THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 11, Issue of March 15, pp. 9512–9519, 2002
Printed in U.S.A.

Crystal Structure of a Bacterial Signal Peptidase Apoenzyme

IMPLICATIONS FOR SIGNAL PEPTIDE BINDING AND THE SER-LYS DYAD MECHANISM*

Received for publication, November 15, 2001
Published, JBC Papers in Press, December 10, 2001, DOI 10.1074/jbc.M110983200

Mark Paetzel‡§, Ross E. Dalbey¶, and Natalie C. J. Strynadka‡‖

From the ‡Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia, V6T 1Z3 Canada and the ¶Department of Chemistry, The Ohio State University, Columbus, Ohio 43210

We report here the x-ray crystal structure of a soluble catalytically active fragment of the *Escherichia coli* type I signal peptidase (SPase-(Δ2–75)) in the absence of inhibitor or substrate (apoenzyme). The structure was solved by molecular replacement and refined to 2.4 Å resolution in a different space group ($P4_12_12$) from that of the previously published acyl-enzyme inhibitor-bound structure ($P2_12_12$) (Paetzel, M., Dalbey, R. E., and Strynadka, N. C. J. (1998) *Nature* 396, 186–190). A comparison with the acyl-enzyme structure shows significant side-chain and main-chain differences in the binding site and active site regions, which result in a smaller S1 binding pocket in the apoenzyme. The apoenzyme structure is consistent with SPase utilizing an unusual oxyanion hole containing one side-chain hydroxyl hydrogen (Ser-88 Oγ*H*) and one main-chain amide hydrogen (Ser-90 NH). Analysis of the apoenzyme active site reveals a potential deacylating water that was displaced by the inhibitor. It has been proposed that SPase utilizes a Ser-Lys dyad mechanism in the cleavage reaction. A similar mechanism has been proposed for the LexA family of proteases. A structural comparison of SPase and members of the LexA family of proteases reveals a difference in the side-chain orientation for the general base lysine, both of which are stabilized by an adjacent hydroxyl group. To gain insight into how signal peptidase recognizes its substrates, we have modeled a signal peptide into the binding site of SPase. The model is built based on the recently solved crystal structure of the analogous enzyme LexA (Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. J. (2001) *Cell* 106, 1–10) with its bound cleavage site region.

Type I signal (leader) peptidase (SPase,[1] EC 3.4.21.89) is the membrane-bound serine endoprotease responsible for cleavage of the amino-terminal signal (or leader) peptide extensions from secretory proteins and some membrane proteins (for recent reviews see Refs. 1 and 2). The type I SPase from *Escherichia coli* has served as the model Gram-negative SPase and is the most thoroughly characterized SPase to date. It has been cloned (3), sequenced (4), overexpressed (5), purified (4, 6, 7), and kinetically (8) and structurally (9) characterized. Site-directed mutagenesis (8, 10) and chemical modification (8, 11) studies are consistent with SPase utilizing a Ser-Lys dyad mechanism, whereby Ser-90 serves as the nucleophile and Lys-145 serves as the general base. *E. coli* SPase (323 amino acids, 35,988 Da, pI 6.9) contains two amino-terminal transmembrane segments (residues 4–28 and 58–76), a small cytoplasmic region (residues 29–58), and a carboxyl-terminal periplasmic catalytic region (residues 77–323). A catalytically active fragment of SPase (SPase-(Δ2–75)), which lacks the two transmembrane segments and the cytoplasmic domain, has been cloned, overexpressed, characterized (12, 13), and crystallized (14). The Δ2–75 construct required detergent or lipid for optimal activity (13) and crystallization (14). The crystal structure of SPase-(Δ2–75) in complex with a covalently bound β-lactam-based inhibitor revealed directly that Ser-90 was the nucleophile and that the position of the Nζ of Lys-145 was consistent with a role as the general base (9) in the catalytic mechanism. The shallow hydrophobic pockets observed adjacent to the catalytic residues helped to explain the Ala-*X*-Ala substrate specificity. The structure also revealed that SPase is an unusual serine protease in that it attacks the substrate scissile amide carbonyl from the *si*-face rather than the *re*-face as seen in most serine proteases (15, 16).

An important component of the catalytic machinery in serine proteases is the oxyanion hole that works by neutralizing the developing negative charge on the scissile carbonyl oxygen during the formation of the tetrahedral intermediates. Typically oxyanion holes are formed by two main-chain amide hydrogens that serve as hydrogen bond donors to the developing oxyanion (17). The inhibitor-bound structure of *E. coli* SPase (9) revealed only one hydrogen bond donor (Ser-90 NH) to the inhibitors acyl carbonyl group. Although there was no other potential main-chain hydrogen bond donor it was suggested that the side-chain hydroxyl group from Ser-88 could fulfill the role if it were not blocked by the bulky thiazolidine ring of the inhibitor. Here we present crystallographic evidence that in agreement with recent mutagenesis data (18) strongly suggests that the hydroxyl group of Ser-88 does contribute to the oxyanion hole in the Gram-negative signal peptidases. A detailed analysis of the active site of SPase-(Δ2–75) in the absence of inhibitor or substrate (apoenzyme) reveals significant differences from the inhibitor-bound structure and reveals a strong candidate for the key deacylating water in the proposed mech-

* This work was supported by the Burroughs Wellcome Foundation New Investigator Award (to N. C. J. S.), the Canadian Bacterial Diseases Network of Excellence, and the Howard Hughes Medical Institute International Scholar Award (to N. C. J. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1KN9) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Funded by a Canadian Institute of Health Research Postdoctoral Fellowship.

‖ Recipient of a Canadian Institute of Health Research Scholarship. To whom correspondence should be addressed: University of British Columbia, Dept. of Biochemistry and Molecular Biology, Faculty of Medicine, 2146 Health Sciences Mall, Vancouver, B. C., V6T 1Z3 Canada. Tel.: 604-822-8032 (laboratory), 604-822-0789 (office); Fax: 604-822-5227; E-mail: natalie@byron.biochem.ubc.ca.

[1] The abbreviations used are: SPase, signal peptidase; Δ2–75, the construct of *E. coli* signal peptidase lacking residues 2–75, which corresponds to the transmembrane segments and the cytoplasmic region; r.m.s.d., root mean-square deviation; UmuD′, the UmuD protein with the N-terminal 24 residues removed.

anism. We also present a structural comparison with other Ser/Lys proteases from the LexA family of proteases. A signal peptide is modeled into the binding sites of SPase based on the structure of the bound cleavage site region in LexA.

