Ultra-high and High Resolution Structures and Mutational Analysis of Monomeric *Streptococcus pyogenes* SpeB Reveal a Functional Role for the Glycine-rich C-terminal Loop

Gonzalo E. González-Páez and Dennis W. Wolan

From the Departments of Molecular and Experimental Medicine and Chemical Physiology, The Scripps Research Institute, La Jolla, California 92037

**Background:** SpeB is a protease secreted from the pathogenic bacterium *Streptococcus pyogenes*. Apo and inhibitor complex structures of SpeB reveal conformational changes associated with activation and binding. The glycine-rich C-terminal active site loop is required for substrate recognition and product release. Crystal structures and mutagenesis afford insights into the specificity of SpeB and the functional role of the active site loops in binding biological substrates.

**Results:** Apo and inhibitor complex structures of SpeB reveal conformational changes associated with activation and binding. The glycine-rich C-terminal active site loop is required for substrate recognition and product release. SpeB is a protease secreted from the pathogenic bacterium *Streptococcus pyogenes* and has been studied as a potential virulence factor since its identification almost 70 years ago. Here, we report the crystal structures of apo mature SpeB to 1.06 Å resolution as well as complexes with the general cysteine protease inhibitor trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane and a novel substrate mimetic peptide inhibitor. These structures uncover conformational changes associated with maturation of SpeB from the inactive zymogen to its active form and identify the residues required for substrate binding. With the use of a newly developed fluorogenic tripeptide substrate to measure SpeB activity, we determined IC$_{50}$ values for trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane and our new peptide inhibitor and the effects of mutations within the C-terminal active site loop. The structures and mutational analysis suggest that the conformational movements of the glycine-rich C-terminal loop are important for the recognition and recruitment of biological substrates and release of hydrolyzed products.

**Conclusion:** The glycine-rich C-terminal active site loop is required for substrate recognition and product release. Cysteine protease SpeB is secreted from *Streptococcus pyogenes* and has been studied as a potential virulence factor since its identification almost 70 years ago. Here, we report the crystal structures of apo mature SpeB to 1.06 Å resolution as well as complexes with the general cysteine protease inhibitor trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane and a novel substrate mimetic peptide inhibitor. These structures uncover conformational changes associated with maturation of SpeB from the inactive zymogen to its active form and identify the residues required for substrate binding. With the use of a newly developed fluorogenic tripeptide substrate to measure SpeB activity, we determined IC$_{50}$ values for trans-epoxysuccinyl-L-leucylamido(4-guanidino)butane and our new peptide inhibitor and the effects of mutations within the C-terminal active site loop. The structures and mutational analysis suggest that the conformational movements of the glycine-rich C-terminal loop are important for the recognition and recruitment of biological substrates and release of hydrolyzed products.

The Gram-positive bacterium *Streptococcus pyogenes* is a highly disseminated human-specific pathogen that causes a wide range of diseases, including pharyngitis, tonsillitis, common skin rashes, rheumatic and scarlet fever, and meningitis. The majority of these bacterial infections remain localized to skin and soft tissue; however, many *S. pyogenes* strains can penetrate deeper into tissues and induce severely invasive diseases, such as sepsis, necrotizing fasciitis, and toxic shock syndrome, with high associated mortality. Despite our current arsenal of antibiotic treatments, *S. pyogenes* remains at the forefront of bacterial pathogens that cause significant morbidity and mortality throughout the world. As such, it is imperative to identify new approaches and methods to control these common, but often fatal, infections. *S. pyogenes* secretes a variety of documented virulence factors into the extracellular space that promote infectivity and down-regulate host immune responses. The cysteine protease SpeB is one such factor that has been extensively studied for its ability to promote *S. pyogenes* infection.

SpeB, or streptopain, is a highly conserved secreted cysteine protease that is found in the majority of *S. pyogenes* strains. Production, secretion, and activation of SpeB is highly regulated, and its proteolytic activity has been linked to immunomodulating activities during infection, including 1) degradation of host immunoglobulins to promote immune system evasion (4); 2) cleavage of the cytokine precursor interleukin-1β (IL-1β) to its mature form, resulting in inflammation and septic shock (5); and 3) release of the peptide hormone bradykinin from the precursor H-kininogen to produce hypotension during septic shock (6). Additionally, SpeB has been implicated in degradation of fibronectin and vitronectin, two host extracellular matrix proteins involved in tissue integrity (7), and in the liberation of proteins tethered to the streptococcal cell surface, including M proteins and C5a peptidase that are proposed to enhance *S. pyogenes* virulence (8). Consideration of the potential host targets and the careful regulation of SpeB proteolytic activity suggest an important role for the protease (9). However, due to conflicting results from human tissue samples and animal models, the relative importance of SpeB to *S. pyogenes* pathogenicity has been highly debated (10).

*S. pyogenes* exclusively infects humans, and active SpeB is secreted from the bacteria in individuals with invasive disease (11). Notwithstanding this protease secretion in human *S. pyogenes* infections and establishment of host and bacterial substrates, some data suggest an inverse relationship between SpeB production and disease severity (12–14). Some recent results have suggested that down-regulation of SpeB expression and inhibition of protease activity promote the accumulation and activation of host protease plasmin on the GAS bacterial surface, thereby promoting infectivity (13). Similarly, a clinical correlation has been reported between *S. pyogenes* invasive dis-
ease severity and diminished SpeB expression (12). Regardless of its potential role as a virulence factor, understanding of the structure and function of the secreted protease SpeB will help in elucidating the biology and chemistry at the interface between human epithelial cells and *S. pyogenes* and potentially provide insight into the role of proteases secreted from other bacteria.

SpeB, like most extracellular proteases, is synthesized as an inactive zymogen to protect the intracellular components of the bacteria from proteolytic activity during protein production. SpeB is composed of an N-terminal signal sequence (residues 1–27), a prodomain (residues 28–145), and a catalytic C-terminal region (residues 146–398). Once secreted into the extracellular milieu, SpeB is susceptible for maturation to the active protease by autocatalysis, mature SpeB, trypsin, and subtilisin (15–18). Mutational analysis and structural studies have already revealed residues essential for catalytic activity (Cys192 and His340) (19) and substrate binding (Trp357 and Trp359) (20–22). These studies, in combination with kinetic interrogation, have identified the preferred peptide sequences targeted by SpeB for cleavage (17). Based on this information and the sequence of the SpeB residues hydrolyzed in its self-activation, we have designed and synthesized a fluorogenic tripeptide substrate, acetyl-Ala-Ile-Lys-7-aminomethylcoumarin (Ac-AIK-CHO). These structures elucidate the residues essential for substrate binding on the nonprime side of the SpeB active site and the conformational changes of the protein associated with binding. Unexpectedly, these crystal structures provide the first view of a “floppy” glycine-rich C-terminal loop, adjacent to the active site, in both the apo and complex isoforms. With this structural information, we performed alanine-scanning mutagenesis on all of the loop glycine residues to afford an understanding of the interactions between the protease and potential substrates as well as how to target the protease for specific small molecule drug discovery.

