Ionization Time-of-Flight Mass Spectrometry Analysis of the Protein and Peptide Molecular Mass. For peptide analysis, a sample was prepared using $ZipTips_{C18}$ with 0.6 μL of C18 resin (Millipore) following the manufacturer's protocol. A sample (100 μL, 1 mg/mL) of thermolysin-treated $SppA_{BS}^{\Delta 1-25}$K199A purified by a Superdex-200 purification step, containing final concentrations of 6 M guanidine hydrochloride and 0.1% trifluoroacetic acid (TFA), was aspirated and dispensed five times with 0.1% TFA pre-equilibrated $ZipTips_{C18}$. The beads in the $ZipTips_{C18}$ were washed twice using 10 μL of 0.1% TFA and 5% methanol. The sample was eluted with 10 μL of 50% acetonitrile (ACN) and 0.1% TFA. The eluted sample was analyzed using an Agilent 6210 Time-Of-Flight LC/MS system (Department of Chemistry, Simon Fraser University). The software used was Agilent MassHunter version B.02.00. Mass spectrometer settings for the peptide mass determination were as follows: ionization mode, positive electrospray ionization (+ESI); gas temperature, 200 °C; gas flow, 7 L/min; nebulizer, 30 psi; V_{Cap}, 3500 V; sample introduced by flow injection.

Mass spectral analyses of $SppA_{BS}^{\Delta 2-54}$(WT) and $SppA_{BS}^{\Delta 2-54}$K199A and other intact protein samples were also performed using an Agilent 6210 Time-Of-Flight LC/MS system (Department of Chemistry, Simon Fraser University). Data were analyzed using Agilent MassHunter BioConfirm. Mass spectrometer settings for the protein mass determination were as follows: ionization mode, positive electrospray ionization (+ESI); gas temperature, 325 °C; gas flow, 7 L/min; nebulizer, 30 psi; capillary voltage, 3500 V; HPLC column, Zorbax 300-C8; particle size, 3.5 μm, 50 mm (length) × 2.1 mm (diameter); solvent A, water with 0.1% formic acid; solvent B, acetonitrile with 0.1% formic acid; linear gradient from 2 to 60% B over 5 min (0.3 mL/min) and then held at 60% for 4 min (0.5 mL/min).

Crystallization. The sitting-drop vapor diffusion method was used to grow crystals of the $SppA_{BS}^{\Delta 1-25}$K199A thermolysin resistant fragment. The drop contained 1 μL of protein and 1 μL of reservoir solution and was covered with paraffin oil. The refined reservoir condition included 23% *tert*-butanol and 0.1 M Tris-HCl (pH 8.5), and the drop was equilibrated against 1 mL of the reservoir solution at room temperature. The crystal grew within 2 weeks at room temperature (~296 K). The cryo solution contained 20% 2-methyl-2,4-pentanediol (MPD), 23% *tert*-butanol, and 0.1 M Tris-HCl (pH 8.5). The crystal was transferred to the cryo solution and flash-cooled in liquid nitrogen.

Collection of Diffraction Data. Diffraction images were collected on beamline 08B1-1 at the Canadian Macromolecular Crystallography Facility (CMCF) of the Canadian Light Source (CLS), using a Rayonix MX300HE X-ray detector and MxDC (Macromolecular Crystallography Data Collector). Each diffraction image was exposed for 0.8 s at a wavelength of 0.9795 Å. An oscillation angle of 0.5° was used, and a total of 360 images were collected at a detector distance of 305 mm. The images were processed using HKL2000.[11] The unit cell dimensions of the crystal were 87.8 Å × 131.0 Å × 207.1 Å, and the space group was $P2_12_12_1$. Eight molecules are found in the asymmetric unit with a Matthews coefficient of 2.5 $Å^3$/Da (50.4% solvent). The $SppA_{BS}$ molecular mass after limited proteolysis (26624.5 Da protein + 3355.7 Da peptide) was used to calculate the Matthews coefficient.[12] See Table 1 for crystal parameters, data collection statistics, and refinement statistics.

Structure Determination and Refinement. The structure was determined through molecular replacement using Phaser,[13] and the search model was created using Chainsaw.[14] The $SppA_{BS}$ homology model was built using the C-terminal domain of $SppA_{EC}$ (PDB entry 3BF0, chain A) as a template. Conserved side chain residues were kept, and nonconserved residues were truncated to the Cβ side chain atom. The missing side chain atoms were built, and the initial structural refinement

dx.doi.org/10.1021/bi4011489 | *Biochemistry* 2013, 52, 8811−8822

Table 1. Crystal Parameters, Data Collection Statistics, and Refinement Statistics

	PDB entry 4KWB
Crystal Parameters	
space group	$P2_12_12_1$
a, *b*, *c* (Å)	87.8, 131.0, 207.1
V_m (Å^3/Da)	2.5
% solvent	50.4
no. of protein molecules (chains) in the asymmetric unit	8
Data Collection Statistics[a]	
source	CLS
beamline	08ID-1
wavelength (Å)	0.9795
resolution (Å)	44.6−2.4 (2.5−2.4)
total no. of reflections	638197
no. of unique reflections	93168 (9085)
R_{merge}[b]	0.060 (0.302)
mean $(I)/\sigma(I)$	50.3 (6.7)
completeness (%)	99.2 (97.5)
redundancy	6.8 (6.3)
Refinement Statistics	
no. of residues	1865
no. of water molecules	221
total no. of atoms	14324
R_{cryst}[c] (%)/R_{free}[d] (%)	20.6/24.0
average *B* factor (Å^2) (all atoms)	50.9
average *B* factor (Å^2) (protein)	51.7
average *B* factor (Å^2) (waters)	44.1
average *B* factor (Å^2) (peptide)	61.5
Wilson plot estimated *B* factor (Å^2)[e]	51.0
correlation coefficient[f]	0.94
root-mean-square deviation for angles (deg)	1.061
root-mean-square deviation for bonds (Å)	0.008
Ramachandran plot[g] (%)	
core region (no. of residues)	1776 (97.7)
allowed region (no. of residues)	41 (2.3)
outliers (no. of residues)	0 (0)

[a]The data collection statistics in parentheses are the values for the highest-resolution shell. [b]$R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection and $\langle I(hkl)\rangle$ is the mean intensity of that reflection. [c]$R_{cryst} = \sum_{hkl}||F_{obs}| - |F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. [d]R_{free} is calculated using 5% of the reflections randomly excluded from refinement. [e]The Wilson plot estimated *B* factor was calculated using Truncate from the CCP4 program suite.[36,37] [f]The correlation coefficient was calculated using Refmac5 from the CCP4 program suite.[17,36] [g]The Ramachandran plot was calculated using Procheck from the CCP4 program suite.[22]

was performed using Autobuild within PHENIX version 1.6.4.[15] Coot[16] was utilized to make manual adjustments to the structure. A final round of restrained refinement, including TLS, was performed using Refmac5[17] in CCP4.[18] TLS was analyzed using the TLS motion determination server.[19,20]

Structural Analysis. PyMOL[21] was used to make the figures. The stereochemistry of the structure was analyzed with PROCHECK.[22] PROMOTIF[23] was used to identify and analyze the secondary structure and motifs within the protein. AreaImol[24] in CCP4[18] and PISA[25] were utilized for the accessible surface area calculation using a probe with a 1.4 Å radius. To calculate the buried surface area between the peptides and the $SppA_{BS}$ octamer, the following equation was used: $B = (A_1 + A_2 - A_3)/2$, where *B* is the buried surface area in square angstroms, A_1 is the accessible surface area (in square angstroms) of the eight peptides alone (unattached to the $SppA_{BS}$ octameric protein), A_2 is the accessible surface area (in square angstroms) of the $SppA_{BS}$ octameric protein alone (unattached to the eight peptides), and A_3 is the accessible surface area (in square angstroms) of the $SppA_{BS}$ octameric protein containing the eight bound peptides.[26] Coot[16] was used for measuring the distances between atoms as well as building the C-terminal polyalanine extension. The superposition of atomic coordinates was performed with PyMOL. The octamer of the $SppA_{BS}$ free active site structure (PDB entry 3RST) was aligned to that of the peptide-bound $SppA_{BS}$ complex structure using PyMOL. Five cycles were conducted. In each cycle, each pair of atoms with a standard deviation from the mean of >2 was rejected, and in the final cycle, 1500 atoms were used to calculate the root-mean-square deviation.

