J|A|C|S
ARTICLES
Published on Web 06/27/2006

Subangstrom Crystallography Reveals that Short Ionic Hydrogen Bonds, and Not a His-Asp Low-Barrier Hydrogen Bond, Stabilize the Transition State in Serine Protease Catalysis

Cynthia N. Fuhrmann, Matthew D. Daugherty, and David A. Agard*

Contribution from the Howard Hughes Medical Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, 600 16th Street, Box 2240, San Francisco, California 94143-2240

Received November 13, 2005; E-mail: agard@msg.ucsf.edu

Abstract: To address questions regarding the mechanism of serine protease catalysis, we have solved two X-ray crystal structures of α-lytic protease (αLP) that mimic aspects of the transition states: αLP at pH 5 (0.82 Å resolution) and αLP bound to the peptidyl boronic acid inhibitor, MeOSuc-Ala-Ala-Pro-boroVal (0.90 Å resolution). Based on these structures, there is no evidence of, or requirement for, histidine-flipping during the acylation step of the reaction. Rather, our data suggests that upon protonation of His57, Ser195 undergoes a conformational change that destabilizes the His57-Ser195 hydrogen bond, preventing the back-reaction. In both structures the His57-Asp102 hydrogen bond in the catalytic triad is a normal ionic hydrogen bond, and not a low-barrier hydrogen bond (LBHB) as previously hypothesized. We propose that the enzyme has evolved a network of relatively short hydrogen bonds that collectively stabilize the transition states. In particular, a short ionic hydrogen bond (SIHB) between His57 Nϵ2 and the substrate's leaving group may promote forward progression of the TI_1-to-acylenzyme reaction. We provide experimental evidence that refutes use of either a short donor−acceptor distance or a downfield ^1H chemical shift as sole indicators of a LBHB.

Introduction

In the past decade, the role of low-barrier hydrogen bonds in transition state stabilization has been debated among enzymologists. Particular attention has focused on members of the serine protease family.

The chymotrypsin family of serine proteases is the largest class of proteases in nature and one of the best-studied families of enzymes. This family is characterized by a double-β barrel fold, with the catalytic triad (Asp102, His57, Ser195) located at the juncture of the two domains. Crystallographic and kinetic studies have established the general mechanism by which chymotrypsin-like serine proteases act.[1,2] In the first stage (acylation, Figure 1a), Ser195 attacks the carbonyl of the substrate's scissile bond; this is facilitated by His57, which deprotonates Ser195 and thereby increases the nucleophilicity of Ser195 Oγ. The resulting first tetrahedral intermediate (TI_1) is stabilized through hydrogen-bonding interactions with backbone amides in the oxyanion hole (Gly193 and Ser195), while the positively charged His57 is stabilized by the nearby Asp102. His57 protonates the amine of the scissile bond, promoting formation of the acylenzyme and subsequent release of the C-terminal portion of the cleaved substrate. The acylation reaction thereby occurs in two steps, passing through two transition states (TS and TS′) and a meta-stable tetrahedral intermediate (TI_1). The second stage of the reaction (deacylation) follows a similar mechanism, with a water molecule deprotonated by His57 becoming the active nucleophile, passing through a second meta-stable tetrahedral intermediate (TI_2) prior to product release.

Although the general mechanism for serine protease catalysis is well established, some critical questions remain. Perhaps the most heated debate has focused on what factors are most critical for transition-state stabilization in the acylation reaction. In 1994, Frey et al. proposed that in the transition states and tetrahedral intermediate for acylation, His57 and Asp102 form an especially strong hydrogen bond, called a low-barrier hydrogen bond (LBHB; Figure 1).[8,9] Because a LBHB can store 10−20 kcal/mol of energy,[10] Frey hypothesized that formation of such a hydrogen bond would promote formation of the transition state by stabilizing the Asp-His dyad and increasing the basicity of His57, enhancing catalysis of the first step of the acylation

(1) Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501−4524.
(2) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W.H. Freeman: New York, 1999.
(3) Fuhrmann, C. N.; Kelch, B. A.; Ota, N.; Agard, D. A. *J. Mol. Biol.* **2004**, *338*, 999−1013.
(4) Garcia-Viloca, M.; Gelabert, R.; Gonzalez-Lafont, A.; Moreno, M.; Lluch, J. M. *J. Phys. Chem. A* **1997**, *101*, 8727−8733.
(5) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem.* **1998**, *273*, 25529−25532.
(6) Merritt, E. A.; Bacon, D. J. *Methods Enzymol.* **1997**, *277*, 505−524.
(7) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.
(8) Frey, P. A.; Whitt, S. A.; Tobin, J. B. *Science* **1994**, *264*, 1927−1930.
(9) Frey, P. A. *J. Phys. Org. Chem.* **2004**, *17*, 511−520.
(10) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887−1890.

10.1021/ja057721o CCC: $33.50

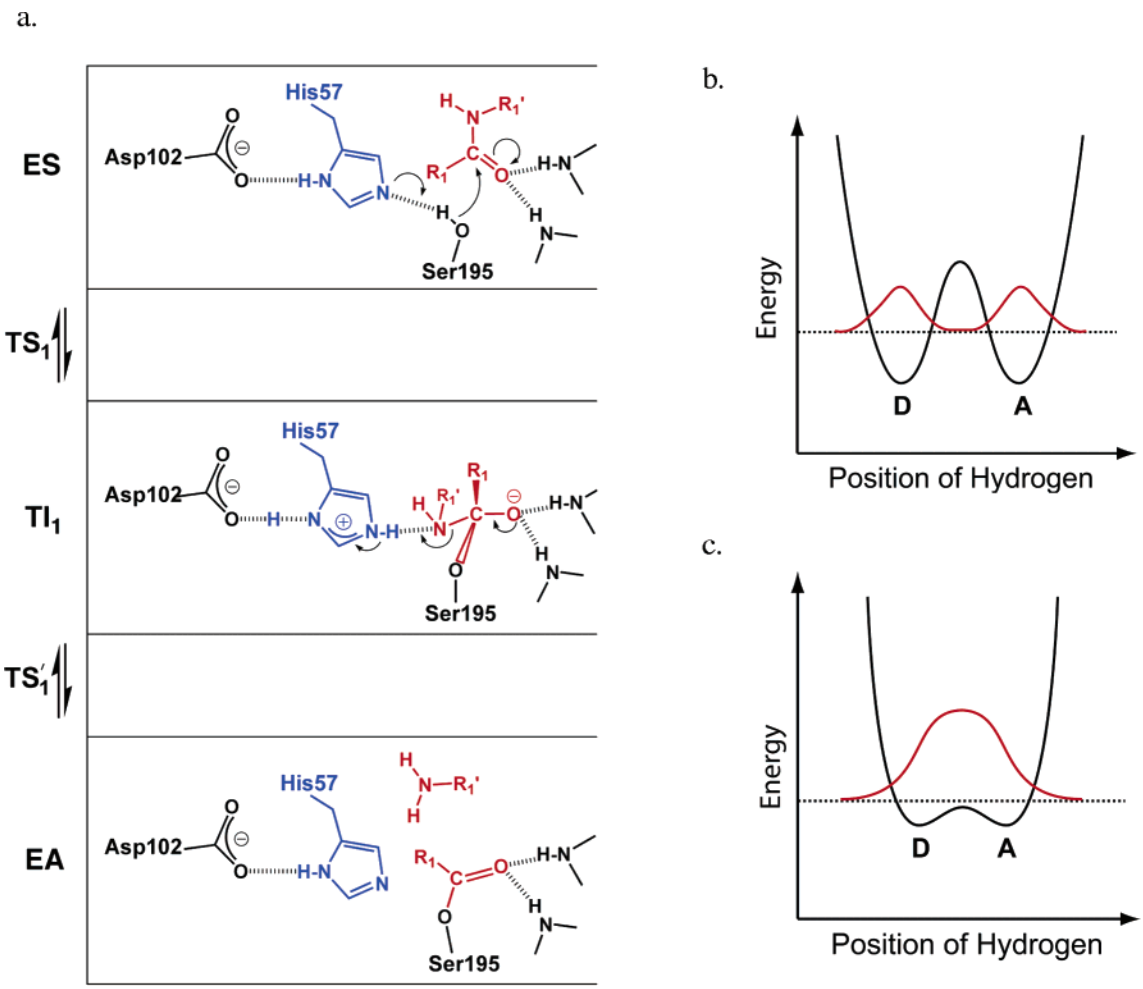

Figure 1. Low-barrier hydrogen bond (LBHB) hypothesis. (a) This simplified schematic illustrates the mechanism for the first stage of proteolytic catalysis, acylation. The general base in the reaction, His57, is highlighted in blue; the substrate is shown in red. The reaction passes through two tetrahedral transition states (TS_1 and TS'_1). The LBHB hypothesis predicts that a LBHB is formed between His57 and Asp102 in the transition states and TI_1. If this were true, the hydrogen atom would be located approximately equidistant between these two residues (illustrated here in TI_1). (b,c) Energetic landscapes (black solid line) for hydrogen transfer between donor (D) and acceptor (A) in a standard, symmetric hydrogen bond (b) and a LBHB (c).[4,5] The dotted line represents the lowest vibrational energy level for the hydrogen atom, and the red line represents its probability density function.

reaction. Others argue against a special role for a LBHB between His57 and Asp102, hypothesizing that a constellation of electrostatic and van der Waals interactions within the active site are primarily responsible for transition-state stabilization.[11−13]

In a standard hydrogen bond, the probability of a hydrogen atom being located on the acceptor atom or the donor atom is determined by the depth of the energy well at each of these positions (see Figure 1b), with an energy barrier restricting hydrogen atom exchange. By contrast, in a LBHB (sometimes called a "short strong hydrogen bond") the energy barrier for hydrogen transfer is less than the first vibrational energy level for the hydrogen.[8,15] Formation of a LBHB requires a donor-to-acceptor distance of less than 2.65 Å for a nitrogen−oxygen pair such as His57 and Asp102, and ΔpK_a at or near zero.[8,14,15] Therefore, in contrast to a standard hydrogen bond where the hydrogen is located on the donor atom, in a LBHB, the hydrogen atom has the highest probability of being located equidistant[4,16] or nearly equidistant[9] between its two parent heavy atoms (illustrated in Figure 1c). This lengthening of the donor−hydrogen bond results in a downfield shift of the ^{1}H NMR signal to the range of 16−20 ppm, a common diagnostic used for identifying LBHBs.[8,9] However, the most widely accepted diagnostic remains the physical characteristics of a short distance between heavy atoms and significant displacement of the hydrogen atom toward the acceptor.[8,15] Because ultrahigh resolution X-ray crystallography can both precisely determine the position of heavy atoms to better than ±0.01 Å and visualize hydrogen atoms in difference electron density maps, this is an excellent technique to use to identify LBHBs.[17−19]

Theories such as the LBHB hypothesis are inherently difficult to address experimentally due to the transient nature of transition states. To gain some structural understanding of these states, analogues mimicking certain aspects of the transition states (or

(11) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035−27038.
(12) Kollman, P. A.; Kuhn, B.; Donini, O.; Perakyla, M.; Stanton, R.; Bakowies, D. *Acc. Chem. Res.* **2001**, *34*, 72−79.
(13) Ash, E. L.; Sudmeier, J. L.; De Fabo, E. C.; Bachovchin, W. W. *Science* **1997**, *278*, 1128−1132.
(14) Because of its wide application in the field of enzymology, we have chosen to define the LBHB and its diagnostics as according to Frey,[8] based on the Hibbert and Emsley[15] definition for strong hydrogen bonds.
(15) Hibbert, F.; Emsley, J. *J. Adv. Phys. Org. Chem.* **1990**, *26*, 255−379.
(16) Schutz, C. N.; Warshel, A. *Proteins* **2004**, *55*, 711−723.
(17) Schmidt, A.; Lamzin, V. S. *Curr. Opin. Struct. Biol.* **2002**, *12*, 698−703.
(18) Dauter, Z. *Method Enzymol.* **2003**, *368*, 288−337.
(19) Harris, T. K.; Zhao, Q.; Mildvan, A. S. *J. Mol. Struct.* **2000**, *552*, 97−109.

the metastable tetrahedral intermediates) are studied.[20] In 1998, Kuhn et al. published the 0.78 Å resolution structure of subtilisin at pH 5.9.[21] At this pH, the catalytic histidine should be doubly protonated and positively charged, providing a partial mimic of the transition state. This structure clearly showed a short distance separating His Nδ1 and Asp Oδ2 (2.62 Å), and electron density for a hydrogen atom 1.2 Å from Nδ1 (1.5 Å from Oδ2). Although dubbed a "catalytic hydrogen bond" by the authors, this work has been cited as a key piece of evidence for the existence of a LBHB in serine protease transition states.[9]

Although low pH does indeed mimic certain aspects of the transition states of the serine protease mechanism, peptide inhibitors that covalently bind Ser195 and mimic the tetrahedral transition intermediate states (TI) provide a more chemically accurate description of the transition states (TS). Until now, subangstrom resolution crystallographic data has not been available for tetrahedral intermediate structures of serine proteases. However, lower (1.4, 1.5 Å)-resolution crystal structures of chymotrypsin bound to trifluoromethyl ketone inhibitors Ac-Phe-CF_3 or Ac-Leu-Phe-CF_3 have shown a distance of 2.6 Å between His57 and Asp102.[22] With an estimated coordinate error of 0.15 Å,[22] these structures do not rule out the possibility that this is, in fact, a standard hydrogen bond. However, combined with ^{1}H NMR chemical shift, deuterium isotope effect, and D/H fractionation factor data, these structures support the LBHB hypothesis.[23,24] It should be noted that although trifluoromethyl ketone inhibitors do offer a good mimic of certain aspects of the TI_2 state (including the tetrahedral adduct and the oxyanion), they are imperfect in that hydrogen bonds that would otherwise form with a substrate amide, ester, or alcohol group would be disrupted entirely, or displaced by ~1.4 Å, when interacting with the F on $-CF_3$.

Peptide boronic acids offer an excellent alternative analogue of the deacylation tetrahedral transition state.[25−27] The specificity profile for inhibition by peptide boronic acids is highly correlated with that of catalytic activity toward substrates.[28] Unlike the corresponding substrates, these inhibitors form a stable boron-centered tetrahedral structure (K_i = 6.4 nM for MeOSuc-Ala-Ala-Pro-boroVal-OH bound to α-lytic protease[25]), making these peptides ideal for structural studies. Furthermore, the boronic acid displays a terminal hydroxyl group in the S1′ pocket, providing a more accurate mimic of this group than offered by trifluoromethyl ketone complexes.

