Neutron Structure of Subtilisin BPN′: Effects of Chemical Environment on Hydrogen-Bonding Geometries and the Pattern of Hydrogen–Deuterium Exchange in Secondary Structure Elements†

Anthony A. Kossiakoff,*,‡,§ Mark Ultsch,‡ Steven White,∥,⊥ and Charles Eigenbrot§,#

Department of Protein Engineering, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, and Department of Biology, Brookhaven National Laboratory, Upton, New York 11973

Received July 23, 1990; Revised Manuscript Received October 19, 1990

ABSTRACT: The neutron structure of subtilisin BPN′ has been refined and analyzed at 2.0-Å resolution. The structure studied was a mutant variant of subtilisin, Met222 → Gln, and was used because large, uninhibited crystals could be grown, which was not the case for the native molecule. Comparison of the structure with that of the native molecule indicated that the two structures are essentially the same. Using the capability of the neutron method to locate hydrogen and deuterium atoms, the protonation states of the six histidine residues were assigned. The active site histidine, His64, was found to be neutral at the pH of the analysis (pH 6.1). This group has an unexpectedly low pK_a compared to assignments made by other techniques. The altered pK_a of the group could result from electrostatic effects of other molecules in the crystal lattice. The dihedral conformations of a majority of the hydroxyl rotors were assigned. The preferred orientation was trans (180°) with the other two low-energy conformers (60°, 300°) about equally populated. For the serines, about 21% of the hydroxyls act exclusively as H-bond acceptors and 37% as H-bond donors, and in 42% the group functions as both. The experimentally observed dihedral conformations were compared to predicted conformations based on calculated energy criteria and showed a strong correspondence. Deviation from low-energy states could usually be explained by local electrostatic effects. The hydrogen exchange pattern of subtilisin identified the β-sheet and α-helix secondary structure elements to be the most resistant to exchange. Fifty-five percent of the peptide amide hydrogens were fully exchanged, 15% unexchanged, and 30% partially exchanged. The largest concentration of unexchanged sites was in the seven-stranded parallel β-sheet, in which there were 11 fully protected groups. Little correlation was found between H-bond length and angle and a peptide group's susceptibility toward exchange. Of the five α-helices the most protected from exchange is the one defined by residues 224–236. The pattern of exchange identifies regions in this helix where the H-bonding regularity is disrupted.

The physical structure as well as the chemical activity of proteins is dominated by the behavior of hydrogen atoms. In the absence of being able to assign hydrogen positions directly, it is difficult to define accurately some of the important chemical interactions taking place among groups in the protein, especially those involving H-bonding. X-ray diffraction has the basic limitation that hydrogen atoms are not observable and their positions must be inferred from stereochemistry on the basis of the heavy atoms to which they are bonded. Consequently, unusual or unexpected stereochemistry for the hydrogens will be overlooked, and of course, this class of hydrogen is the one most important to providing new structural insights.

In contrast to its X-ray counterpart, neutron diffraction has the capability of directly locating hydrogen atoms due to their relatively large neutron scattering cross section. Additionally, the technique can readily distinguish hydrogen from deuterium because of the difference in sign of their scattering lengths. These attributes enable neutron diffraction to provide unique biophysical data and to address structural issues not attainable by other techniques. The application of the technique has mainly focused on four research areas: (1) protein function and mechanism, (2) protein dynamics and flexibility employing the hydrogen exchange (H/D) technique, (3) stereochemistry of hydrogen-bonding interactions, and (4) protein solvation. The neutron analysis of subtilisin presented here describes structural data bearing on issues relating to the inherent flexibility of secondary structure elements and H-bonding stereochemistry. Subtilisin presents an extremely good model system for these types of study because its structure, kinetic properties, and stability have been characterized in great detail (Wright et al., 1969; Drenth et al., 1972; Markland & Smith, 1971; Polgar & Bender 1969; Kraut, 1972; Wells et al., 1987; Estell et al., 1987; Pantoliano et al., 1989). Thus, the findings of the neutron structure can draw on a variety of other results to facilitate interpretation. It is also the largest molecule to be studied by neutrons to date and contains a diverse set of secondary and tertiary structure types, making it possible to analyze in one protein structure the comparative breathing characteristics of segments of sheets, helices, loops, and turns.

It was desired to study subtilisin under conditions where there was no inhibitor in the active site to provide information about the catalytically important residues in the "free" enzyme state. However, growing crystals of the native enzyme required

† The work was supported by NIH Grant GM33571 and Genentech, Inc.