$SppA_{BS}$ Self-Processing (intracomplex vs intercomplex). $SppA_{BS}^{\Delta 2-54}$ (16 μg) and $SppA_{BS}^{\Delta 2-54}$K199A (8 μg) alone or together were incubated at 37 °C in buffer A (total reaction volume of 40 μL). Aliquots (10 μL) were taken at 0, 1, and 24 h. To each aliquot was added an equal volume of 2× SDS loading dye, followed by boiling for 10 min. The samples were then run on a 13.5% sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) gel and stained with PageBlue stain (Fermentas). The $SppA_{BS}$ enzymes in this experiment were not exposed to thermolysin.

$SppA_{BS}$ Self-Processing (effect of the Y331A mutation). $SppA_{BS}^{\Delta 2-54}$K199A (15 μg) and $SppA_{BS}^{\Delta 2-54}$K199A/Y331A (15 μg), in separate tubes, were incubated at room temperature in buffer A (total reaction volumes of 50 μL each). Aliquots (10 μL) were taken from each sample at 0, 4, 5, and 6 days. To each aliquot was added an equal volume of 2× SDS loading dye, followed by boiling for 10 min. The samples were then run on a 13.5% SDS−PAGE gel and stained with PageBlue stain (Fermentas). The $SppA_{BS}$ enzymes in this experiment were not exposed to thermolysin.

Size-Exclusion Chromatographic Analysis of $SppA_{BS}$ and $SppA_{BS}$ C-Terminal Truncations. $SppA_{BS}^{\Delta 1-25}$K199A and each of the C-terminally truncated constructs ($SppA_{BS}^{\Delta 1-25,\ \Delta 329-335}$K199A, $SppA_{BS}^{\Delta 1-25,\ \Delta 307-335}$K199A, and $SppA_{BS}^{\Delta 1-25,\ \Delta 295-335}$K199A) were incubated with thermolysin (500:1 molar ratio) overnight at room temperature. The samples were then applied to a buffer A-equilibrated Superdex-200 size-exclusion chromatography column connected to an Amersham ÄKTA FPLC system that was run at a flow rate of 0.5 mL/min. The column was calibrated with gel-filtration calibration kit standards (GE Healthcare Life Sciences). Each eluted peak was analyzed by SDS−PAGE.

$SppA_{BS}$ Kinetics and Inhibition Analysis Using a Fluorogenic Peptide Assay. Each reaction was conducted at 23 °C in a total volume of 100 μL using a SpectraMax M5Multi-Mode Microplate Reader (Molecular Devices). A black 96-well micro array plate with a clear base (Greiner Bio-One) was used for all the reactions. The fluorogenic peptide substrate dodecanoyl-NGEVAKA-MCA (MCA, 4-methyl 7-cumaryl amide) and the peptide NH_3-SPRMMYLYAK-COOH were each dissolved in 100% dimethyl sulfoxide (DMSO) to make 10 mM stock solutions. Dodecanoyl-NGEVAKA-MCA was synthesized by CanPeptide Inc., and NH_3-SPRMMYLYAK-COOH was synthesized by Chinapeptide Inc. Each reaction mixture contained 0.625−60 μM dodecanoyl-NGEVAKA-MCA and 45 nM $SppA_{BS}^{\Delta 2-54}$ in buffer A with 2%

dx.doi.org/10.1021/bi4011489 | Biochemistry 2013, 52, 8811−8822

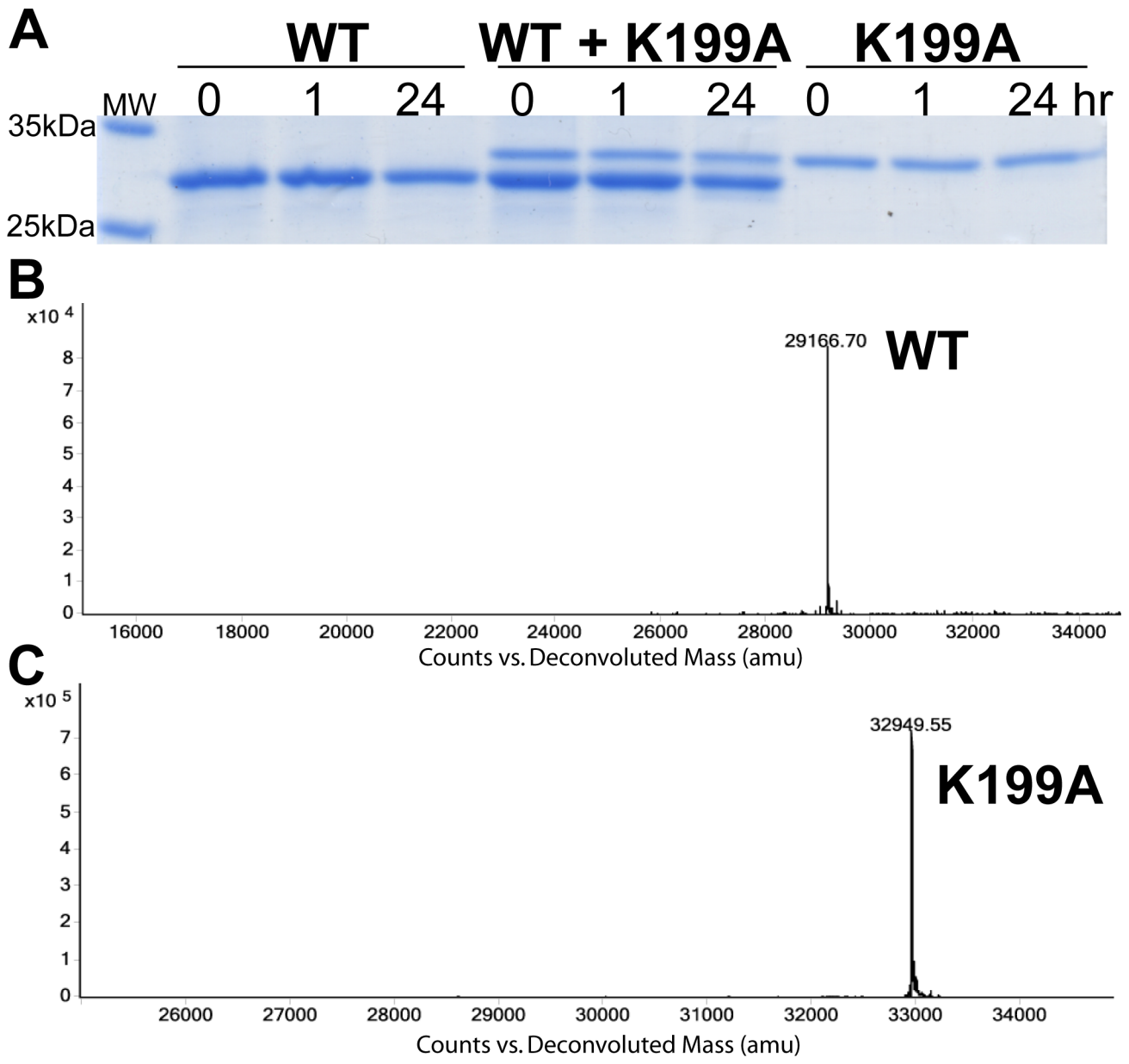


Figure 1. C-Terminal self-cleavage occurs within the $SppA_{BS}$ chamber. (A) Samples of $SppA_{BS}^{\Delta 2-54}$ alone (left), $SppA_{BS}^{\Delta 2-54}$ and $SppA_{BS}^{\Delta 2-54}$K199A together (middle), and $SppA_{BS}^{\Delta 2-54}$K199A alone (right) were incubated at 37 °C, and aliquots were taken at 0, 1, and 24 h. Loading dye was added to each aliquot and then the mixture boiled to stop the self-processing reactions. Samples were run on a 13.5% SDS−PAGE gel, followed by staining with PageBlue stain. The far left lane shows the protein molecular mass markers. (B) Electrospray ionization mass spectrometry results for $SppA_{BS}^{\Delta 2-54}$ revealing a major peak at 29166.70 amu. (C) Electrospray ionization mass spectrometry results for $SppA_{BS}^{\Delta 2-54}$K199A revealing a major peak at 32949.55 amu. Note that these $SppA_{BS}$ enzymes were not exposed to thermolysin treatment.

DMSO. The excitation and emission wavelengths were 380 and 460 nm, respectively. Data points, collected in triplicate, were fit to the Michaelis−Menten equation to determine the k_{cat} and K_M values using Prism5 (GraphPad). $SppA_{BS}$ inhibition assays utilizing a synthetic peptide corresponding to the C-terminus (NH_3-SPRMMYLYAK-COOH) of $SppA_{BS}$ were performed using 45 nM $SppA_{BS}^{\Delta 2-54}$ incubated with NH_3-SPRMMYLY-AK-COOH at concentrations of 625 nM, 1.25 μM, 2.5 μM, 5 μM, 10 μM, and 20 μM for 30 min at room temperature in buffer A. A substrate concentration range of 0.625−60 μM was used. The data points were fit, employing the competitive inhibitor method to determine the K_i values using Prism5 (GraphPad).