TABLE I
Crystallographic data

Data Collection[a]	
Space group	$P4_12_12$ (92)
Unit cell dimensions a,b,c (Å)	112.4 × 112.4 × 198.7
Molecules in a.s.u.	4
Vm (Å^3/Da)	2.78
Resolution (Å)	50.0 − 2.40
Total observed reflections	277,201
Unique reflections	49,984
% possible	98.8 (100.0)
I/σ(I) (%)	25.4 (5.8)
R_{merge}[b]	4.1 (29.6)
Refinement	
Residues	907
Atoms	6912
Waters	258
R[c]	23.9
R_{free}[d]	27.9
r.m.s. deviations	
Bonds (Å)	0.0066
Angles (°)	1.4524
Overall B (Å^2) for protein atoms	45.1
Overall B (Å^2) for water atoms	40.9

[a] The data collection statistics in parentheses are the values for the highest resolution shell (2.49 to 2.40 Å).

[b] $R_{merge} = \Sigma|\,|I_{o,i}| - |I_{ave,i}|\,|/\Sigma|I_{ave,i}|$, where $I_{ave,i}$ is the average structure factor amplitude of reflection I, and $I_{o,i}$ represents the individual measurements of reflection I and its symmetry equivalent reflection.

[c] $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$ (on all data 2.4 − 50.0 Å).

[d] $R_{free} = \Sigma_{hkl\ c\ T}(|F_o| - |F_c|)^2/\Sigma_{hkl\ \subset\ T}|F_o|^2$, where $\Sigma_{hkl\ \delta\ T}$ are reflections belonging to a test set of 10% of the data, and F_o and F_c are the observed and calculated structure factors, respectively.

EXPERIMENTAL PROCEDURES

Materials—The SPase-(Δ2–75) protein (relative molecular weight (M_r) 27,952 by electrospray ionization mass spectrometry analysis, Ref. 12, 249 amino acid residues, pI 5.6 by isoelectric focusing analysis, Ref. 13) was expressed and purified as described previously (14). The protein (10 mg/ml) was solubilized in 20 mM Tris-HCl, pH 7.4, 0.5% Triton X-100.

Crystallization—Subtle adjustments to the previously published crystallization conditions (14) resulted in the formation of a new crystal form of Δ2–75. The crystals of apoenzyme SPase-(Δ2–75) were grown by sitting drop vapor diffusion with a reservoir consisting of 0.70 M ammonium dihydrogen phosphate, 0.1 M sodium citrate, pH 5.4 and a drop consisting of 4 μl of protein and 8 μl of reservoir solution. The crystals formed in drops that had been left undisturbed for 6 weeks at a consistent temperature of 22 °C. The crystals belong to the tetragonal space group $P4_12_12$ with unit cell dimensions of a = 112.4 Å, b = 112.4 Å, c = 198.7 Å. Given the unit cell dimensions and the molecular weight, the specific volume (V_m) (19) is 2.78 Å^3/Da for four molecules in the asymmetric unit. The x-ray diffraction intensities were measured at 100 K on beamline BL7–1 (Stanford Synchrotron Radiation Laboratory) at a wavelength of 1.08000 Å. The data were processed using the programs DENZO and SCALEPACK (20).

Phasing, Model Building, Refinement, and Structural Analysis—A molecular replacement solution was found using the program AmoRe (21), and the atomic coordinates of the previously published crystal structure of the SPase-(Δ2–75) inhibitor complex (PDB, 1B12). Rebuilding of some of the flexible regions was required because of differences resulting from crystal packing. Model building and analysis was performed with the program XFIT within the suite XTALVIEW (22). Refinement of the structure was carried out using the program CNS (23). The secondary structural analysis was performed with the program PROMOTIF (24). The utilities Lsq_explicit, Lsq_improve, and Lsq_molecule within the program O (25) were used to superimpose molecules. The binding pocket analysis was performed using the CAST server (26) (cast.engr.uic.edu/cast/).

Modeling of a Signal Peptide Bound to SPase—A model of a signal peptide bound to the binding sites of SPase was prepared by using the structure of the covalently bound penem inhibitor complex to guide the

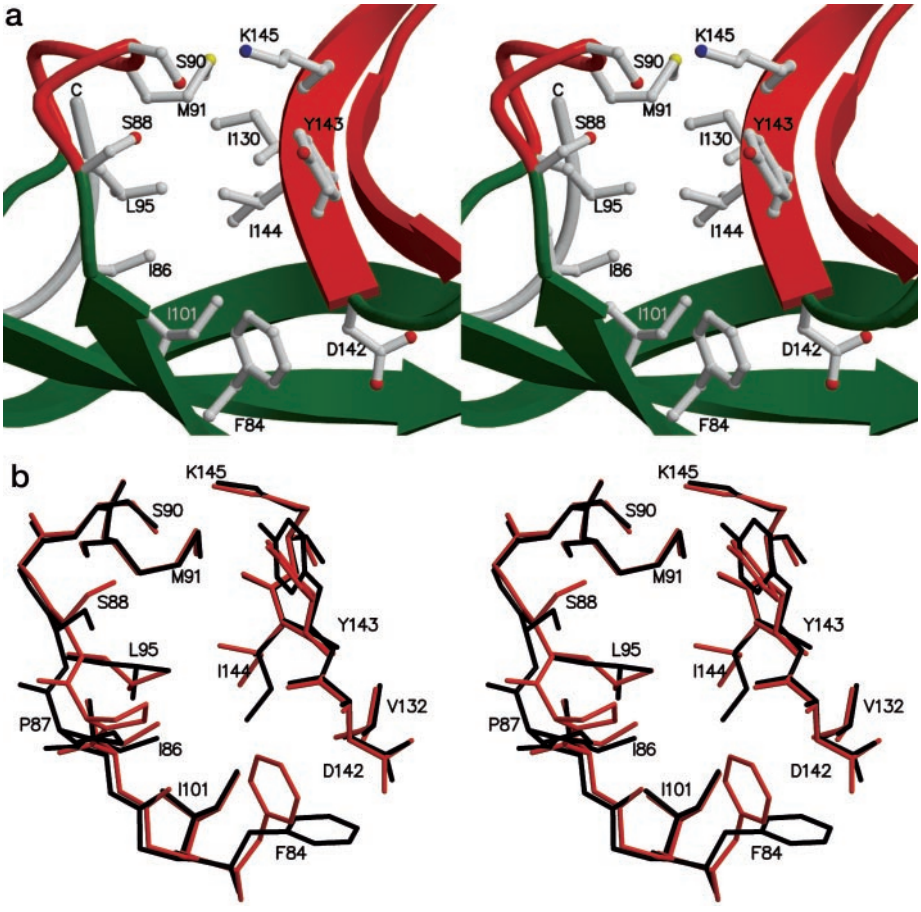

FIG. 1. **The active site of *E. coli* SPase apoenzyme.** *a*, stereographic rendering of the active site of *E. coli* SPase in the absence of inhibitor or substrate. The regions of the model that are colored *red* correspond to the boxes of conserved amino acid sequence in SPase. *b*, a comparison of the active site residues of *E. coli* SPase with and without a bound inhibitor. The apoenzyme structure is shown in *red* and the acyl-enzyme inhibitor-bound structure is shown in *black*. The inhibitor has been removed for clarity.

position of the P1 Ala residue of the substrate. The P3 Ala position to the P6 substrate position was guided by the structure of the cleavage site domain of the analogous enzyme LexA (PDB code, 1JHE). The model was prepared using the program O (25). The model was energy-minimized using the program CNS (23).