α-Lytic protease (αLP), an extracellular bacterial protease secreted by *Lysobacter enzymogenes*, has for many decades served as a model representative in both mechanistic and structural studies for the chymotrypsin family of serine proteases.[29] αLP's single histidine residue, located in the catalytic triad, marked it as an ideal protease for optical spectroscopic and NMR studies.[30] As a result, numerous kinetic, spectroscopic, X-ray crystallographic, and theoretical studies have focused on this enzyme.[13,28,31−37] Previously, we described the structure of αLP at pH 8 at 0.83 Å resolution.[3]

Here we present two structures of αLP at ultrahigh resolution, each mimicking an aspect of the transition state: first, as a comparison to subtilisin, αLP at pH 5 (αLP_{pH5}); second, αLP bound to the transition-state mimetic inhibitor, MeOSuc-Ala-Ala-Pro-boroVal-OH (αLP+boroVal(gol)). Interestingly, the covalent binding of a glycerol molecule to the boronate in the latter structure has provided us with, to our knowledge, the best available mimic of the acylation tetrahedral intermediate. Although ultrahigh-resolution crystal structures have been described for chymotrypsin-like serine proteases in the apo-enyzme,[3,38] acylenzyme,[39,40] and peptide-associated[41,42] states, this is the first ultrahigh-resolution crystal structure of a tetrahedral intermediate/transition state mimic for this family. This structure has allowed us to address the existence of a LBHB in the catalytic triad and other questions regarding the precise mechanism of acylation during proteolytic hydrolysis by chymotrypsin-like serine proteases.

Results

Data for the structure of αLP at pH 5.1 (αLP_{pH5}) were obtained to 0.82 Å resolution, providing the highest resolution structure of αLP to date. The structure of αLP bound to MeOSuc-Ala-Ala-Pro-boroVal-OH (αLP+boroVal(gol)) was solved to 0.90 Å resolution, the highest-resolution structure of any serine protease bound to a transition-state mimetic inhibitor. Statistics for data collection and model refinement are provided in Table 1. The final models were determined to extremely high accuracy, as indicated by the very low *R* factors (R/R_{free} of 8.06/9.27% and 8.07/9.15%, respectively). Both models exhibit a low level of thermal disorder, with an average *B* factor for all protein atoms of 5.5 and 6.1 Å^2 for αLP_{pH5} and αLP+boroVal-(gol), respectively. As described previously,[3] distinct hydrogen atoms were visible on the polar atoms of several Ser, Thr, and Tyr residues; on most aliphatic and aromatic carbons (including methyl groups); and on some distinctly ordered waters. Additionally, in both structures the side chain of Phe228[32,43] is significantly distorted. This deformation, first observed in the 0.83 Å structure of apo-αLP at pH 8, is thought to be an evolved structural element important for deriving the protein's unusual kinetic stability.[3]

(20) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334−338.
(21) Kuhn, P.; Knapp, M.; Soltis, S. M.; Ganshaw, G.; Thoene, M.; Bott, R. *Biochemistry* **1998**, *37*, 13446−13452.
(22) Neidhart, D.; Wei, Y.; Cassidy, C.; Lin, J.; Cleland, W. W.; Frey, P. A. *Biochemistry* **2001**, *40*, 2439−2447.
(23) Lin, J.; Cassidy, C. S.; Frey, P. A. *Biochemistry* **1998**, *37*, 11940−11948.
(24) Lin, J.; Westler, W. M.; Cleland, W. W.; Markley, J. L.; Frey, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14664−14668.
(25) Kettner, C. A.; Bone, R.; Agard, D. A.; Bachovchin, W. W. *Biochemistry* **1988**, *27*, 7682−7688.
(26) Kettner, C. A.; Shenvi, A. B. *J. Biol. Chem.* **1984**, *259*, 15106−15114.
(27) Bone, R.; Shenvi, A. B.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1987**, *26*, 7609−7614.
(28) Bone, R.; Fujishige, A.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1991**, *30*, 10388−10398.
(29) Fuhrmann, C. N.; Ota, N.; Rader, S. D.; Agard, D. A. In *Handbook of Proteolytic Enzymes*, 2nd ed.; Barrett, A., Rawlings, N. D., Woessner, J. F., Eds.; Academic Press: Cambridge, 2004; Vol. 2.
(30) Bachovchin, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 7948−7951.
(31) Brayer, G. D.; Delbaere, L. T.; James, M. N. *J. Mol. Biol.* **1979**, *131*, 743−775.
(32) Fujinaga, M.; Delbaere, L. T.; Brayer, G. D.; James, M. N. *J. Mol. Biol.* **1985**, *184*, 479−502.
(33) Bone, R.; Silen, J. L.; Agard, D. A. *Nature* **1989**, *339*, 191−195.
(34) Mace, J. E.; Wilk, B. J.; Agard, D. A. *J Mol Biol* **1995**, *251*, 116−134.
(35) Rader, S. D.; Agard, D. A. *Protein Sci.* **1997**, *6*, 1375−1386.
(36) Davis, J. H.; Agard, D. A. *Biochemistry* **1998**, *37*, 7696−7707.
(37) Sohl, J. L.; Jaswal, S. S.; Agard, D. A. *Nature* **1998**, *395*, 817−819.
(38) Wurtele, M.; Hahn, M.; Hilpert, K.; Hohne, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*(Pt 4), 520−523.
(39) Wilmouth, R. C.; Edman, K.; Neutze, R.; Wright, P. A.; Clifton, I. J.; Schneider, T. R.; Schofield, C. J.; Hajdu, J. *Nat. Struct. Biol.* **2001**, *8*, 689−694.
(40) Katona, G.; Wilmouth, R. C.; Wright, P. A.; Berglund, G. I.; Hajdu, J.; Neutze, R.; Schofield, C. J. *J. Biol. Chem.* **2002**, *277*, 21962−21970.
(41) Rypniewski, W. R.; Ostergaard, P. R.; Norregaard-Madsen, M.; Dauter, M.; Wilson, K. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 8−19.
(42) Schmidt, A.; Jelsch, C.; Ostergaard, P.; Rypniewski, W.; Lamzin, V. S. *J. Biol. Chem.* **2003**, *278*, 43357−43362.
(43) αLP residues are numbered based on homology to chymotrypsin, as described in ref 32.

Table 1. Data Collection Statistics

	data set	
	αLP (pH 5.1)	αLP + boroVal(gol)
A. Data Statistics		
space group	P 3_2 2 1	P 3_2 2 1
unit cell (*a*, *c*; Å)	65.7, 79.6	65.8, 79.6
mosaicity (°)	∼ 0.2	∼0.17
limiting resolution (Å)	0.82	0.90
total no. reflections	1,405,443	1,279,938
no. unique reflections	189,525	147,100
$I/\sigma(I)$	38.7 (3.0)[a]	27.9 (4.8)[b]
% completeness	97.9 (84.5)[a]	99.9 (99.5)[b]
R_{merge} (Å)	4.6 (37.5)[a,c]	8.5 (44.8)[b,c]
B. Structure Refinement		
resolution range (Å)	20−0.82	20−0.90
R (no. reflections in working set)	8.06 (179,637)	8.07 (139,610)
R_{free} (no. reflections in test set)	9.27 (9467)	9.15 (7352)
RMS bond lengths[d] (Å)	0.025	0.016
RMS angle distances[d] (Å)	0.045	0.045
protein residues	198	202
residues with alternate conformations	37	21
SO_4^{2-}/glycerol molecules	11/2	8/1
water, 100% occupancy	148	160
water, partial occupancy	435	389
average isotropic *B* factors		
protein atoms (Å^2)	5.5	6.1
solvent atoms (Å^2)	14.8	19.7

[a] Highest resolution bin for αLP (apo-enzyme, pH 5.1) = 0.83−0.82 Å. [b] Highest resolution bin for αLP + boroVal(gol) = 0.91−0.90 Å. [c] R_{merge} as calculated by Scalepack.[44] [d] RMS values as defined by SHELXL-97.[48]

Data Collection Strategy for αLP+boroVal(gol). Earlier crystallographic studies of αLP bound to boronic acid inhibitors at ultrahigh resolution revealed that the Ser195-boron adduct adopts alternate conformations as a result of radiation damage at the high X-ray dosage necessary for such resolution.[45] Taking advantage of the large size of the αLP crystal (∼0.3 mm × 0.3 mm × 0.5 mm), it was possible to spread the ultrahigh-resolution data collection over nine positions on the crystal, minimizing X-ray exposure accumulated at any one position. A complete moderate-resolution data set (1.6 Å), having a total exposure equal to one to two frames of ultrahigh-resolution data, was collected at each position before and after the ultrahigh-resolution data collection. Several tests confirmed that no damage had occurred during the ultrahigh-resolution data collection (see Methods). High-resolution data sets from each position were scaled and merged with moderate-resolution data, resulting in a complete data set to 0.90 Å resolution (Table 1).

General Structural Features of αLP at pH 5. The structure of αLP at pH 5 (αLP_{pH5}) was determined to 0.82 Å resolution. Ultrahigh resolution data was collected at only two positions on the crystal. The crystal's extended exposure to X-rays resulted in visible radiation damage including partial reduction of αLP's three disulfide bonds, similar to previous observations for αLP_{pH8}.[3]

At 0.82 Å resolution, precise features of the structure can be modeled, including hydrogen atoms and low-occupancy alternate conformations of protein atoms. Thirty-seven residues were modeled with alternate conformations (including the six partially reduced cysteine residues). Four areas of the protein displayed discrete disorder continuing along the backbone atoms of two or more residues: 59−59A[43] (2 residues), 66−82 (4 residues), 201A-202 (2 residues), and 219A-219B (2 residues). Smeared $2F_o - F_c$ electron density and $F_o - F_c$ difference density (contoured at 3σ) also clearly indicated the presence of multiple conformers for the carbonyl oxygen of Leu41 in αLP_{pH5}, which hydrogen bonds to a water in the S1′ pocket of the enzyme.[46,47] This anisotropy was modeled with a second conformer of Leu41 C═O, which refined to an occupancy of ∼26%. Because the major conformation of Leu41 O is sterically occluded by the side chain of $Cys42_b$ (with a distance of 2.34 Å separating $Leu41_a$ O from $Cys42_b$ Sγ), Leu41 must occupy the minor conformation when the Cys42/Cys58 disulfide bond is reduced by radiation damage. Radiation damage is not required for this conformational heterogeneity, however, because similar anisotropy is also observed for Leu41 in αLP+boroVal(gol) (where Cys42/Cys58 occupy only one conformation). Similar smearing of $2F_o - F_c$ electron density was observed for Leu41 in αLP_{pH8}; therefore, we are unable to assess a correlation between dynamic motion of Leu41 and protease activity.

Comparing αLP_{pH5} and αLP_{pH8}. The general structure of αLP_{pH5} is nearly identical to that of αLP_{pH8} (RMSD for main-chain atoms 0.066 Å; RMSD for all non-hydrogen atoms 0.259 Å).[3] As in αLP_{pH8}, αLP_{pH5} contains two solvent molecules in its binding pockets: a sulfate ion (Sul202) in the active site and a glycerol (Gol210) in the S2 pocket (see Figure 2a). Sul202 is present in two conformations ($Sul202_a$, $Sul202_b$), with a total occupancy of ∼88%. This sulfate ion forms a hydrogen-bonding network in the active site, including hydrogen bonds to His57, Ser195 Oγ, Gly141 N in the oxyanion hole, Gol210, Arg192, and Arg122 (in a crystal contact). Atoms O2 and O4 of $Sul202_a$ (occupancy ∼75%) form electrostatic interactions with positively charged residues His57 and Arg192/Arg122, respectively. Therefore, it would be expected that the two negative charges of $Sul202_a$ reside primarily on these two oxygens. Indeed, the S−O2 and S−O4 bond lengths are longer than those of S−O1 and S−O3 (1.48 and 1.49 Å, compared to 1.45 and 1.43 Å), despite restraints applied during the refinement to promote a bond length of 1.42 Å for all sulfate bonds. We have therefore assigned partial negative charges to $Sul202_a$ O2 and O4 (Figure 4a). In the absence of Sul202, two water molecules (Wat668 and Wat669) maintain hydrogen-bonding interactions with Arg192 and Arg122, and with Ser195 and the oxyanion hole, respectively (Figure 2a).

Although αLP_{pH5} and αLP_{pH8} are nearly identical in their overall structures, notable differences do exist in the vicinity of the active site. At pH 5, His57 should be protonated at Nϵ2 in all αLP molecules of the crystal. Indeed, in αLP_{pH5} a peak corresponding to Nϵ2−H is visible in σ_A-weighted $F_o - F_c$ difference maps (Figure 4a). Several structural changes occur in the vicinity of the active site in response to the protonation of His57 Nϵ2. One primary effect is that Ser195 occupies only one conformation at pH 5 (see Figure 2b and Discussion). In αLP_{pH5}, Sul202 is present at higher occupancy than in αLP_{pH8};

(44) Fox, G. C.; Holmes, K. C. *Acta Crystallogr.* **1966**, *20*, 886−891.
(45) Fuhrmann, C. N. Ph.D. Dissertation, University of California, San Francisco, San Francisco, CA, 2005.
(46) Substrate residues are labeled according to the nomenclature of Schechter and Berger, 1967, where residues extending N-terminal from the cleaved bond are P1, P2, etc.; those extending C-terminally are P1′, P2′, etc. The corresponding enzyme binding pockets are labeled S1, S1′, etc.
(47) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun*. **1967**, *27*, 157−162.
(48) Sheldrick, G. M.; Schneider, T. R. *Methods Enzymol*. **1997**, *277*, 319−343.