RESULTS

$SppA_{BS}$ Removes Its Own C-Terminus. Overexpression of the full-length (membrane-bound) wild-type $SppA_{BS}$ or the $SppA_{BS}$ construct that lacks the predicted transmembrane segment ($SppA_{BS}^{\Delta 1-25}$) resulted in cells with a low growth rate and an undetectable expression level. However, overexpression of a construct truncated further into the N-terminus ($SppA_{BS}^{\Delta 2-54}$) gave sufficient purification yields for activity assays to be performed. SDS−PAGE analysis of N-terminal six-His affinity tag purified $SppA_{BS}^{\Delta 2-54}$ and an active site mutant (K199A) of $SppA_{BS}^{\Delta 2-54}$, which is much less active, reveals that $SppA_{BS}^{\Delta 2-54}$ self-processes its own C-terminus (Figure 1). The processing of the C-terminus occurs very quickly and is complete before the enzyme can be purified. Electrospray

dx.doi.org/10.1021/bi4011489 | Biochemistry 2013, 52, 8811−8822

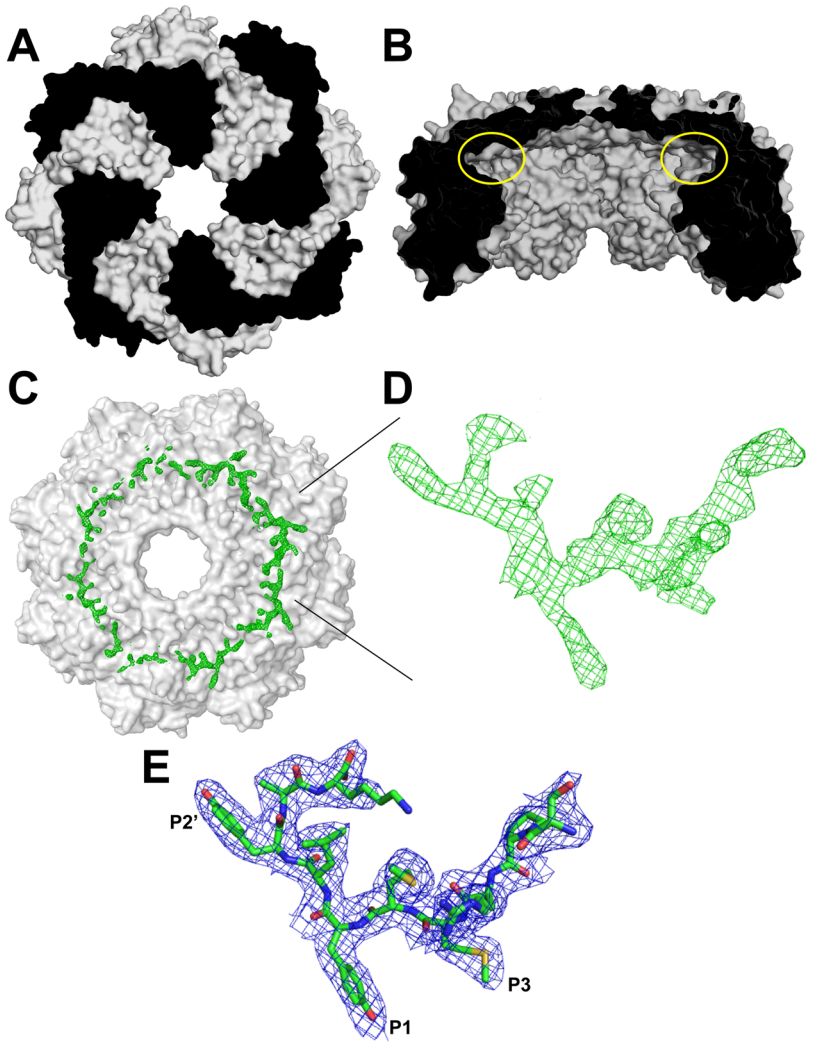


Figure 2. Electron density corresponding to the $SppA_{BS}$ C-terminal peptide bound in the substrate-binding groove of $SppA_{BS}$. (A) Octameric assembly of $SppA_{BS}$ shown in surface representation, each protomer colored black and white in alternating fashion. (B) Cross-sectional view of the $SppA_{BS}$ octamer showing the concave groove (indicated by yellow circles) where the peptides bind. (C) Circular shape of the positive difference electron density map ($F_o - F_c$) observed in the substrate-binding grooves of the octameric $SppA_{BS}$ (shown in gray surface representation). (D) Positive difference electron density map ($F_o - F_c$) that is observed in each of the eight substrate-binding grooves of $SppA_{BS}$. (E) $SppA_{BS}$ C-terminal peptide, ^{326}SPRMMYLYAK^{335}, shown with the surrounding $2F_o - F_c$ map (blue mesh). The $F_o - F_c$ electron density maps (green mesh) are contoured at the 3σ level, and the $2F_o - F_c$ electron density map (blue mesh) is contoured at the 1σ level.

ionization mass spectrometry results show that the measured value of $SppA_{BS}^{\Delta 2-54}$ is 29166.70 Da and that of $SppA_{BS}^{\Delta 2-54}$K199A is 32949.55 Da. Incubating the wild-type and active site mutant $SppA_{BS}$ constructs together does not result in *trans* processing of the K199A mutated enzymes' C-termini.

Purification, Crystallization, and Structure Solution of a Protease Resistant Fragment of $SppA_{BS}^{\Delta 1-25}$K199A. Unlike the $SppA_{BS}^{\Delta 1-25}$ construct with a native active site, an enzyme with the lysine general base mutated to an alanine ($SppA_{BS}^{\Delta 1-25}$K199A) produced adequate amounts of protein to pursue crystallization trials. This protein was subjected to limited proteolysis using thermolysin to optimize crystallization. Electrospray ionization mass spectrometry analysis of the protease resistant fragment reveals a molecular mass of 26624.50 Da, which is consistent with the $SppA_{BS}$ sequence Leu51−Gly295 (Figure S1 of the Supporting Information). It was confirmed by amino-terminal sequencing analysis that the starting residue of the thermolysin-treated $SppA_{BS}$ (that was crystallized) is Leu51.[7] The protease resistant fragment produced ordered crystals within 2 weeks at room temperature (~296 K), under paraffin oil when using *tert*-butanol as the precipitant. The structure solution was obtained utilizing molecular replacement methods. The crystal belongs to space group $P2_12_12_1$ with eight molecules in the asymmetric unit [Matthews coefficient of 2.5 Å^3/Da (50.4% solvent)]. Strong electron density was observed for residues 56−295, except within a loop region (residues 73−81). The final refined 2.4 Å resolution structure contains 1865 residues with R_{cryst} and R_{free} values of 20.6 and 24.0, respectively. See Table 1 for crystal parameters, data collection statistics, and refinement statistics. For the sake of brevity, the protease resistant fragment of $SppA_{BS}^{\Delta 1-25}$K199A, which resulted in the crystal structure reported here, will henceforth be termed $SppA_{BS}$.

$SppA_{BS}$ Binds Its Own C-Termini within Its Active Sites. Analysis of the $SppA_{BS}$ electron density map revealed a continuous positive difference density forming a continuous circle around the concave binding groove of $SppA_{BS}$. Peptides