Figure Preparation—Figs. 1, 3*b*, 4, and 5 were prepared using the programs Molscript (27) and Raster3D (28). Fig. 3*a* was prepared using the programs XFIT (22) and Raster3D (28). Fig. 2 was prepared using the programs GRASP, UNGRASP (29), and Raster3D (28).

Accession Numbers—Atomic coordinates for the apoenzyme SPase-(Δ2–75) structure have been deposited with the RCSB Protein Data Bank under the accession code 1KN9 (30). Atomic coordinates for the SPase-(Δ2–75) penem inhibitor complex structure (9) are under the accession code 1B12.

RESULTS AND DISCUSSION

The catalytic domain of *E. coli* type I signal peptidase (SPase-(Δ2–75)) has been crystallized in the absence of inhibitor or substrate (apoenzyme). Earlier attempts to crystallize the apoenzyme of SPase-(Δ2–75) resulted in poorly diffracting crystals that belonged to the tetragonal space group $P4_22_12$ (14). Adjustments to the previously published crystallization conditions have resulted in improved crystals of the apoenzyme that belong to the tetragonal space group $P4_12_12$ with diffraction to 2.4 Å resolution (see "Experimental Procedures" for details). Typical of many enzyme systems, the apoenzyme form is apparently less ordered than that of the inhibitor bound form. Orthorhombic ($P2_12_12$) crystals of the *E. coli* SPase acyl-enzyme inhibitor complex were shown to diffract to beyond 1.9 Å resolution (9). The structure of the apoenzyme was solved by molecular replacement and refined to 2.4 Å resolution with an *R*-factor of 0.239 and an *R*-free value of 0.279 (see Table I for complete refinement statistics).

Structural Difference between the Apoenzyme and the Acyl-enzyme—Both the apoenzyme structure and the acyl-enzyme inhibitor-bound structure reveal two different paths taken for the N-terminal region (residues 80–85) just preceding the nucleophilic Ser-90. In two of the four molecules in the asymmetric unit this region is in a β-strand conformation and in the other two there is a bulge in this region. This difference in the N-terminal region appears to be the result of crystal-packing forces. If we compare molecules with similar N-terminal region conformations we see little change in the overall structure upon formation of the inhibitor acyl-enzyme complex. The root mean-square deviation (r.m.s.d.) upon superposition of the apoenzyme structure (molecule A) and the acyl-enzyme inhibitor-bound structure (molecule A) is 0.558 Å for 234 common Cα atoms. Most of the differences arise from the extended loops and hairpins near the solvent surface (108–124, 170–176, 198–202, and 304–313). Comparing molecules with dissimilar N-terminal region conformations gives a significantly larger r.m.s.d.; for example superimposing molecule B from the apoenzyme structure onto molecule A in the inhibitor-bound acyl-enzyme structure gives an r.m.s.d. of 0.974.

Comparing the apoenzyme and acyl-enzyme active site residues reveals very little change in the position of the nucleophilic Ser-90 and general base Lys-145 (Fig. 1). There are however significant differences in the substrate-binding pockets. The shallow and hydrophobic substrate-binding pockets of *E. coli* SPase are consistent with the substrate specificity for small and neutral residues at the P1 and P3 position (P1 and P3 (31) correspond to the −1 and −3 position relative to the cleavage site) (32). The S1 pocket is made up of nonpolar atoms from residues 86–88, 90–91, 95, and 143–145. The S3 pocket is made up of nonpolar atoms from residues 84, 86, 101, 132, 142, and 144. In the apoenzyme structure the main-chain of Pro-87 and Ser-88 is positioned such that the S1 binding pocket is narrower than that seen in the acyl-enzyme inhibitor-bound structure (Figs. 1 and 2). The side chain of Phe-84 has also

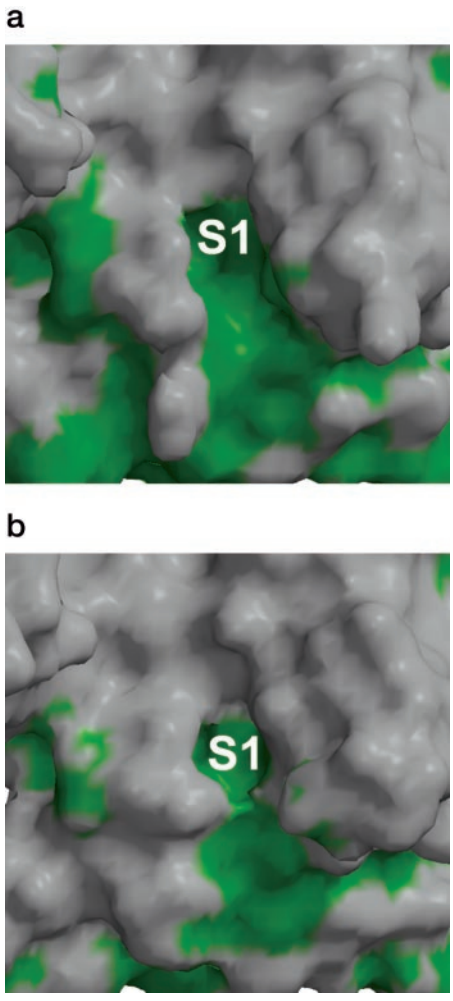

FIG. 2. **The molecular surface of the *E. coli* SPase binding site.** *a*, the molecular surface of the acyl-enzyme inhibitor-bound *E. coli* SPase. The inhibitor has been removed for clarity. *b*, the molecular surface of the apoenzyme of *E. coli* SPase. The hydrophobic surface is shown in *green*. The S1 substrate binding pocket is indicated.

rotated ~80° about χ_1, toward the direction of the binding pocket (from −83° in the inhibitor-bound structure to −160° in the apoenzyme structure). Differences are also seen in the side-chain positions of Ile-144, Leu-95, and Tyr-143 (Fig. 1), which also result in a more narrow and shallow binding pocket (Fig. 2). Molecular surface analysis of the binding pocket region (26) confirms quantitatively that both the area and the volume of the S1 binding pocket are smaller in the apoenzyme structure. The S1 binding site in the acyl-enzyme structure is the largest pocket on the SPase molecular surface with an area of 156.8 $Å^2$ and a volume of 224.1 $Å^3$. The S1 binding site in the apoenzyme structure is the third largest pocket on the SPase molecular surface with an area of 145.5 $Å^2$ and a volume of 129.3 $Å^3$.