* Correspondence should be addressed to this author at Genentech, Inc.

‡ Genentech, Inc.

§ University of California.

∥ Brookhaven National Laboratory.

⊥ Present address: Department of Microbiology, Duke University Medical Center, Durham, NC 27710.

Present address: Genentech, Inc., 460 Point San Bruno Blvd., South San Francisco, CA 94080.

0006-2960/91/0430-1211$02.50/0 © 1991 American Chemical Society

the addition of an active site directed inhibitor to eliminate the problem of autolysis (Bott et al., 1988). It was determined that several active variant subtilisin molecules could be crystallized under conditions where autolysis was virtually eliminated and thus did not require added inhibitor. The subtilisin structure reported here is that of a variant where the methonine residue at position 222 was replaced by a glutamine. This M222Q variant was grown to the size required for neutron diffraction (=3 mm^3) and showed few structural differences from the native molecule. Consequently, it is believed that conclusions drawn from the M222Q analysis can be extrapolated to the native enzyme.

In this report two stereochemical phenomena are analyzed that can only be addressed by direct knowledge of hydrogen and deuterium atom positions: the H-bonding geometries of hydroxyl groups and the H/D exchange characteristics of the protein's secondary structure elements. Hydroxyl rotors ($O\gamma$–$D\gamma$) possess rotational freedom to align optimally within their electrostatic environment. Thus, their character provides data as to how the energetic components are partitioned between nonbonded and torsional effects and under what circumstances each of these effects dominate the observed geometry. It was found that the observed conformations coincide closely with those calculated by bonded and nonbonded energy criteria. This finding provides experimental validation for the accuracy of this type of calculation applied to short-range interactions.

The second area emphasized here is the analysis of how the hydrogen exchange (H/D) properties differ among secondary structure types when correlated with solvent accessibility, H-bonding, and packing effects. The neutron structure of trypsin showed that antiparallel β-sheet structures afford significant protection toward exchange (Kossiakoff, 1982). Subtilisin has an extensive parallel β-sheet structure that is surrounded by several α-helices. This sheet structure was found to be more susceptible to exchange than was the case for the antiparallel sheet in trypsin. Because there are distinct differences in the H-bonding and conformational regularity between parallel and antiparallel sheets (Salemme & Weatherford, 1981), it was of interest to identify the important elements that differentiate the H/D characteristics of these extensive H-bonding structures. The overall H/D exchange pattern of the groups contained in secondary structure interactions suggests that the breathing motions characteristic of the molecule involve relatively small segments of the structures. Within the H-bonding structures, departure from the regularity of the H-bond interactions is usually marked by a significant increase in exchange.

Experimental Procedures

Data Collection. The M222Q crystals of subtilisin used for data collection were grown from a solution of 25% $(NH_4)_2SO_4$, 10 mM $CaCl_2$, and 10 mg/mL protein at pH 5.6 by vapor diffusion. The crystals were of space group $P2_12_12_1$ with cell parameters a = 39.3 Å, b = 72.8 Å, and c = 75.2 Å. It was not possible to grow crystals to the size required for a neutron analysis (>1 mm^3) without using macroseeding methods (Thaller et al., 1985). Crystals that ceased growing were washed in a mother liquor solution and transferred to wells with solutions containing protein concentrations slightly lower than that which would initiate crystallization. After several days of equilibration, crystal growth again commenced and would usually continue for several weeks. Normally, it took four macroseeding steps to produce crystals greater than 1 mm^3. The crystal used in this analysis was of dimensions 1.2 × 0.9 × 2.5 mm^3, giving an approximate volume of 2.7 mm^3.

The crystal was soaked for approximately 1 month in a solution containing a D_2O–$(ND_4)_2SO_4$ solution at pH 6.1 (corrected for deuterium isotope effects, ΔpH = 0.4). The crystal was placed on a four-circle goniostat with the b^* axis parallel to the instrument ϕ axis.

Data were collected at the Brookhaven National Laboratory High Flux Beam Reactor (HFBR). The neutron beam was monochromatized and focused by using a bent pyrolytic graphite crystal. The beam flux produced at the crystal was approximately 5 × 10^7 neutrons cm^{-2} s^{-1}. A wavelength of 1.61 Å was used for most of the data. The beam was passed through a $\lambda/2$ filter to eliminate the diffraction artifacts due to the $\lambda/2$ wavelength contamination in the direct beam.