dx.doi.org/10.1021/bi4011489 | *Biochemistry* 2013, 52, 8811−8822

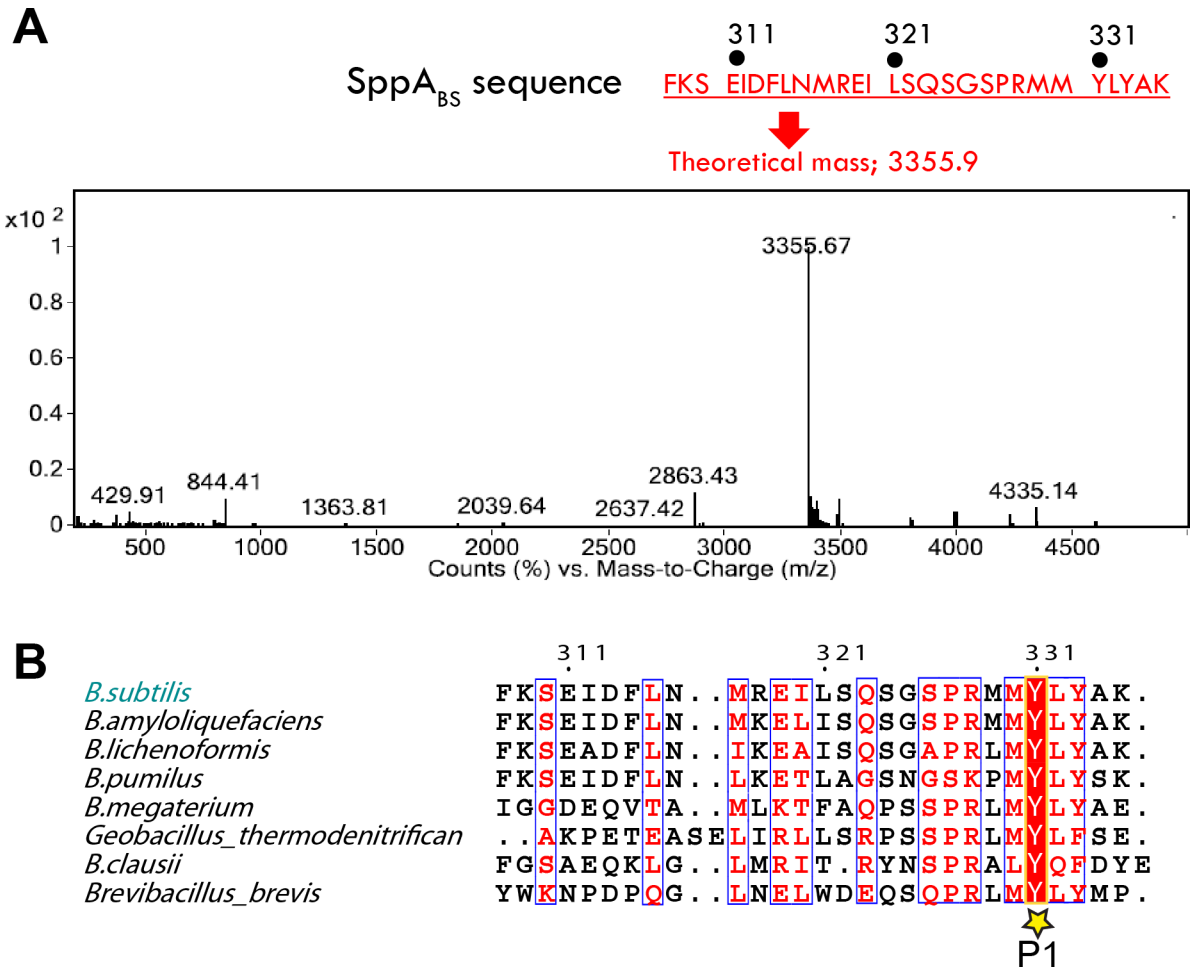


Figure 3. Identification of the SppA$_{BS}$-bound peptide. (A) Thermolysin-treated SppA$_{BS}$ was purified by size-exclusion chromatography and then analyzed using electrospray ionization mass spectrometry. The peak with the highest counts has a mass-to-charge ratio (m/z) of 3355.67. The theoretical molecular mass for the C-terminal peptide colored red is 3355.9 Da. (B) Alignment of the C-termini of SppA from different *Bacillus* species. Tyr331, bound in the S1 pocket of SppA$_{BS}$, is conserved among all *Bacillus* species. UniProt entries were as follows: O34525 for *B. subtilis*, A7Z7M8 for *Bacillus amyloliquefaciens*, Q65G59 for *Bacillus licheniformis*, A8FG77 for *Bacillus pumilus*, D5DMZ1 for *Bacillus megaterium*, A4IRT5 for *Geobacillus thermodenitrificans*, Q5WEC8 for *Bacillus clausii*, and C0ZL24 for *Brevibacillus brevis*. The conserved Tyr331 (red box) is marked with a star and labeled as P1.

were built into the density in each of the eight binding grooves (Figure 2).

Because no substrate was added during the purification or crystallization, the peptide seen inside the binding site was hypothesized to be an SppA$_{BS}$ fragment resulting from thermolysin digestion or self-processing. The strongest difference density was located near the SppA$_{BS}$ catalytic residues. Side chain electron density was compared to the sequence for the N-terminal and C-terminal regions of SppA$_{BS}$, and the C-terminal ^{331}YLY^{333} sequence was found to be a good candidate for the region of density near the catalytic residues (substrate positions P1−P2′). Electrospray ionization mass spectrometry was utilized to determine the mass of the bound peptide. The most abundant peak corresponds to m/z 3355.67 (Figure 3). The C-terminal sequence of SppA$_{BS}$, "^{308}FKSEIDFLNMREILSQSGSPRMMYLYAK^{335}", has a theoretical mass of 3355.9 Da (Lys335 is the last residue of SppA$_{BS}$). We confirmed the identity of the peptide by omit map analysis (Figure S3 of the Supporting Information) and by liquid chromatography−mass spectrometry−mass spectrometry fragmentation analysis (Figure S4 of the Supporting Information). Eight peptides of the C-terminal sequence (residues 326−335; underlined portion of the sequence shown above) were built into the difference electron density map within each of the eight active sites of SppA$_{BS}$. The modeled peptides fit and refined well within the electron density (Figure 2). The position of the bound peptides relative to the catalytic residues reveals that position P1 corresponds to Tyr331, position P3 is Met329, and position P2′ is Tyr333 (Figure 4). The side chains for these residues make intimate contact with substrate specificity pockets S1, S3, and S2′ on the SppA$_{BS}$ surface (Figure 4A). A protein sequence alignment shows that Tyr331 is highly conserved among *Bacillus* species (Figure 3B).

S1, S3, and S2′ Substrate Specificity Pockets in SppA$_{BS}$. The deep S1 substrate specificity pocket has hydrophobic walls and a polar base and thus accommodates nicely the aromatic side chain of the P1 residue, Tyr331. The S1 pocket is formed by atoms from residues Gly113, Gly114, Ser119, Ser147, Gly148, Gly171, Phe227, Tyr151, Val172, Val116, Ser223, and Glu164 (Figure 4A). The side chain of the P3 residue, Met329, is located within the S3 pocket formed by atoms from residues Leu166, Val220, Ile173, Met174, Asp252, Met219, and also Glu164, Ser223, Val116, and Val172. The last four residues also contribute to the S1 pocket. There is a