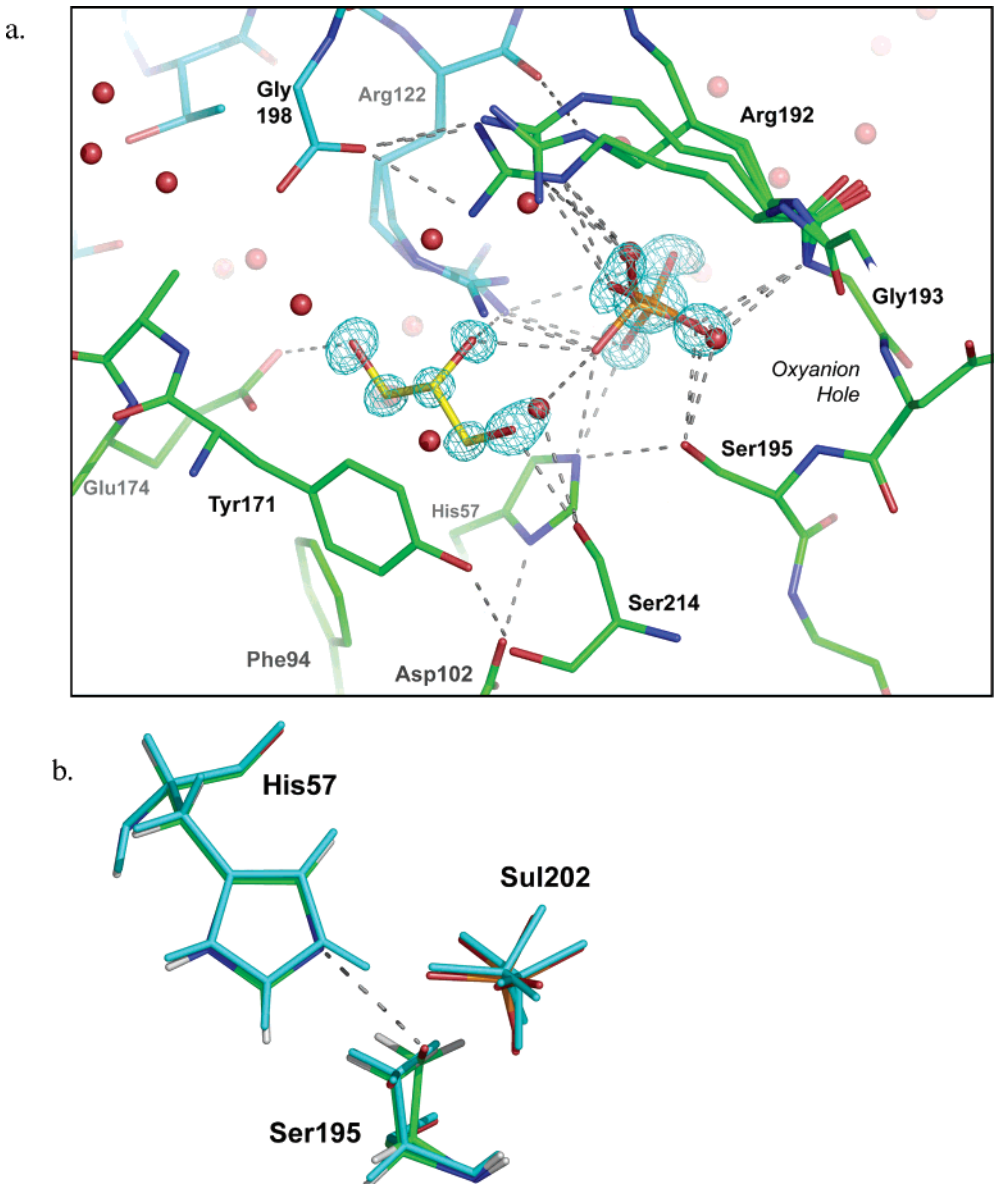

Figure 2. Active site of αLP_{pH5}. (a) $2F_o - F_c$ electron density (contoured at 2σ) shows the position of two solvent molecules in the substrate binding pockets of αLP. Sulfate 202 occupies the region of the active site and is stabilized by hydrogen-bonding interactions (dashed lines) with Ser195, His57, the oxyanion hole, and nearby arginine residues. A glycerol molecule occupies the S2 binding pocket. A neighboring molecule in the crystal lattice is represented in cyan, and water molecules, by red spheres. (b) Overlay of αLP_{pH8} (atoms colored by element) and αLP_{pH5} (cyan); the αLP_{pH5} structure has been offset slightly to allow visualization of both structures. In αLP_{pH8}, we previously hypothesized that Ser195 occupies one of two different positions, depending on the protonation state of His57.[3] The structure of αLP at pH 5, in which all αLP molecules are protonated at His57 Nϵ2, confirms this hypothesis. Ser195 occupies one conformation, which corresponds to the conformation previously predicted in αLP_{pH8} to be that associated with His57 in its protonated state.

correspondingly, Arg192 and Arg122 both exhibit a higher occupancy for conformations that hydrogen bond to Sul202. These primary structural changes cause subtle differences in backbone positions of surrounding residues. These structural observations are confirmed in $F_{o,\alpha LP,pH5} - F_{o,\alpha LP,pH8}$ maps (data not shown).

General Structural Features of αLP+boroVal(gol). The boronic acid inhibitor MeOSuc-Ala-Ala-Pro-boroVal-OH inhibits αLP proteolysis by forming a covalent attachment to Ser195, providing a structural mimic of the proteolytic transition state. During initial refinement of αLP+boroVal(gol), both positive electron density and σ_A-weighted $2F_o - F_c$ density (extending from boroValP1[46,47] to the C-terminal carbonyl of MeOSucP5) indicated successful binding of the inhibitor in the binding pocket. Covalent linkage of the boronate to Ser195 was evident by a short distance (1.47 Å) between Ser 195 Oγ and the boron of the inhibitor, boroValP1 B, as well as continuous electron density extending from Ser195 Oγ to boroValP1 B. In this structure, the side chain of Ser195 occupies a conformation similar to that observed in αLP_{pH5}. The inhibitor, bound at complete occupancy, was initially modeled as a single conformation, but difference maps indicated a second conformation for AlaP3−AlaP4 (occupancy ∼22%). Three hydrogen bonds (Ser214 O···boroValP1 N, Gly216 N···AlaP3 O, and AlaP3 N···Gly216 O) stabilize the inhibitor−enzyme complex, forming a three-stranded β-sheet (see Figure 3b, Table 2).

The short atom−atom distance (1.4 Å) and difference electron density maps clearly indicated the covalent addition of a glycerol

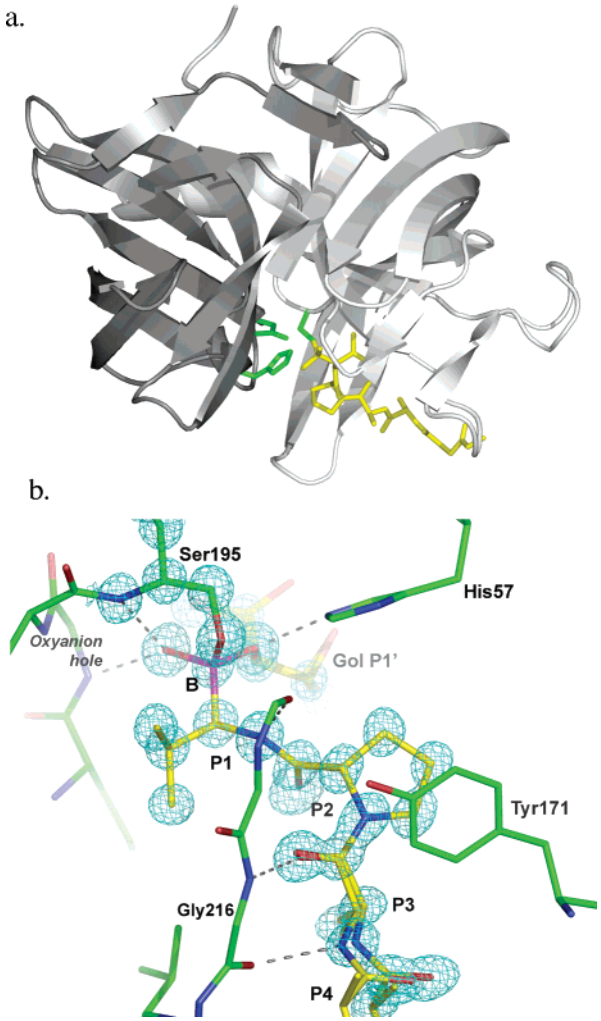

Figure 3. Boronic acid inhibitor, MeOSuc-Ala-Ala-Pro-boroVal-OH, binds Ser195 in the active site of αLP, mimicking the tetrahedral intermediate. (a) Ribbon diagram of αLP, showing the position of the inhibitor (yellow) in the enzyme's binding pocket. Side chains for the catalytic triad (Asp102, His57, Ser195) are shown in green. For simplicity, glycerol and water molecules are not shown. (b) σ_A-weighted $2F_o - F_c$ electron density maps (cyan; contoured at 2σ) illustrate the covalent attachment of the boronic acid inhibitor to Ser195. αLP interacts with the boronic ester directly through three hydrogen bonds and with the length of the inhibitor through three additional backbone hydrogen bonds. A glycerol molecule (GolP1′) covalently attached to the boronate mimics the leaving group, providing a model of TI_1.

molecule to the boronate. This glycerol (GolP1′), introduced to the crystal as a component of the cryosolvent, is present in three conformations with a total occupancy of ∼66%. In any one of the three conformations, the glycerol is bound via a secondary ester linkage to boroValP1 O2 (Figures 4, 5). The terminal oxygen O1 of two of the conformations for GolP1′ forms hydrogen bonds with Leu41 O and boroValP1 O1 in the oxyanion hole (Table 2). In the absence of this glycerol molecule, a water maintains both of these hydrogen bonds. A similar water forms this contact in the apo-enzyme structures (αLP_{pH8} [3] and αLP_{pH5}). Although GolP1′ does not satisfy the typical P1′ hydrogen bonds expected of a natural substrate,[49] it does lie in the S1′ binding pocket of the enzyme and is chemically bonded to the tetrahedral adduct (with a boroValP1 O2···GolP1′ C2 distance of 1.4 Å), providing a mimic of the substrate's leaving group. One can therefore consider αLP+boroVal(gol) as a model for the acylation tetrahedral intermediate (TI_1).

The αLP+boroVal(gol) model is nearly identical to the 0.83 Å structure of the apo-enzyme (αLP_{pH8}),[3] with an RMSD of 0.177 Å for main chain atoms, and 0.530 Å for all non-hydrogen protein atoms. Except for residues in symmetry-related molecules involved in crystal contacts with the binding pocket, significant structural differences are located primarily in the vicinity of the binding pocket. Binding of the inhibitor displaces the sulfate ion (Sul202) and glycerol molecule (GolP2) occupying the binding pocket in apo-αLP, causing shifts in local water networks. Main-chain atoms of adjacent residues 171 and 173, forming a β-hairpin adjacent to the S2 binding pocket, move by as much as ∼0.8 Å upon binding the inhibitor. This is likely due to the loss of the hydrogen-bond network present in the apo structure, which links GolP2 to the backbone nitrogen of Ala173 through hydrogen bonds with a water molecule.

Interestingly, upon inhibitor binding, the β-strand corresponding to residues 214−217 shifts ∼0.8 Å toward the inhibitor. Motion of this strand has been characterized previously for the apo- and bound forms of the enzyme, and its dynamics have been hypothesized to regulate substrate specificity.[28,33−35,50,51] This collapse of the binding pocket is likely driven by formation of three hydrogen bonds between the backbone of this strand and the inhibitor (Figure 3b). Two of these hydrogen bonds replace interactions in the apo-enzyme to a glycerol (GolP2) and partially occupied water, yielding a gain of one hydrogen bond in the peptide-bound state. The observed strand displacement becomes significantly reduced at Ser214 ($C\alpha_{pH8}\cdots C\alpha_{boroVal}$ displacement of ∼0.2 Å). Nearby residues in the catalytic triad reflect a correspondingly slight shift of 0.15−0.3 Å; therefore, local hydrogen bonds within the catalytic triad have similar geometries between the apo and TI_1 states (Table 2). Finally, Gly193 (forming the oxyanion hole) shifts away from the active site by ∼0.4 Å relative to αLP_{pH8}, likely due to steric interactions with boroValP1 O1.

The Ser195-Boronate Adduct. At ultrahigh resolution, it is possible to decisively describe the geometry of the boroValP1−Ser195 adduct. In αLP+boroVal(gol), the adduct is of essentially tetrahedral geometry, with angles about the boron ranging from 105° to 112° (Table 3). As expected, all three B−O bonds are approximately the same length (average 1.49 ± 0.03 Å) and are shorter than the less polar B−Cα bond (1.61 Å; Table 3). These bond lengths are nearly identical to those seen in structures of tetrahedral boron-centered compounds solved by small molecule crystallography.[52,53]

Active-Site Hydrogen Atoms in the Transition-State Mimic Structures. As expected for these highly resolved structures, hydrogen atoms were visible in σ_A-weighted $F_o - F_c$ density contoured at 2σ on 89%/90% of Cα atoms and 87%/83% of backbone amides in αLP_{pH5}/αLP+boroVal, respectively.

(49) Bone, R.; Sampson, N. S.; Bartlett, P. A.; Agard, D. A. *Biochemistry* **1991**, *30*, 2263−2272.
(50) Ota, N.; Agard, D. A. *Protein Sci.* **2001**, *10*, 1403−1414.
(51) Miller, D. W.; Agard, D. A. *J. Mol. Biol*. **1999**, *286*, 267−278.
(52) Allen, F. H. *Acta Crystallogr. Sect., B* **2002**, *58*, 380−388. CSD accession numbers JIQHUM, JIQJEY, VUSTUY, HEMBOQ.
(53) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. *Acta Crystallogr., Sect. B* **2002**, *58*, 389−397.

