# The 0.78 Å Structure of a Serine Protease: *Bacillus lentus* Subtilisin[†,‡]

Peter Kuhn,[§] Mark Knapp,[∥] S. Michael Soltis,[§] Grant Ganshaw,[⊥] Michael Thoene,[⊥,#] and Richard Bott*,[⊥]

*Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford California 94309, Lawrence Livermore National Laboratory, Livermore, California, and Genencor International, Palo Alto, California 94304*

*Received June 12, 1998; Revised Manuscript Received August 13, 1998*

ABSTRACT: Ultrahigh-resolution X-ray diffraction data from cryo-cooled, *B. lentus* subtilisin crystals has been collected to a resolution of 0.78 Å. The refined model coordinates have a rms deviation of 0.22 Å relative to the same structure determined at room temperature and 2.0 Å resolution. Several regions of main-chain and side-chain disorder have been identified for 21 out of 269 residues in one polypeptide chain. Hydrogen atoms appear as significant peaks in the $F_o - F_c$ difference electron density map, and carbon, nitrogen, and oxygen atoms can be differentiated. The estimated standard deviation (ESD) for all main-chain non-hydrogen bond lengths is 0.009 Å and 0.5° for bond angles based on an unrestrained full-matrix least-squares refinement. Hydrogen bonds are resolved in the serine protease catalytic triad (Ser-His-Asp). Electron density is observed for an unusual, short hydrogen bond between aspartic acid and histidine in the catalytic triad. The hydrogen atom, identified by NMR in numerous serine proteases, appears to be shared by the heteroatoms in the bond. This represents the first reported correlation between detailed chemical features identified by NMR and those in a cryo-cooled crystallographic structure determination at ultrahigh resolution. The short hydrogen bond, designated "catalytic hydrogen bond", occurs as part of an elaborate hydrogen bond network, involving Asp of the catalytic triad. While unusual, these features appear to have conserved analogues in other serine protease families although specific details differ from family to family.

Serine proteases are found in all organisms, functioning in digestion, posttranslational processing of secreted proteins, neurotransmitters and hormones, blood coagulation, and complement fixation (*1*). The serine proteases occur in three distinct structural families, represented by chymotrypsin (*2*), subtilisin BPN′ (*3*), and wheat germ serine carboxypeptidase II (*4*). All three serine protease families share a common reactive center composed of an oxyanion hole (*5*) and a catalytic triad (*6*), although the sequence of serine, histidine, and aspartic acid differ. Apart from this similarity, they share no other structural homology. Chymotrypsin and subtilisin class enzymes bind substrate in an analogous manner, forming an antiparallel $\beta$-structure between the polypeptide backbone of the substrate and a structurally conserved segment of the enzyme (*7*). Because of this, all three classes of serine protease are presumed to share a common mechanism of action. In this mechanism, Ser functions as the primary nucleophile and His plays a dual role as proton acceptor and donor at different steps in the reaction. The role of Asp is thought to bring the His residue in the correct orientation to facilitate nucleophilic attack by Ser.

The functional importance of the catalytic triad and oxyanion hole in catalysis has been confirmed by site-directed mutagenesis; removing an element of the oxyanion hole (*8*) or Asp to Ala replacement (*9*) reduces $k_{cat}$ by $10^4$, while replacing His or Ser (*9*) with Ala reduces activity by $10^6$. The protonation states of the catalytic histidine and aspartic acid residues in serine proteases have been studied by NMR (*10*) and neutron diffraction (*11*). NMR studies of both the trypsin and subtilisin classes of enzymes show a strongly downfield shifted proton assigned to histidine in the catalytic triad (*12*−*14*). It has been proposed that the downfield shifted proton is associated with a low-barrier hydrogen bond (LBHB)[1] that forms during the formation of the transition-state intermediate (*12*). However, the exact nature of the downfield-shifted proton, whether it represents a LBHB and its role in enzyme catalysis, has been debated.

*Bacillus lentus* subtilisin (BLS) differs from subtilisin BPN′ (BPN) at 103 of a possible 269 residues, including deletion of six residues in three segments (*15*). BLS and BPN share a conserved tertiary fold, and the residue numbering used here is by homology with BPN. In this numbering, the catalytic triad is formed by Asp 32, His 64, and Ser 221. NMR studies report a similar downfield shift

[†] P.K. and M.K. contributed equally to this work. This work is based upon research conducted at the Stanford Synchrotron Radiation Laboratory (SSRL), which is funded by the Department of Energy, Office of Basic Energy Sciences. The Biotechnology Program is supported by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the Department of Energy, Office of Biological and Environmental Research. The work described was also performed in part under the auspices of the Department of Energy at Lawrence Livermore National Laboratory.

[‡] PDB submission code: 1GCI.

* To whom correspondence should be addressed.

[§] Stanford University.

[∥] Lawrence Livermore National Laboratory.

[⊥] Genencor International.

[#] Present address: 811 Riley Drive, Albany CA 94706.

[1] Abbreviations: LBHB, low barrier hydrogen bond; BLS, *Bacillus lentus* subtilisin; BPN, subtilisin BPN′; SSRL, Stanford Synchrotron Radiation Laboratory; NMR, nuclear magnetic resonance; PDB, Brookhaven Protein Data Bank.