Structures of the SpeB zymogen and mature SpeB have been previously determined by both x-ray crystallography and NMR. The protease has a canonical papain-like fold (20–22). Here, based on this information and the sequence of the SpeB residues hydrolyzed in its self-activation, we have designed and synthesized a fluorogenic tripeptide substrate, acetyl-Ala-Ile-Lys-7-aminomethylcoumarin (Ac-AIK-CHO), which demonstrates robust catalytic turnover by SpeB, and a specific peptide inhibitor acetyl-Ala-Gln-Ile-(S)-2,6-diaminohexanal (Ac-AEIK-CHO). We then determined the crystal structure of mature SpeB with and without our peptide inhibitor. This structural information would help to elucidate any conformational changes associated with active site binding and afford an understanding of the interactions between the protease and potential substrates as well as how to target the protease for specific small molecule drug discovery.

2The abbreviations used are: AMC, 7-aminomethylcoumarin; CHO, (S)-2,6-diaminohexanal; E64, trans-epoxysuccinyl-l-leucylamido(4-guanidino)butane; r.m.s., root mean square; Fmoc, N-(9-fluorenyl)methoxycarbonyl; TLS, translation/libration/scramble.

**EXPERIMENTAL PROCEDURES**

**Protein Preparation and Purification**—The *S. pyogenes* 10782 zymogen SpeB clone (residues 28–398) was generated without the N-terminal secretion leader sequence (residues 1–27) using standard PCR-based cloning and verified via double-stranded plasmid sequencing. SpeB is overexpressed as a C-terminal His_{6} tag fusion (additional amino acids for the affinity tag include LEH_{6}, residues 399–406) from *Escherichia coli* BL21DE3pLysS (Stratagene) in a pET23b vector (Novagen). Cells were grown in 2× YT medium supplemented with 200 μg/ml ampicillin and 50 μg/ml chloramphenicol at 37 °C to an *A_{600 nm}* of 0.8–1.0. Flasks were then transferred to 25 °C, and protein expression was induced with 0.5 mm isopropyl 1-thio-β-d-galactopyranoside for 5 h. SpeB self-maturates during expression and results in the fully active protease (residues 146–398). Cells were immediately harvested and resuspended in ice-cold 100 mm Tris, pH 8.0, 100 mm NaCl (buffer A) and subjected to three cycles of lysis by microfluidization (Microfluidics). The cell lysate was clarified by centrifugation at 45,000 × *g* for 30 min at 4 °C, and soluble fractions were loaded onto a 1-ml HisTrap HP nickel-nitritotriacetic acid affinity column (GE Healthcare) pre-equilibrated with buffer A and eluted with buffer A containing 250 mm imidazole. The eluted protein was immediately diluted 5-fold with buffer B (50 mm NaOAc, pH 4.5) and purified by cation exchange chromatography (HiTrap SP HP, GE Healthcare) with a 20-column volume gradient to 50% of buffer B containing 1 mm NaCl. Fractions corresponding to SpeB were pooled and dialyzed overnight at 4 °C in phosphate-buffered saline (PBS), pH 7.4, and 1 mm DTT. If required, pure SpeB protein was further subjected to gel filtration chromatography (Superdex 75, GE Healthcare) in PBS and 1 mm DTT to remove any remaining contaminants. Mature SpeB was concentrated to ~10 mg/ml using Millipore Ultrafree-15 devices with a molecular weight cut-off of 3500 and immediately stored at ~70 °C. Pure SpeB yields are ~10–20 mg/liter of culture with >95% purity, as assessed by SDS-PAGE (supplemental Fig. 1A).

**Mutations of Glycine Residues within C-terminal Loop**—SpeB point mutations were generated using QuikChange (Qiagen) mutagenesis. SpeB C-terminal loop mutants were constructed using forward primer and complement 5’-CAG CTA CAG GAG CCG TTG CTA CTT CTA CTG GTG-3’ for C192A, 5’-CTT CAG CTC TGG CTA CTT GGT GTG TGG-3’ for G378A, 5’-CTT CAG CTC TTG CTA CTT GGT GTG TGG-3’ for G380A, 5’-CTT CAG CTC TTG CTA CTT GGT GTG TGG-3’ for G381A, 5’-CTT CAG CTC TTG CTA CTT GGT GTG TGG-3’ for G384A, 5’-CTT CAG CTC TTG CTA CTT GGT GTG TGG-3’ for G385A. Purification was performed exactly as for wild type SpeB. The point mutations did not affect structural stability of any of the recombinant expressed SpeB proteins as determined by differential scanning fluorimetry (data not shown) (23).
Apo and Peptide Complex Structures of Monomeric SpeB

Crystallization and X-ray Data Collection—Crystals of mature apo-SpeB were grown by sitting drop vapor diffusion by mixing equal volumes (2 μl) of SpeB (10 mg/ml) and reservoir solution consisting of 0.15 M sodium nitrate and 15% PEG 3500 at 4 °C. Inhibitors E64 (G-Biosciences) and Ac-AEIK-CHO were added in a 3-fold molar excess to SpeB and incubated for 2 h at 25 °C and immediately used for co-crystallization experiments. Crystals of SpeB in complex with E64 and Ac-AEIK-CHO were grown from solutions consisting of 0.2 M magnesium nitrate with 20% PEG 3500 at 4 °C and 0.15 M sodium nitrate with 26% PEG 3500 at 25 °C, respectively. The His6 tag was not removed because the protein crystallized readily.

Data for all three SpeB x-ray structures were collected on single, flash-cooled crystals at 100 K in a cryoprotectant consisting of mother liquor and 20% PEG 400 and were processed with HKL2000 (24) in orthorhombic space group P21212 (Table 1). The calculated Matthews coefficient (V_m = 2.1 Å^3 Da^-1) suggested one monomer per asymmetric unit with a solvent content of 42%. X-ray data for the apo and E64 complex SpeB structures were collected to 1.06 and 1.50 Å resolution, respectively, on beamline 11.1 at the Stanford Synchrotron Radiation Lightsource (Menlo Park, CA). For the SpeB-Ac-AEIK-CHO complex, data were collected to 1.37 Å resolution on the Advanced Photon Source beamline 23-ID-B at the Argonne National Laboratory (Argonne, IL). Data collection and processing statistics are summarized in Table 1.

Structure Solution and Refinement—The apo-SpeB structure was determined by molecular replacement with Phaser (25, 26), using the previously published mature SpeB structure (2U2J) (21) as the initial search model. The apo-SpeB structure then served as the search model for the E64 and Ac-AEIK-CHO SpeB complex structures. All structures were manually built with Coot (27) and iteratively refined using Phenix (28) with anisotropic B-factor refinement was included, and hydrogen atoms were included in the model coordinates. For the SpeB complex structures, TLS B-factor refinement was carried out in the last round of refinement with the SpeB monomer split into nine TLS groups as determined by TLS Motion Determination in Phenix (28). TLS refinement resulted in improved electron density maps with a minimal change in R_cryt and R_free. The electron density maps clearly identified that inhibitors E64 and Ac-AEIK-CHO were covalently attached to Cys192 within the active site. Water molecules were automatically positioned by Phenix using a 2.5σ cut-off in F_o − F_c maps and manually inspected. All three structures revealed a nitrate molecule (from the buffer) bound to residues located within the C-terminal loop. For the apo-SpeB structure, the final R_cryt and R_free are 14.4 and 16.0%, respectively; for the E64 complex structure, the final R_cryt and R_free are 15.8 and 18.8%, respectively; and for the Ac-AEIK-CHO complex structure, the final R_cryt and R_free are 16.4 and 19.1%, respectively (Table 1).