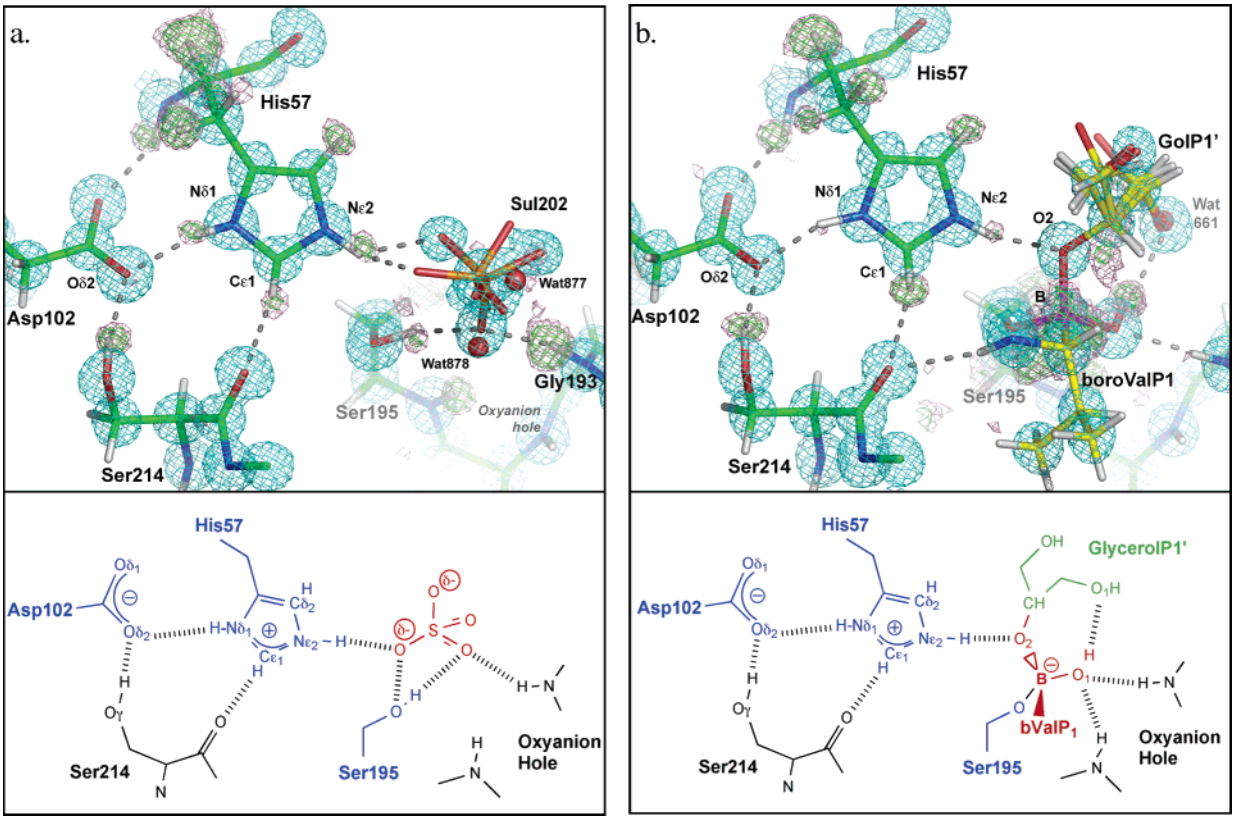

Figure 4. Experimentally observed hydrogen atoms in the active-site region. An illustrative schematic along with the refined model and electron density maps are shown for the active sites of (a) αLP_{pH5} and (b) αLP+boroVal(gol). For both models, σ_A-weighted $2F_o - F_c$ electron density maps are drawn at 2σ (cyan), and σ_A-weighted $F_o - F_c$ maps at 3σ (green) and 2.5σ (pink). $F_o - F_c$ difference electron density maps were calculated prior to addition of active-site hydrogens to the model (see Methods). Sticks represent the final refined models, with the positions of key hydrogen atoms refined (see Table 2). Dashed lines illustrate hydrogen-bonding interactions.

Because of the potential for bias to be introduced when calculating omit maps, active-site hydrogen atoms were added to the models late in refinement. Most mechanistic protein hydrogens were visible in σ_A-weighted $F_o - F_c$ difference electron density maps at contour levels at or greater than 3σ before their addition to the model (see Figures 4, 5). The positions of particular active-site hydrogen atoms were refined without restraints; the resulting donor-to-hydrogen distances are listed in Table 2.

Ser195 in αLP_{pH5}. Interestingly, a hydrogen atom was visible on Ser195 Oγ in σ_A-weighted $F_o - F_c$ difference electron density maps of αLP_{pH5} (Figure 4a). This hydrogen atom had not been visible on the corresponding conformation of Ser195 in αLP_{pH8} ($Ser195_b$) because it overlapped with electron density for $Ser195_a$ Oγ in that structure.[3] Not surprisingly, Ser195 Oγ−H is positioned in a plane ideal to hydrogen bond with Sul202. Visible at a contour of 3σ in $F_o - F_c$ difference maps, this peak is ∼1.04 Å from Ser195 Oγ, displaced toward Sul202 likely because of electrostatic effects.

Hydrogen Atoms in the Oxyanion Hole. One of the prominent structural features of the serine protease mechanism is a conserved pocket designed to stabilize the oxyanion formed in the tetrahedral transition states. In αLP, this "oxyanion hole" is composed of the backbone amides of Gly193 and Ser195, which participate in hydrogen-bonding interactions with the substrate's oxyanion. In both αLP_{pH5} and αLP+boroVal(gol), hydrogen atoms were visualized as distinct peaks on both amides in σ_A-weighted $F_o - F_c$ maps calculated prior to the addition of hydrogen atoms to the model. To our knowledge, this is the first time oxyanion hole hydrogen atoms have been visible by X-ray crystallography in a transition-state analogue.

In αLP_{pH5}, Gly193 N (but not Ser195 N) is involved in hydrogen-bonding interactions, primarily with Sul202 (Figure 2, Table 2). Both oxyanion hydrogen atoms were visible in σ_A-weighted $F_o - F_c$ maps contoured at 4σ (Figure 4a). Gly193 N−H refined to a position 1.06 ± 0.05 Å from its parent atom in agreement with the position of $F_o - F_c$ density observed off Gly193 N. This is in contrast to the previously solved structure of αLP_{pH8}, where the position of 193 N−H refined to be 1.37 ± 0.06 Å from its parent.[3] A peak of $F_o - F_c$ density was observed ∼0.89 Å from Ser195 N, in good agreement with the expected position of a hydrogen based on amide library values[48] (see Methods).

In αLP+boroVal(gol), O1 of the tetrahedral boronate lies in the oxyanion hole. Interestingly, σ_A-weighted $F_o - F_c$ difference electron density (contoured at 3σ) was visible off boroValP1 O1 (Figure 5). Addition of a restrained hydroxyl-hydrogen atom to O1 resulted in the disappearance of this density. Upon release of restraints, this hydrogen refined to a position 1.00 ± 0.06 Å from O1. The O1−H bond is aligned to facilitate hydrogen bonding to the terminal hydroxyl group of glycerol P1′, with a donor−acceptor distance of 2.74 ± 0.01 Å. Additionally, boroValP1 O1 participates in hydrogen bonding with Gly193 N−H and Ser195 N−H of the oxyanion hole, with N···O1 distances of 2.733 ± 0.005 and 2.847 ± 0.005 Å, respectively (Table 2). In early σ_A-weighted $F_o - F_c$ maps calculated prior

Table 2. Hydrogen-Bonding Geometries

	apo-αLP, pH 8[a]			apo-αLP, pH 5			αLP + boroVal(gol)		
interaction (D···A)	*d*(D···A) (Å)	*d*(D–H) (Å)[b]	D–H···A (deg)[b]	*d*(D···A) (Å)	*d*(D–H) (Å)[b]	D–H···A (deg)[b]	*d*(D···A) (Å)	*d*(D–H) (Å)[b]	D–H···A (deg)[b]
Gly193 N···Sul202 $_a$ O1	2.820(9)	1.37(6)	162(5)	2.812(7)	1.06(5)	158(4)	–	–	–
Gly193 N···Sul202 $_a$ O3	–	–	–	3.280(8)	1.06(5)	136(4)	–	–	–
Gly193 N···Sul202 $_b$ O3	2.85(3)	1.37(6)	154(5)	2.879(19)	1.06(5)	166(4)	–	–	–
Gly193 N···Wat760$_a$	2.90(3)	1.37(6)	142(4)	2.739(4)	1.06(5)	145(4)	–	–	–
Gly193 N···Wat765	2.78(4)	1.37(6)	115(4)	–	–	–	–	–	–
Ser195 N···Wat760$_a$	2.96(3)	*n/d*	141	–	–	–	–	–	–
Gly193 N···boroValP1 O1	–	–	–	–	–	–	2.733(5)[c]	0.87(5)	147(5)
Ser195 N···boroValP1 O1	–	–	–	–	–	–	2.847(5)[c]	0.76(5)	157(5)
Ser195$_b$ Oγ···Sul202$_a$O1	2.64	*n/d*	*n/d*	2.647(6)	*n/d*	170	–	–	–
Ser195$_b$ Oγ···Sul202$_b$O3	2.77	*n/d*	*n/d*	2.86(2)	*n/d*	166	–	–	–
His57 N···Asp102 Oδ1	2.820	*n/d*	148	2.804(5)	*n/d*	149	2.818[c]	*n/d*	149
Ser214 Oγ···Asp102 Oδ2	2.698(4)	0.98(7)	174(6)	2.705(4)	0.97(6)	169(5)	2.688(5)[c]	0.98(7)	168(6)
His57 Nδ1···Asp102 Oδ2	**2.770(5)**	**0.83(6)**	**170(2)**	**2.755(5)**	**0.88(5)**	**164(5)**	**2.734(5)[c]**	**1.01(5)**	**159(4)**
His57 Cϵ1···Ser214 O	3.033(6)	1.00(6)	126(3)	3.015(5)	0.85(5)	128(4)	3.089(6)	0.87(6)	140(5)
His57 Nϵ2···Ser195$_a$ Oγ	2.826(15)	*n/d*	139	–	–	–	–	–	–
His57 Nϵ2···Ser195$_b$ Oγ	2.962(5)	*n/d*	120	3.002(4)	1.00(5)	131(4)	3.011(5)	0.91(6)	127(4)
His57 Nϵ2···boroValP1 O2	–	–	–	–	–	–	**2.643(4)[c]**	**0.91(6)**	**165(5)**
His57 Nϵ2···Sul202$_a$ O2	2.80	*n/d*	*n/d*	2.760(5)	1.00(5)	131(4)	–	–	–
Glycerol$_a$ O1···Leu41$_a$ O	–	–	–	–	–	–	2.620(11)	*n/d*	*n/d*
Glycerol$_a$ O1···Leu41$_b$ O	–	–	–	–	–	–	2.91	*n/d*	*n/d*
boroValP1 O1··· glycerol$_a$ O1	–	–	–	–	–	–	2.744(12)	1.00(6)	146(5)
boroValP1 N···Ser214 O	–	–	–	–	–	–	3.110(5)[c]	*n/d*	164
Gly216 N···AlaP3$_a$ O	–	–	–	–	–	–	2.936(10)[c]	*n/d*	160
AlaP3$_a$ N···Gly216 O	–	–	–	–	–	–	3.099(11)[c]	*n/d*	164

[a] See ref 3. [b] D−H bond lengths are stated in cases where the position of the hydrogen atom was refined without restraints. *n/d* indicates that the position of the hydrogen atom was not refined against the data. The value provided for D−H···A angle geometry in these cases is based on a hydrogen atom whose position was fixed at optimal geometry by SHELXL-97.[48] It should be noted that X-ray crystallographic values for donor−hydrogen distances are shorter than true internuclear distances due to displacement of the hydrogen electron density toward the donor atom, and librational effects.[48] [c] Hydrogen bond illustrated in Figure 8.

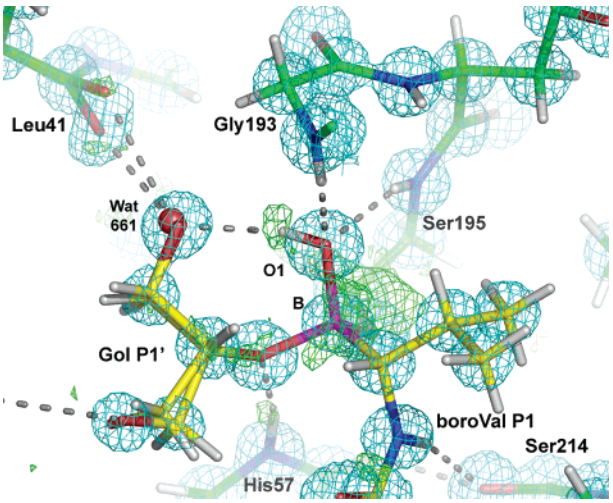

Figure 5. Boronic acid is protonated in the oxyanion hole. $2F_o - F_c$ electron density is contoured at 2σ (cyan), and σ_A-weighted $F_o - F_c$ density contoured at 3σ (green). Even at this high contour level, a clear difference peak is visible off boroValP1 O1. O1 forms hydrogen-bonding interactions with Gly193 and Ser195 in the oxyanion hole, and with the terminal oxygen of GolP1′. Positive $F_o - F_c$ density on the boron is consistent with the negative charge of the complex being located on the boron, as is expected for a protonated boronic ester. σ_A-weighted $F_o - F_c$ maps were calculated prior to addition of the boroValP1 O1 or His57 Nϵ2 hydrogen atoms to the model. The stick model shows the final structure of αLP+boroVal(gol), in which the position of O1−H has been refined.

to the addition of hydrogen atoms to the model, a strong peak was visible off Ser195 N (at contour level 4σ) and a weaker peak off Gly193 N (at contour level 2σ). The positions of Gly193 N−H and Ser195 N−H refined to be 0.87 ± 0.05 Å and 0.76 ± 0.05 Å from their donor atoms, respectively.

Table 3. Boronate Geometries for αLP + boroVal(gol)

	length (Å)		angle (deg)
B−Oγ	1.466(6)	Oγ−B−O1	112.3(4)
B−O1	1.485(6)	Oγ−B−O2	104.7(4)
B−O2	1.520(7)	Oγ−B− Cα	108.6(4)
B−Cα	1.608(8)	O1−B−O2	107.5(4)
		O1−B−Cα	111.7(4)
		O2−B−Cα	111.8(4)

As discussed below, protonation of boroValP1 O1 suggests that the negative charge of the tetrahedral adduct in boronic acid transition-state analogues lies on the boron. Indeed, in the final stages of refinement, positive σ_A-weighted $F_o - F_c$ difference electron density on boroValP1 B indicated that the five electrons defined for a boron at this position were not sufficient (Figure 5). To address this, the occupancy of boron was allowed to refine. The occupancy refined to ∼89% of a carbon atom, indicating that the true density of B in boroValP1 is ∼5.3 electrons. Whether this value represents a partial negative charge or is an underestimate of the six electrons expected for a full formal charge is unknown.

Ser214 Positions the Catalytic Triad. Ser214, a conserved residue among serine proteases, has in the past been referred to as the fourth member of the catalytic triad due to the conserved hydrogen bonds that it forms to the side chains of Asp102, His57, and the backbone amide of the substrate's P1 residue.[54] The Ser214 Oγ···Asp102 Oδ2 hydrogen bond maintains essentially identical geometry in αLP$_{pH8}$, αLP$_{pH5}$, and αLP+boroVal(gol), with an average donor−acceptor distance of 2.70 ± 0.01 Å in the three structures (Table 2). In

(54) Krem, M. M.; Di Cera, E. *EMBO J.* **2001**, *20*, 3036−3045.