Another major difference in the active site is seen in the side-chain position of Ser-88, where χ_1 is rotated by ~180° from its position in the acyl-enzyme structure. The position of the Ser-88 hydroxyl group in the apoenzyme is consistent with its role in the formation of the SPase oxyanion hole and stabilization of the tetrahedral oxyanion intermediate. Paetzel *et al.* (9) had previously proposed that the Ser-88 χ_1 angle in the inhibitor-bound structure was a result of a clash with the inhibitors thiazoladine ring (9) and that in the absence of the bulky inhibitor, the Ser-88 β-hydroxyl may be able to contribute to the oxyanion stabilization. These structural findings are consistent with recently published mutagenesis and kinetic stud-

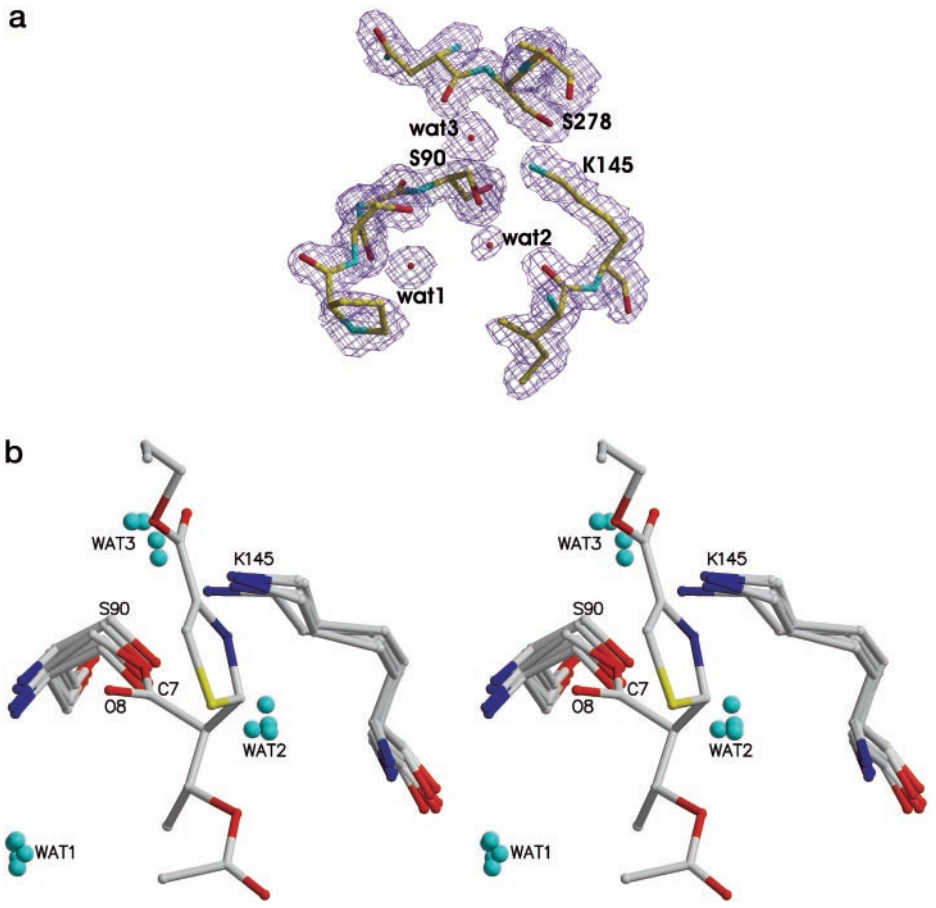

FIG. 3. **Active site waters in *E. coli* SPase.** *a*, a cross-validated sigmaa 2Fo-Fc electron density map contoured at 1.5σ surrounding the active site residues of *E. coli* SPase apoenzyme. *b*, each molecule in the asymmetric unit of the SPase apoenzyme structure has been superimposed to show the positions of the conserved active site waters in SPase. The active site of the penem inhibitor-bound structure is also superimposed on the apoenzyme structure to show how the penem would have displaced WAT2 and WAT3. Note that WAT1 is also observed in the inhibitor-bound structure. The ester carbonyl carbon (C7) and oxygen (O8), which mimics the ester of the acyl-intermediate, are labeled.

ies (18), which showed that the S88T mutant maintained near wild-type activity whereas the S88C and S88A mutations resulted in a 740-fold and >2440-fold reduction in activity, respectively. Sequence analysis reveals that SPases from Gram-negative bacteria and those from the eukaryotic endoplasmic reticulum contain a Ser at the 88 position (*E. coli* numbering). There is some precedent for a serine hydroxyl acting as a component of an oxyanion hole (in lipolytic enzymes such as cutinase, Ref. 33), but SPase is the first characterized serine protease to utilize a side-chain hydroxyl in this manner.

Active Site Waters—The crystal structure of the apoenzyme of SPase-(Δ2–75) reveals electron density and appropriate coordination for three water molecules near the active site region conserved in each of the four molecules in the asymmetric unit (Fig. 3*a*). The water designated WAT1 is coordinated by Met-91 NH, Ser-88 O, and Leu-95 O. The water designated WAT2 is coordinated by Ile-144 NH. The water designated WAT3 is coordinated by Lys-145 Nζ. WAT1 was also present in the SPase acyl-enzyme inhibitor-bound structure (9). WAT2 AND WAT3 were not present in the acyl-enzyme inhibitor-bound structure as they would have been sterically displaced by the presence of the inhibitor (Fig. 3*b*). This suggests either WAT2 or WAT3 as the deacylating water, with the long lived stability of the penem inhibitor acyl-enzyme complex a result of displacement of the deacylating water. Inhibitors for TEM-1 β-lactamase have been designed based on this principle (34, 35), and analysis of serine protease acyl-enzyme complexes have revealed similar effects (36, 37).

For a water to function as the deacylating (also called the "catalytic," "hydrolytic," or "nucleophilic") water during the deacylation step of the hydrolytic reaction it needs to be positioned such that it is within hydrogen bonding distance to a general base for activation as well as at a suitable angle of approach with respect to the carbonyl of the ester intermediate (the so-called Bürgi angle of ~107°; Ref. 38). By superimposing the apoenzyme structure and the inhibitor-bound acyl-enzyme structure we can approximate the distance from WAT2 and WAT3 to the ester carbonyl carbon (C7, O8 in Fig. 3*b*) as well as the angle subtended by the water oxygen (WAT2,3), the ester carbonyl carbon (C7), and the carbonyl oxygen (O8) (Table II and Fig. 3*b*). WAT2 would most likely be displaced by the binding of the signal peptide because of its β-sheet style hydrogen-bonding interactions with the strand containing Ile-144, which coordinates WAT2 (Fig. 5*a*). Of the three water molecules seen near the active site of the SPase apoenzyme, WAT3 would be the most likely candidate for a deacylating water. It is approximately equidistant between Lys-145-Nζ (3.3 Å) and the ester carbonyl carbon (3.8 Å) of the inhibitor (C7) and has an angle of approach to the ester carbonyl carbon (Burgi angle) of 78°.