All models were analyzed and validated with PROCHECK (26, 29), WHATCHECK (30), and Molprobity (31) on the Joint Center for Structural Genomics (JCSG) Web server. Analysis of backbone dihedral angles with the program PROCHECK (29) indicated that all residues for the three structures are located in the most favorable and additionally allowed regions in the Ramachandran plot with no outliers. Coordinates and structure factors have been deposited in the Protein Data Bank (32), with accession entries 4D8B (apo), 4D8E (E64 complex), and 4D8I (Ac-AEIK-CHO complex). Structure refinement statistics are shown in Table 1.

SpeB Fluorogenic Substrate Design and Kinetic Assays—We synthesized an N-terminal acetylated, fluorogenic SpeB tripeptide substrate Ac-AIK-AMC, based on previously determined optimal cleavage sequences (17). This substrate was used to determine the catalytic activity of native SpeB and effects of point mutations on catalytic efficiency and to establish IC_{50} values for E64 and Ac-AEIK-CHO. Unless otherwise stated, SpeB was incubated at 50 nm in 50 μl of PBS, pH 7.4, 0.1 mM EDTA, 10 mM DTT, 0.1% CHAPS at 25 °C with the activity measurement initiated by introduction of 50 μM Ac-AIK-AMC. The C-terminal His6 tag was not removed from the SpeB protein preparations because the presence of the tag would not interfere with in vitro kinetics. The His6 tag is positioned on the opposite side of the protein surface more than 50 Å from the active site (supplemental Fig. 1B). The increase in fluorescence due to substrate hydrolysis was measured every 30 s for a 10-min duration in 96-well plates on a PerkinElmer EnVision plate reader (excitation 355 nm, emission 460 nm). All components of the assay are stored as frozen aliquots and thawed immediately prior to the assay. Michaelis-Menten (k_m and k_cat) and IC_{50} values were determined using GraphPad Prism software (GraphPad, Inc.).

SpeB IC_{50} Determination—SpeB was incubated at 50 nm in the presence of increasing amounts of E64 or Ac-AEIK-CHO (10 nm to 10 μM) in a reaction buffer consisting of PBS, pH 7.4, 0.1 mM EDTA, 10 mM DTT, and 0.1% CHAPS and incubated for 20 min at 25 °C. 50 μM Ac-AIK-AMC was subsequently added, and the rate of substrate hydrolysis was measured for 10 min by the increase in fluorescence.

Cleavage of Inactive C192A Zymogen SpeB—C192A SpeB was incubated at 10 μM with 200 nm active wild type or SpeB mutants in a reaction buffer consisting of PBS, pH 7.4, 0.1 mM EDTA, 10 mM DTT, and 0.1% CHAPS. The mixtures were agitated at 37 °C, and aliquots were removed every 10–30 min. Reactions were immediately quenched with the addition of 1% SDS-reducing gel-loading buffer and boiled. Samples were subjected to SDS-PAGE to determine the extent of substrate cleavage in the inactive C192A SpeB.

Ac-AIK-AMC Substrate Synthesis—The SpeB Ac-AIK-AMC peptide substrate was synthesized using standard Fmoc solid phase synthesis chemistry starting with Fmoc-Lys(carbamate)-AMC Wang resin (EMD Biosciences). After completion of the peptide synthesis and its N-terminal acetylation with acetic anhydride and diisopropylethylamine in dichloromethane, the substrate was released from the resin with a mixture of trifluoroacetic acid (TFA), triisopropylsilane, and water (95%, 2.5%, 2.5%). Crude Ac-AIK-AMC was purified by reverse-phase HPLC using a C18 Xterra column (Waters), and all fractions containing the desired product were lyophilized. The final purity of Ac-AIK-AMC exceeded 95% and was verified by mass spectrometry: expected m/z 530.3, LC/MS (electrospray ionization) m/z 530.4 (MH^+). Michaelis-Menten kinetic measure-
ments for the hydrolysis of Ac-AIK-AMC by a 50 nM concentration of the wild type mature SpeB yielded $k_{\text{cat}} = 0.69 \text{ s}^{-1}$, $K_m = 246.6 \mu\text{M}$, and $k_{\text{cat}}/K_m = 2.8 \times 10^3 \text{ s}^{-1} \text{ M}^{-1}$ (Table 2 and supplemental Fig. 2).

**Synthesis of Ac-AEIK-CHO Inhibitor**—An N-terminal acetylated, SpeB peptide inhibitor Ac-AEIK-CHO was designed based on the nonprime side cleavage sequence required for self-removal of the zymogen prodomain (residues 28–145). The Ac-AEIK-CHO peptide inhibitor was synthesized using standard Fmoc solid phase synthesis chemistry starting with N-$t$-butoxycarbonyl-L-lysine-2-chlorotrityl resin (AnaSpec). Release from the resin with a mixture of acetic acid, trifluoroethanol, and dichloromethane (10%, 10%, 80%) afforded a C-terminal carboxylic acid peptide with ethyl and tert-butyloxycarbonyl-protected lysine side chains (33) (1; supplemental Fig. 3). The resulting peptide was dried under vacuum and purified by reverse-phase HPLC and lyophilization.

Compound 1 was dissolved in acetonitrile with equal molar concentrations of N-methylmorpholine hydrochloride and isobutyl chloroformate added in rapid succession at room temperature for 30 min to yield the mixed anhydride (2; supplemental Fig. 3), which was dried under vacuum, dissolved in water, and placed on ice. Sodium borohydride (NaBH$_4$) was added to 2 at a 1.1 molar ratio, brought to room temperature, and reacted for 30 min to produce the C-terminal alcohol (34) (3; supplemental Fig. 3), which was extracted with ethyl acetate and dried under vacuum. Compound 3 was dissolved in acetonitrile and a 1.1 molar ratio of Dess-Martin periodinane, dissolved in an equal volume of acetonitrile, and slowly added dropwise at room temperature and reacted overnight to yield the resulting C-terminal aldehyde (4; supplemental Fig. 3), which was purified by standard aqueous workup and dried under vacuum. Deprotection of the peptide aldehyde 4 was performed with a mixture of TFA, trisopropylsilane, and water (95%, 2.5%, 2.5%). Crude Ac-AEIK-CHO was purified by reverse-phase HPLC using a C18 Xterra column (Waters), and the desired product fractions were lyophilized to produce the final product (5; supplemental Fig. 3). The purity of Ac-AEIK-CHO exceeded 90% and was verified by mass spectrometry: expected m/z 487.6, LC/MS (electrospray ionization) m/z 488.2 (MH$^+$). All intermediates and reactions were monitored by LC/MS.

**RESULTS**

**Apo-SpeB Structure**—The SpeB zymogen (residues 28–398), without its N-terminal secretion sequence (residues 1–27), was cloned from *S. pyogenes* strain 10782 and overexpressed in *E. coli*. The protease fully matures during expression with self-removal of the N-terminal prodomain (residues 28–145) and purifies as a monomer as evidenced by gel filtration (supplemental Fig. 1A). The apo-SpeB structure was determined by molecular replacement using the mature SpeB structure (Protein Data Bank entry 2UZJ) (21) as the initial search model. The final $R_{\text{cryst}}$ and $R_{\text{free}}$ values for the apo structure were 14.4 and 16.0%, respectively, with 97.6% of the residues in the most favored region of the Ramachandran plot (Table 1).