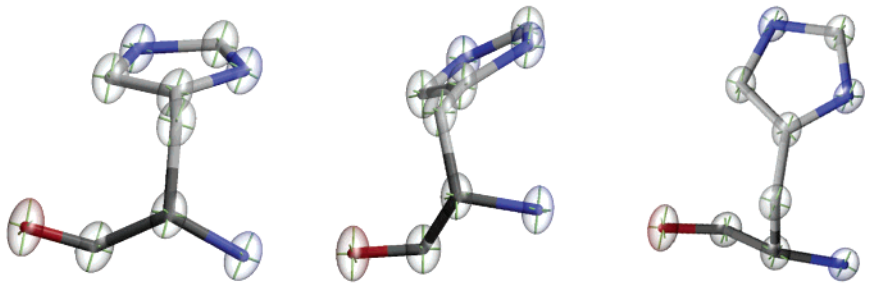

Figure 6. Anisotropic motion of His57 in αLP+boroVal(gol). His57 is shown in three orientations, with atomic anisotropy represented by ellipsoids drawn at the 50% probability level. Anisotropic motion for all atoms in His57 are of comparable intensity and direction, with no indication of imidazole ring-flipping. This figure was made using Raster3D/rastep[6] and Molscript.[7]

all three structures, a hydrogen atom was visible in positive σ_A-weighted $F_o - F_c$ electron density maps on Oγ, positioned in line with this hydrogen bond (see Figure 4).[3] This hydrogen atom refined to identical positions in all three structures, with an Oγ−H distance of 0.97 ± 0.06 Å in αLP_{pH5} and 0.98 ± 0.07 Å in αLP+boroVal(gol).

The backbone carbonyl of Ser214 acts as acceptor for an unusual C−H···O hydrogen bond with His57 Cϵ1-H. This conserved hydrogen bond was previously recognized and proposed to play a key role in the mechanism of serine proteases.[55] In αLP+boroVal(gol), the donor−acceptor distance for this hydrogen bond (3.089 ± 0.006 Å) is slightly longer than that observed in the apo-enzyme at pH 5 (3.015 ± 0.005 Å) or pH 8 (3.033 ± 0.006 Å).[3] Positive σ_A-weighted $F_o - F_c$ electron density is clearly visible in a distinct peak off His57 Cϵ1 in all three structures (Figure 4), with a refined Cϵ1−H bond length similar to the expected value of 0.95 Å[48,56] (see Table 2).

Hydrogen Atoms on His57. According to the general mechanism for serine protease catalysis, His57 is expected to be protonated at both Nδ1 and Nϵ2 in all of the tetrahedral intermediate and transition states. Solved at a pH below the pK_a of His57, the structure αLP_{pH5} should mimic this aspect of the transition state. Indeed, in αLP_{pH5} all hydrogen atoms on His57 were visible in σ_A-weighted $F_o - F_c$ electron density maps, including Hϵ2 (which refined to an Nϵ2−H bond length of 1.00 ± 0.05 Å), confirming the protonation state of His57 (Figure 4a). Availability of the Nϵ2−H hydrogen atom allows hydrogen bonding to both conformations of Sul202 and to Ser195 Oγ (Table 2).

Two of the most famous controversies in enzymology have focused on the precise character of the His57 Nδ1···Asp102 Oδ2 hydrogen bond in the transition state.[8,57] In αLP_{pH5}, the Nδ1···Oδ2 hydrogen bond is 2.755 ± 0.005 Å, which is only slightly shorter than the distance observed in αLP_{pH8} (2.770 ± 0.005 Å).[3] In αLP+boroVal, which is a better mimic of the transition state, this hydrogen bond distance is 2.734 ± 0.005 Å. Therefore, in both αLP_{pH5} and αLP+boroVal(gol), this hydrogen bond is too long to qualify as a short, strong or low-barrier hydrogen bond (see Discussion). In αLP_{pH5}, a clean peak of difference electron density contoured at 3σ indicated the position of a hydrogen atom on Nδ1 (Figure 4a). This hydrogen refined to a distance 0.88 ± 0.05 Å away from its parent atom (Nδ1). In αLP+boroVal(gol), a strong difference electron density peak (3σ) was observed in early stages of refinement at a position expected for a hydrogen atom on Nδ1. Surprisingly, the height of this peak decreased in the later stages of refinement (Figure 4b). Several lines of evidence, however, suggest that this is a bona fide hydrogen atom. First, the peak we observe is not due to dynamic motion of His57 or Asp102, which have very low average *B* factors (3.9 ± 0.4 for atoms in these residues), or of His57 Nδ1 in particular (which displays anisotropic motion primarily in a direction perpendicular to the Nδ1−H axis (Figure 6)). Even at reduced strength, the peak height is similar to those observed for numerous amide hydrogen atoms in the structure. A hydrogen atom in this position refines to 1.01 ± 0.05 Å from Nδ1 (1.77 Å from Asp102 Oδ2). Therefore, both αLP_{pH5} and αLP+boroVal(gol) show that this hydrogen is clearly located on His57 Nδ1 when His57 is doubly protonated, and not shared between His57 and Asp102 as would be predicted by the LBHB hypothesis.

During the acylation reaction, His57 donates its Nϵ2−H hydrogen atom to the nitrogen of the substrate's leaving group, allowing cleavage of the substrate and production of the acylenzyme. In the tetrahedral intermediate mimetic structure, αLP+boroVal(gol), His57 Nϵ2 lies at a favorable distance (2.643 ± 0.004 Å) and angle (165° ± 5°) from boroValP1 O2 for hydrogen bonding (Table 2). A peak is visible in σ_A-weighted $F_o - F_c$ difference electron density maps contoured at 3σ (Figure 4b). This hydrogen refines to a position 0.91 ± 0.06 Å from its parent atom, Nϵ2, and 1.75 Å from the acceptor atom, O2. In contrast to the Nϵ2···O2 interaction, His57 Nϵ2 forms only a weak hydrogen bond with Ser195 Oγ, with a donor−acceptor distance of 3.011 ± 0.005 Å, and a D−H···A angle of 127° ± 4°.

Discussion

We solved the structures of α-lytic protease at pH 5.1 (αLP_{pH5}) and bound to a transition-state mimetic inhibitor, MeOSuc-Ala-Ala-Pro-Val boronic acid (αLP+boroVal(gol)) at subangstrom resolution (0.82 Å and 0.90 Å, respectively). Both structures were determined to very high accuracy, with the lowest *R* factors reported to date for a protein larger than 100 residues (Table 1). Although the refined structures are nearly identical to the corresponding structures solved at lower resolution according to RMSD (see Results), information regarding the precise positioning of atoms has emerged at ultrahigh resolution that allows us to address long-standing questions regarding the mechanism of proteolytic catalysis.

(55) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241*, 83−93.
(56) X-ray crystallographic values for donor-hydrogen distances are shorter than true internuclear distances due to displacement of the hydrogen electron density toward the donor atom, and librational effects.
(57) Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969**, *221*, 337−340.

Structural Effects of the Protonation of His57 in the Apo-Enzyme. In the previously solved 0.83 Å resolution crystal structure of αLP at pH 8 (αLP_{pH8}) it had been observed that Ser195 was present in two conformations.[3] One of these conformations ($Ser195_a$) formed a hydrogen bond with His57 Nϵ2; in the other conformation, $Ser195_b$ Oγ was shifted by ~1 Å, increasing its distance to His57 Nϵ2 and bringing it to an angle that was less favorable for hydrogen bonding (see Figure 2b). It was hypothesized that this latter conformation of Ser195 corresponded with a population of αLP molecules within the crystal for which His57 was protonated at Nϵ2, implying that the pK_a of His57 had shifted from 7 to 8.8 in αLP crystals (likely due to the close proximity of a sulfate ion in the active site). Protonation at Nϵ2 would cause steric clash with the canonical conformation of Ser195, causing it to shift to the second conformation.

The structure of αLP at pH 5 confirms this proposition. According to this hypothesis, at pH 5 100% of αLP molecules would be protonated at His57 Nϵ2, and therefore, only the second conformation of Ser195 would be present in the structure. As illustrated in Figure 2b, this is indeed the case. In αLP_{pH5}, Ser195 occupies one conformation, which is essentially identical to that of $Ser195_b$ in αLP_{pH8}. αLP_{pH5} thereby confirms the hypothesis presented in Fuhrmann et al. (2004) and justifies the pK_a calculation performed in that study.

Modeling the Acylation Transition-State Intermediate. The αLP+boroVal(gol) structure is the best mimic of the tetrahedral intermediate (and therefore the TS and TS′ states) for the acylation reaction determined at subangstrom resolution. As is known, the MeOSuc-Ala-Ala-Pro-boroVal-OH inhibitor forms a stable tetrahedral adduct with Ser195, mimicking the transition-state intermediate of the deacylation reaction.[27] Surprisingly, we found a glycerol molecule covalently bound to the boronic acid at ~66% occupancy. The glycerol (GolP1′) occupies the S1′ pocket, mimicking the leaving group of the substrate and therefore providing a model of the tetrahedral intermediate for the acylation reaction (TI_1, and by extrapolation TS_1 and TS_1'). Components of the cryosolvent, these glycerol molecules entered the crystal lattice during the 45-s soak of the crystal prior to mounting and freezing under the cryostream. Attack of the boronate by a glycerol molecule has been observed previously in αLP crystals in cryosolvent soaks as short as 10 s (C. Fuhrmann and D. Agard, unpublished results), and reflects the reactivity of B−O to alkoxide groups. As expected, no hydrogen atom was observed on the B−O glycerol-bridging oxygen atom.

This 0.90 Å resolution structure, the first high-occupancy structure of a tetrahedral transition-state analogue solved at ultrahigh resolution, allows precise analysis of the geometry of the Ser195-inhibitor adduct. The geometry about the boron deviates only slightly from tetrahedral, with angles ranging from 104.7° to 112.3° (Table 3). Similar angles have been observed in structures of small-molecule compounds containing a central tetrahedral boron.[52,53] Therefore, we conclude that the ±5° deviation from perfect tetrahedral geometry in αLP+boroVal-(gol) is typical for a boronic acid in a heterogeneous environment such as the enzyme active site.

Peptide boronic acids have been used for decades to mimic the transition states for serine protease catalysis.[25,26] Although it was thought that the boronate oxygen in the oxyanion hole might be protonated, this ultrahigh-resolution crystal structure provides the first opportunity to observe this hydrogen atom (Figure 5). This, coupled with an excess of electron density at boroValP1 B, confirms that the negative charge in boronic acid transition-state analogues lies at least partially on the central boron. Although an excellent mimic of the tetrahedral intermediate, a boronic acid-bound protease differs from a true TI or TS state, primarily in the presence of negative charge on this central atom (rather than in the oxyanion hole) and the presence of a hydrogen atom on the oxyanion mimic. In this regard, some other transition-state analogues, including the trifluoromethyl ketones, are better mimics of the oxyanion interaction.

The TI_1 → EA transition: Movement of Ser195 Makes His57 Ring Flipping Unnecessary. It has been proposed that a "reaction-driven" ring flip of His57 is required for His57 to act both as a general base (in the formation of the first tetrahedral intermediate) and as an acid (donating a proton to the substrate leaving group to promote formation of the acylenzyme).[58] Because ultrahigh-resolution crystallographic data provides information regarding both the precise positioning of atoms and their directions of motion, the structures presented here—both of which mimic aspects of the TI and TS states—allow us to address this proposal.

In the first step of the mechanism (as modeled by αLP_{pH8}), neutral His57 is positioned to abstract the proton from Ser195 to facilitate nucleophilic attack on the scissile peptide carbon.[3] In the second step of the proteolytic reaction—breakdown of the tetrahedral intermediate into the acylenzyme complex—it was postulated that His57 ring flipping is necessary to preferentially promote protonation of the leaving group over back-protonation of Ser195.[58] However, we see that this is not the case. In our tetrahedral acylation transition-state analogue structure, αLP+boroVal(gol), His57 is optimally positioned to hydrogen bond (through Nϵ2) to the mimic of the leaving-group nitrogen, boroValP1 O2 (see later discussion). Additionally, Nϵ2 hydrogen bonding with Ser195 Oγ is disfavored (Table 2). Therefore, in our structure, flipping of His57 is unnecessary, as His57 is prepositioned to favor formation of the acylenzyme complex and to prevent back-protonation of Ser195.

Our experimental observation of His57 Nϵ2 hydrogen bonding to the leaving group in αLP+boroVal(gol) was predicted by QM/MM simulations performed by Ishida, et al. that calculated the energy profile for the acylation reaction.[59] These theoretical studies predicted that upon becoming tetrahedral, the scissile bond would rotate to a position where the −NH leaving group hydrogen-bonded to His57 Nϵ2. The calculations showed that rotation about the scissile bond would occur prior to formation of TI_1, leaving the substrate in a position favorable for formation of the EA complex.

Our data also suggests that movement of Ser195 occurs in the ES → TI_1 transition. In both αLP_{pH5} and αLP+boroVal-(gol), the side chain of Ser195 has moved such that Oγ is located ~1 Å away from its position in αLP_{pH8} where His57 is neutral. A comparison of αLP_{pH8} and αLP_{pH5} (Figure 2b) shows that this repositioning of Ser195 occurs upon protonation of His57 at Nϵ2. On the basis of these data, we hypothesize that abstraction of the Oγ proton and movement of the Ser195 side chain occur in concert during the ES → TI_1 transition.

(58) Ash, E. L.; Sudmeier, J. L.; Day, R. M.; Vincent, M.; Torchilin, E. V.; Haddad, K. C.; Bradshaw, E. M.; Sanford, D. G.; Bachovchin, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10371−10376.
(59) Ishida, T.; Kato, S. *J. Am. Chem. Soc.* **2003**, *125*, 12035−12048.