dx.doi.org/10.1021/bi4011489 | *Biochemistry* 2013, 52, 8811−8822

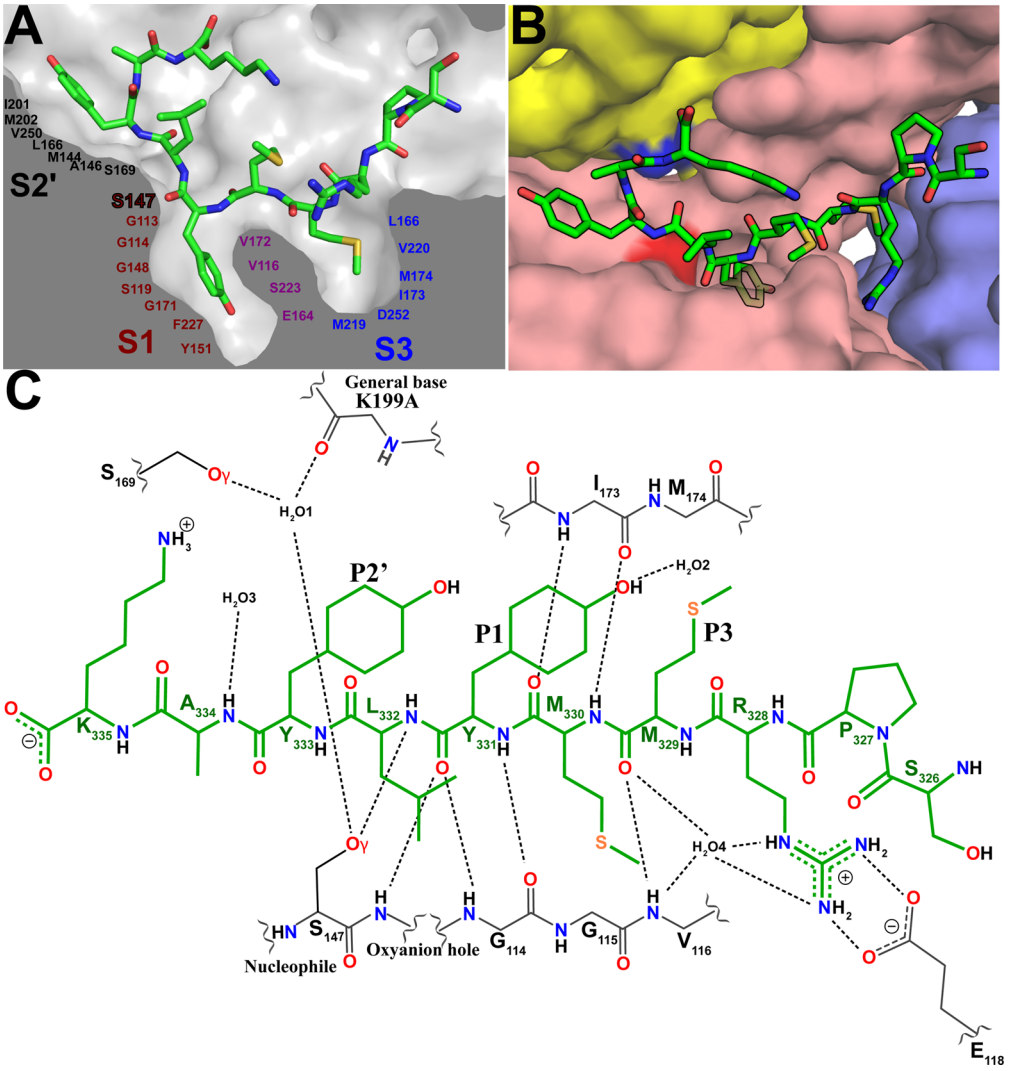


Figure 4. Interactions between the bound C-terminal peptide and the substrate-binding groove of $SppA_{BS}$. (A) Cross-sectional view of the $SppA_{BS}$ substrate binding pocket shown in molecular surface representation. The C-terminal peptide is shown as sticks (carbons colored green, oxygens red, sulfurs yellow, and nitrogens blue). Residues involved in making the S2′ pocket are labeled in black. The S1 pocket is labeled in red and the S3 pocket in blue. Residues that contribute to both the S1 and S3 pockets are labeled in purple. (B) C-Terminal peptide bound to the substrate-binding groove and active site of $SppA_{BS}$, which is created by three protomers. The yellow-colored protomer donates the general base, lysine (blue); the salmon-colored protomer provides the S1 substrate specificity pocket with the nucleophile, serine (red); and the light blue- and salmon-colored protomers come together to form the S3 substrate specificity pocket. (C) Schematic of the interactions between the $SppA_{BS}$ C-terminal peptide (green) and the $SppA_{BS}$ substrate-binding groove residues (black). The general base (K199A), nucleophile (S147), oxyanion hole, water (H_2O), and P1, P3, and P2′ positions are labeled in black. Dashed lines represent hydrogen bonds. Oxygens are colored red, nitrogens blue, and sulfurs yellow.

shallow hydrophobic depression, the S2′ substrate specificity pocket, created by atoms from residues Ile201, Met202, Val250, Leu166, Met144, Ala146, and Ser169 that accommodates one side of the aromatic side chain of the P2′ residue, Tyr333.

Interactions of the C-Terminal Peptide with the $SppA_{BS}$ Substrate-Binding Groove. Three protomers come together to form each active site within the $SppA_{BS}$ structure, and the C-terminal peptides interact with all three protomers (Figure 4B). The yellow-colored protomer in Figure 4B provides the general base (K199A), while the salmon-colored protomer provides the nucleophile Ser147 and forms the S1 and S2′, substrate specificity pockets as well as most of the hydrogen bonding interactions with the bound substrate. The S3 pocket is located at the interface between the light blue- and salmon-colored protomers. The P3 (Met329) and P4 (Arg328) residues of the peptide are located at this interface. The side chains of the P2 (Met330) and P1′ (Leu332) residues face the solvent.

The eight peptides occlude 6886 Å^2 of surface area within the $SppA_{BS}$ substrate-binding grooves, ~861 Å^2 in each binding groove. The peptides are bound in an antiparallel β-sheet fashion with respect to the residues that line the substrate-binding groove (Figure 4C and Figure S5 of the Supporting Information). Oγ of nucleophile Ser147 is in the proximity of the would-be scissile bond. It is within 2.8 Å of the carbonyl carbon of the P1 residue, Tyr331, and within 3.0 Å of the main chain NH groups of the P1′ residue. The reported distances are an average from the eight active sites. The carbonyl oxygen of Tyr331 points into the oxyanion hole created by the main chain NH groups of Gly114 and Gly148. Gly148 is located just after the serine nucleophile (Ser147), but Gly114 is located within the β-strand that forms an antiparallel interaction with the bound peptide (substrate).

dx.doi.org/10.1021/bi4011489 | *Biochemistry* 2013, 52, 8811−8822

There are 11 direct hydrogen bonding interactions between the C-terminal peptide and the substrate binding site, including the oxyanion hole interaction and Oγ of the nucleophile interacting with the NH group of the P1 residue (Table 2).

Table 2. Direct Hydrogen Bonding Interactions between the C-Terminal Peptide and the Substrate-Binding Groove of $SppA_{BS}$

C-terminal peptide atom	SppA atom	distance[a](Å)
Tyr331 O	Gly148 N	2.9
Tyr331 O	Gly114 N	2.9
Tyr331 N	Gly114 O	2.9
Tyr331 Oη	Ser223 O	2.8
Met329 O	Val116 N	3.3
Met330 O	Ile173 N	3.0
Met330 N	Ile173 O	3.0
Leu332 N	Ser147 Oγ	3.0
Tyr333 N	Pro112 O	3.0
Arg328 $Nη_1$	Glu118 $Oε_1$	2.8
Arg328 $Nη_2$	Glu118 $Oε_2$	2.9

[a]Average value from the eight bound peptides.

There is a salt bridge between the P4 residue, Arg328 Nη1 and Nη2, of the peptide and Glu118 Oε1 and Oε2 of the protomer. Oη of Tyr331 is within hydrogen bonding distance of the carbonyl oxygen of Ser223. Oη of Tyr333 is within hydrogen bonding distance of the carbonyl oxygen of Pro327, which belongs to the peptide that is bound within the neighboring binding groove. There are four conserved waters present in all eight active sites. As seen in Figure 4C and Figure S5 of the Supporting Information, the waters create a bridge between the peptide and protomer (Table 3).

Table 3. Hydrogen Bonding Interactions of Conserved Waters within the Substrate-Binding Groove of the $SppA_{BS}$ C-Terminal Peptide Complex

water	SppA atom and C-terminal peptide atom	distance[a](Å)
water 1	Lys199Ala O	3.0
	Ser147 Oγ	3.6
	Ser169 Oγ	3.1
water 2	Tyr331 Oη	2.7
	Glu164 $Oε_2$	2.8
	Tyr151 Oη	2.6
water 3	Lys199Ala O	3.4
	Ile201 O	2.6
	Ala334 N	3.0
water 4	Tyr117 N	3.2
	Arg328 $Nη_2$	2.7
	Arg328 Nε	3.8
	Met329 O	3.1
	Val116 N	3.3

[a]Average value from the eight bound peptides.

Structural Changes within $SppA_{BS}$ upon Binding of a Peptide. The superposition of the $SppA_{BS}$ structures, with and without bound peptides within the active sites (octamer to octamer), has a root-mean-square deviation of 0.26 Å. Differences were observed in β-strand 5 and α-helix 4 of the extension region that forms the roof of the octameric dome (Figure 5). There are also differences observed within the active site (Figure 5C). First, in the $SppA_{BS}$ structure with the bound

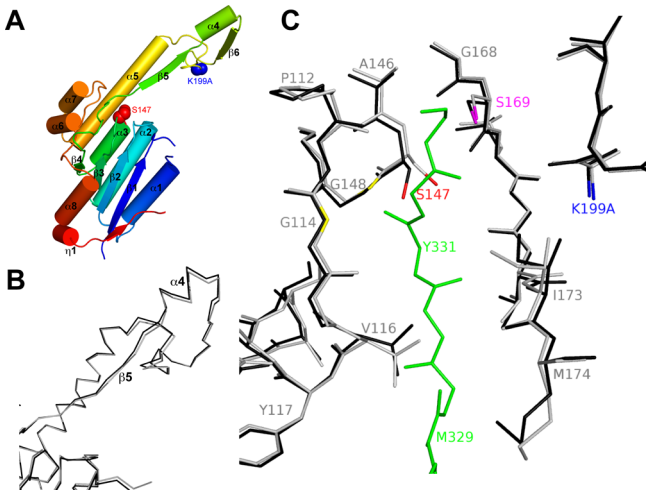


Figure 5. Comparison of $SppA_{BS}$ structures with and without C-terminal peptides bound within the active sites. (A) Cartoon representation of an $SppA_{BS}$ monomer with the secondary structural elements labeled. The protein is colored from blue at the N-terminus to red at the C-terminus. (B) Cα trace view of the $SppA_{BS}$ structure with the free active site (black) superimposed on the C-terminal peptide-bound $SppA_{BS}$ structure (gray). The figure focuses on the extension region of $SppA_{BS}$. (C) C-Terminal peptide-bound $SppA_{BS}$ structure (gray) and free active site structure of $SppA_{BS}$ (black). The C-terminal peptide (only the main chain shown for the sake of clarity) is colored green. Oγ of the nucleophile Ser147 is colored red; Cβ of the general base K199A is colored blue, and the general base orienting Oγ of residue Ser167 is colored magenta. The NH groups of the oxyanion hole residues, Gly114 and Gly148, are colored yellow on the gray structure.