S0006-2960(98)01398-1 CCC: $15.00 © 1998 American Chemical Society
Published on Web 09/11/1998

Table 1: Summary of Crystallization, Data Collection, and Data Processing of the 0.78 Å Resolution Diffraction Data of *B. lentus* Subtilisin

| | high1 | high2 | medium | low | merge |
|---|---|---|---|---|---|
| resolution (Å) | 1.5−0.78 | 1.5−0.78 | 4.0−1.3 | 35.0−2.7 | 35.0−0.78 |
| observations | 1 101 776 | 684 375 | 180 705 | 28 790 | 1 995 646 |
| reflections | 536 641 | 143 985 | 97 091 | 17 964 | 795 681 |
| unique reflections | 210 870 | 140 309 | 36 681 | 6349 | 257 583 |
| completeness (%) | 79.9 (79.9)[a] | 82.2 (85.1)[a] | 99.1 (93.5)[a] | 93.1 (90.2)[a] | 97.3 (92.7)[a] |
| multiplicity | 2.5 | 1.8 | 2.6 | 2.8 | 3.8 |
| $R_{merge}$[b] | 6.4 (25.3)[a] | 6.0 (15.9)[a] | 3.3 (6.4)[a] | 3.1 (2.6)[a] | 4.1 (29)[a] |
| $\langle I\rangle/\langle\sigma(I)\rangle$ | 8.4 (2.9)[a] | 8.8 (4.6)[a] | 16.8 (11.2)[a] | 18.3 (18.4)[a] | 5.6 (2.6)[a] |

[a] Numbers in parentheses correspond to the data in the last resolution shell. [b] $R_{merge} = \sum_\eta\sum_\iota |I(h) - I(h)_I|/\sum\eta\sum_\iota|I(h)_i$ with $I(h)$ = mean intensity.

for the N$\delta$1 proton of His 64 in BLS (*13*). Like BPN, BLS crystallizes in a tightly packed crystal complex with a Matthews coefficient (*16*) of $V_m = 2.2$, corresponding to 45% solvent content. Very intense X-rays from synchrotron sources yield diffraction to a resolution of 0.75 Å when crystals have been flash-cooled and maintained at 100 K.

Atomic resolution (1.2−0.95 Å) macromolecular structures have been reported to have much higher accuracy, relative to earlier studies of the same enzyme, as measured from inverted least-squares refinement matrixes (*17*). Refinement of such structures routinely incorporates the refinement of anisotropic temperature factors which aid in the delineation of loop motion and far-ranging disorder, the recognition of hydrogen atoms, and the identification of protonation states. At ultrahigh resolution (<0.9 Å), it has been reported that charge distributions can be recognized in the electron density map of the polypeptide Crambin at 0.86 Å resolution (*18*).

Ultrahigh resolution provides increased definition to visualize hydrogen atoms. Hydrogen atoms comprise 5−10% of a protein's mass and exceed the total number of heavy atoms (C, N, O, and S) in a typical enzyme. The disposition of hydrogen is critical to further our understanding of structure and function in enzymes. Here, we report the results of the refinement using a complete, high quality crystallographic data set to a resolution of 0.78 Å and present the detailed structure of the BLS active site. We find that the catalytic triad Asp participates in an elaborate hydrogen bond network that includes an unusual short hydrogen bond to His in the catalytic triad. This level of detail should now be possible to achieve for numerous enzymes due to advanced beam line design that exploits short wavelength, high-intensity X-rays.

## MATERIAL AND METHODS

*Crystallization, Data Collection, and Data Processing*. *B. lentus* subtilisin has been purified from commercial samples of this enzyme (Purafect, Genencor International). Crystals were grown using hanging drop vapor diffusion in two stages. (1) Seed crystals were grown at 37 °C from a solution composed of 7.5 $\mu$L of 25 mg/mL protein in buffer, 50 mM sodium acetate (pH 5.9), 10 mM $CaCl_2$ saturated with the inhibitor, phenylmethanesulfonyl fluoride, and 7.5 $\mu$L of the reservoir composed of 30−33% saturated $(NH_4)_2SO_4$ and buffer. (2) Seed crystals were transferred at room temperature into drops containing 2.5 $\mu$L of a 25 mg/mL protein solution saturated with a synthetic substrate, succinyl-alanyl-alanyl-prolyl-phenylalanine-*p*-nitro-anilide, and 7.5 $\mu$L of the reservoir buffer, 22−24% saturated $(NH_4)_2SO_4$. The mother liquor, and crystals to a limited extent, turned yellow, indicating turnover of substrate. Prior to data collection, crystals were submerged for 15−60 s in a cryo-protectant solution containing 25% glycerol and 75% crystallization reservoir buffer. Crystals belong to the orthorhombic space group $P2_12_12_1$ with unit cell dimensions of 52.65 Å × 61.25 Å × 74.75 Å. General methods and strategy for ultrahigh-resolution data collection and processing, recently developed at SSRL, were employed (*19*). All diffraction data were measured at a wavelength of 0.77 Å (16.1 keV) at SSRL beam line 9-1 with a 300 mm diameter MAR-Research imaging plate detector. Complete data were measured to a resolution of 0.78 Å in three resolution passes; high, medium, and low, respectively (Table 1). The mosaicity of the crystal refined to a value of 0.35°. To achieve maximum completeness in the highest resolution shells, the crystal was reoriented in situ during data collection by bending the cryo-loop mounting wire. The oscillation angles, ranges, and crystal reorientation were optimized with MOSFLM (*20*). All data were indexed and integrated with MOSFLM version 5.40 and scaled with the CCP4 program SCALA version 2.2.3 (*21*). The diffraction quality was continuously monitored in real-time, however, no significant decay was detected, as judged by $I/\sigma_I$ over the course of data collection. At the end of data collection, the crystal was exposed to the X-ray beam for a long exposure of 5 min and diffraction was recorded to a resolution of 0.75 Å. Nearly 2 million observations were measured and the merged data of 257 583 unique reflections were 97% complete, yielding an overall $R_{merge}$ of 4.1% (29% in the last resolution shell).