If His57 ring flipping were a mechanism required for transfer of the proton from Nϵ2 to the leaving group, this phenomenon would have to occur either in the TI_1 state or in the transition state for the $TI_1 \rightarrow$ EA reaction. To date, only indirect experimental evidence has been used to promote the His57 ring-flipping model.[58,60] Most recently, NMR studies of αLP published by Haddad et al. suggested that His57 is more mobile in apo-αLP at low pH than at high pH.[60] The authors stated that such movement could encompass motions as small as a wobble or as large as a ring-flip, making their observation consistent with the ring-flipping hypothesis. Our ultrahigh-resolution crystal structures of αLP at low pH (αLP_{pH5}) and as an analogue of TI_1 (αLP+boroVal(gol)) allow us to characterize the motion of His57 in the doubly protonated state, and to test for ring-flipping or wobbling. At subangstrom resolution, atoms around the His ring can be readily identified in σ_A-weighted $2F_o - F_c$ maps at high contour levels (data not shown), where nitrogen atoms appear distinctly larger than carbon atoms. More precisely, refined occupancies and *B* factors directly report on electron content. If two sets of superimposed conformations of His57 were present in the crystals (one flipped, one not), this would be readily apparent as deviations of atomic occupancies and *B* factors, which are not observed in either αLP_{pH5} or αLP+boroVal(gol).

Haddad et al. argue that the flipped conformation of His57 may not be observed in crystal structures, since only a small fraction of molecules are required to be "flipped" to make the mechanism effective.[60] Even in this case, if flipping is to occur, one would expect to observe atomic motions indicative of wobble about the Cβ−Cγ bond in the TS and TI_1 states. If wobbling occurred, each atom of the ring should display anisotropy perpendicular to the ring, and Cβ and Cγ should be more isotropic than Nδ1, Cδ2, Cϵ2, and Nϵ2. However, refined atomic anisotropic *B* factors for His57 in αLP_{pH5} and in αLP+boroVal(gol) suggest a rigid body motion of the entire residue diagonal to the plane of the ring (Figure 6). This rigid body motion of His57, rather than histidine flipping, may have been the motion previously observed by NMR.[60]

In summary, based on the αLP+boroVal(gol) complex structure, which is the best structure of the acylation tetrahedral intermediate mimic to date, there is no indication that His57 flipping is necessary or that it occurs in αLP. As is evident in both apo-αLP and αLP+boroVal(gol), when His57 Nϵ2 is protonated, Ser195 occupies a different conformation, forming at best a weak hydrogen bond to His57. Furthermore, as evidenced by αLP+boroVal(gol), the movement of Ser195 occurs prior to formation of TI_1. His57 is then already in an optimal position in TI_1 to donate a proton to the substrate leaving group, making a histidine-flipping mechanism unnecessary and even detrimental to progression of the reaction. These results are consistent with other ultrahigh-resolution crystal structures of serine proteases at low pH,[40] predictions by QM/MM modeling of the acylation reaction,[59] and NMR data previously used to promote the ring-flip hypothesis.[58,60]

His57 Nδ1−Asp102 Oδ2 : Debating the Low-Barrier Hydrogen Bond. According to the low-barrier hydrogen bond (LBHB) hypothesis for serine protease catalysis, a LBHB forms between His57 Nδ1 and Asp102 Oδ2 in states TS, TI, and TS′ during the acylation reaction (Figure 1a). In principle, this LBHB would stabilize the His-Asp pair, facilitating deprotonation of Ser195 Oγ by His57 and thereby catalyzing the first step of the acylation reaction.[5,8,9] Rather, in αLP+boroVal(gol) which mimics the acylation tetrahedral intermediate, we observe that Nδ1 and Oδ2 are separated by 2.734 ± 0.005 Å, a distance favorable for a normal H-bond but not compatible with a LBHB. Additionally, the Nδ1−H hydrogen atom is not displaced from its donor atom as required by the LBHB hypothesis (Table 2 and Figure 4).

In addition to inhibitor-bound transition-state analogues, samples of the apo-enzyme buffered at a pH well below the pK_a for His57 are also frequently studied as mimics of the transition state due to the positively charged nature of His57. In one such structure, that of subtilisin solved at pH 5.9 (0.78 Å resolution), the His Nδ1···Asp Oδ2 distance was 2.62 Å, and a peak for the hydrogen atom was observed nearly equidistant between Nδ1 and Oδ2; these results were highly suggestive of a LBHB.[21] However, in the structure of αLP at pH 5 presented here, the His57···Asp102 hydrogen bond has geometry similar to that of αLP+boroVal(gol). The donor−acceptor distance (2.755 ± 0.005 Å) and donor hydrogen-bond length (0.88 ± 0.05 Å) definitively show that the His57···Asp102 hydrogen bond is not a LBHB. Interpretation of the discrepancy between the low-pH αLP and subtilisin results is complicated by each crystal's active-site environment. In αLP_{pH5}, His57 may be drawn away from Asp102 by its electrostatic interactions with Sul202. In the subtilisin structure, the presence of peptide proteolysis products may have caused crowding in the active site. Another possibility is that a LBHB is present in subtilisin family members but not in chymotrypsin homologues.

Transition-state analogues are also imperfect mimics of the transition states. As described in the Introduction, commonly used peptidyl trifluoromethyl ketone (TFK) inhibitors contain a $-CF_3$ group which precludes proper formation of a hydrogen bond to His57 Nϵ2. One result of this might be the positioning of His57 away from $-CF_3$ and toward Asp102, falsely creating a structural environment favorable for formation of a LBHB. Conversely, the negative charge on the boron in peptidyl boronic acid inhibitors may attract positively charged His57, thereby increasing the His57···Asp102 distance. Additional studies, including theoretical analyses and structural studies of αLP in the presence of TFK inhibitors or in the absence of sulfate, should be performed to address these issues.

Despite these caveats, our results agree with observations from studies of other serine proteases. Importantly, Kossiakoff and Spencer's seminal neutron diffraction study of deuterated bovine trypsin bound to a transition-state analogue (monoisopropylphosphoryl group) also showed that the hydrogen atom is located on His57, as is evident in both of our structures.[61] Low- and moderate-resolution crystallographic studies have reported both short[22,62−64] and long[65] His-Asp distances; how-

(60) Haddad, K. C.; Sudmeier, J. L.; Bachovchin, D. A.; Bachovchin, W. W. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 1006−1011.
(61) Kossiakoff, A. A.; Spencer, S. A. *Nature* **1980**, *288*, 414−416.
(62) Rehse, P. H.; Steinmetzer, T.; Li, Y.; Konishi, Y.; Cygler, M. *Biochemistry* **1995**, *34*, 11537−11544.
(63) Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. *Biochemistry* **1994**, *33*, 7326−7336.
(64) Dixon, M. M.; Brennan, R. G.; Matthews, B. W. *Int. J. Biol. Macromol.* **1991**, *13*, 89−96.
(65) Lange, G.; Betzel, C.; Branner, S.; Wilson, K. S. *Eur. J. Biochem.* **1994**, *224*, 507−518.

ever, the coordinate errors at those resolutions are too large to unambiguously support or disprove the LBHB hypothesis. Recent crystal structures of serine proteases at atomic resolution suggest with higher confidence that this hydrogen bond is longer than 2.65 Å.[40,66,67] Additionally, preliminary high-resolution data for non-boronic acid tetrahedral mimic structures of two αLP homologues show His···Asp distances similar to those observed in αLP (Kelch, B. A. and Agard, D. A., personal communication). These data collectively argue against the hypothesis that the His57-Asp102 hydrogen bond is a LBHB.

Kirsch and colleagues tested the functional importance of the putative His-Asp LBHB in subtilisin BPN′ by mutating the catalytic Asp to a Cys.[68] As evidenced by the apparent loss at 11 °C of the downfield NMR signal assigned to His Nδ1−H, this mutant lacked the putative LBHB. On the basis of the observed 50-fold decrease in k_{cat}, the authors concluded that if it does exist, the LBHB contributes much less to transition-state stabilization (2.2 kcal/mol) than originally predicted (5−10 kcal/mol).[68] Indeed, arguments against the LBHB hypothesis state that the covalent nature of a LBHB would make it less effective in stabilizing the transition state than the electrostatic effects of a normal hydrogen bond.[69−73] Together with our observations and those of others, these results suggest that a LBHB is not necessary for efficient catalysis by serine proteases.

His57 Nϵ2 Forms a Short, Ionic Hydrogen Bond to the Leaving Group Mimic. In αLP+boroVal(gol), the protease-boronic acid adduct is capable of hydrogen bonding directly to His57, providing a mimic of the His57-leaving group interaction predicted for the acylation transition state. Although short hydrogen bonds are rare in this structure, His57 is positioned to form a short hydrogen bond to boroValP1 O2 (distance = 2.643 ± 0.004 Å) with ideal geometry (Table 2). Indeed, this is by far the shortest N···O hydrogen bond in the major conformation of αLP+boroVal(gol) (Figure 8). With a short donor−acceptor distance and a ^{1}H NMR chemical shift of 16.5 ppm,[74] this hydrogen bond fulfills two common diagnostics for a LBHB. However, unlike a LBHB (where the hydrogen atom is equidistant[4,5,16,72] or approximately equidistant[9] between donor and acceptor), the hydrogen in αLP+boroVal(gol) is clearly located on His57 Nϵ2 (Figure 4b).[75] We therefore call this hydrogen bond a short ionic hydrogen bond (SIHB), defined as a hydrogen bond fulfilling the donor−acceptor distance requirement of a LBHB but for which the hydrogen atom is localized on the donor atom. A description of how such a bond could exist, and the energetic implications of a SIHB at this site, are discussed below.

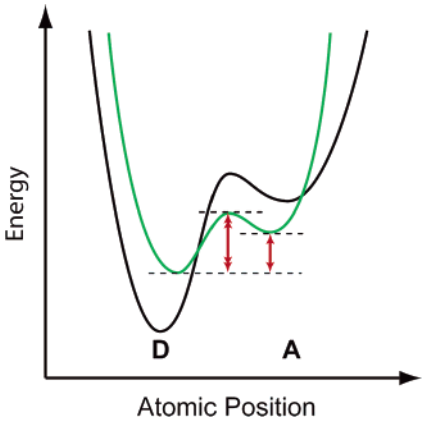

Figure 7. Energetic effects of reducing *d*(D···A). The free energy profile for proton transfer is illustrated for a hypothetical asymmetric hydrogen bond (black line). As the distance between donor (D) and acceptor (A) decreases (with new energy profile illustrated by the green line), the barrier for proton transfer decreases (as labeled with red double arrows), and the energy minima become closer in energy (red single arrows).

On Diagnosis of a LBHB. Traditionally, the major diagnostics for a LBHB have been an N···O distance of less than 2.65 Å and/or a ^{1}H NMR chemical shift downfield of 16 ppm.[5,8] Our observation of a SIHB between His57 Nϵ2 and the leaving group mimetic demonstrates experimentally that a short donor−acceptor distance is not a characteristic exclusive to LBHBs. This is evident when one considers the energy characteristics of these hydrogen bonds. Beginning with a conventional ionic hydrogen bond, as the distance between donor and acceptor decreases, the energetic landscape of proton transfer is changed in two ways (illustrated in Figure 7). First, the energy barrier to proton transfer decreases.[5,76] Second, the pK_a's of the donor/acceptor equalize, reflected in equalization of energy minima. It is important to note that the critical distance at which ΔpK_a equals zero (and a LBHB is formed)[16,72,77] is influenced by the environment surrounding the hydrogen bond.[16] This allows ionic hydrogen bonds to exist at short donor−acceptor distances and challenges the use of a threshold distance as a singular definition for identifying LBHBs.

The downfield chemical shift characteristic of a LBHB is thought to be caused by lengthening of the N−H bond, which decreases shielding of the proton by its donor atom.[9,78] Lengthening of the donor−hydrogen bond is frequently assumed to be concomitant with shortening of the donor−acceptor distance;[9,15,78] in support of this, ^{1}H chemical shift data has been shown to correlate with O···O[77−79] and N···O[80] distances. However, our results demonstrate that these correlations are not universal. The chemical shift measured for His57 Nδ1−H when αLP is bound to MeOSuc-Ala-Ala-Pro-boroVal is 16.0 ppm,[74] which would predict a His57···Asp102 distance of 2.64 ± 0.04 Å.[80,81] This is a significant underestimation of the true

(66) Fodor, K.; Harmat, V.; Neutze, R.; Szilagyi, L.; Graf, L.; Katona, G. *Biochemistry* **2006**, *45*, 2114−2121.
(67) Betzel, C.; Gourinath, S.; Kumar, P.; Kaur, P.; Perbandt, M.; Eschenburg, S.; Singh, T. P. *Biochemistry* **2001**, *40*, 3080−3088.
(68) Stratton, J. R.; Pelton, J. G.; Kirsch, J. F. *Biochemistry* **2001**, *40*, 10411−10416.
(69) Warshel, A.; Papazyan, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13665−13670.
(70) Shan, S. O.; Loh, S.; Herschlag, D. *Science* **1996**, *272*, 97−101.
(71) Shan, S. O.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14474−14479.
(72) Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Phys. Chem. A* **1997**, *101*, 3880−3886.
(73) Kim, Y.; Lim, S.; Kim, Y. *J. Phys. Chem. A* **1999**, *103*, 6632−6637.
(74) Bachovchin, W. W.; Wong, W. Y.; Farr-Jones, S.; Shenvi, A. B.; Kettner, C. A. *Biochemistry* **1988**, *27*, 7689−7697.
(75) There has been controversy regarding whether the hydrogen atom in a LBHB is located equidistant between donor and acceptor (in a symmetric hydrogen bond[16]) or nearly equidistant (in a slightly asymmetric hydrogen bond[9]). Regardless, the His57-boroVal hydrogen atom in our structure is located on His57 and is not at all displaced towards boroVal; therefore, the His57-boroVal hydrogen bond is not a LBHB by either definition.
(76) Cleland, W. W. *Biochemistry* **1992**, *31*, 317−319.
(77) Steiner, T. *Angew. Chem., Int. Ed.* **2002**, *41*, 49−76.
(78) Perrin, C. L.; Nielson, J. B. *Annu. Rev. Phys. Chem.* **1997**, *48*, 511−544.
(79) Zhao, Q.; Abeygunawardana, C.; Gittis, A. G.; Mildvan, A. S. *Biochemistry* **1997**, *36*, 14616−14626.
(80) Viragh, C.; Harris, T. K.; Reddy, P. M.; Massiah, M. A.; Mildvan, A. S.; Kovach, I. M. *Biochemistry* **2000**, *39*, 16200−16205.
(81) Mildvan, A. S.; Massiah, M. A.; Harris, T. K.; Marks, G. T.; Harrison, D. H. T.; Viragh, C.; Reddy, P. M.; Kovach, I. M. *J. Mol. Struct.* **2002**, *615*, 163−175.