C-terminus, the hydroxyl group of Ser147 is rotated approximately 70° away from the peptide bound in the active site. Second, in the structure with the free active site, alternate rotamer conformations were observed for Glu164 and Arg254, which are part of the S3 substrate specificity pocket, but only one conformation is observed for these residues in the complex structure.[7] Third, in the $SppA_{BS}$ structure with the C-terminus bound, residues Ala146−Gly148, Pro112−Val116, Ile173, and Met174 have moved such that the substrate-binding groove is narrower than the free active site structure. Each of these residues is involved in hydrogen bonding interactions with the bound peptide (Figure 4C).

Potential Role of Tyr331 in C-Terminal Recognition and SppA Stability. As seen in Figure 1, and as was seen in activity assays using fluorometric peptides,[7] mutating the lysine general base (Lys199) to an alanine creates an enzyme with significantly suppressed activity, yet when the K199A mutant enzyme is incubated for several days at room temperature, the enzyme self-degrades (Figure 6). We observe that mutating Tyr331 (the residue that occupies the S1 binding pocket within the complex) to an alanine prevents degradation, even after 6 days at room temperature.

The C-Terminus of $SppA_{BS}$ Is Not Essential for Oligomerization. To investigate whether the C-terminus of $SppA_{BS}$ plays a role in oligomeric assembly, we prepared three different C-terminal truncations of $SppA_{BS}$. The Δ329−335 truncation ends before the conserved ^{329}MMYLYK^{335} sequence; the Δ307−335 truncation is based on the peptide identified by mass spectrometry, and the Δ295−335 truncation is based on the last residue with clear electron density in the structure. To determine the oligomeric state, the three C-

dx.doi.org/10.1021/bi4011489 | Biochemistry 2013, 52, 8811−8822

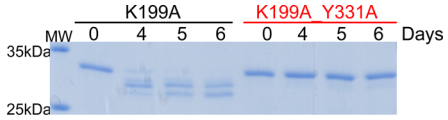


Figure 6. Mutating Tyr331 to alanine increases the stability of $SppA_{BS}$. $SppA_{BS}^{\Delta 2-54}$K199A (K199A) and $SppA_{BS}^{\Delta 2-54}$K199A/Y331A (K199A_Y331A) proteins were each incubated at room temperature, and aliquots were collected after 0, 4, 5, and 6 days. Loading dye was added to each aliquot and then the mixture boiled to stop the self-processing reactions. Samples were run on a 13.5% SDS–PAGE gel, followed by staining with PageBlue stain. The far left lane shows the protein molecular mass markers. Note that these $SppA_{BS}$ enzymes did not undergo thermolysin treatment.

terminally truncated constructs were subjected to limited proteolysis following the protocol that resulted in the $SppA_{BS}^{\Delta 1-25}$K199A's octameric structure. All three C-terminal truncations showed approximately the same size-exclusion chromatographic elution profile as thermolysin-treated octameric $SppA_{BS}^{\Delta 1-25}$K199A, which suggests that the C-terminus is not essential for oligomerization of $SppA_{BS}$ (Figure 7). The

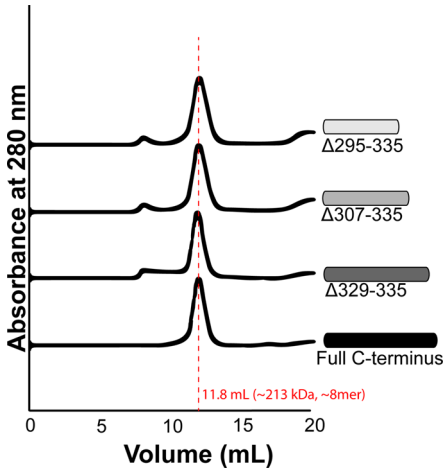


Figure 7. Truncation analysis suggests that the C-terminus of $SppA_{BS}$ is not essential for oligomeric assembly. A size-exclusion chromatographic (SEC) analysis of thermolysin-treated $SppA_{BS}$ and thermolysin-treated $SppA_{BS}$ with C-terminal truncations is shown. The chromatogram for each construct is shown as elution volume vs absorbance at 280 nm (see Figure S7 of the Supporting Information for the SEC standard calibration curve, Table S2 of the Supporting Information for a list of elution volumes and molecular masses, and Figure S8 of the Supporting Information for the SDS–PAGE gel analysis of the SEC elution fractions). A schematic of each construct is shown beside the chromatogram. Each construct lacks the N-terminal transmembrane segment (residues 1–25) and has the catalytic lysine mutated to alanine (K199A). The red dashed line shows the peak elution at 11.8 mL.

stability of the $SppA_{BS}$ C-terminal truncation constructs appeared to be the same as that of the $SppA_{BS}$ construct with the full C-terminus. A very small population of aggregate protein is observed in the chromatograms for the truncation mutants.

The Synthetic Peptide NH_3-SPRMMYLYAK-COOH Can Function as an $SppA_{BS}$ Competitive Inhibitor. Kinetic constants for $SppA_{BS}$ were determined using a fluorogenic lipopeptide substrate (dodecanoyl-NGEVAKA-MCA). The measured values of K_M and k_{cat} are 17.9 ± 3.2 μM and (7.8 ± 0.5) × 10^{-2} s^{-1}, respectively. A peptide (NH_3-SPRMMYLYAK-COOH) was synthesized corresponding to the sequence of the modeled C-terminal peptide bound in the $SppA_{BS}$ substrate-binding groove. Kinetic analysis shows that this peptide is capable of functioning as an $SppA_{BS}$ competitive inhibitor. Using six different peptide concentrations, the V_{max} and k_{cat} at each inhibitor concentration remained approximately the same, whereas the K_M values increased as the inhibitor concentration increased (Figure 8). The K_i value was determined to be 2.1 ± 0.4 μM.

DISCUSSION

In this study, we have shown that $SppA_{BS}$ with native active sites processes its own C-termini, and that an active site mutation (K199A) slows this C-terminal self-processing (Figure 1). Incubating the wild-type active site and mutant enzymes together does not result in processing of the K199A mutated enzymes' C-termini, suggesting that the processing occurs within the enzyme's own octameric chamber (*cis*), and not in an intercomplex (*trans*) fashion. It also appears that the C-terminus is processed quite soon after the octameric protease is assembled, given that $SppA_{BS}$ with native active sites is cleaved before the enzyme can be purified.

Substrate-Binding Groove Interactions Provide Order to the Last 10 Residues of a Flexible C-Terminus. Adjustments to the purification (conducting limited proteolysis while $SppA_{BS}$ is bound to the Ni^{2+}-NTA column instead of after elution) and the crystallization procedure (growing crystals under paraffin oil without detergent or MPD) versus those used previously for $SppA_{BS}$[7] have led to a structure that reveals how this self-compartmentalized membrane-bound enzyme binds its own C-termini within its eight active sites. A continuous circle of positive difference electron density was observed in the $SppA_{BS}$ substrate-binding grooves inside the octameric complex (Figure 2C,D). The density in each binding groove is consistent with the sequence for the last 10 residues at the C-terminus of $SppA_{BS}$ (residues 326–335) (Figure 2E), but the modeled peptides are shorter than the peptide we identified by mass spectrometry (residues 308–335) (Figure 3). It is likely that the N-termini of the bound peptides (residues 308–325) reside within the cavity, but a lack of specific noncovalent interactions with $SppA_{BS}$ may result in disorder and therefore a lack of observable electron density. This enzyme's substrate-binding groove is fairly short and completely occupied with electron density from the 10 residues closest to the enzyme's C-termini (residues 326–335).

The structure directly identifies the S1, S3, and S2′ substrate specificity binding sites, with Tyr331, Met329, and Tyr333, respectively, occupying these pockets (Figure 4). Previous activity assays with a series of fluorogenic peptide substrates reveal that tyrosine is one of most preferred residues at the P1 position.[7] Interestingly, we find that mutating Tyr331 to alanine slows the long-term degradation of the K199A mutant enzyme (Figure 6). This result suggests that the enzyme lacking the general base still has measurable catalytic activity and is capable of slow self-processing, and that Tyr331 is important for C-terminal recognition and processing.