A Structural Comparison of Ser/Lys Proteases—Based on our structural evidence, the SPase and the LexA families of proteases belong to the same evolutionary clan (SF) of serine proteases. A structural comparison of SPase and UmuD′ protease revealed that these enzymes share the same protein fold in the catalytic core (39). The crystal structures of two other members of the LexA family of self-cleaving proteases (LexA, Ref. 40 and λ phage cI, Ref. 41) have also been recently solved. The central catalytic domain of SPase can be superimposed on the structures of UmuD′ (42), LexA (40), and cIλ (41) with r.m.s. deviations of 1.54, 1.61, and 1.54 Å for 68, 71, and 74 common Cα atoms respectively (Fig. 4*a*). Analysis of the active sites of these superimposed structures reveal that these enzymes share a common hydrogen-bonding environment for the Lys general base in that it is hydrogen bonded not only to the Ser nucleophile Oγ but also to another side-chain hydroxyl Oγ

TABLE II
Active site water molecules in the E. coli SPase apoenzyme structure

WAT[a]	B-factor	Coordination	Distance to Lys-145 Nζ	Distance to penem ester carbonyl carbon C7	Bürgi angle[b]
	Å^2	*Å*	*Å*	*Å*	*degrees*[c]
1[d]	27.8	Met-91NH(2.7) Ser-88 O(2.7) Leu-95 O(2.9)	8.0	5.1	73
2[e]	37.2	Ile-144NH(3.3)	4.5	3.0	112
3[f]	40.5	Lys-145 Nζ(3.2)	3.3	3.8	78

[a] The values cited are an average of the values observed in each of the four molecules in the asymmetric unit.
[b] Bürgi angle, the angle subtended by the water oxygen, the ester carbonyl carbon, and the ester carbonyl oxygen ($O_{wat} - C_{acyl} - O_{acyl}$) (38).
[c] Based on superposition of apoenzyme structure onto acyl-enzyme inhibitor complex structure.
[d] WAT 1, solvent atoms; 1S, 4S, 7S, and 10S, which coordinate with molecules A, B, C, and D, respectively, in the PDB file 1KN9.
[e] WAT 2, solvent atoms 2S, 5S, 8S and 11S, which coordinate with molecules A, B, C, and D, respectively, in the PDB file 1KN9.
[f] WAT 3, solvent atoms 3S, 6S, 9S, and 12S, which coordinate with molecules A, B, C, and D, respectively, in the PDB file 1KN9.

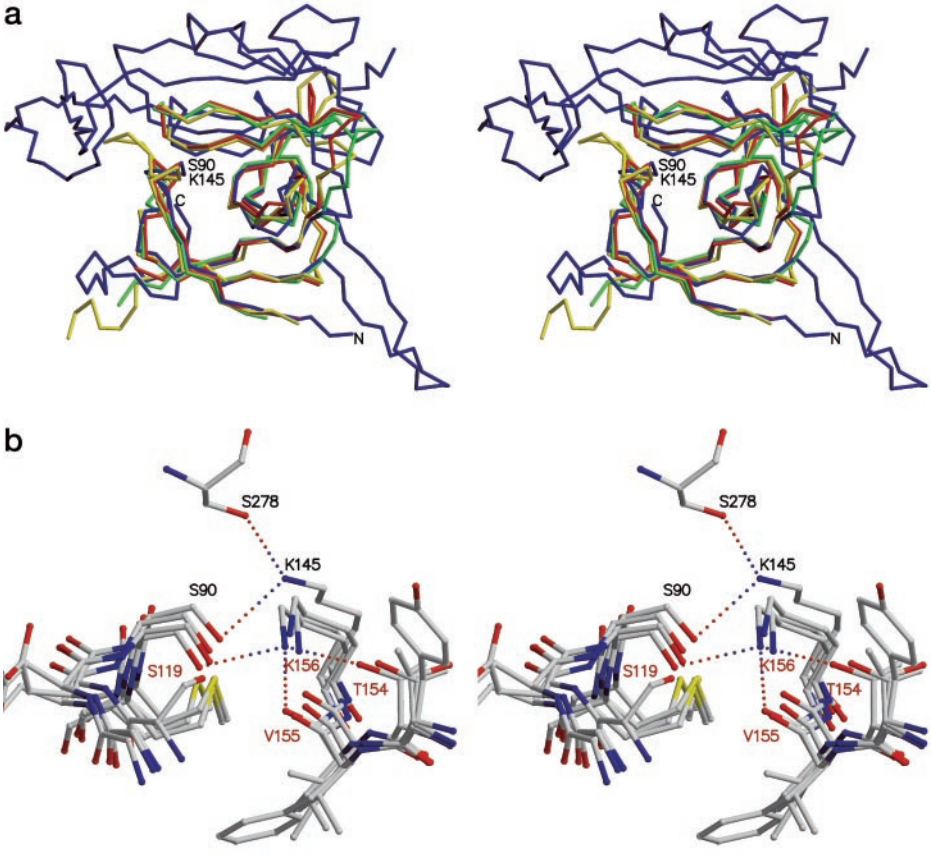

FIG. 4. **The structural comparison of Ser/Lys proteases.** *a*, Cα trace of SPase (*blue*) (9) overlapped on the C-terminal protease domains of UmuD′ (*red*) (42), cI-λ (*yellow*) (41), and LexA (*green*) (40). *b*, a *ball-and-stick* rendering of the overlapped active sites of SPase, UmuD′, cIλ, and LexA. The hydrogen bond interactions for the Nζ of the general base lysine for SPase (Lys-145) and LexA (Lys-156) are shown.

(Fig. 4*b*). The position of the second coordinating hydroxyl Oγ is different in SPase from that of the LexA family of proteases. The Lys-145 Nζ of SPase is hydrogen-bonded to Ser-278 Oγ and Ser-90 Oγ. The Lys-156 Nζ of LexA is coordinated by Thr-154 Oγ and Ser-119 Oγ. In SPase the non-nucleophilic Ser-278 Oγ resides on a β-strand that lays adjacent but perpendicular to the β-strands that contains the general base Lys-145. The similar hydroxyl Oγ in the LexA-like proteases (Thr-154 in Lex A) resides on the same β-strand as the lysine general base. The different orientations for the second coordinating hydroxyl Oγ results in slightly different orientations (side-chain χ_4 angle) for the ϵ-amino group of the Lys general base. For example, the χ_4 angle for Lys-145 in the apoSPase structure is 178.3° (an average for the four molecules in the asymmetric unit) and the χ_4 angle for Lys-97 in the UmuD′ structure (1umu) is −75.5° (an average for the two molecules in the asymmetric unit). It should be noted that although the Thr residue is conserved in the LexA family, the Thr Oγ in the UmuD′ structure is pointing in the opposite direction to that seen in the LexA and cIλ structures.