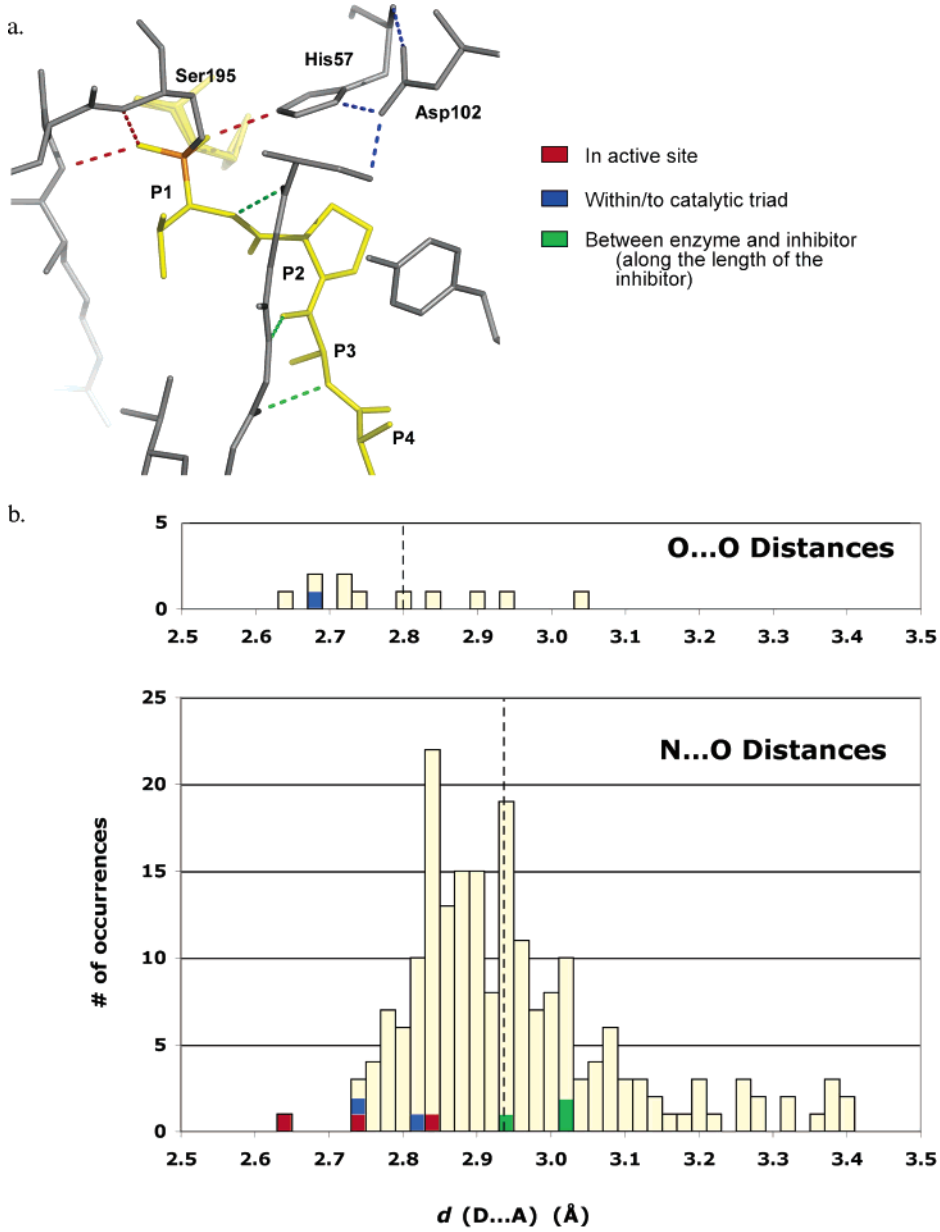

Figure 8. Distribution of hydrogen-bond lengths in αLP+boroVal(gol). The histogram (b) shows the distribution of donor−acceptor distances for all hydrogen bonds in αLP+boroVal(gol) between nitrogen and oxygen atoms (N···O) and two oxygen atoms (O···O), with the average D···A distance denoted by a dashed line. Hydrogen bonds formed between the enzyme and inhibitor in the active site (i.e., to the boronic acid moiety directly) are colored red, and along the length of the inhibitor backbone are colored green. Enzyme intramolecular hydrogen-bonding interactions found within or to the catalytic triad are colored in blue. The position of these hydrogen bonds within the structure are illustrated in (a), where the inhibitor is represented by yellow sticks, and the enzyme in gray. Hydrogen bonds identified in each of these three categories are indicated by footnote c in Table 2. Hydrogen bonds involving minor alternate conformations are not included.

His57···Asp102 distance (2.755 ± 0.005 Å). In keeping with our results, multiple theoretical studies have predicted that ionic as well as low-barrier hydrogen bonds can produce downfield proton chemical shifts.[4,69,82,83] Therefore, while a minimal ^{1}H chemical shift and long donor−acceptor distance rule out the existence of a LBHB, these measurements should not be used as exclusive tests to positively identify LBHB's. Rather, an accumulation of experimental data, including D/H fractionation factors, isotope effects on NMR and infrared spectra, hydrogen bond strength, pK_a matching, and the precise location of the hydrogen atom, is required to prove the existence of a LBHB.[9,15]

αLP+boroVal(gol) allows prediction of the energetic landscape for acylation. We have described two key observations based on the ultrahigh-resolution crystal structures of αLP. First, as discussed above, upon formation of the TI_1 state, Ser195 has already undergone a conformational change that prevents the $TI_1 \rightarrow$ ES back-reaction by weakening the His57-Ser195 hydrogen bond. Second, in αLP+boroVal(gol) His57 Nϵ2 forms a SIHB to the leaving group mimic (boroValP1 O2). This hydrogen bond has been alluded to in papers scattered throughout the history of serine proteases as playing an important role in catalysis.[13,74,84,85] Indeed, as described above (Figure 7), the short donor−acceptor distance observed here

(82) Guthrie, J. P. *Chem. Biol.* **1996**, *3*, 163−170.
(83) Shokhen, M.; Albeck, A. *Proteins* **2004**, *54*, 468−477.
(84) Robillard, G.; Shulman, R. G. *J. Mol. Biol.* **1974**, *86*, 541−558.

suggests that Hε2 has a reduced energetic barrier to proton transfer relative to a standard hydrogen bond. In the proteolytic mechanism, this would have the important effect of promoting protonation and release of the leaving group.

These two observations suggest that the acylation tetrahedral intermediate structure we have isolated is biased toward the acylenzyme complex (and significantly away from the Michaelis complex) along the reaction coordinate. Interestingly, two quantum mechanical studies of the acylation reaction have independently predicted that the lowest-energy tetrahedral intermediate conformation is that which has already undergone a conformational change to disrupt the His57 Nε2−Ser195 Oγ hydrogen bond and promote hydrogen bonding to the leaving group.[59,86] All of these data suggest that TI_1 more closely resembles TS_1' than TS_1 structurally. Therefore, by the Hammond postulate[20] we extrapolate that TI_1 is also closer in energy to TS_1' than to TS_1. The short length of the hydrogen bond between His57 Nε2 and the leaving group would further reduce the $TI_1 \rightarrow TS_1'$ energy barrier.

A Network of Short Hydrogen Bonds Stabilizes the Catalytic Triad. Interestingly, three of the four shortest N···O hydrogen bonds in the αLP+boroVal(gol) structure form interactions directly in the active site or within the catalytic triad (see Figure 8, Table 2). Figure 8 shows the distribution of hydrogen-bond distances within αLP+boroVal(gol), and the positions of active site, catalytic triad, and enzyme−inhibitor hydrogen bonds within this distribution. All hydrogen bonds formed between enzyme residues within or to the catalytic triad (blue squares), or within the active site to the inhibitor (red squares), are among the shortest 25% of hydrogen bonds found within the structure. The only exception is the His57 Nε2···Ser195 Oγ hydrogen bond, which we have argued is weaker in TI_1 to promote forward progression of the reaction (see Discussion above). Adding to this hydrogen-bond network is the conserved C−H···O hydrogen bond connecting the backbone of Ser214 to His57. In contrast to hydrogen bonds found at or near the active site, the three hydrogen bonds that form between the enzyme and the inhibitor backbone are weaker hydrogen bonds (d(D···A) = 2.93−3.11 Å). We propose that the enzyme has evolved to stabilize the tetrahedral intermediate and transition states by creating a preorganized active site that is optimized for transition-state stabilization. Residues are optimally positioned by a collection of relatively strong ionic hydrogen bonds in the vicinity of the active site. Furthermore, the fact that hydrogen bonds involved in positioning the substrate in the enzyme binding pocket are of average donor−acceptor distances suggests that the enzyme's evolution favored maximizing k_{cat} rather than decreasing K_m. This phenomenon has been inferred from kinetic data demonstrating the "k_{cat} effect": that shortening peptide substrates from three to one amino acids long decreases k_{cat} 2500-fold, but K_m is only increased 3-fold (R. Bone and D. Agard, unpublished results; similar effects observed for other proteases[87,88]). Optimizing catalysis over substrate binding would benefit the enzyme in the final steps of proteolysis by facilitating product release. This may be a general strategy used by serine proteases (and enzymes in general) to optimize both catalysis and enzyme turnover.

Conclusion

We have solved two ultrahigh-resolution crystal structures of α-lytic protease that mimic aspects of the catalytic transition states: αLP at pH 5 and αLP bound to a peptidyl boronic acid (αLP+boroVal(gol)). The unexpected covalent binding of a glycerol molecule to the tetrahedral boronate in αLP+boroVal(gol) provided the best current model of the tetrahedral intermediate for acylation, allowing us to address questions of catalysis that have existed in the serine protease field for decades. Our crystallographic structures show that histidine flipping is not required for forward progression through the reaction mechanism and suggest, rather, that movement of Ser195 prior to the formation of TI_1 destabilizes hydrogen bonding to His57 and thereby prevents the $TI_1 \rightarrow$ ES back-reaction. On the basis of precise atomic positions obtained at subangstrom resolution, we have concluded that a low-barrier hydrogen bond between Asp102 and His57 does not exist in the acylation transition state or intermediate. Although a His-Asp LBHB may have been observed in subtilisin,[21] it is not required as a general mechanism for serine proteases. Instead, transition-state stabilization by chymotrypsin-like serine proteases is likely achieved through a *network* of optimized hydrogen bonds that position the catalytic triad and stabilize the Ser195-substrate tetrahedral adduct. In particular, we identified a short ionic hydrogen bond (SIHB) between His57 and the amide of the leaving group that may play the primary role in catalyzing the second step of the acylation reaction. Combined, our observations contradict the common use of a downfield ^{1}H chemical shift or short donor-to-acceptor distance as stand-alone litmus tests for a LBHB, and provide experimental evidence for a preorganized collection of relatively strong ionic hydrogen bonds in the vicinity of the active site that likely provide the driving force for proteolytic catalysis.

Materials and Methods

Crystal Preparation. αLP protein was expressed and purified from its native bacterium, *Lysobacter enzymogenes*.[45] As for previous experiments published by these authors, αLP was never lyophilized during the process of protein purification. The crystal used in the determination of αLP+boroVal(gol) was grown by methods described previously, using a precipitant solution composed of 1.3 M Li_2SO_4 and 20 mM Tris, buffered to pH 8.[3] Low-pH crystals were grown under similar conditions, but equilibrated with precipitant solution at pH 4.25. These crystals grew much more slowly than those equilibrated at pH 8, but retained the ability to diffract to subangstrom resolution.

The boronic acid inhibitor (MeOSuc-Ala-Ala-Pro-boroVal-OH) was prepared by dissolving MeOSuc-Ala-Ala-Pro-(D,L)-boroVal-OH in 20 mM Tris-sulfate (pH 7.5) to a final concentration of 0.14 M. One microliter of this solution was added to a drop containing αLP crystals. The drop equilibrated for 17.5 h prior to data collection, allowing formation of the enzyme−inhibitor complex. Although a racemic mixture of the D and L forms of the inhibitor was used, only the L form was visible in electron density (likely due to the preferential binding of this form by at least 100-fold[25]).

Data Collection: αLP at pH 5.1. Data were collected at the Advanced Light Source (HHMI beamline 8.2.1) on a single crystal of approximate dimensions 0.2 mm × 0.2 mm × 0.4 mm. Prior to mounting on the cryostream, the crystal was whisked through a cryoprotectant solution composed of precipitant solution (1.3 M Li_2-

(85) Steitz, T. A.; Shulman, R. G. *Annu. Rev. Biophys. Bioeng.* **1982**, *11*, 419−444.
(86) Daggett, V.; Schroder, S.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 8926−8935.
(87) Case, A.; Stein, R. L. *Biochemistry* **2003**, *42*, 3335−3348.
(88) Thompson, R. C.; Blout, E. R. *Biochemistry* **1973**, *12*, 57−65.

SO_4 and 20 mM Tris) and 15% glycerol. At the time of data collection, the cryoprotectant was measured to be pH 5.1.

As described previously,[3] ultrahigh resolution data collection requires a short detector-to-crystal distance coupled with a short wavelength (to ensure that the outer high-resolution reflections will be captured on the detector) and long exposure times (in order to detect these weakly diffracted reflections). Because low-resolution reflections tend to become oversaturated during these long exposure times, additional data sets must be collected to optimally capture these "low-resolution" reflections. Here, we collected high-resolution data at two positions on the crystal, and later merged these data to obtain a complete "high-resolution" data set (to 0.86 Å resolution). We then collected two complete data sets using shorter exposure times, providing a "medium resolution" set (to 1.15 Å resolution) and a "low resolution" set (to 1.24 Å resolution). Because the crystal continued to display strong diffraction potential, we collected an additional, "ultrahigh resolution," data set with the detector offset along 2-θ, providing data to 0.78 Å resolution. All four data sets were integrated using Denzo[89] and scaled and merged in Scalepack,[44] providing a single data set that was 100% complete and highly redundant (see Table 1). The final resolution was determined to be 0.82 Å to ensure a median signal-to-noise ratio greater than two and an R_{merge} less than 50% for the highest-resolution data.

Data Collection: αLP bound to mAAP-boroVal···glycerol. Data were collected at the Advanced Light Source (HHMI beamline 8.2.2), using X-rays of wavelength 0.75 Å, on a single hexagonal crystal of dimensions ~0.3 mm × 0.3 mm × 0.5 mm. Prior to mounting on the goniometer, the crystal was equilibrated in a cryoprotectant containing 15% glycerol for 45 s. The crystal was then frozen directly in the stream of gaseous nitrogen, which was maintained at a temperature of 100 K throughout data collection.