Diffracted intensities were collected on a two-dimensional position-sensitive detector (Schoenborn, 1984). The active region of this detector was 17 × 20 cm, and the resolution was 1 mm in the horizontal and 1.7 mm in the vertical direction (560 by 280 data channels). The crystal to detector distance was normally 610 cm (for about 20% of the data the detector was set at 800 cm), providing a range of about 17° in 2θ for each detector setting. The data were collected in two crystal orientations by using normal beam rotation geometry. At χ = 0° the crystal was rotated around the ϕ axis (crystal b*) in steps of 0.05° over a range of 90° ± 2θ. To obtain data in the so-called blind region, the crystal was set at χ = 90° and rotated around the ω axis (crystal c*). Because of instrument limitations the scan range was restricted to ±25°. To collect all the available data, 18 detector settings were required.

Data Reduction. Characteristically, neutron diffraction data from protein crystals have poor signal to noise ratios because of the low beam flux (about 6 orders of magnitude less than can be produced by an X-ray rotating anode generator) coupled with the inherently high backgrounds that result from incoherent, inelastic scattering of neutrons by hydrogen atoms in the sample (Bacon, 1975). In order to improve the accuracy of the data, a method was used employing a peak fitting algorithm (Spencer & Kossiakoff, 1980). This algorithm defines the peak boundary accurately to eliminate extraneous background and improve the statistical quality of the data.

During data reduction a systematic error in the data from four consecutive detector settings was detected. Although the source of the error was not readily apparent, the reflection profiles had characteristics similar to those that might be expected if, through crystal misalignment, the percentage of the crystal being bathed in the beam changed as a function of rotation angle. The effect to the data was determined to be a systematic function that was correctable by using a local scaling procedure that scaled reflection intensities in bins using scale factors determined from ratios between the measured data and structure factor amplitudes calculated from a subtilisin model containing the heavy atoms and sterically constrained hydrogens. Such a procedure helps to eliminate data artifacts without introducing bias into the phasing model. A total of 10380 unique intensities >3σ were obtained. The completeness of the data was 68% to 10.0–2.0-Å resolution. The shell of data from 2.15- to 2.00-Å resolution was about 32% complete. The overall R_{merge} for the data was 10% based on intensity.

Initial Phasing Model. The initial phasing model was calculated by applying the appropriate scattering lengths (Table I) to the refined X-ray coordinates. The coordinate set used was that obtained by Bott et al. (1988) for the structure of subtilisin BPN′ refined at 1.8 Å to a residual of R = 0.152. Hydrogen atoms that are stereochemically con-

Table I: Neutron Scattering Lengths[a]

atom type	scattering length (F) ($\times 10^{-13}$ cm)	atom type	scattering length (F) ($\times 10^{-13}$ cm)
H	−3.74	O	5.80
D	6.67	S	2.80
C	6.65	Ca	4.70
N	9.40		

[a] From Bacon (1975).

Table II: Refinement of M222Q Subtilisin at 2.0 Å

model	source	statistics
initial model	subtilisin BPN′ X-ray model;[a] hydrogens placed by stereochemistry; no deuterium or D_2O	$R = 0.255$[b]
refinement		
function	approach	statistics
atom shifts	difference maps (curvature/gradient)[c]	
reidealization of coordinates	model building and energy refinement[d]	
current model	five cycles of refinement; 57 D_2O included	$R = 0.192$
	rms deviations from ideal bond lengths and angles	
bond lengths (Å)	0.013	
bond angles (deg)	3.5	
torsion angles (deg)	4.2	

[a] X-ray model refined at 1.8 Å resolution ($R = 0.152$) (Bott, personal communication). [b] $R = \sum|F_o - F_c|/\sum|F_o|$. [c] Freer et al. (1975) and Chambers and Stroud (1977). [d] Hermans and McQueen (1974).

strained (i.e., methylene hydrogens, ring hydrogens, etc.) were included with their temperature factors assigned as $B' = 1.5B$, where B is the temperature factor of the parent atom. The R factor of the model with no deuterium atoms included (expected atom type for labile hydrogens) or water (D_2O) was $R = 0.255$ (Table II).