The LexA family of proteases have an additional hydrogen bond to the Lys general base that SPase does not have. The Lys Nζ is within hydrogen bonding distance to the main-chain oxygen of the residue preceding the general base (Fig. 4*b*). In the case of LexA for example, the neutral ϵ-amino group of Lys-156 (the deprotonated state is a requirement for it serving as the general base) would have two hydrogen bond acceptors (Val-155 O and Thr-154 Oγ) and one hydrogen bond donor (Ser-119 Oγ).

Photosystem II D1 C-terminal processing protease (CtpA) has no sequence or structural similarity to SPase or the LexA family of proteases but utilizes a similar Ser-Lys dyad mechanism. The recently solved crystal structure of CtpA shows a non-nucleophilic hydroxyl group (Thr-168) within hydrogen-bonding distance to the proposed Lys general base (Lys-397) (43). For a list of the structures used in this comparison see Table III. The class A β-lactamases, which utilize a Ser-Lys dyad mechanism, also have a second Ser hydroxyl hydrogen bond to the general base Lys Nζ (44).

TABLE III
Ser/Lys protease structures

Enzymes	PDB	Reference	Res. Å	R/R_{free}	Mol/asu[a]	Ser nucleophile	Lys general base
SPase-acyl inhibitor[b]	1B12	(9)	1.9	0.220/0.246	4	90	145
SPase-apo	1KN9	This work	2.4	0.239/0.279	4	90	145
UmuD' Se[c]	1UMU	(42)	2.5	0.207/0.303	2	60	97
UmuD' WT[d]	1AY9	(50)	3.0	0.218/0.287	2	60	97
UmuD' NMR	1I4V	(51)	—	—	—	60	97
cI-λ	1F39	(41)	1.9	0.228/0.257	2	149	192
LexA Δ1–67[e] S119A	1JHC	(40)	2.0	0.230/0.278	1	119	156
LexA, S119A	1JHH	(40)	2.1	0.230/0.273	2	119	156
LexA, G85D	1JHF	(40)	1.8	0.240/0.263	2	119	156
LexA, Δ1–67, QM	1JHE	(40)	2.5	0.220/0.284	2	119	156
CtpA Se-Met/L153M/L210M	1FC6	(43)	1.8	0.168/0.237	1	372	397
CtpA[f] Native C2A	1FC7	(43)	2.0	0.187/0.267	1	372	397
CtpA Native C2B	1FC9	(43)	1.9	0.191/0.275	1	372	397
CtpA Native R32	1FCF	(43)	2.1	0.220/0.318	1	372	397

[a] Mol/asu, molecules per asymmetric unit.
[b] SPase, *E. coli* type I signal peptidase.
[c] UmuD' Se, UmuD' with seleno-methionine in place of the methionines.
[d] UmuD' WT, wild-type UmuD'.
[e] LexA, Δ1–67, QM, LexA repressor deletion construct, which lacks residues 1–67 and contains the mutations (L89P, Q92W, E152A, K156A).
[f] CtpA, Photosystem II D1 C-terminal processing protease.

It appears from this structural comparison that the Ser-Lys dyad mechanism requires a third player (Ser/Thr Oγ) for optimal activity, much like the Asp in the Ser-His-Asp catalytic triad mechanism. The second hydroxyl hydrogen bond to the Lys Nζ in the Ser-Lys dyad proteases possibly functions to help align the general base Lys to the nucleophilic Ser hydroxyl. Mutagenesis data show that the SPase Ser-278 is essential for optimal activity (45). The three hydrogen-bonding contacts with the general base Lys in the LexA family of proteases suggest that the Lys Nζ in these enzymes would be very firmly held in place.

Model of a Signal Peptide Bound to a Signal Peptidase—Recently the crystal structure of the self-cleaving LexA repressor was solved with its cleavage site region bound in the active site (40). Using the acyl-enzyme inhibitor-bound SPase structure, along with the structure of the LexA cleavage site, we have modeled the signal peptide from *E. coli* outer membrane protein A (OmpA) bound to the *E. coli* SPase (Fig. 5). The OmpA signal peptide was chosen for this modeling study because *E. coli* SPase shows the best catalytic constants using the preprotein substrate pro-OmpA-nucleaseA (PONA) (46). PONA is a hybrid protein of the OmpA signal peptide attached to the *Staphylococcus aureus* nuclease A. The model shows the signal peptide making a number of β-sheet type of hydrogen-bonding interactions with the β-strands (81–85, 142–149) that line the S1 and S3 substrate binding pockets (Fig. 5*a*). The hydrogen bonds include Gln-20 (P2) O···Ile-144 NH, Gln-20 (P2) NH···Asp-142 O, Ala-19 (P3) NH···Gln-85 O, Thr-17 (P5) O···Gln-85 NH, Ala-16 NH (P6)···Pro-83 O, and Gly-14 NH (P8)···Tyr-81 O. This model suggests possible substrate-enzyme contact points all the way to the P8 position of the signal peptide. The P2, P4, and P5 side-chains are mostly solvent exposed, which is consistent with the lack of specificity at these positions. The model does however suggest the side chains of the P6 and P7 residues may be in position to make significant contact with SPase. Interestingly, the P7 residue (Phe-15) comes in close proximity to Trp-300, which has been shown by site-directed mutagenesis (47) and chemical modification (48) to be important for optimal activity. Many signal peptides contain a Gly or Pro residue at approximately the P6 residue, suggesting that this may represent the transition point from the β-strand region (C-region) to the α-helical region (H-region). By extending the signal peptide β-sheet type of interaction to the P6–P8 position, the H-region is presented in close proximity to the presumed position of the transmembrane segments of SPase (Fig. 5*b*).