Previous research on classical serine proteases such as subtilisin and trypsin shows that mutating the general base to alanine results in a large decrease in activity (10^6-fold) but not in a completely inactive enzyme.[27−29] A recent example in the literature is nucleoporin, which undergoes autoproteolysis even after the general base is mutated to alanine, albeit at a much slower rate.[30]

dx.doi.org/10.1021/bi4011489 | *Biochemistry* 2013, 52, 8811−8822

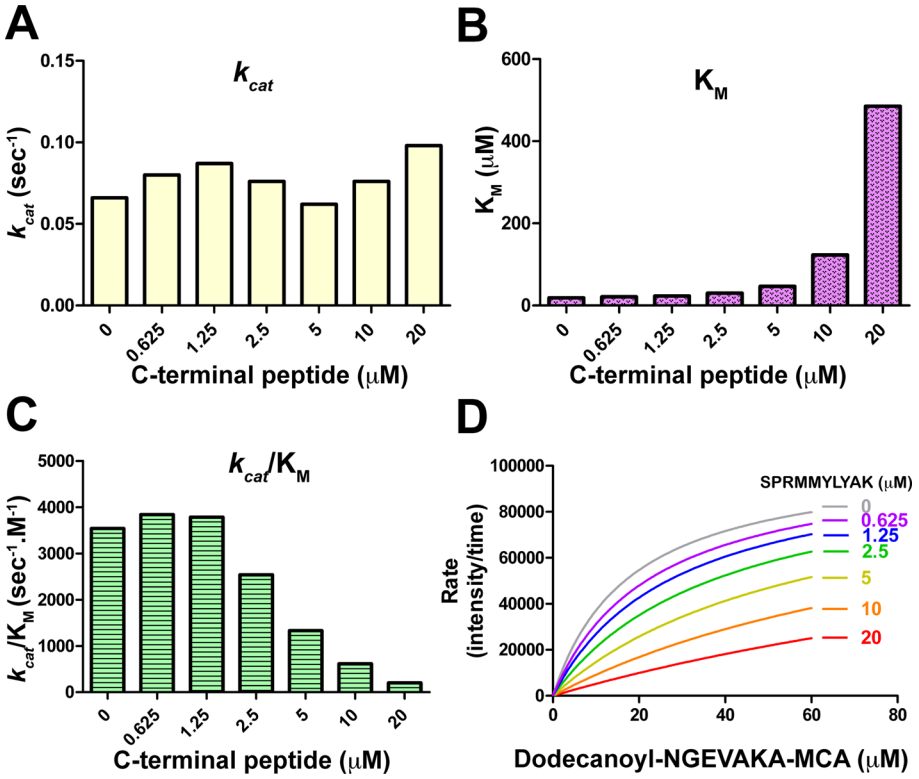


Figure 8. Synthetic peptide based on the $SppA_{BS}$ C-terminal sequence that can compete for the $SppA_{BS}$ substrate binding sites. The $SppA_{BS}$-catalyzed cleavage of the fluorogenic substrate dodecanoyl-NGEVAKA-MCA was measured in the presence of different concentrations of the synthetic peptide ^{326}SPRMMYLYAK^{335}, which corresponds to the C-terminus of $SppA_{BS}$. The bar graphs show (A) k_{cat}, (B) K_M, and (C) k_{cat}/K_M values at different concentrations of the $SppA_{BS}$ C-terminal peptide. (D) Substrate cleavage rate (fluorescence intensity over time) plotted vs substrate concentration. The colored graph lines represent the fluorescence measured at different concentrations of the C-terminal peptide (competitive inhibitor).

Could another residue within the $SppA_{BS}$ active site function as a general base in the absence of the ε-amino group of Lys199? From the crystal structure, we can see that there is no other titratable residue within hydrogen bonding distance of the serine nucleophile that would be capable of functioning as a general base, but there is an ordered water molecule (water 1, Figure 4C) positioned within the active site such that it could possibly be functioning as a base, albeit a poor one.

Is the C-Terminus of $SppA_{BS}$ a Type II Intramolecular Chaperone? Many proteases are synthesized with propeptides (or pro-segments) that are cleaved off from the mature folded enzyme. The pro-segment is often located at one of the termini. For example, secreted proteases in the *Bacillus* genus, such as subtilisin, have a propeptide at the N-terminus that is not present in the final mature form of the enzyme.[31] Propeptides often function as intramolecular chaperones that facilitate protein folding.[32] Type I intramolecular chaperones are located at the N-terminus and function in protein folding at the tertiary structure level, while type II intramolecular chaperones are located at the C-terminus and function at the quaternary structure level (oligomeric assembly). The proteasome is an example of a self-compartmentalized (chambered) protease that utilizes the C-termini in oligomerization (a type II intramolecular chaperone).[33] We have made a number of $SppA_{BS}$ C-terminal truncations to investigate whether the C-terminus of $SppA_{BS}$ is essential for oligomeric assembly. Each of the truncated constructs was capable of forming the octameric complex as accessed by analytical size-exclusion chromatography, and therefore, it appears that the C-terminus of $SppA_{BS}$ does not function as a type II intramolecular chaperone (Figure 7).

Does the C-Terminus of $SppA_{BS}$ Play a Regulatory Role? Zymogens (or pro-enzymes) are synthesized as inactive precursors and are activated by the cleavage of a pro-segment of the enzyme, usually when the enzyme reaches its intended final destination and milieu.[34] Carboxypeptidase Y, from the yeast *Saccharomyces cerevisiae*, is a Ser/His/Asp catalytic triad-utilizing protease that is synthesized as a precursor form with an N-terminal propeptide extension.[35] It is inhibited, in a competitive manner, by a nine-residue peptide corresponding to the N-terminal propeptide region. Similarly, we showed that a synthetic peptide, corresponding to the C-terminal end of $SppA_{BS}$, acts as a competitive inhibitor (Figure 8). This suggests that the C-terminus of $SppA_{BS}$ may function to regulate the activity of this compartmentalized membrane-bound protease until it is properly assembled within the bacterial membrane. This is likely a critical regulatory feature of this enzyme in that our previous studies have shown that $SppA_{BS}$ is capable of digesting proteins as well as peptides.[7] The structure of the $SppA_{BS}$ complex can be used as a starting point for SppA inhibitor design.

Modeling Studies Suggest That the C-Termini Can Reach the Substrate-Binding Grooves Prior to Limited Proteolysis. Limited proteolysis was required for the optimization of $SppA_{BS}$ crystallization. This resulted in trimming at both the N- and C-termini of $SppA_{BS}$. One might then ask whether the C-termini can reach the substrate-binding grooves without being released by limited proteolysis.

dx.doi.org/10.1021/bi4011489 | *Biochemistry* 2013, 52, 8811−8822

Our observation of self-processing in *cis* (Figure 1) suggests that they can. In addition, our structural analysis suggests the C-termini of $SppA_{BS}$ are more than long enough to reach the binding sites. The peptides (residues 326−335) that we built into the active site electron density correspond to the C-terminus of each protomer. The last residue with clear electron density in the globular domain of each protomer is Gly295. This residue is located at the base of the $SppA_{BS}$ molecule, near the wide opening at the membrane association surface (Figure 9). The first residue that shows clear electron density within the bound C-terminal peptide is Ser326. On the basis of our modeling studies, the minimal number of residues that are required to reach the S1 substrate specificity pocket from Gly295 is 13 residues, assuming that there is no helical structure in the C-terminus after Gly295. Thus, 30 residues is more than enough for the C-terminus to reach the binding groove that sits at the protomer interface with either its left or right neighbor. The direct distance between Gly295 at the base of the complex and Ser326 within the bound peptide is slightly shorter (32 Å vs 44 Å) for the path to the left neighbor (Figure 9). Because of the high local effective concentration, it is most likely the C-termini were bound within the active sites before the limited proteolysis treatment, and therefore, the bound peptides are most likely those originating from a neighboring protomer chain.