Structure Refinement. (*a*) *Coordinate Refinement.* Refinement of atomic coordinates was carried out in a two-step procedure using the constrained difference Fourier technique (Chambers & Stroud, 1977). In the first step coordinate shifts are applied to atoms according to density gradients in the difference map at atomic centers. After each cycle of coordinate shifts the structure was rebuilt to idealize bond lengths and bond angles and to minimize global energy according to an energy minimization routine developed by Hermans et al. (Hermans & McQueen, 1974). The nonbonded energy function used consisted of two terms equivalent to the Lennard-Jones "6–12" potential and a third term that expressed the electrostatic interaction (dielectric constant = 3) between the atoms calculated as the monopole potential between partial charges (Poland & Scheraga, 1967).

During the first several cycles of refinement, the nonbonded and torsional components were down-weighted compared to the terms involved in the idealization of the bond lengths and angles. This ensured that large shifts were not made during the first stages of refinement. In subsequent cycles, the weights of the nonbonded and torsional terms were gradually increased, eliminating persistent high-energy interactions. Convergence to a low-energy structure was usually achieved in five cycles of the energy minimization step.

(*b*) *Temperature Factor Refinement.* Stereochemically constrained hydrogens bonded to the same parent atom should have similar temperature factor values. Therefore, in the refinement, when two or more hydrogens (or deuteriums) were bound to the same parent, their individual B factors were adjusted toward parity. This was done by calculating the average B of the set of atoms and then adjusting the value of each B factor by $^2/_3$ of the difference between the average and its original value. This type of adjustment allows some influence of the independently determined atomic B factors and, when extended over several cycles, constituted a more effective method for arriving at a stabilized set of average B values than one that simply sets all values to the average in one step.

There were a significant number of partially exchanged amide peptide hydrogens in the protein. Because of the effective cancellation of the diffraction signal in this circumstance, the measurable signal is usually small. For about half of the partially exchanged hydrogens the scattering contribution was too small to attempt to refine. However, in approximately 30 cases, where the peptide hydrogen had exchanged to greater than 60% (+2.5 F)[1] or less than 15% (−2.5 F), there was sufficient signal to refine the atoms. In the refinement, these atoms were given occupancies of 0.6 of either a hydrogen or deuterium atom, depending on whether the observed density in a difference map was positive (D) or negative (H). In subsequent cycles, the occupancies were fixed, with the temperature factor being the adjustable parameter.

(*c*) *Limitations in Structure Refinement.* A refinement of a neutron structure presents a number of unique problems not encountered in an equivalent X-ray analysis (Kossiakoff, 1983; Kossiakoff & Spencer, 1981). Because of the inclusion of hydrogen atoms in a neutron structure there is a significant decrease in the ratio of the number of data observations (*hkl*'s) compared to adjustable parameters (i.e., x, y, and z coordinates and temperature factors) used in the refinement. This has to be compensated for by incorporating strong restraints toward ideal geometry and reasonable temperature factor differences between adjacent bonded atoms (see above). A problem also arises from the close proximity of hydrogen atoms to their parent heavy atom (C, N, O). At less than atomic resolution, the difference density generated from positional errors of one atom overlaps that of an adjacent atom. Because the scattering length of hydrogen is negative, an error in a hydrogen position is corrected by shifting it down the gradient, which is opposite to the direction required for adjusting parent atom positions.

Considering the above, the starting phasing model must be calculated from an accurate set of X-ray coordinates. Test cases for determining the range of convergence using 2.2-Å neutron data indicated that dependable shifts could only be generated if the coordinate errors were less than 0.3 Å (Kossiakoff & Spencer, 1981). Neutron refinement at the resolution of this analysis will rarely improve the accuracy of the heavy atom positions, especially those with high temperature factors. The power of the method is rather in the ability to determine hydrogen atom positions and to discriminate between hydrogens and deuteriums.

Refinement Results. Starting with an agreement residual of $R = 0.255$, the structure was refined for six cycles to give a residual of $R = 0.192$ for all heavy atoms and hydrogens (or deuteriums) and 57 water molecules. The resulting structure was highly restrained toward ideal bond lengths and angles. The average deviation in bond lengths was 0.013 Å, bond angles 3.5°, and torsional angles 4.2°. In addition, the structure contains no high-energy contacts over 3.2 kcal/mol (and only 10 over 2.0 kcal/mol). The final average shift of all non-hydrogen atoms from the initial X-ray model of the wild-type (wt)[2] protein was 0.20 Å.

[1] "F" denotes a fermi, which is a measure of the scattering magnitude (scattering length) of an atom type. It is in units of 10^{-13} cm.