Concluding Remarks—There is very little sequence identity among signal peptides except in the cleavage site region (C-region) where there is a preference for small and neutral residues at the P1 and P3 positions. By far the most common residue at these positions is Ala, although Gly and Ser are sometimes observed at the P1 position and Val, Ser, and Leu are sometimes seen at the P3 position. By comparing the apoenzyme and acyl-enzyme structures of SPase we can see that the flexibility in the shallow and hydrophobic binding site region is consistent with the relatively broad substrate specificity of SPase. Although the β-lactam style SPase inhibitor mimics the substrate amide bond it is not clear whether the same changes seen at the active site and binding site upon formation of the inhibitor acyl-enzyme are reflective of the changes that would be observed with the substrate acyl-enzyme. Unpublished results from a crystal structure of a peptide-based inhibitor non-covalently bound to SPase show approximately the same Pro-87 and Ser-88 main-chain position as the acyl-enzyme inhibitor-bound structure. Conversely the peptide-based non-covalently bound inhibitor structure shows a similar Ser-88 χ_1 angle as that in the apoenzyme structure, which would be consistent with SPase utilizing a Ser-88 hydroxyl in its oxyanion hole.[2]

Structural comparisons with the related enzymes from the LexA family of proteases reveal similarities in the hydrogen-bonding network for the general base Lys, yet differences in the orientation of the Lys Nζ supported by the network. It may be that the Ser-Lys dyad mechanism would be more accurately described as a Ser-Lys-(Ser/Thr) triad mechanism.

Because SPase is an essential enzyme found at the bacterial membrane surface it has been actively studied as a potential antibacterial target (49). The crystal structure of the apoenzyme form of *E. coli* SPase and our model of the OmpA signal peptide bound to SPase may be helpful in the rational design of inhibitors and artificial substrates as well as kinetic and mutagenesis experiments. Efforts are now underway to crystallize the full-length *E. coli* SPase bound to a preprotein substrate.

Acknowledgments—We thank Dr. Henry Bellamy for allowing us to use beam-line BL1–5 at the Stanford Synchrotron Radiation Laboratory (SSRL). We thank Dr. Yu Luo for helpful discussions on the structure of LexA.

[2] M. Paetzel, unpublished data.

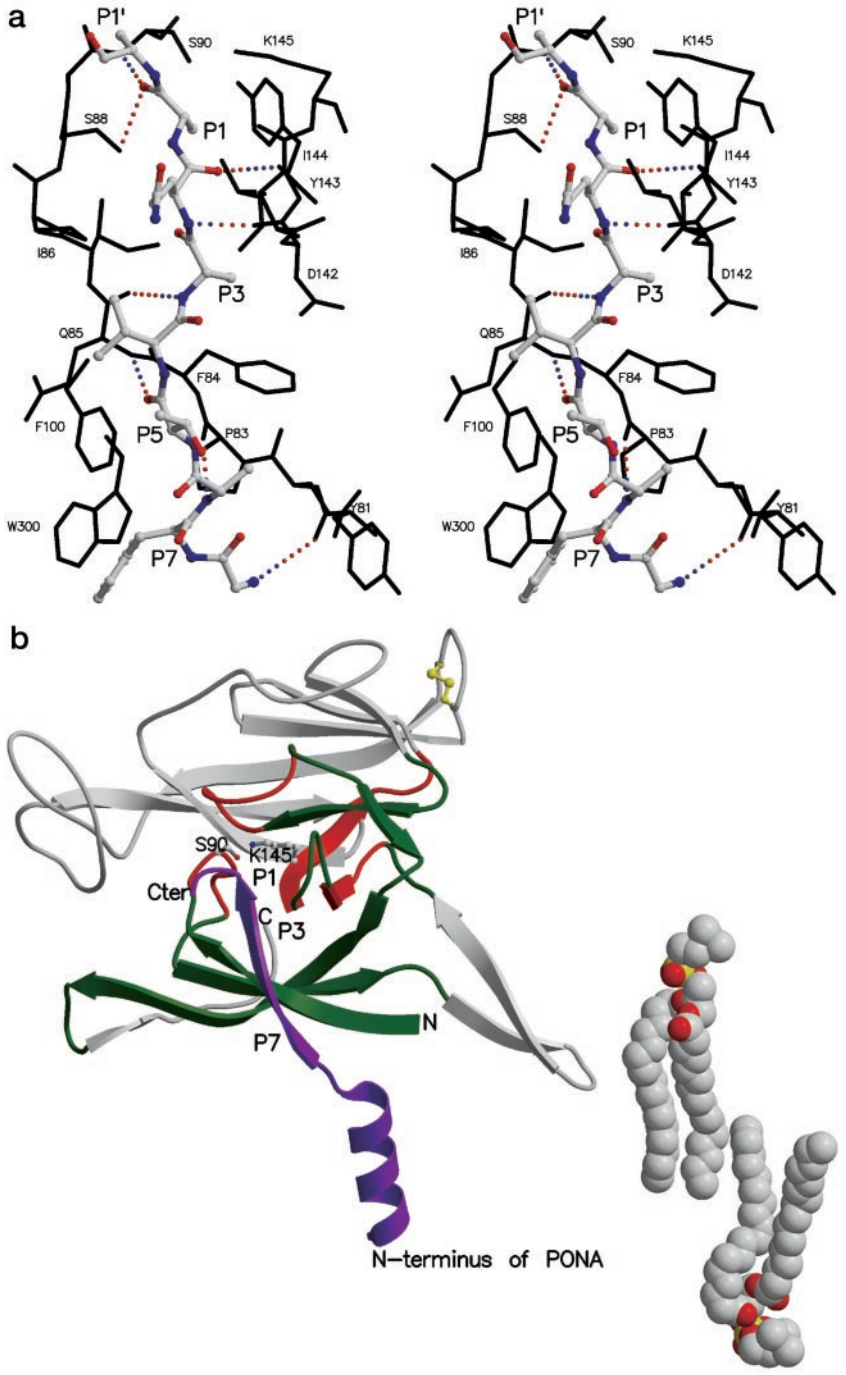

FIG. 5. **A model of the OmpA signal peptide bound to the active site of the *E. coli* SPase.** *a*, the P8 to P1′ residues (Gly-14–Ala-22) of the OmpA signal peptide have been modeled into the substrate binding pockets of *E. coli* SPase. The model is based on the LexA cleavage site (40). The protein is shown in *black stick* and the modeled peptide is in *ball-and-stick*; *gray*, carbon; *blue*, nitrogen; and *red*, oxygen. Potential hydrogen-bonding interactions between the signal peptide and SPase are shown. *b, ribbon diagram* of the OmpA signal peptide modeled into the binding site of *E. coli* SPase-(Δ2–75). The signal peptide is shown in *purple*. The positions of the P1, P3, and P7 residues are indicated. Ser-90 and Lys-145 are shown in *ball-and-stick*. Two phospholipid molecules rendered in CPK are positioned next to the complex to give a perspective of the thickness of the phospholipid bilayer relative to the signal peptidase/signal peptide complex. The phospholipid molecules were taken from a molecular dynamics simulation of a bilayer (52). The regions of *E. coli* SPase that are present in all type I SPases are shown in *red* and *green*. The *red* regions correspond to the conserved boxes of sequence.