The structure includes residues 146–398, an ordered leucine residue from the C-terminal His$_6$ tag (LEH$_6$), a nitrate ion, and 548 water molecules. The nitrate was introduced via the crystallization conditions and is not required for biological activity. The bound nitrate has a significantly higher $B$-value in comparison with the protein, most likely due to incomplete occupancy (Table 1). Three regions of unaccounted for positive electron density within solvent channels were not modeled in the final $F_o - F_c$ maps and are most likely cryoprotectant PEG 400 at low occupancy.

The crystals of apo-SpeB were obtained in conditions with a biologically relevant pH of 7.0, which resulted in a different
crystal lattice and space group in comparison with the mature SpeB x-ray structure 2UZJ crystallized at pH 4.0 (21). Due to crystal packing interactions in the 2UZJ structure, SpeB was proposed to exist in a homodimeric arrangement with the interface situated along a crystallographic axis. This arrangement in the crystal lattice resulted in residue Ser182 (our numbering system) adopting a cis-peptide bond and contributing to the oxyanion hole of the opposite subunit as well as a key hydrogen bond to an unidentified ligand that was bound within the active site at the dimer interface (21). This interface does not exist in the crystals grown at neutral pH. Our gel filtration purification profile and x-ray structure strongly support a monomer as the biologically relevant form of SpeB. Our results are consistent with the x-ray structure of the zymogen form of SpeB (Protein Data Bank entry 1DKI) at pH 4.2 (20) and the NMR structure of mature SpeB at pH 6.0 (Protein Data Bank entry 2JTC) (22).

Crystal structures of mature SpeB reveal an ordered C-terminal Gly-rich loop. A, schematic representations of apo (4D8B; cyan), E64-bound (4D8E; green), and Ac-AEIK-CHO-bound (4D8I; rose) mature SpeB show the expected canonical papain-like fold. The N-terminal α-helical domain supplies the four α-helices that sandwich the C-terminal four-stranded antiparallel β-sheet. Residues 376–385 of the C-terminal loop, E64, and Ac-AEIK-CHO are shown in stick representations (yellow, carbon; red, oxygen; blue, nitrogen). A blue Fo−Fc simulated annealing omit map contoured at 2.5σ provides compelling evidence for an ordered structure of the C-terminal loop that was not observed in previous x-ray structures of SpeB (20, 21). B, chemical structure and atom numbering of the general cysteine protease inhibitor E64 and of the covalent complex formed upon nucleophilic attack by the active site residue Cys 192. C, chemical structure of a new Ac-AEIK-CHO inhibitor based on the recognition sequence for zymogen prodomain removal during self-activation. Nucleophilic attack by the active site residue Cys 192 results in covalent attachment of the peptide and protonation of the aldehyde oxygen to form a hydroxyl.
19.1%, respectively, for the Ac-AEIK-CHO complex. Both structures have greater than 97.2% of the residues in the most favored region of the Ramachandran plot (Table 1). As with the apo structure, both complex structures crystallized with identical crystal packing and had continuous electron density for residues 146–398, an ordered leucine residue from the C-terminal His8 tag, a bound nitrate, and water molecules. After the initial refinement, clear residual density was present in $F_{o} - F_{c}$ maps corresponding to E64 or Ac-AEIK-CHO covalently attached to the active site residue Cys192 (Fig. 2A). For both structures, the compounds were refined with 100% occupancy and have B-values similar to adjacent active site residues. The overall conformation of SpeB is highly conserved between the E64-bound and Ac-AEIK-CHO-bound structures, as evidenced by an r.m.s. deviation of 0.13 Å for all Ca atoms with a maximum deviation of 0.34 Å (r.m.s. deviation of 0.26 Å with maximum deviation of 2.69 Å for all atoms).

**E64-bound SpeB Interactions**—The SpeB active site, as with the majority of conserved papain fold proteases, consists of a catalytic dyad composed of residues Cys192 and His340, which exist as a zwitterionic pair (35). The nucleophilicity of the negatively charged Cys192 is greatly stabilized by both its location at the N terminus of the central $\alpha$-helix (residues 192–204), to exploit the dipole moment, and the interaction with His340 (35, 36) (Fig. 2A). His340 also serves as the acid/base catalyst during substrate cleavage. The majority of specific interactions provided by the SpeB active site residues are concentrated near the carboxyl end of E64 (Fig. 2A). Opening of the E64 epoxide ring occurs upon nucleophilic attack by SpeB Cys192 and results in covalent attachment of the C2 atom of E64 to the Cys192 thiol with a bond distance of 1.9 Å (Figs. 1B and 2B). The Glu162 side-chain Ne2 and main-chain amide of Cys192 comprise the oxyanion hole that stabilizes the interaction to one of the E64 carboxyl oxygens. The other E64 carboxyl oxygen forms potential hydrogen bonds with the Nδ1 of the catalytic His340 imidazole and a bound water molecule that is also bridged to the side-chain Ne1 of Trp359 (Fig. 2B). The E64 carbonyl (C4 and O4) is stabilized by interactions with the main-chain amide of Ser282 and a highly ordered water molecule, which is positioned via hydrogen bonds to the main-chain amides of both Cys192 and Val193. The E64 leucine side chain is inserted into the hydrophobic S2 binding pocket formed by Val334, Gly339, and Ala341 (Fig. 2C). The interactions between SpeB surface and the E64 amino-4-guanidinobutane moiety are mediated through ordered water molecules. This region of E64 has less well defined electron density and higher B-values, indicative of increased flexibility (Fig. 2B). E64 is largely exposed to bulk solvent with only 30% of its surface buried by the SpeB active site (Fig. 2C). Overall, the interactions between E64 and SpeB are highly conserved with previously determined E64-bound papain-like proteases, including *Ervatamia coronaria* erva-matin-A (Protein Data Bank entry 3BCN) (r.m.s. deviation of 2.04 Å for 68 conserved core Ca atoms) and erva-matin-C (Protein Data Bank entry 2PRE) (r.m.s. deviation of 2.03 Å for 68 conserved core Ca atoms) (37), and *Staphylococcus aureus* staphopain (1CV8) (r.m.s. deviation of 1.57 Å for 62 conserved core Ca atoms) (38).

**Ac-AEIK-CHO-bound SpeB Structure**—We synthesized a specific peptide C-terminal aldehyde inhibitor Ac-AEIK-CHO based on the nonprime side recognition sequence forzymogen prodomain removal during self-activation (Ala142-Glu143-Lle144-Lys145-Gln146). Standard Michaelis-Menten kinetics revealed that Ac-AEIK-CHO inhibits mature SpeB activity with an IC$_{50}$ of 205 nM (supplemental Fig. 4). The C1 atom of the peptide C-terminal aldehyde is susceptible to nucleophilic attack by SpeB and results in formation of a covalent adduct between Ac-AEIK-CHO C1 and the Cys192 thiol with a bond length of 1.9 Å (Figs. 1C and 2B). As a consequence, the aldehyde oxygen is reduced to a hydroxyl. An ordered water molecule that is positioned by the side-chain Oε1 of residue Gln162 is a likely candidate to supply the proton to the aldehyde oxygen (Fig. 2B). This reduction occurs concomitant with sp$^3$ hybridization of the C1 atom (Fig. 2A). The reacted peptide inhibitor hydroxyl group is positioned in close proximity to the oxyanion hole, created by residues Gln162 and Cys192, and the electron-withdrawing properties of these residues stabilize the oxyanion intermediate during catalysis. With iterative rounds of model building and refinement of the Ac-AEIK-CHO-bound structure, additional electron density indicated that the reacted peptide hydroxyl had an alternate conformation positioned toward the side-chain Nδ1 of His340 (Fig. 2A). Notwithstanding, the peptide inhibitor is not entirely representative of a biological substrate that would include residues to bind on the prime side of the SpeB active site. Under these circumstances, protein substrates with trans-peptide bonds would position a carbonyl oxygen toward the oxyanion hole. Additional ordered water molecules are coordinated near the peptide inhibitor C-terminal hydroxyl moiety that do not contribute to peptide binding or electrostatic perturbations of SpeB active site residues (Fig. 2B). These waters would most likely be displaced or absent after binding of a biologically relevant protein substrate, which would provide interactions with the prime side region of the SpeB active site.