*Refinement.* For a starting model, phases were calculated from the 2.0 Å room-temperature structure (PDB filename 1JEA). Initial least-squares refinement was carried out in PROLSQ (*22*) in the 10−0.85 Å resolution range. All temperature factors were isotropic and the refinement converged with an *R*-factor of 15.7% (Table 2). After rebuilding the model using FRODO and XtalView (*23*), standard conjugate gradient refinement (CGLS) was carried out using SHELX-97 (*24*), which converged at $R = 15.4\%$. At this stage, 21 amino acids were modeled as disordered residues to interpret features of positive density greater than $3\sigma$ in an $F_o - F_c$ difference electron density map, occurring within 1 Å of existing model coordinates. In eight cases, the disorder involved both main-chain and side-chain atoms. The model (model A) was then refined using anisotropic temperature factors in SHELX-97 which converged at $R = 11.5\%$. Occupancies of the solvent molecules were refined using PROSLQ and remained constant in subsequent SHELXL refinement. For further analysis of the quality of the model, an unconstrained refinement (except for the disordered

Table 2: Structure Refinement and Statistics for *B. lentus* Subtilisin at 0.78 Å

| | isotropic temperature factors (model I)[a] | anisotropic temperature factors (model A)[a] | anisotropic temperature factors and hydrogens[b] (model H)[a] |
|---|---|---|---|
| no. of parameters | 8837 | 19 852 | 20 361 |
| no. of restraints | 8039 | 24 185 | 24 234 |
| no. of riding hydrogens | 0 | 0 | 1920 |
| no. of solvent molecules | 269 | 383 | 383 |
| no. of hetero atoms | 1887 | 1887 | 1887 |
| $R_{work\ (\%)}$ | 15.4 | 11.5 | 9.9 |
| $R_{free\ (\%)}$ | 16.0 | 12.5 | 10.3 |
| $R$ ($F > 4\sigma$) (%) | 14.6 | 10.9 | 9.4 |
| $B_{ave}$ (protein)[c] | 7.94 | 7.84 | 9.30 |
| $B_{ave}$ (solvent)[c] | 16.54 | 20.85 | 21.4 |
| rms bond length (Å) | 0.012 | 0.011 | 0.012 |
| rms bond angles (deg) | 1.87 | 1.70 | 1.68 |

[a] A total of 244 527 unique reflections were used for refinement. [b] Hydrogen atoms for the catalytic triad, Asp 60, Gly 63, Thr 66, and solvent molecules were not included in any model. [c] Including hydrogens for model H.

residues) was carried out in SHELX-97 (10:1 ratio of observations to parameters). The rms difference between the constrained model (CGLS) and the full-matrix, unconstrained model (L.S.1, Block 1, without SIMU, DELU) is 0.064 Å with no significant change in the *R*-factor or the electron density map. In the final refinement, riding hydrogen atoms were included in the model (model H), except for those specified in Table 2. The final *R*-factor was 9.9% for all data.

## RESULTS

*High-Resolution Features.* By exploiting shorter wavelength X-rays (0.77 Å), we have collected a 97% complete data set to 0.78 Å resolution. At this resolution, there is a dramatic increase in the quality and detail of the electron density map relative to the same structure determined at room temperature and 2.0 Å resolution. Most notably, individual heavy atoms can be differentiated and electron density for hydrogen atoms can be discerned. Introduction of disordered segments, followed by refinement of anisotropic temperature factors, improved detail in the electron density maps. Density was observed for the hydrogen atoms on Cα, C$\beta$, and for carbon atoms in aromatic and heterocyclic side chains. Figure 1A displays the $2F_o - F_c$ ($F_o$ and $F_c$ are the observed and calculated structure factors, respectively) electron density map contoured at 1 and 4$\sigma$ for residues, Glu 54 and Lys 94, which form the only conserved salt bridge in subtilisin. Within a local environment, electron density for C, N, and O atoms can be distinguished based on the volume of the density contoured at 4$\sigma$. The electron density for nitrogen atom (N$\zeta$) is larger than the density for carbon atoms in Lys 94 and smaller than the density for the carboxyl oxygen atoms (O$\delta$1 and O$\delta$2) and carbonyl oxygen of Glu 54. The carboxylate group in this salt bridge would be expected to have a charge localized on O$\epsilon$2 of Glu 54. O$\epsilon$2 should be singly bonded to the carbon atom (C$\gamma$) and hydrogen bonded to the nitrogen atom (N$\zeta$) of Lys 94. The electron density for the O$\epsilon$2 is discrete at the 4$\sigma$ level, while the density for the O$\epsilon$1−C$\delta$ bond is continuous, having comparable density with carbonyl (C═O) atoms of Glu 54 and Lys 94.