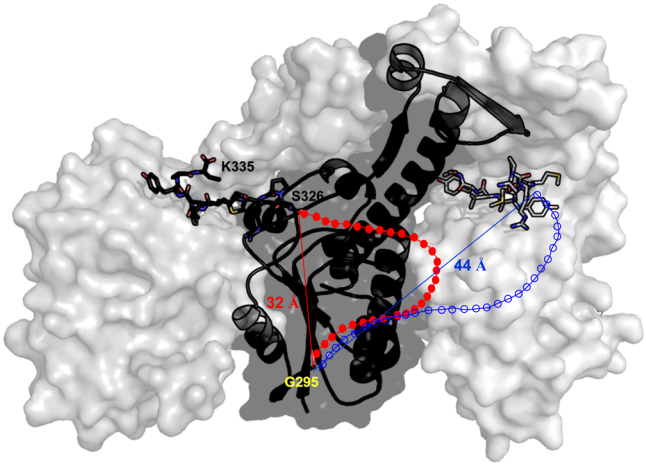


Figure 9. Proposed model for how the C-termini of $SppA_{BS}$ occupy the active sites of the $SppA_{BS}$ protomers. Only three of the eight $SppA_{BS}$ protomers are shown for the sake of clarity. The view is from inside of the $SppA_{BS}$ chamber. The C-terminus of the black $SppA_{BS}$ protomer is bound within the substrate-binding groove that sits at the interface of each adjacent protomer (gray). Loops colored red and blue are drawn to connect the C-terminal peptides. Each sphere represents one residue. Ser326 and Lys335 are the N- and C-terminal residues of the peptide, respectively. Gly295 (yellow) is the last residue within the core region of $SppA_{BS}$ that has clear electron density. The distance between Gly295 and Ser326 is shown.

Does the C-Terminus Play the Same Role in SppA from All Bacteria? There is a significant amount of conservation within the C-termini among SppA from Gram-negative bacteria, although it is difficult to identify a uniformly conserved C-terminal residue such as the conserved tyrosine that we see in the *Bacillus* genus (Tyr331, in *B. subtilis*). Previous kinetic analysis using a series of short fluorogenic peptide substrates showed that $SppA_{BS}$ is capable of cleaving peptides with a variety of residues at the P1 position, tyrosine being one of the most preferred residues. $SppA_{EC}$ showed a significantly narrower substrate preference, preferring large aliphatic residues at the P1 position.[7] Future studies will investigate if SppA from other bacterial species also self-process their C-termini.

Conclusion. We have shown that $SppA_{BS}$ processes its own C-termini. We have determined a crystal structure of $SppA_{BS}$ that reveals the C-termini bound within the substrate-binding grooves, identifying its S1, S3, and S2′ specificity binding pockets. The P1 and P3 residues of the bound peptide agree with the observed substrate preference in previous $SppA_{BS}$ kinetic assays.[7] We have used C-terminal truncation mutants to show that the C-terminus is not essential for $SppA_{BS}$ oligomeric assembly. We show that a synthetic peptide based on the bound peptide is able to compete with a fluorogenic peptide substrate for the $SppA_{BS}$ active sites, suggesting the C-termini may function in the regulation of proteolysis and that this peptide could be used as a starting point for inhibitor development.

■ ASSOCIATED CONTENT

Supporting Information

Electrospray ionization mass spectrometry data (Figures S1 and S6 and Table S1), liquid chromatography−tandem mass spectrometry fragmentation analysis data (Figure S4), size-exclusion chromatography data (Figures S7 and S8 and Table S2), electron density omit map analysis (Figures S2 and S3), and a stereoview of the active site (Figure S5). This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

Atomic coordinates and structure factors have been deposited in the Protein Data Bank as entry 4KWB.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: mpaetzel@sfu.ca. Telephone: (778) 782-4230. Fax: (778) 782-5583.

Funding

This work was supported in part by the Canadian Institute of Health Research (to M.P.) and the National Science and Engineering Research Council of Canada (to M.P.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the staff at beamline 08ID-1 at the Canadian Light Source, Saskatoon, SK (especially Shaunivan Labiuk and Julien Cotelesage), for their help with data collection. The Canadian Light Source is supported by NSERC, NRC, CIHR, and the University of Saskatchewan. We also thank Hongwen Chen from the spectroscopy facility of the Department of Chemistry of Simon Fraser University for his help with the mass spectrometry analysis. We also thank Matt Willetts from Bruker Daltonics Inc. for his advice on the mass spectroscopy analysis.

■ ABBREVIATIONS

PDB, Protein Data Bank; $SppA_{BS}$, protease resistant fragment of signal peptide peptidase A from *B. subtilis* ($SppA_{BS}^{\Delta 1-25}$K199A) that was used for crystallization; $SppA_{EC}$, signal peptide peptidase A from *E. coli*.

REFERENCES

(1) von Heijne, G. (1990) The signal peptide. *J. Membr. Biol. 115*, 195−201.

(2) du Plessis, D. J., Nouwen, N., and Driessen, A. J. (2011) The sec translocase. *Biochim. Biophys. Acta 1808*, 851−865.

(3) Paetzel, M., Karla, A., Strynadka, N. C., and Dalbey, R. E. (2002) Signal peptidases. *Chem. Rev. 102*, 4549−4580.

(4) Ichihara, S., Beppu, N., and Mizushima, S. (1984) Protease IV, a cytoplasmic membrane protein of *Escherichia coli*, has signal peptide peptidase activity. *J. Biol. Chem. 259*, 9853−9857.

(5) Bolhuis, A., Matzen, A., Hyyrylainen, H. L., Kontinen, V. P., Meima, R., Chapuis, J., Venema, G., Bron, S., Freudl, R., and van Dijl, J. M. (1999) Signal peptide peptidase- and ClpP-like proteins of *Bacillus subtilis* required for efficient translocation and processing of secretory proteins. *J. Biol. Chem. 274*, 24585−24592.

(6) Kim, A. C., Oliver, D. C., and Paetzel, M. (2008) Crystal structure of a bacterial signal peptide peptidase. *J. Mol. Biol. 376*, 352−366.

(7) Nam, S. E., Kim, A. C., and Paetzel, M. (2012) Crystal structure of *Bacillus subtilis* signal peptide peptidase A. *J. Mol. Biol. 419*, 347−358.

(8) Wang, P., Shim, E., Cravatt, B., Jacobsen, R., Schoeniger, J., Kim, A. C., Paetzel, M., and Dalbey, R. E. (2008) *Escherichia coli* signal peptide peptidase A is a serine-lysine protease with a lysine recruited to the nonconserved amino-terminal domain in the S49 protease family. *Biochemistry 47*, 6361−6369.

(9) Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. papain. *Biochem. Biophys. Res. Commun. 27*, 157−162.

(10) Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook* (Walker, J. M., Ed.) pp 571−607, Humana Press, Totowa, NJ.

(11) Otwinowski, Z., and Minor, W. (1993) Denzo and Scalepack. In *International Tables for Crystallography* (Sawyer, N., Isaacs, N., and Baily, S., Eds.) Vol. *F*, pp 56−62, Daresbury Laboratory, Warrington, U.K.

(12) Kantardjieff, K. A., and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci. 12*, 1865−1871.

(13) McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr. 61*, 458−464.

(14) Stein, N. (2008) CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr. 41*, 641−643.

(15) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213−221.

(16) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. 60*, 2126−2132.

(17) Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D67*, 355−367.

(18) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D67*, 235−242.

(19) Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D62*, 439−450.

(20) Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol. 374*, 300−321.

(21) DeLano, W. L. (2002) *The PyMOL molecular graphics system*, DeLano Scientific, San Carlos, CA.

(22) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst. 26*, 283−291.

(23) Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF: A program to identify and analyze structural motifs in proteins. *Protein Sci. 5*, 212−220.

(24) Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol. 55*, 379−400.

(25) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774−797.

(26) Richards, F. M. (1977) Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng. 6*, 151−176.

(27) Ekici, O. D., Paetzel, M., and Dalbey, R. E. (2008) Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci. 17*, 2023−2037.

(28) Corey, D. R., and Craik, C. S. (1992) An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J. Am. Chem. Soc. 114*, 1784−1790.

(29) Carter, P., and Wells, J. A. (1988) Dissecting the catalytic triad of a serine protease. *Nature 332*, 564−568.

(30) Hodel, A. E., Hodel, M. R., Griffis, E. R., Hennig, K. A., Ratner, G. A., Xu, S., and Powers, M. A. (2002) The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. *Mol. Cell 10*, 347−358.

(31) Sone, M., Falzon, L., and Inouye, M. (2005) The role of tryptophan residues in the autoprocessing of prosubtilisin E. *Biochim. Biophys. Acta 1749*, 15−22.

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

(33) Murata, S., Yashiroda, H., and Tanaka, K. (2009) Molecular mechanisms of proteasome assembly. *Nat. Rev. Mol. Cell Biol. 10*, 104−115.

(34) Khan, A. R., and James, M. N. (1998) Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci. 7*, 815−836.

(35) Nagayama, M., Kuroda, K., and Ueda, M. (2012) Identification of interaction site of propeptide toward mature carboxypeptidase Y (mCPY) based on the similarity between propeptide and CPY inhibitor (IC). *Biosci., Biotechnol., Biochem. 76*, 153−156.

(36) Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760−763.

(37) French, S., and Wilson, K. (1978) On the treatment of negative intensity observations. *Acta Crystallogr. A34*, 517−525.