In comparison with the E64 complex structure, many of the same residues provide key interactions for binding and positioning of the Ac-AEIK-CHO (Fig. 2B). The peptide P1 Lys side chain is positioned via interaction of the Nε2 with the side-chain hydroxyl from Ser279 and has elevated B-values and reduced electron density in comparison with the rest of the bound peptide, suggesting flexibility within this region (Fig. 2A). SpeB has been shown to preferentially cleave biological protein substrates with a variety of residues at P1, including Arg, Leu, Asn, and Phe (18, 39, 40). Based on the Ac-AEIK-CHO structure, the S1 pocket can accommodate a variety of substrate P1 residues and prefers basic side chains as evidenced by the strong electro-negative potential of the SpeB S1 pocket (Fig. 2C). Similar to E64, the main-chain carbonyl of the peptide inhibitor P2 Ile is anchored into a highly electropositive patch provided by the main-chain amide of Ser282 and a water molecule ordered by the main-chain amide of Val193 (Fig. 2B). The main-chain carbonyl of Ser282 provides a key interaction to the peptide inhibitor main-chain amide of P2 Ile (Fig. 2B) that is not observed in the E64-bound SpeB structure where the E64 carbonyl (C11 and O5) superimposes on the peptide P2 Ile amide. The peptide inhibitor P2 Ile side chain is anchored into the S2 hydrophobic
pocket similarly to the E64 Leu moiety. Despite the difference in Ile and Leu side-chain branching of Ac-AEIK-CHO and E64, respectively, the S2 pocket residues are conformationally conserved. This is surprising because the preference of Ile over Leu at P2, in otherwise identical substrates, has been reported to result in an 86-fold increase in catalytic efficiency (41). As with
other papain-like proteases, the specificity of the S2 pocket of SpeB is a primary determinant in substrate recognition (17, 42). Our SpeB-Ac-AEIK-CHO structure illustrates an ability to accommodate a number of branched hydrophobic side chains in the S2 pocket, including Ile, Tyr, Met, and Val, which have been identified in a number of biological SpeB substrates, and accounts for the relatively broad specificity of SpeB (18, 39, 40). As with E64, there is a lack of specific interactions from the SpeB active site to the P3 Glu and P4 Ala residues of Ac-AEIK-CHO with 57% of the peptide inhibitor surface exposed to bulk solvent (Fig. 2C). The Ac-AEIK-CHO P3 glutamate is positioned by water molecules that are ordered by SpeB surface residues (Fig. 2B). For most papain-like proteases, the S3 pocket is not well defined and is located on the surface of the protease (37). Our structure supports the substrate specificity profiles of SpeB that have suggested that three amino acids on the nonprime side of the substrate are sufficient for recognition and that alteration of residues at the P3 or P1 position does not drastically influence catalytic efficiency by the protease (17). Additional prime side specificities have been established (18, 39, 40), but cannot be validated, with the current x-ray structures.

Conformational Movements Associated with Activation and Active Site Binding—The SpeB prodomain (residues 28–145) preserves the protease in a catalytically inactive state by the insertion of an α-helix into the active site that displaces both the C-terminal (residues 368–390) and catalytic (residues 332–341) loops (Fig. 3A) (20–22). Activation of the protease via removal of the prodomain allows the C-terminal loop to flip ~150–170° to adopt the conformation observed in the apo structure (Fig. 3A). Superposition of thezymogen (1DKI) and active apo-SpeB structures, not including the C-terminal loop, indicate a highly conserved structure as evidenced by an r.m.s. deviation of 1.61 Å for all Ca atoms with a maximum deviation of 8.5 Å (compared with 1.41 Å with a maximum deviation of 10.7 Å for all atoms). The additional large conformational movement is found in the catalytic loop (residues 332–341), which contains His340 (Fig. 3A). Upon removal of the prodomain, the catalytic loop is free to move and results in conformational alignment of His340 with a catalytically active orientation and promotion of Cys392 nucleophilicity (Fig. 3A) (21, 22). Importantly, movement of the catalytic loop also exposes and forms the hydrophobic patch critical to substrate recognition.

Comparison of the apo-SpeB with the E64- and Ac-AEIK-CHO-bound structures reveals several large conformational changes in loops surrounding the active site upon ligand binding. As with the transition from zymogen to active apo, the most significant movements occur within both the C-terminal loop (residues 368–390) and the catalytic loop (residues 332–341) with a maximum Ca displacement of 17.5 and 4.5 Å, respectively (Figs. 1 and 3A). Ala376 and Gly385 represent the C-terminal loop hinge points where the loop swings open and away from the active site by 23.6 and 51.0°, respectively, upon binding (Fig. 3B). This movement abolishes the type II β-turn (Gly378-Thr379-Gly380-Gly381) in the apo structure and increases the solvent accessibility of the entire protein surface by 1063 Å² and the active site cleft by ~450 Å². Despite the proper alignment of His340 for catalysis in the apo active site, the catalytic loop shifts an additional 4.5 Å to accommodate substrate binding. This conformational transition would not be permitted without the concerted relocation of the C-terminal loop upon substrate binding because clashes would occur between the two loops (Fig. 3A). For the rest of the protein, minimal conformational rearrangements occur upon substrate interaction as evidenced by an r.m.s. deviation of 0.42 Å with a maximum deviation of 1.90 Å for all protein Ca atoms (compared with 0.58 Å with a maximum deviation of 7.49 Å for all atoms). These loop movements associated with active site binding would not be possible if SpeB existed in the proposed homodimeric arrangement and do not favor the formation of a transitory homodimeric SpeB as previously suggested (21). Furthermore, gel filtration analysis of SpeB in the presence of 5 μM E64, 50-fold above the IC50 of 98 nM (supplemental Fig. 4), shows that the presence of E64 does not affect the monomeric state (supplemental Fig. 1).

Mutational Analysis of Glycine Residues within the C-terminal Loop—The C-terminal loop (residues 368–390) contains six Gly residues, all of which adopt different ϕ/ψ angles between the apo and bound structures of SpeB (Fig. 4A). These six Gly residues were mutated to Ala to explore the effects on SpeB substrate binding and catalysis with respect to restricted glycine ϕ/ψ angles that promote either the apo-SpeB loop (occluded state) or the bound SpeB loop (open state) conformations (Fig. 3B). Activity of the SpeB Ala mutations was measured by hydrolysis of the small tripeptide fluorogenic substrate Ac-AIK-AMC. G378 is located at n + 2 from the hinge residue Ala376, and mutation to Ala would restrict the loop to an occluded state because the ϕ/ψ angles in the apo and bound structures are ~99°/~163 and 100/123, respectively. Forcing the loop to remain positioned over the active site would block access of substrates and result in a deleterious effect on catalysis (Fig. 4B). As predicted, the G378A mutation caused an inhibition of catalytic turnover (kcat) that results in a modest decrease in catalytic efficiency (kcat/Km) of ~34% relative to wild type SpeB (Table 2).