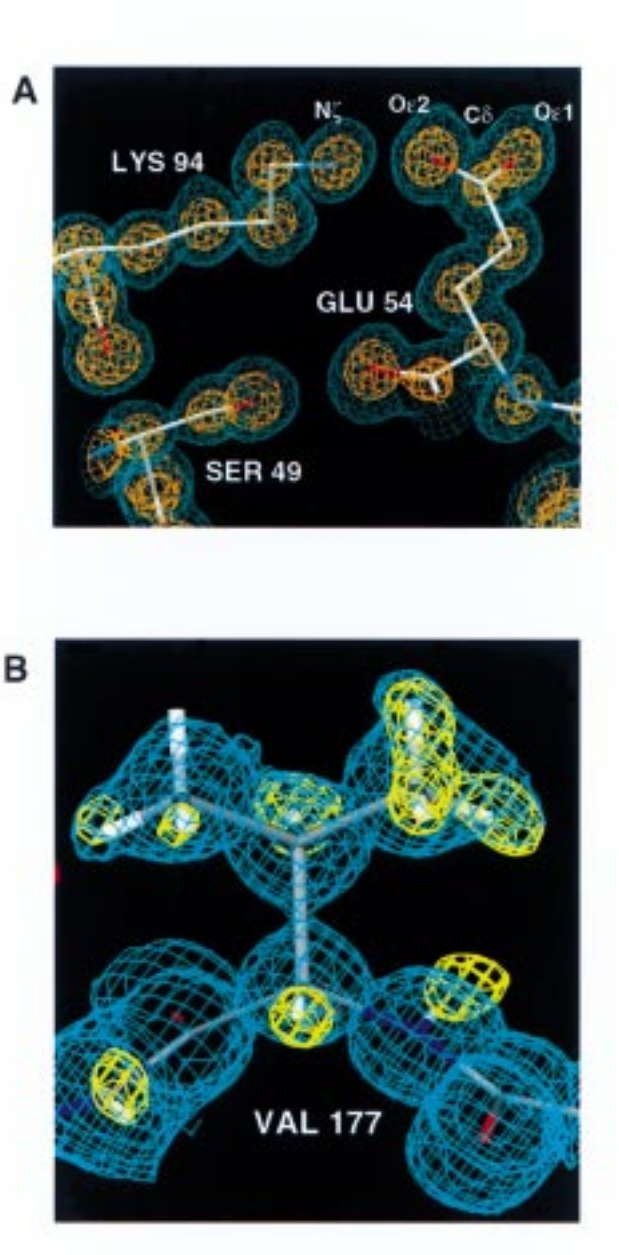


FIGURE 1: (A) Salt bridge between Glu 54 and Lys 94. Electron density ($2F_o - F_c$ map, model A phases) is contoured at the 1$\sigma$ (aqua) and 4$\sigma$ (gold). C, N, and O atoms can be differentiated. At 1$\sigma$, indications for hydrogen atoms are visible for Lys 94 C$\beta$ and C$\delta$. (B) Hydrogen atoms on Val 177. $F_o - F_c$ difference map (model A phases) before any hydrogen atoms were included in the model contoured at 2.5$\sigma$ (yellow) and $2F_o - F_c$ (model H phases) after the inclusion of hydrogen atoms contoured at 1$\sigma$ (blue).

A total of 383 positive peaks above 3$\sigma$ in the $F_o - F_c$ electron density difference is located at idealized hydrogen positions. These peaks, representing approximately 14% of the 2643 hydrogen atoms expected in the asymmetric unit, are clustered predominantly in the internal helices (residues 7−11, 63−72, and 226−232) and in the eight stranded central $\beta$-sheet. Peaks were observed for all C$\beta$ hydrogen atoms in 42 out of 200 residues, excluding Gly and Ala residues. On the basis of these, we included a majority of riding hydrogen atoms in our phasing model (model H, Table 1). This was done with the expectation that the new phasing model would improve the signal-to-noise in the subsequent difference electron density for hydrogen atoms, particularly

in the area of the catalytic triad. In Figure 1B, the $F_o - F_c$ (model A, 2.5$\sigma$) and $2F_o - F_c$ (model H, 1$\sigma$) electron density maps are shown for Val 177. A complete triad of peaks was not observed for any Ala C$\beta$ but were observed at the 3$\sigma$ level for other residues such as the methyl carbon atom (C$\gamma$2) in Val 177. Only two peaks are observed at the 2.5$\sigma$ level for the other methyl group, which are confirmed in the $2F_o - F_c$ density map calculated with phases from model H. From this we conclude that peaks at the 3$\sigma$ level correspond to well-ordered hydrogen atoms.

In addition to the 3$\sigma$ peaks, we observed peaks at the 2$\sigma$ level, which accounted for an estimated 65% of all hydrogen positions. This also supported including hydrogen atoms in a subsequent refinement, excluding the catalytic triad, solvent atoms, and the residues interacting with Asp 60, with the expectation that these peaks might become stronger and more reliably interpreted. Inclusion of hydrogen atoms (model H) gave enhanced detail of other hydrogen atoms in regions of interest that had not been modeled or were not detectable, falling below the 3$\sigma$ threshold in previous difference density maps. One such example is the electron density for Asp 60. Figure 2A displays the $F_o - F_c$ electron density map superimposed on the density of the $2F_o - F_c$ map for Asp 60 which forms hydrogen bonds with the oxygen atom (O$\gamma$1) of Thr 66 and the nitrogen atom (N) of Gly 63. The hydrogen atoms involved in hydrogen bonds between Asp 60 and both Thr 66 and Gly 63, are seen at the 3$\sigma$ level. In the comparable map, using model A, only the hydrogen atom of Thr 66 O$\gamma$1 is observed. The density for the carbon−oxygen bonds (C$\gamma$−O$\delta$1 and C$\gamma$−O$\delta$2) of Asp 60 is equally distributed as would be expected for a charged Asp at pH 5.9. Similar electron density distributions are seen for other Asp, Glu, and Arg side chains.

A complete water molecule ($H_2O$) is observed in the internal solvent channel (Figure 2B). Peaks in the $F_o - F_c$ electron density difference map near oxygen atom (O1024) are associated with a hydrogen bond network extending to the oxygen atom (O$\gamma$1) of Thr 71. These water molecules are well ordered as evidenced by relatively low temperature factors and are highly conserved in subtilisins (*3*). These two peaks of the $H_2O$ molecule (O1024) are planar and fit well into the expected pattern of alternating proton-donor and acceptor atoms, in a zigzag tetrahedral relationship. In contrast to the special hydrogen bond described below, all of these hydrogen atoms have the expected bond length.