Differences between Wild-Type and M222Q Molecules. Except for the local region around the site of mutation, position 222, the X-ray models of the wt and M222Q variant molecule are essentially identical (R. Bott, personal communication). The rms deviation between their main-chain atoms is 0.10 Å and 0.24 Å when all atoms are included. The most notable change is for the side chain of His64, which shifts away from the side chain of Gln222 by about 0.3 Å. The importance of this movement of the catalytically important His64 side chain is difficult to assess because the relative activity of the enzyme compared to wt is dependent on the substrate used (D. Estell, personal communication). The H-bond of the Nδ1 of the imidazole to Oδ2 of the active site Asp32 is not affected (2.8 Å in wt vs 2.9 Å in M222Q). In neither enzyme is a H-bond made to the Oγ of Ser221; in both cases His Nε2 is about 3.4 Å from the Oγ.

The Gln222 side chain in the neutron structure is reasonably well-ordered (*B* factors < 15 Å^2); however, the configuration of the χ^3 side-chain torsion angle, which defines the positions of the Nε2 and Oε1 atom, cannot be determined from the density map. This suggests a disorder whereby the side-chain χ^3 angle is 2-fold degenerate (χ^3 = 179°). Another small structural change observed near Gln222 involves the side chain of Asn155. The magnitude of this movement (0.25 Å) is at the limit of significance of this analysis, but it is mentioned here because the side chain plays an important role in the mechanism of the enzyme by assisting in stabilizing the transition-state structure. As in the wt enzyme, the side chain of 155 is oriented such that Oδ1 H-bonds to the Thr220 HN;[3] this suggests that the side-chain function in the variant is probably the same.

Establishing Criteria of Confidence for the Phasing Model. Features in density maps are much more dependent on phase information than on the experimentally measured structure factor amplitudes. Because the phasing model for a neutron analysis is not experimentally determined, but is derived from the X-ray model, it is essential to develop a set of structural criteria by which to judge the quality of the model and identify potential artifacts that may arise from phasing bias.

Such criteria are readily developed because a neutron Fourier map contains a set of structural characteristics resulting from the existence of observable density due to scattering by hydrogen and deuterium atoms (that were not part of the X-ray model). A direct qualitative indicator of the initial structure is the presence in the maps of highly ordered water molecules. Since no water structure was included in the primary phasing model, density peaks that corresponded to waters in the X-ray structure necessarily resulted solely from their scattering contributions in the measured neutron data. In the initial maps approximately the top 30 water molecules (low *B* factors) from the X-ray analysis were clearly discernible. Other features that are most readily evaluated are those which are effected by hydrogen–deuterium (H/D) exchange. Such exchange occurs for labile hydrogens on nitrogens and oxygens, particularly those involving side-chain hydroxyl groups (Ser, Thr, Tyr), amide groups (Asn, Gln), imidazole (His), lysine (Lys), and arginine (Arg) (Figure 1). The orientations of virtually all the well-ordered Asn's and Gln's were apparent in the first difference map as were the protonation states of the histidine imidazole side chains (Figure 1b).