As opposed to the negative effects mutation to Gly378 has on SpeB catalysis, mutation of Gly385 to Ala would promote the open state observed in the SpeB complex structures as these residues adopt ϕ/ψ angles unique to glycine only in the apo (occluded) structure (Fig. 4B). Such mutations should stimulate activity of SpeB by affording substrates unrestricted access to the active site. As expected, locking the loop in the open state via mutation of both Gly384 and Gly385 to Ala resulted in an increase of ~1.6- and ~3.1-fold, respectively, in the kcat/Km for Ac-AIK-AMC hydrolysis (Fig. 4B). For Gly384, the increase in efficiency is a result of improved substrate binding because the Km increases almost 2.1-fold over the wild type (Table 2). The G385A mutation had no influence on substrate binding but did result in an ~3.1-fold improvement in kcat, suggesting that the loop is not required for catalysis. Gly380, Gly381, and Gly382 form the type II β-turn in the apo structure (Fig. 3B). Mutations at these positions are not predicted to result in significant consequences on SpeB activity because these residues are located near the tip of the C-terminal loop away from the hinge residues and would have limited
effects on its overall flexibility and positioning. Mutation of Gly\textsuperscript{381} and Gly\textsuperscript{382} resulted in anticipated negligible changes in both substrate binding and rates of Ac-AIK-AMC hydrolysis (Table 2). However, despite a catalytic efficiency similar to that of wild type, for G380A, $k_{\text{cat}}$ was reduced by 44%, and $K_m$ improved by 1.8-fold (Table 2). The addition of a methyl group...
Apo and Peptide Complex Structures of Monomeric SpeB

FIGURE 3. Conformational changes associated with activation and active site binding. A, zymogen SpeB (1DKI) is shown as a schematic representation with the static regions of the protein colored gray, zymogen prodomain in light blue, and the RGD motif in rose with active site residues Cys192 and His340 represented as sticks and colored purple. The highly flexible catalytic (residues 332–341) and C-terminal (residues 368–390) loops are colored purple and orange, respectively, and labeled accordingly. The two ends of the C-terminal loop with missing residues in the zymogen x-ray structure are connected with an orange line. B, apo-SpeB structure is colored as in A. The catalytic loop moves ~8.5 Å, relative to the zymogen structure, and orients the catalytic His340 for catalysis. Removal of the prodomain allows the C-terminal loop to flip 150–170° to form a type II β-turn over the active site. C, structure of SpeB in complex with E64 reveals the additional loop movements that occur upon the transition from apo to bound forms. The catalytic loop and C-terminal loop move 4.5 and 17.5 Å, respectively upon binding E64 or Ac-AEIK-CHO. The E64 complex is colored as in A. D, superposition of the apo-SpeB (carbon, cyan) and E64-bound (carbon, green) structures reveals the large conformational movement of the C-terminal loop to accommodate substrate binding. Hinge residues Ala376 and Gly385 have a 23.6 and 51.0° rotation, respectively, from the active site occluded C-terminal loop conformation in the apo structure to the open orientation depicted in the E64- and Ac-AEIK-CHO-bound structures. Residues with conserved orientations between the apo and bound states are labeled in black.

FIGURE 4. Effects of Gly-rich C-terminal loop alanine-scanning mutagenesis. A, Ramachandran plot of changes in the ϕ/ψ angles of the six Gly residues found in the C-terminal Gly-rich loop in the apo and E64-bound SpeB structures reveals conformational isomerism between two well defined conformers. B, table of changes in the ϕ/ψ angles in the apo versus bound state for residues in the C-terminal loop and effects of mutations on catalytic efficiency upon the hydrolysis of the tripeptide fluorogenic substrate Ac-AIK-AMC. Mutation of Gly378 to Ala results in locking the loop into the active site occluded orientation and blocks accessibility to substrates. Mutation of the hinge residue Gly385 to Ala favors the open conformation and results in an approximately 3.1-fold increase in SpeB catalysis. C, prodomain removal in a catalytically inactive zymogen C192A SpeB and effect of mutations in the C-terminal Gly-rich loop. Mutation of C-terminal Gly residues to Ala restricted the C-terminal loop and reduced zymogen prodomain removal considerably. These results suggest that a flexible C-terminal loop is critical to aid in the removal of hydrolyzed products of natural protein targets.
to Gly380 would provide a potential van der Waals interaction with a bound substrate only if the C-terminal loop assumed the occluded apo conformation. This result suggests that the C-terminal loop could sample this occluded position upon substrate binding to the active site. The G380A mutation, however, does negatively affect the catalytic rate, suggesting that reduced flexibility of the C-terminal loop hinders product release.

Additional mutations included Thr379 to Ala and Val to explore the structural and/or functional role of the Thr379 side-chain hydroxyl and Gly384 to Asp because this mutation has been observed in allelic variants of SpeB and significantly reduces protease activity (22). The Thr379 hydroxyl does not contribute significantly to substrate binding or catalysis as observed by the insignificant consequences that mutation to either Ala or Val has on SpeB activity (Fig. 4B). The allelic variant G384D, which was reported to have a 12-fold decrease in proteolytic activity of SpeB as measured by azocasein hydrolysis (22), results in a minimal decrease of ~33% in Ac-AIK-AMC catalysis.

Overall, the results of the alanine-scanning mutagenesis of Gly residues within the C-terminal loop illustrate a conserved net effect on $k_{cat}$ of Ac-AIK-AMC turnover notwithstanding the residue location within the loop. Two alanine mutations (Gly380 and Gly384) did improve the $K_m$ of substrate binding, which was countered by a decrease in $k_{cat}$ (Table 2). However, substrate binding is assumed to be much slower than hydrolysis under these steady state in vitro conditions, and the observed changes in $k_{cat}$ can alter the rate constants that comprise $K_m$ without altering the affinity of SpeB for Ac-AIK-AMC. Taken together, our mutational analysis results suggest that the C-terminal loop does not significantly contribute to substrate affinity, but its structural flexibility is the limiting factor in the catalytic efficiency with respect to the recognition of the small tripeptide substrate and release of the hydrolyzed product.