*Catalytic Site.* At pH 5.9, BLS is nearly inactive with only 10% of its maximum activity. Subtilisin is readily capable of degrading itself and is routinely inhibited prior to crystallization by either an inhibitor or enzyme product (T. Graycar & D. Estell, personal communication). In this study, a synthetic substrate, succinyl-alanyl-alanyl-prolyl-phenylalanine-*p*-nitro-anilide, was allowed to diffuse into a BLS crystal. It was hydrolyzed, leaving the product succinyl-alanyl-alanyl-prolyl-phenylalanine (sAAPF). This product was expected to remain bound. There are numerous features of disconnected electron density within the substrate-binding cleft, formed by six loops on the surface of the molecule (residues 60−64, 99−104, 126−129, 152−166, 185−189, and 209−219). These features are observed at positions that can be occupied by the tetra-peptide product or by solvent. Modeled tetra-peptide did not refine well and the features of discontinuous density did not coalesce into

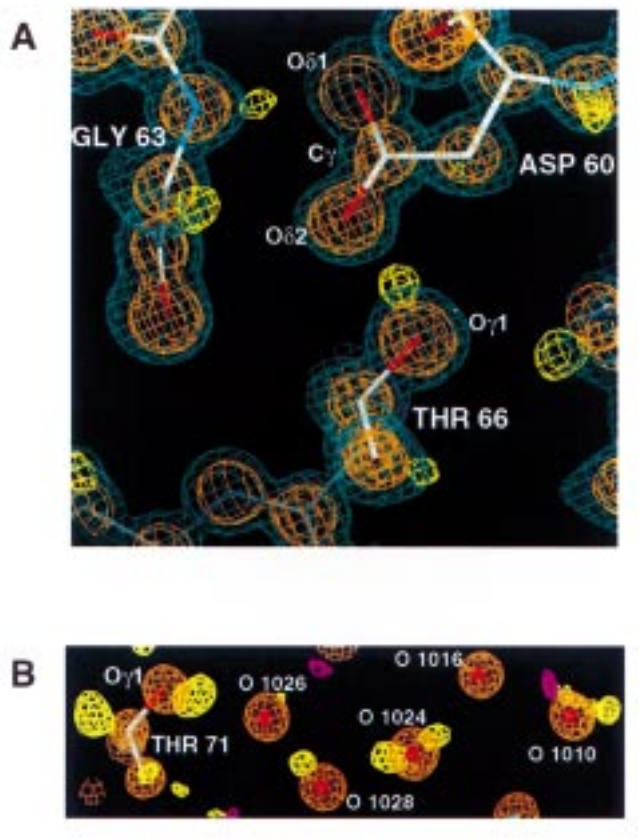


FIGURE 2: (A) Electron density for Asp 60. $F_o - F_c$ difference electron density map, contoured at +2.5$\sigma$ (yellow) and −2.5$\sigma$ (purple), is superimposed on a $2F_o - F_c$ electron density map contoured at 1$\sigma$ (aqua) and 4$\sigma$ (gold), based on model H (Table 2). No hydrogen atoms were included for residues Asp 60, Gly 63, Thr 66, and solvent. At the 4$\sigma$ contour, the electron density between C$\gamma$ and both O$\delta$1 and O$\delta$2 are equivalent as expected. (B) $H_2O$ hydrogen bonding in an internal water channel. Electron density; $2F_o - F_c$ (model H phases) contoured at 4$\sigma$ (gold) level is superposed on $F_o - F_c$ (model H phases) difference electron density contoured at +2.5$\sigma$ (yellow) and −2.5$\sigma$ (purple) level. Peaks for both hydrogen atoms on solvent O1024 are seen along with hydrogen atoms of neighboring solvent and Thr 71 O$\gamma$1. A zigzag pattern of alternating proton donors and acceptors can be seen extending from Thr 71 O$\gamma$1 in the interior, through a chain of four solvent molecules ending at the surface of the enzyme (Figure 1).

connected density, recognizable as a tetra-peptide when the product was included in the phasing model. The present model, therefore, represents predominantly uncomplexed, enzyme with solvent in the substrate-binding cleft.

The catalytic triad is in the expected conformation as shown in the stereo diagram of Figure 3A and the thermal ellipsoid representation of the same view in Figure 3B. Ser 221, His 64, and Asp 32 are reasonably ordered, and the oxygen atom (O$\gamma$) of Ser 221 is within hydrogen bond distance (3.1 Å) of the nitrogen atom (N$\epsilon$2) of His 64. The thermal motions of Ser 221 O$\gamma$ and the imidizole ring of His 64 are larger than those of Asp 32. Electron density for the hydrogen atom of Ser 221 O$\gamma$ is not observed, nor is the density for the His 64 N$\epsilon$2 hydrogen atom.

At pH 5.9, histidine should be protonated; however, we observe, at most, one difference electron density peak at N$\epsilon$2 or N$\delta$1 for any of the seven histidines in BLS. We are not able to distinguish single and double bonds in histidine side-chain atoms, presumably because the bonds are delocalized