REFERENCES

1. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (2000) *Pharmacol. Ther.* **87,** 27–49
2. Carlos, J. L., Paetzel, M., Klenotic, P. A., Strynadka, N. C. J., and Dalbey, R. E. (2001) *The Enzymes* **22,** 27–55
3. Date, T., and Wickner, W. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78,** 6106–6110
4. Wolfe, P. B., Wickner, W., and Goodman, J. M. (1983) *J. Biol. Chem.* **258,** 12073–12080
5. Dalbey, R. E., and Wickner, W. (1985) *J. Biol. Chem.* **260,** 15925–15931
6. Wolfe, P. B., Silver, P., and Wickner, W. (1982) *J. Biol. Chem.* **257,** 7898–7902
7. Tschantz, W. R., and Dalbey, R. E. (1994) *Methods Enzymol.* **244,** 285–301
8. Tschantz, W. R., Sung, M., Delgado-Partin, V. M., and Dalbey, R. E. (1993) *J. Biol. Chem.* **268,** 27349–27354
9. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (1998) *Nature* **396,** 186–190
10. Sung, M., and Dalbey, R. E. (1992) *J. Biol. Chem.* **267,** 13154–13159
11. Paetzel, M., Strynadka, N. C., Tschantz, W. R., Casareno, R., Bullinger, P. R., and Dalbey, R. E. (1997) *J. Biol. Chem.* **272,** 9994–10003
12. Kuo, D. W., Chan, H. K., Wilson, C. J., Griffin, P. R., Williams, H., and Knight, W. B. (1993) *Arch Biochem. Biophys.* **303,** 274–280
13. Tschantz, W. R., Paetzel, M., Cao, G., Suciu, D., Inouye, M., and Dalbey, R. E. (1995) *Biochemistry* **34,** 3935–3941
14. Paetzel, M., Chernaia, M., Strynadka, N., Tschantz, W., Cao, G., Dalbey, R. E., and James, M. N. (1995) *Proteins* **23,** 122–125
15. Bullock, T. L., Breddam, K., and Remington, S. J. (1996) *J. Mol. Biol.* **255,** 714–725
16. James, M. N. (1994) in *Proteolysis and Protein Turnover* (Bond, J. S., and Barrett, A. J., eds), pp. 1–8, Portland, Brookfield, VT
17. Menard, R., and Storer, A. C. (1992) *Biol. Chem. Hoppe Seyler* **373,** 393–400
18. Carlos, J. L., Klenotic, P. A., Paetzel, M., Strynadka, N. C., and Dalbey, R. E. (2000) *Biochemistry* **39,** 7276–7283
19. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
20. Otwinowski, Z. (1993) in *Denzo* (Sawyer, L., Isaacs, N., and Baily, S., eds), pp. 56–62, SERC Daresbury Laboratory
21. Collaborative Computational Project, No. 4 (1994) *Acta Crystallogr Sect. D* **50,** 760–763
22. McRee, D. E. (1999) *J. Struct. Biol.* **125,** 156–165
23. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr Sect. D Biol.* **54,** 905–921
24. Hutchinson, E. G., and Thornton, J. M. (1996) *Protein Sci.* **5,** 212–220
25. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kieldgaard, M. (1991) *Acta Crystallogr Sect. A* **47,** 110–119
26. Liang, J., Edelsbrunner, H., and Woodward, C. (1998) *Protein Sci.* **7,**

1884–1897
27. Kraulis, P. G. (1991) *J. Appl. Crystallogr.* **24,** 946–950
28. Meritt, E. A., and Bacon, D. J. (1997) *Methods Enzymol.* **277,** 505–524
29. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
30. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28,** 235–242
31. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162
32. von Heijne, G. (1985) *J. Mol. Biol.* **184,** 99–105
33. Nicolas, A., Egmond, M., Verrips, C. T., de Vlieg, J., Longhi, S., Cambillau, C., and Martinez, C. (1996) *Biochemistry* **35,** 398–410
34. Strynadka, N. C., Martin, R., Jensen, S. E., Gold, M., and Jones, J. B. (1996) *Nat. Struct. Biol.* **3,** 688–695
35. Maveyraud, L., Massova, I., Birck, C., Miyashita, K., Samama, J. P., and Maobashery, S. (1996) *J. Am. Chem. Soc.* **118,** 7435–7440
36. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) *Nat. Struct. Biol.* **4,** 456–462
37. Wilmouth, R. C., Westwood, N. J., Anderson, K., Brownlee, W., Claridge, T. D., Clifton, I. J., Pritchard, G. J., Aplin, R. T., and Schofield, C. J. (1998) *Biochemistry* **37,** 17506–17513
38. Bürgi, H. B., Dunitz, J. D., and Shefter, E. (1973) *J. Am. Chem. Soc.* **95,** 5065–5067
39. Paetzel, M., and Strynadka, N. C. (1999) *Protein Sci.* **8,** 2533–2536
40. Luo, Y., Pfuetzner, R. A., Mosimann, S., Paetzel, M., Frey, E. A., Cherney, M., Kim, B., Little, J. W., and Strynadka, N. C. J. (2001) *Cell* **106,** 1–10
41. Bell, C. E., Frescura, P., Hochschild, A., and Lewis, M. (2000) *Cell* **101,** 801–811
42. Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R., and Hendrickson, W. A. (1996) *Nature* **380,** 727–730
43. Liao, D. I., Qian, J., Chisholm, D. A., Jordan, D. B., and Diner, B. A. (2000) *Nat. Struct. Biol.* **7,** 749–753
44. Strynadka, N. C., Adachi, H., Jensen, S. E., Johns, K., Sielecki, A., Betzel, C., Sutoh, K., and James, M. N. (1992) *Nature* **359,** 700–705
45. Klenotic, P. A., Carlos, J. L., Samuelson, J. C., Schuenemann, T. A., Tschantz, W. R., Paetzel, M., Strynadka, N. C., and Dalbey, R. E. (2000) *J. Biol. Chem.* **275,** 6490–6498
46. Chatterjee, S., Suciu, D., Dalbey, R. E., Kahn, P. C., and Inouye, M. (1995) *J. Mol. Biol.* **245,** 311–314
47. Kim, Y. T., Muramatsu, T., and Takahashi, K. (1995) *Eur. J. Biochem.* **234,** 358–362
48. Kim, Y. T., Muramatsu, T., and Takahashi, K. (1995) *J. Biochem.* (*Tokyo*) **117,** 535–544
49. Black, M. T., and Bruton, G. (1998) *Curr. Pharm. Des.* **4,** 133–154
50. Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R., and Hendrickson, W. A. (1996) *Structure* **4,** 1401–1412
51. Ferentz, A. E., Walker, G. C., and Wagner, G. (2001) *EMBO J.* **20,** 4287–4298
52. Heller, H., Scaefer, M., and Schulten, K. (1993) *J. Phys. Chem.* **97,** 8343–8360