The C-terminal loop Ala mutants were also assessed for their ability to remove the prodomain of the inactive C192A SpeB zymogen and to determine if the mutations also altered specificity profiles for a biologically relevant substrate. Due to limitations in detection, the zymogen was added to a final concentration of 10 μM, well above the predicted $K_m$, and introduced product release as a potential rate-limiting step. Overall, the individual C-terminal loop Gly mutations to sterically hindered Ala residues slowed hydrolysis of the zymogen considerably. Therefore, a flexible C-terminal loop seems critical to aid in the removal of hydrolyzed products (Table 2). Moreover, residues not important to Ac-AIK-AMC hydrolysis had a significant impact on protein substrate catalysis. For instance, mutation of Thr379 to Ala, which was not significant for Ac-AIK-AMC turnover, resulted in a drastically reduced hydrolysis of prodomain, whereas mutation to Val did not substantially affect the rate of zymogen activation (Fig. 4C). From these results, we surmise that Thr379, particularly its branched side chain, is most likely involved in substrate recognition and/or product release. G380A, G384A, and G384D mutants had an improved $K_m$ of Ac-AIK-AMC binding with a concomitant decrease in catalytic turnover. With respect to a biological substrate, these mutations have similar zymogen maturation intermediate profiles; however, the accumulated amounts of intermediates at various time points differ among the SpeB mutants (Fig. 4C). The individual cleavage sites are therefore cleaved at different rates upon mutation of residues in the glycine-rich active site loop, which suggests that the mutations affect the propensity to recognize and hydrolyze certain cleavage sequences (39). Mutation of Gly381 and Gly382 to Ala had negligible effects on $k_{cat}$ and $K_m$ for Ac-AIK-AMC hydrolysis but had a deleterious effect on prodomain removal, most likely due to decreased loop flexibility. The G385A mutant, with a 3.1-fold increase in $k_{cat}$, has a similar rate of zymogen prodomain removal as wild type protein (Fig. 4C). Because these measurements were carried out in saturating conditions, we cannot assess increased efficiency of this mutation under these experimental conditions because product release is the rate-limiting step.

**DISCUSSION**

The crystal structures of the *S. pyogenes* secreted cysteine protease SpeB in the mature apo form and in complexes with E64 and with Ac-AEI5-CHO provide a detailed view of the conformational changes that occur upon activation and substrate binding. These structural results, in combination with the development and application of a newly synthesized robust fluorogenic substrate to probe SpeB activity, the design of a new specific peptide inhibitor, and alanine-scanning mutagenesis, offer significant detailed insights into the broad substrate specificity of SpeB and the functional role of the active site loops.

**Comparison with Previous SpeB Structures**—Interactions between the SpeB active site and E64 have been previously characterized by NMR and mutational analysis (22). Likewise, an E64-like molecule was co-purified and crystallized in the mature SpeB X-ray structure 2UZI; however, the identity of the compound was not confirmed (21). Specific interactions between E64 and SpeB residues, as indicated by chemical shift perturbations in the NMR studies, are well conserved in our X-ray structure. In the NMR study, residues that caused the most significant chemical shifts, including Gly281 and Val334, were mutated to alanine, and the effects on azocasein hydrolysis and prodomain removal of a catalytically inactive C192S mutant of the SpeB zymogen were examined (22). Mutation of Gly281 to Ala resulted in a 15.3–46-fold decrease in catalytic efficiency, and the authors concluded that the region of the protein containing Gly281 (residues 279–281) directly interacted with E64. Based on our E64 complex structure, the addition of a methyl group via the G281A mutation would introduce steric van der Waals clashes and occlude a portion of the active site from binding E64. Mutation of the hydrophobic S2 pocket
residue Val^{334} to alanine precipitously reduced SpeB activity by 131–420-fold (22). The V334A mutation would not drastically change the volume of the S2 pocket; however, because this residue is the primary source of hydrophobicity within the pocket, reduction to an alanine side chain would greatly alter the electrostatic surface potential and modify the substrate specificity of SpeB (Fig. 2C). Both G281A and V334A mutants proteolyze the inactive zymogen SpeB with altered digestion intermediates, supporting the likelihood of changes in substrate specificity due to both mutations (22).

*Role of C-terminal Loop in SpeB Function*—In the NMR study of mature SpeB, the C-terminal loop had increased flexibility and dynamic motion upon binding E64 (22). The conformational exchange was measured in the picosecond to nanosecond time scale range and not on the slower microsecond to millisecond time scale typical for standard active site loops associated with enzyme catalysis (43, 44). In this study, increased hydrolysis of the fluorogenic substrate for the G384A and G385A mutants, which lock the C-terminal loop in the open conformation, provide evidence that the C-terminal loop is not required for substrate hydrolysis by SpeB. Consistent with the NMR findings, the C-terminal loop does not conform to the induced fit theory, whereby active site loops are proposed to be essential for the protection of bound substrates and reactive intermediates and/or help to stabilize transition state intermediates during the course of a reaction. Many kinases, including protein kinase A (PKA) and MEK1, contain glycine-rich active site loops that position ATP for γ-phosphate transfer and are similar in length to the C-terminal loop of SpeB (45). However, unlike SpeB, these kinase loops exhibit induced fit and partially close over the active site upon ATP binding (45–48). It is therefore probable that the primary biological function of the C-terminal loop is not associated with the regulation of SpeB hydrolysis but is involved in substrate recognition and product release.

*Targeting the SpeB Active Site for Inhibitor Design*—SpeB has been studied as a virulence factor for decades, although its importance for *S. pyogenes* pathogenesis, as assessed by genetic knock-out models, has been highly debated. The development of specific small molecules directed toward SpeB would promote the freedom to apply reversible, rapid, and dosable chemical methodologies to the functional studies of SpeB under native bacterial conditions and infection models. The broad substrate specificity that SpeB shares with other papain fold cysteine proteases presents a challenge in the development of specific small molecule active site-directed inhibitors. Furthermore, high affinity compounds typically contain an irreversible reactive warhead, such as the epoxide moiety of E64, that exploits the nucleophilicity of the protease active site residues and, in turn, tends to diminish specificity toward individual target proteins (49). Highly ordered waters identified in the apo-SpeB active site provide insight into the potential design of novel active site-directed inhibitors because these atoms can serve as a foundation on which a compound scaffold can be constructed (Fig. 5). For example, superposition of E64 in complex with SpeB shows that several of the apo-SpeB active site-bound waters are conserved with E64 atoms. Redesign of E64 to accommodate hydrogen bond acceptors and/or donors at sites where water is found would help to promote specificity and affinity toward SpeB. Surprisingly, two highly ordered water molecules within the apo-SpeB structure superimpose with the E64 Ile side-chain methyl groups located within the hydrophobic S2 pocket (Fig. 5). These ordered waters suggest that hydrogen bond donors and acceptor moieties should also be explored in the design of any small molecules targeting the SpeB S2 binding pocket.

*Comparison of SpeB with Other Proteins with Glycine-rich Active Site Loops*—Along with SpeB, the C10 peptidase family includes *Porphyromonas gingivalis* PrtT peptidase, *P. gingivalis* periodontain, and *Prevotella intermedia* interpain A (40). Sequence alignments show that PrtT peptidase and periodontain share a conserved C-terminal loop similar in length and glycine content with SpeB. Interpain A, the only other member of the family with structural information, consists of a truncated C-terminal loop but has an extended catalytic loop that may serve a similar function to the SpeB C-terminal loop (Fig. 6). Unfortunately, interpain A and the other C10 family members have not been extensively studied, and there is little information on substrate specificity and kinetics.

Superposition of a number of papain-like proteases shows that the structural core of the proteases is conserved but with very little structural similarity in the loop regions (Fig. 6). Similar to SpeB, *S. aureus* staphopain hydrolyzes bradykinin (50) and fibronectin (51) and prefers branched hydrophobic residues at P2 and Arg residues at P1. Despite these similarities in both substrate preferences and structure, staphopain belongs to the C47 peptidase family and shares only 11% sequence homology with SpeB. As with interpain, staphopain has a shortened C-terminal loop and an elongated catalytic loop (Fig. 6). Members of the large C1 peptidase family, including *Carica
These x-ray structures, the use of a new fluorogenic peptide substrate and substrate mimic peptide inhibitor, and alanine-scanning mutagenesis have shed light on conformational changes that take place during the life cycle of activation and substrate binding to the active site of SpeB. Our findings demonstrate that the C-terminal loop has a significant role in the recognition and recruitment of substrates and release of products but is not necessary or sufficient for substrate hydrolysis. These insights and the new peptides will provide excellent opportunities and valuable tools that will aid in further probing the function and biological role of SpeB during S. pyogenes infections.