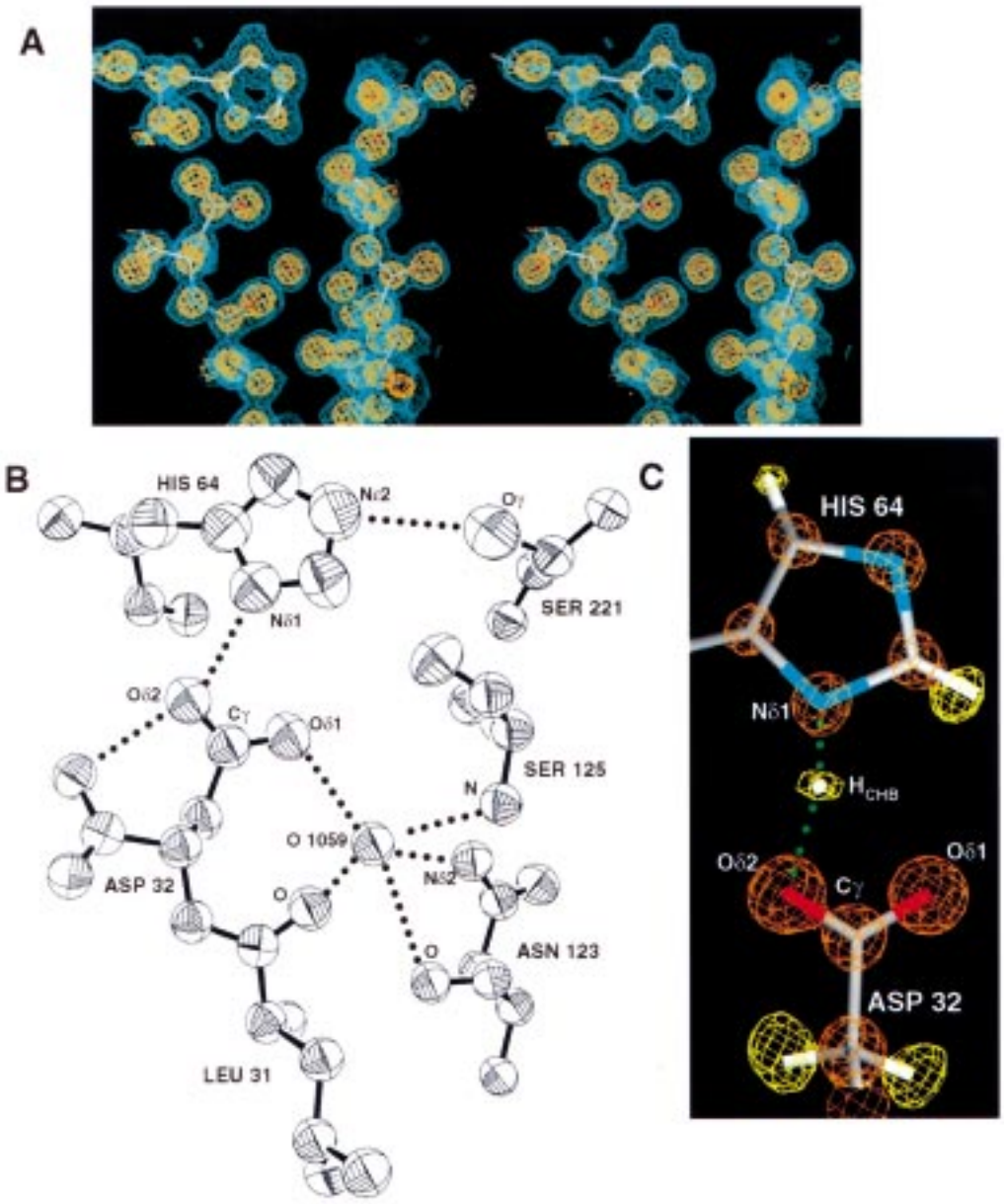


FIGURE 3: The catalytic triad. (A) Stereoview displaying Model H superimposed on the $2F_o - F_c$ (model H phases) at $1\sigma$ (aqua) and $4\sigma$ (gold). The densities for C and N in His 64 are weaker than in Asp 32. The Asp 32 C$\gamma$−O$\delta$2 bond at $4\sigma$ is continuous, while the density for the C$\gamma$ and O$\delta$1 are resolved. (B) Schematic of the catalytic residues and hydrogen bonded neighbors with thermal ellipsoid representation countered at 50% probability (*29*). Catalytic triad residues Ser 221 and His 64 show larger thermal motion than the Asp 32. Solvent O1059 appears to be a relatively rigid and integral part of the enzyme structure. (C) Catalytic hydrogen bond (CHB). A $F_o - F_c$ (model H phases) difference map contoured at $+2.5\sigma$ (yellow) and $-2.5\sigma$ (red) and a $2F_o - F_c$ (model H phases) electron density map contoured at $4\sigma$ (gold). The position of the short hydrogen atom (labeled $H_{CHB}$) in the CHB is positioned in the positive electron density present between His 64 N$\delta$1 and Asp 32 O$\delta$2.

due to resonance. However, the orientation of the His 64 side chain is unambiguous, based on the relative size of the $4\sigma$ contours for side chain atoms. N$\delta$1 and N$\epsilon$1 are easily distinguished from C$\delta$2 and C$\epsilon$1. Electron density, contoured at $4\sigma$, links the oxygen (O$\delta$2) and carbon (C$\gamma$) atoms in Asp 32, and it is comparable to other doubly bonded carbonyl atoms, while the density for the oxygen atom (O$\delta$1) is discrete (Figure 3C). Despite hydrogen bonds formed between O$\delta$2 and His 62 and Thr 33, the O$\delta$2−C$\gamma$ bond has more double bond character than the O$\delta$1−C$\gamma$ bond. However, a significant difference in bond length is not detected, precluding a definitive statement as to the charge on O$\delta$1.

An internal water molecule (O1059) is located in the active site. The analysis of the electron density in the vicinity of this solvent molecule allows the characterization of the hydrogen bond network. Two proton donors (Ser 125 and Asn 123 N$\delta$2), two proton acceptors (Ile 31, Asn 123), and Asp 32 are present, providing five potential ligands to O1059 (Figure 3B).