In previous studies of αLP bound to peptidyl boronic acid inhibitors, we observed conformational changes at the boron-Ser195 adduct due to damage caused by the long X-ray exposures necessary to obtain ultrahigh-resolution data.[45] Therefore, in this experiment, we took special care to minimize the radiation exposure at any one position of the crystal during collection of high-resolution data. Data were collected on a single crystal of dimensions ~0.3 mm × 0.3 mm × 0.5 mm. A 100-μm collimator was used, reducing the size of the beam relative to the crystal and thereby allowing more translations along the crystal. On the first position of the crystal, a complete data set was collected at moderate resolution (1.2 Å). This data set represented data with no detectable effects from radiation damage (abbreviated "no damage"). At this crystal position, a second, higher-resolution, data set was also collected (~1.04 Å resolution). For collection of ultrahigh resolution data, the crystal was translated nine times along the length or width of the crystal, with only 15−17 of the total ~100 frames of high-resolution data collected at any one position. To verify that no radiation damage had occurred during high-resolution data collection, a moderate-resolution data set (to ~1.6 Å resolution) was collected prior to and after each high-resolution data set. Therefore, at any given position on the crystal, three data sets were collected: (1) a complete moderate-resolution data set ("pre-high"), (2) ~12° of high-resolution data (~20% of a complete data set), and (3) a complete moderate-resolution data set ("post-high"). Each high-resolution data set overlapped by at least 1 frame of data (0.7°) with the previous high-resolution data set, allowing accurate merging of the data sets. $F_{o(post-high)} - F_{o(no\ damage)}$ maps were calculated for each position of the crystal to test for detectable damage, such as reduction of disulfide bonds, which may have occurred during the high-resolution data collection (data not shown). As another test, $F_{o(post-high)} - F_{o(pre-high)}$ maps were also calculated. Finally, an αLP model was refined using a 1.6 Å resolution data set collected "post-high" at one position (data not shown). In this model, σ_A-weighted $F_o - F_c$ maps also showed no indication of X-ray radiation damage. In these tests, reduction of disulfide bonds Cys42/Cys58 and Cys189/Cys220A was used as an indicator for X-ray radiation damage. Even when such damage is present in as few as 10% of the molecules, this is clearly visible in $F_o - F_c$ difference maps at high contour levels (3σ), providing an easy test for X-ray radiation damage at the molecular level.

Data sets were indexed using HKL2000.[89] High-resolution data sets were scaled and merged with the two moderate-resolution data sets (from the first position on the crystal); scaling was performed using Scalepack.[44] The resulting data set was limited to a resolution of 0.90 Å (see Table 3.1).

Model Refinement. Both αLP_{pH5} and αLP+boroVal(gol) were refined in SHELXL-97[48] using strategies similar to those used to solve the 0.83 Å structure of αLP_{pH8}.[3] Throughout the refinement of each model, 5% of the data were set aside as a test set for calculation of R_{free} as validation of the model.[90]

Model Refinement of αLP at pH 5.1. Refinement was begun using the previously determined structure of αLP at pH 8 (1SSX.pdb, 0.83 Å resolution).[3] To minimize model bias, the starting model was stripped of disordered waters (i.e., if partially occupied or with an isotropic temperature factor greater or equal to 50) and of minor alternate conformations of amino acids (with occupancies of major conformations returned to 100%). Anisotropic *B* factors and hydrogen atoms were not included in the starting model, and sulfate and glycerol molecules were removed. Following rigid body refinement and minimization, the starting model had an R/R_{free} of 19.6%/21.6%. During the early stages of refinement, major changes were made to the model while minimizing against data to 1.2 Å resolution, including editing of waters, the addition of three sulfate ions and one glycerol molecule, and inclusion of nine amino acid alternate conformations (R/R_{free} = 16.7%/18.2%). At this point, data was added to 0.9 Å resolution, and all atoms modeled with anisotropic temperature factors, resulting in a reduction of R/R_{free} to 12.9%/14.3%. Addition of 32 whole and 85 partially occupied waters, one glycerol molecule, and six amino acid alternate conformations reduced R/R_{free} to 11.2%/12.6%. Hydrogen atoms at this stage were included as "riding hydrogens," with fixed geometry as defined by SHELXL-97,[48] and the model minimized against data to 0.82 Å resolution (R/R_{free} = 10.9%/12.1%). Because positive difference electron density maps showed unique positions for hydrogen atoms on methyl, primary amine, and hydroxyl groups in residues with a single conformation, these hydrogens were added to the model using the AFIX 137 and AFIX 147 instructions, which refine torsion angles to position hydrogen atoms in maximum electron density.

To ensure no model bias for analysis of difference electron density maps, hydrogen atoms were not added to key atoms in the catalytic triad, Ser214, or oxyanion hole until the final stages of model building. After most of the model was complete (R/R_{free} = 8.35%/9.64%), hydrogens were added to His57, Ser195 Oγ and N, Ser214 Oγ, and Gly193 N as riding hydrogens. After five cycles of minimization, restraints were removed from these hydrogen atoms and their positions allowed to refine throughout the rest of the refinement process. At this stage, restraints were also removed from all non-hydrogen atoms in residues of a single conformation. This resulted in a decrease in R and R_{free} of 0.016% and 0.002%, respectively.

Two of the active-site hydrogen atoms refined to positions away from their peak $F_o - F_c$ density. In the oxyanion hole, Ser195 N−H, which does not participate in a hydrogen bond in the free enzyme, refined to a position just 0.65 Å from its parent atom. This was likely an artifact due to the strong distribution of electron density at Ser195 N. Therefore, in the final model of aLP_{pH5} the Ser195 N−H hydrogen atom was constrained to the canonical position for amide hydrogens. The $F_o - F_c$ peak corresponding to Ser195 Oγ−H was anomalously shaped (likely due to either conformational heterogeneity or noise), resulting in refinement of the Oγ−H hydrogen atom away from the

(89) Otwinowski, Z.; Minor, W. In *Methods Enzymol.* **1997**, *276*, 307−326.

(90) Brunger, A. T. *Nature* **1992**, *355*, 472−475.

primary $F_o - F_c$ peak. Therefore, the hydrogen was constrained to the canonical Oγ−H bond length (0.84 Å) and torsion angle (109.5°).[48] The position of the constrained hydrogen atom was ∼0.2 Å from the $F_o - F_c$ peak, which is displaced toward Sul202.

The final model (R/R_{free} = 8.06%/ 9.27% for all data; 7.17%/8.37% for data with $F_o > 4\sigma(F_o)$) contained 37 residues with alternate conformations (including four residues with three or more conformations), 148 full-occupancy waters, 435 partially occupied waters, 11 sulfate ions, and 2 glycerol molecules. The coordinates are available from the Protein Data Bank (2H5C.pdb).

$F_{o,pH5} - F_{o,pH8}$ difference maps were calculated to compare αLP_{pH5} to αLP_{pH8} (1SSX.pdb[3]). Using CNS,[91] data from each structure were scaled and then maps calculated using phases from 1SSX.pdb.

Model Refinement of αLP bound to mAAP-boroVal···glycerol. The previously solved structure of αLP bound to methoxysuccinyl-Ala-Ala-Pro-boroAla (with glycerol bound; unpublished data) was used as a starting model for refinement. To minimize model bias, all hydrogen atoms, 114 water molecules (partially occupied or having temperature factor greater than 50), one partially occupied glycerol molecule, and one partially occupied sulfate ion were removed prior to refinement. Twenty-two alternate conformations of side chains were also removed, and the occupancy of major conformations reset to unity. The inhibitor was not included in this initial model.

Refinement of the model was performed entirely in ShelxL-97, following strategies described previously for the 0.83 Å resolution structure of αLP (pH 8).[3] Rigid body refinement and subsequent minimization of the initial model showed an excellent fit to the data, with R/R_{free} of 18.4%/20.4%. Early in the refinement, data was restricted to 1.2 Å resolution to allow efficient modeling of waters and alternate conformations. Temperature factors were described using an isotropic model initially. Both σ_A-weighted $2F_o - F_c$ and difference electron density maps clearly indicated that the inhibitor was covalently bound to the protein at Ser195 Oγ. Electron density was visible as punctate atoms along the length of the inhibitor from boroVal P1 to the methoxysuccinyl-Ala peptide bond. Therefore, the inhibitor was modeled as Ala-Ala-Pro-boroVal, with a resulting R/R_{free} of 16.4%/18.2%. Next, anisotropic B factors were added to the model for all atoms, and minimized against data to 1.0 Å resolution (R/R_{free} = 12.4%/14.1%). This was followed by additional modeling of waters and side chain alternate conformations (R/R_{free} = 11.7%/13.3%). Difference electron density corresponding to hydrogen atoms was visible for the majority of hydrogens in the structure. Hydrogen atoms were added to the model as "riding hydrogens," meaning that their position was defined by geometry and not allowed to refine. Hydrogen atoms were not added to methyl, primary amine, or hydroxyl groups, or to key atoms in the active site (including all atoms of the inhibitor and His57, atoms in the oxyanion hole (Gly193 N and Ser 195 N), Ser214 Oγ, and Gly216 N). At this stage, R/R_{free} were 11.2%/12.5%. Positive difference electron density (contoured to 3.0σ) suggested the presence of a glycerol molecule bound to the boronic acid of the inhibitor. This glycerol was modeled as a secondary adduct to boroVal P1 O2. At later stages of refinement, it became evident that this glycerol molecule occupied two other conformations, both secondary adducts to boroVal P1 O2, with a total occupancy of 66%. Electron density extending off the N-terminal end of the inhibitor was modeled with the full-length methoxysuccinyl group (R/R_{free} = 9.6%/10.7%). This group refined to an occupancy of 43%, suggesting that other conformations of the methoxysuccinate are disordered. At this time, riding hydrogens were added to methyl, primary amine, and hydroxyl groups (using the AFIX 137 and 147 commands), and to atoms in the inhibitor. In later stages of refinement, it became apparent from positive difference maps that residues MeOSuc P5 through Ala P3 were present in a second conformation, which refined to an occupancy of 22%. Finally, all geometric restraints were removed from residues with single conformations, allowing atomic positions to refine without bias from canonical geometries. "Unrestrained" minimization resulted in a decrease in R by 0.037%, and of R_{free} by 0.049%.

In the last step of refinement, hydrogen atoms were added to key atoms in the active site, and their positions allowed to refine freely. The final model contained 21 residues with alternate conformations, 160 fully- and 389 partially occupied waters, eight sulfate ions, and one glycerol molecule. The model is among the most accurate solved to date, with an R of 8.07% and R_{free} of 9.15% (R/R_{free} = 7.24%/8.25% for data with $F_o > 4\sigma(F_o)$). The coordinates are available from the Protein Data Bank (2H5D.pdb).

$F_{o,\alpha LP+bVal(gol)} - F_{o,1SSX.pdb}$ electron density maps were calculated using methods similar to those for αLP_{pH5} to directly observe differences between the structures of αLP+boroVal(gol) and the apo-enzyme at pH 8.[3]

All models were aligned using LSQMAN,[92] and reported RMSDs were calculated for main chain atoms (Cα, C, N) except where otherwise noted. Structural figures were made using Pymol[93] unless otherwise indicated; chemical schematics were drawn in ChemDraw.

Determination of Structure Precision. One advantage of ultrahigh resolution crystallography is that estimated standard errors for atomic positions and other parameters (including atom−atom distances and angles) can be calculated directly via inversion of the least-squares matrix (coordinates only) in the final stage of refinement.[18,48,94−96] Here, blocked least-squares refinement was performed on the final model for each structure. In this stage, all non-hydrogen atoms were unrestrained except those with alternate conformations (see above). Additionally, selected hydrogen atoms were also unrestrained, as listed in Table 2. For αLP+boroVal(gol), the structure was divided into three blocks: residues 1−100, 98−211, and 209−730. In αLP_{pH5}, the three blocks were defined as residues 1−100, 98−212, and 207−742. Estimated standard deviations listed in Table 2 for atom−atom distances and hydrogen bond angles were output by SHELXL-97 using the RTAB/HTAB commands. These estimated errors are small, as expected for structures of ultrahigh resolution, high completeness, and low R factors.[94,97] Estimated errors for donor-hydrogen bond lengths are larger than for other atom−atom distances, which is unsurprising due to the weaker scattering power of hydrogen (and its correspondingly higher positional uncertainty).

Acknowledgment. Thanks are given to Dr. C. Kettner for providing the boronic acid inhibitor, Caleb Bashor for aiding in protein expression and purification for growth of low-pH crystals, and Dr. R. C. Rizzo for access to the Cambridge Structural Database. C.N.F. thanks Dr. M. Garcia-Viloca for originally spurring her interest in the LBHB debate, and Drs. M. Jacobson, A. Mildvan, J. Kirsch, D. Truhlar, and J. Gao for helpful discussions of hydrogen-bond theory. We greatly appreciate Dr. T. Scanlan and reviewers for critical reading of the manuscript. C.N.F. is a National Science Foundation Predoctoral Scholar, and M.D.D., a Howard Hughes Medical Institute Predoctoral Fellow. Portions of this research were carried out at the Advanced Light Source, which is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence

(91) Brünger, 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.; Warren, G. L. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *D54*(Pt 5), 905−921.

(92) Kleywegt, G. J.; Jones, T. A. *CCP4/ESF-EACBM Newsletter on Protein Crystallography* **1994**, *31*, 9−14.

(93) Delano, W. L.; DeLano Scientific: San Carlos, CA, 2002.

(94) Cruickshank, D. W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*(Pt 3), 583−601.

(95) Cruickshank, D. W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1108.

(96) Sands, D. E. *Acta Crystallogr.* **1966**, *21*, 868−872.

(97) Dauter, Z.; Lamzin, V. S.; Wilson, K. S. *Curr. Opin. Struct. Biol.* **1995**, *5*, 784−790.

Berkeley National Laboratory. This work was funded through the Howard Hughes Medical Institute.

Note Added in Proof. Results from a study by Radisky et al. (PNAS, Vol. 113, 6835−6840) published during production of this article are consistent with the major conclusions of this paper.

Supporting Information Available: Supporting Information is available for review purposes upon request. See any current masthead page for ordering information and Web access instructions.

JA057721O