Acknowledgments—We thank Prof. Ian Wilson for critical reading of the manuscript, Drs. Robyn Stanfield and Marc Elsliger for helpful suggestions and computational assistance, Dr. Marc Deller for assistance with protein stability measurements, Prof. Hugh Rosen and James Paulson for instrumentation, and the staffs of Stanford Synchrotron Radiation Lightsource beamline 11.1 and Advanced Photon Source beamline 23-ID-B.

REFERENCES

1. Potempa, J., Golonka, E., Filipek, R., and Shaw, L. N. (2005) Fighting an enemy within. Cytoplasmic inhibitors of bacterial cysteine proteases. Mol. Microbiol. 57, 605–610
2. Musser, J., and DeLeo, F. (2005) Toward a genome-wide systems biology analysis of host-pathogen interactions in group A Streptococcus. Am. J. Pathol. 167, 1461
3. Yu, C. E., and Ferretti, J. J. (1991) Frequency of the erythrogenic toxin B and C genes (speB and speC) among clinical isolates of group A streptococci. Infect. Immun. 59, 211–215
4. Collin, M., and Olsén, A. (2001) EndoS, a novel secreted protein from Streptococcus pyogenes with endoglycosidase activity on human IgG. EMBO J. 20, 3046–3055
5. Kapur, V., Majesky, M. W., Li, L. L., Black, R. A., and Musser, J. M. (1993) Cleavage of interleukin 1β (IL-1β) precursor to produce active IL-1β by a conserved extracellular cysteine protease from Streptococcus pyogenes. Proc. Natl. Acad. Sci. U.S.A. 90, 7676–7680
6. Herwald, H., Collin, M., Müller-Esterl, W., and Björck, L. (1996) Streptococcal cysteine proteinase releases kinins. A virulence mechanism. J. Exp. Med. 184, 665–673
7. Kapur, V., Topouzis, S., Majesky, M. W., Li, L. L., Hamrick, M. R., Hamill, R. J., Patti, J. M., and Musser, J. M. (1993) A conserved Streptococcus pyogenes extracellular cysteine protease cleaves human fibronectin and degrades vitronectin. Microb. Pathog. 15, 327–346
8. Berge, A., and Björck, L. (1995) Streptococcal cysteine proteinase releases biologically active fragments of streptococcal surface proteins. J. Biol. Chem. 270, 9862–9867
9. Nelson, D. C., Garbe, J., and Collin, M. (2011) Cysteine proteinase SpeB from Streptococcus pyogenes. A potent modifer of immunologically important host and bacterial proteins. Biol. Chem. 392, 1077–1088
10. Collin, M., and Olsén, A. (2003) Extracellular enzymes with immunomodulating activities. Variations on a theme in Streptococcus pyogenes. Infect. Immun. 71, 2983–2992
11. Gubba, S., Low, D. E., and Musser, J. M. (1998) Expression and characterization of group A Streptococcus extracellular cysteine proteinase recombinant mutant proteins and documentation of seroconversion during human invasive disease episodes. Infect. Immun. 66, 765–770
12. Kansal, R. G., McGeer, A., Low, D. E., Norbury-Teglund, A., and Koth, M. (2000) Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive group A streptococcal infection cases. Infect. Immun. 68, 6362–6369
13. Walker, M. J., Hollands, A., Sanderson-Smith, M. L., Cole, J. N., Kirk, J. K.,
Henningsham, A., McArthur, J. D., Dinkla, K., Aziz, R. K., Kansal, R. G., Simpson, A. J., Buchanan, J. T., Chhatwal, G. S., Koth, M., and Nizet, V. (2007) DNase Sda1 provides selection pressure for a switch to invasive group A streptococcal infection. Nat. Med. 13, 981–985

14. Cole, J. N., McFarland, J. D., McKay, F. C., Sanders-Smith, M. L., Cork, A. J., Ranson, M., Rohde, M., Itzek, A., Sun, H., Ginsburg, D., Koth, M., Nizet, V., Chhatwal, G. S., and Walker, M. J. (2006) Trigger for group A streptococcal M1T1 invasive disease. FASEB J. 20, 1745–1747

15. Liu, T. Y., and Elliott, S. D. (1965) Activation of streptococcal proteinase and itszymogen by bacterial cell walls. Nature 206, 33–34

16. Hauser, A. R., and Schlievert, P. M. (1990) Nucleotide sequence of the streptococcal pyrogenic exotoxin type B gene and relationship between the toxin and the streptococcal proteinase precursor. J. Bacteriol. 172, 4536–4542

17. Nomizu, M., Pietrzynski, G., Kato, T., Lachance, P., Menard, R., and Ziomek, E. (2001) Substrate specificity of the streptococcal cysteine protease. J. Biol. Chem. 276, 44551–44556

18. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Morkiani, N. W., Oefner, R., Reed, 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

20. Ortonowski, Z., and Minow, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326

21. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674

22. Collaborative Computational Project, Number 4 (1994) The CCP4 suite. Programs for protein crystallography. Acta Crystallogr. D 50, 760–763

23. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D 66, 486–501

24. Adams, P. D., Afonine, P. V., Bunkóczi, 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., Oefner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2009) PHENIX. A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D 66, 213–221

25. 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. Appl. Crystallogr. 26, 283–291

26. Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) Errors in protein structures. Nature 381, 272

27. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity. All-atom structure validation for macromolecular crystallography. Acta Crystallogr. D 66, 12–21

28. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. Nucleic Acids Res. 28, 235–242

30. Musser, J. M., Stockbauer, K., Kapur, V., and Rudgers, G. W. (1993) PROCHECK. A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 213–221

31. Albericio, F. (2004) Developments in peptide and amide synthesis. Curr. Opin. Chem. Biol. 8, 211–221

32. Rodriguez, M., Llaneres, M., Doulut, S., Heitz, A., and Martinez, J. (1991) A facile synthesis of chiral n-protected β-amino alcohols. Tetrahedron Lett. 32, 923–926

33. Albericio, F. (2004) Developments in peptide and amide synthesis. Curr. Opin. Chem. Biol. 8, 211–221
54. Brak, K., Kerr, I. D., Barrett, K. T., Fuchi, N., Debnath, M., Ang, K., Engel, J. C., McKerrow, J. H., Doyle, P. S., Brinen, L. S., and Ellman, J. A. (2010) Nonpeptidic tetrafluorophenoxymethyl ketone cruzain inhibitors as promising new leads for Chagas disease chemotherapy. *J. Med. Chem.* **53**, 1763–1773

55. Alphey, M. S., and Hunter, W. N. (2006) High-resolution complex of papain with remnants of a cysteine protease inhibitor derived from *Trypanosoma brucei*. *Acta Crystallogr.* **F** **62**, 504–508

56. Guncar, G., Klemencic, I., Turk, B., Turk, V., Karaoglanovic-Carmona, A., Juliano, L., and Turk, D. (2000) Crystal structure of cathepsin X. A flip-flop of the ring of His-23 allows carboxymonopeptidase and carboxydipeptidase activity of the protease. *Structure* **8**, 305–313