The distance between the nitrogen atom (N$\delta$1) of His 64 and the oxygen atom (O$\delta$2) of Asp 32 is 2.62 Å. Moreover, a $3\sigma$ peak in the $F_o - F_c$ electron density map is found between these atoms (contoured at $\pm 2.5\sigma$) in Figure 3C. The peak is situated 1.2 Å from N$\delta$1 of His 64 and 1.5 Å from O$\delta$2 of Asp 32. The distance of 1.2 Å is longer than those

found for other N−H and C−H bonds in the structure which cluster at the expected distance for hydrogen atoms. It is also 0.3 Å closer to Asp O$\delta$2 than would be typical. The hydrogen atom participating in this short hydrogen bond within the catalytic triad is in the plane of His 64 and Asp 32 side chains and is slightly skewed toward O$\delta$1 of Asp 32. The N$\delta$1−$H_{CHB}$−O$\delta$2 bond angle is 155°.

## DISCUSSION

The delineation of carboxylate isomers and the visualization and characterization of hydrogen bonds are possible only at ultrahigh resolution and provide a new level of detail in enzyme structures. At 100 K and 0.78 Å resolution, we are able to observe the hydrogen atom for the short hydrogen bond identified by NMR in several classes of serine proteases. The bonding distance is 2.6 Å between Asp 32 O$\delta$2 and His 64 N$\epsilon$2, and the hydrogen atom is extended 1.2 Å from His and 1.5 Å from Asp 32. As such, it is shared to some extent between His and Asp. Interestingly, the peak is not in line between the N$\delta$1 and O$\delta$2 atoms but, rather, skewed slightly toward O$\delta$1 of Asp 32. This bond is not shielded from solvent; instead Asp O$\delta$1 forms a hydrogen bond with solvent O1059. Moreover, the p$K_a$s of His N$\delta$1 and Asp 32 O$\delta$2 would appear to be unmatched. We designate this hydrogen bond within the catalytic triad as a catalytic hydrogen bond (CHB) rather than LBHB because, while it is short and partially shared, it is not isolated and does not form between equivalent atoms. In addition, the LBHB between Asp and His was originally postulated to form only when His was protonated, which may or may not be the case under our experimental conditions.

The short bond distance between Asp 32 and His 64, is the same in the 2.0 Å room-temperature structure and therefore not an artifact of cryo-cooling. Further, there are many structures of serine protease from different families that contain an analogous short bonding distance between the Asp and His of the catalytic triad. For example, 11 instances are documented for trypsin structures both with and without inhibitors bound (*25*), as well as in the structure of the homodimeric serine carboxypeptidase II from wheat (CPDW II) (*4*) and in BPN (*3*). While these distances are short, suggesting that a similar CHB is formed, the interactions between the catalytic Asp and neighboring residues differ. For example, in trypsin, Asp 102 O$\delta$2 makes a second hydrogen bond to Ser 214, and Asp 102 O$\delta$1 is within hydrogen bond distance of Ala 56. In CPDW II, Asp 338 O$\delta$2 is positioned to make two additional hydrogen bonds to proton donors, Asn 176 and Val 342.

Trypsin, in complex with the transition state inhibitor monoisopropyl-phosphoryl-trypsin, has been studied using neutron diffraction at 2.2 Å resolution to determine the hydrogen-bonding distances and protonation states within the catalytic triad at pH 6.2 (*11*). A deuterium atom was observed between Asp 102 and His 57; however, in contrast to NMR experiments and our findings, the position of the deuterium was interpreted to be associated solely with His and not with Asp. This might be explained by the fact that, in short hydrogen bonds, deuterium can have a zero-point energy below the energy barrier between the two hydrogen-bonding minima and thus can be confined to one position (*26*). Hydrogen on the other hand, can have a zero-point energy above the barrier between the two bonding minima allowing the hydrogen atom to be shared (*26*).

High-quality structures of related enzymes have been reported which do not exhibit a short hydrogen bond between Asp and His in the catalytic triad. For example, in the atomic resolution structure of cutinase from *Fusarium solani pisi* (*27*), an $\alpha/\beta$ hydrolase, there is no short hydrogen bond, although the carboxylate forms a ligand to an internal solvent molecule akin to the solvent O1059 in BLS. A compilation, including several subtilisins, has been reported where several additional examples of normal length hydrogen bond distances were found (*28*).

Frey et al. (*12*) postulated that the downfield-shifted proton at the His of the catalytic triad represents a LBHB and propose a mechanism whereby the LBHB serves to stabilize the transition-state intermediate. The LBHB would occur during the formation of the acyl enzyme intermediate, after His abstracts the proton from Ser 221. However, the downfield-shifted proton associated with a LBHB can be detected under conditions where the p$K_a$s of the participating atoms are unequal and when the bond in question is unshielded from solvent and as such is not a true LBHB (*12*). The presence of a preformed, short hydrogen bond in BLS with some of the characteristics of a LBHB seems to corroborate this conclusion. The presence of a preformed CHB in the free, inactive state, does not preclude the possibility that a stronger hydrogen bond may form during peptide bond hydrolysis.

In preceding papers, it was reasonably assumed that the negative charge of Asp would be associated with the short hydrogen bond we designate CHB. However, our electron density map indicates that Asp 32 O$\delta$2 is only partially charged or neutral. Moreover, O$\delta$2 forms a second hydrogen bond of normal length to Thr 33. A similar situation occurs in trypsin; O$\delta$2, which participates in the homologous short hydrogen bond to His 57, forms a second hydrogen bond to Ser 214.

The catalytic Ser 221 O$\gamma$ is 3.1 Å from His 64 N$\epsilon$2 in BLS and appears to be relatively mobile based on the anisotropic displacement parameters. We do not see any peak associated with the O$\gamma$ hydrogen atom, suggesting that this hydrogen is not well ordered. Close examination of the catalytic triad reveals that rotation of the $\chi_1$ side-chain torsion angle of Ser 221 by 120° would position O$\gamma$ in a much more favorable position to have its hydrogen atom abstracted by His 64. However, in such a conformation, O$\gamma$ would be within the van der Waal radii of two additional atoms, C$\epsilon$1 of His 64 and C$\beta$ of Ser 125. There is no prominent density for a hydrogen atom bonded to His 64 N$\epsilon$2 even though under these conditions it is presumed that the lower activity is the consequence of His 64 being protonated.

The availability of high-resolution data has provided a much more detailed picture of the serine protease catalytic triad. While general features are conserved, the detailed interaction for the catalytic aspartic acid residue can vary in different classes of serine proteases. We are able to confirm the existence of a special hydrogen bond that had been previously identified by NMR. This hydrogen bond is highly unusual in that the hydrogen atom is shared between heteroatoms and, in addition, a second hydrogen bond of normal length is present.

The details of the generally accepted serine protease mechanism of action, proton abstraction from Ser to His, and in a subsequent step, nucleophilic attack of Ser on substrate are complicated. While it is clear that the ultrahigh-resolution structure has provided additional insight into the nature of the Asp−His interaction, additional structures of intermediate states and mutants are required. These will provide the structural underpinning necessary for dissection of the mechanism of action in subtilisin, a workhorse enzyme in nature.

## ACKNOWLEDGMENT

M.K. and R.B. wish to acknowledge Dr. Edwin Westbrook who allowed us access to the SBC beamline at the Advanced Photon Source ANL that provided us with the first glimpse of the diffracting power of the BLS crystals. We wish to thank Douglas Crabb, Keith Hodgson, Peg Horn, James Kellis, Christian Paech, Andrew Shaw, and Walter Weyler for their critical reading of the manuscript.

## REFERENCES

1. Neurath, H. (1986) *J. Cell. Biochem. 32*, 35−49.
2. Birktoft, J. J., and Blow, D. M. (1972) *J. Mol. Biol. 68*, 187−240.
3. Bott R., Ultsch, M., Kossiakoff, A., Graycar, T., Katz, B., and Power, S. (1988) *J. Biol. Chem. 263*, 7895−7906.
4. Liao D. I., Breddam, K., Sweet, R. M., Bullock, T., and Remington, S. J. (1992) *Biochemistry 31*, 9796−9812.
5. Obertus, J. D., Kraut, J., Alden, R. A., and Birktoft, J. J. (1972) *Biochemistry 11*, 4293−4303.
6. Blow, D. M., Birktoft, J. J., and Hartley, B. S. (1969) *Nature 221*, 337−340.
7. Robertus, J. D., Alden, R. A., Birktoft, J. J., Kraut, J., Powers, J. C., and Wilcox, P. E. (1972) *Biochemistry 11*, 2439−2449.
8. Wells, J. A., Cunningham, B. C., Graycar, T. P., and Estell, D. A. (1986) *Philos. Trans. R. Soc. London A 317*, 415−423.
9. Carter, P., and Wells, J. A. (1988) *Nature 332*, 564−568.
10. Bachovin, W. W., and Roberts, J. D. (1978) *J. Am. Chem. Soc. 100*, 8041−8047.
11. Kossiakoff A. A., and Spencer, S. A. (1981) *Biochemistry 20*, 6462−6474.
12. Frey P. A, Whitt S. A., and Tobin, J. B. (1994) *Science 264*, 1927−1930.
13. Halkides C. J., Wu, Y. Q., and Murray, C. J. (1996) *Biochemistry 35*, 15941−15948.
14. Ash, E. L., Sudmeier, J. L., De Fabo, E. C., and Bachovchin, W. W. (1997) *Science 278*, 1128−1132.
15. Graycar T. P., Bott R. R., and Wilson, L. J. (1995) WOP Application 95/10615 to Genencor.
16. Matthews B. W. (1968) *J. Mol. Biol. 33*, 491−497.
17. Dauter, Z., Lamzin, V. S., and Wilson, K. S. (1997) *Curr. Opin. Struct. Biol. 7*, 681−688.
18. Stec, B., Zhou, R., and Teeter, M. M. (1995) *Acta Crystallogr., Sect. D 51*, 663−681.
19. Genick, U.K., Soltis, S. M., Kuhn, P., Canestrelli, I. L., and Getzoff, E. D. (1998) *Nature 392*, 206−209.
20. A. G. W. Leslie (1992) *CCP4 Newsletter on Protein Crystallography*, no. 26.
21. Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr., Sect. D 50*, 760−767.
22. Henrickson, W. A., Konnert, J., and Srinivasan, R. (1981) in *Biomolecular Structure Conformation Function and Evolution* (Srinivasan, R., Ed.) Vol. 1, pp 43−47, Pergammon Press, Oxford.
23. McRee, D. E. (1993) *Practical Protein Crystallography*, Academic Press, San Diego.
24. Sheldrick, G., and Schneider, T. (1997) *Methods in Enzymology*, Vol. 277, pp 319−343, Academic Press.
25. Marquart, M., Walter, J., Deisenhofer, J., Bode, W., and Huber, R. (1983) *Acta Crystallogr., Sect. B 39*, 480−490.
26. Hibbert, F., and Elmsley, J. (1990) *Adv. Phys. Org. Chem. 26*, 255−379.
27. Longhi, S., Czjzek, M. Lamzin, V. Nicolas, A., and Cambillau, C. (1997) *J. Mol. Biol. 268*, 779−799.
28. Lange, G., Betzel, C., Branner, S., and Wilson, K. S. (1994) *Eur. J. Biochem. 224*, 507−518.
29. Burnett, M. N., and Johnson, C. K. (1996) *ORTEP−III*, Oak-Ridge National Laboratory.

BI9813983