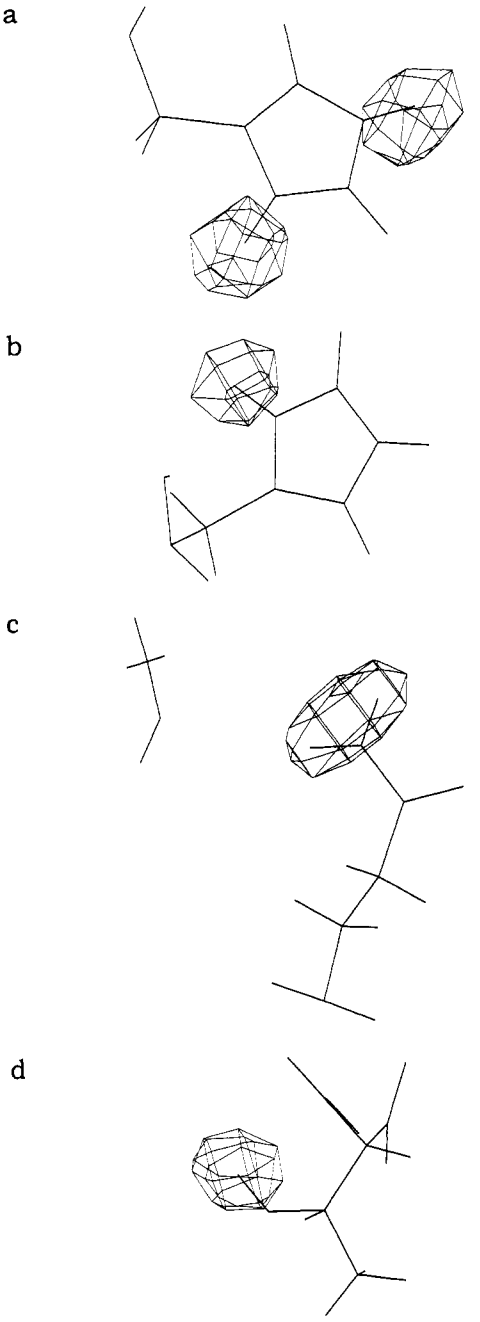

FIGURE 1: Difference maps of side-chain groups with exchangeable hydrogens. (a) His17; density indicates imidazole is protonated. (b) His64; single deuterium on ND1 indicates the imidazole is neutral. (c) Gln2; the presence of the deuteriums clearly identifies the side-chain orientation. (d) Thr208 hydroxyl.

It is noted that some exchanged hydrogens in the neutron structure were unexpectedly absent or weak, on the basis of the well-defined density of the their parent atom. Listed by side-chain type, these are located on Arg Nε (residues 186 and 247), Asn Nδ2 (residues 61 and 62), and Trp Nε1 (residues 106 and 113). Of the nine ordered tyrosine rings, assignment of the hydroxyl deuterium (DO) conformer can be made in five cases. Although it is possible that in the other cases (Tyr's 167, 171, 262, and 263) the hydroxyl is disordered between low-energy conformations of 0° and 180°, thereby reducing the expected density, analysis of hydroxyl orientations of Tyr in trypsin clearly showed that usually there was a marked preference for a single conformer (Kossiakoff et al., 1990).

[2] "wt" designates the wild-type (or native) enzyme structure.

[3] Although many of the peptide amide hydrogens have been exchanged for deuterium in the molecule, to keep a consistent nomenclature they are designated as HN irrespective of whether they are exchanged. Conversely, all labile hydrogens on side chains are designated as deuteriums, i.e., Dγ for a hydroxyl hydrogen of a serine.

The neutron structure of subtilisin reported here was developed from a very good phasing model applied to intermediate resolution data. The resolution of these data is very important in positioning of hydrogen atoms because of the short bond length to the parent atom. Resolution is somewhat less important when assigning H/D exchange features or water sites. The good phasing model intermediate resolution data situation almost exactly describes the initial trypsin analysis, which was done at 2.2-Å resolution (Kossiakoff & Spencer, 1981). The trypsin data were subsequently collected to 1.8 Å, and the structure was rerefined and analyzed. The comparison of the two structures showed that a significant majority of the assignments of waters, H/D exchange properties, and orientation of side chains made in the 2.2-Å structure were basically correct. Thus, the trypsin system offered a good model for characterizing the extent and nature of changes in structural interpretations as a function of two different resolutions and established that detailed structural features can be interpreted with reasonable confidence from 2.0-Å data.

Results and Discussion

Hydroxyl Groups. Hydroxyl groups play a versatile role in protein structure by virtue of their ability to act as either hydrogen bond donors or acceptors or both. As a result, serine, threonine, and tyrosine are found in diverse electrostatic environments. In the absence of local effects from the chemical environment, the orientation of a hydroxyl group (sometimes referred to as a hydroxyl rotor) has a 3-fold degeneracy (2-fold for Tyr) defined by the intrinsic barrier to rotation reflecting steric repulsion of adjacent bonded atoms. This energy barrier ranges from 1.3 kcal/mol for serine and threonine to 3.5 kcal/mol for tyrosine (Weiner et al., 1984). In a majority of cases, however, a hydroxyl group is influenced by the neighboring environment through either a direct H-bond or some longer range electrostatic effects. It is of general interest to evaluate how complementary the electrostatic and steric forces are in determining the perferred orientation and, in those cases where they are in opposition, which force dominates (Kossiakoff et al., 1990).

Hydroxyl hydrogens are highly susceptible to rapid exchange when in a solvated environment, and even when "buried" in the molecule, they are almost invariably found to be exchanged (Kossiakoff et al., 1990). In situations where the hydroxyl appears to have no H-bonding preference, especially when interactions with water are involved, the hydroxyl probably transiently interchanges between being a donor and acceptor, resulting in very diffuse or no density at all in the neutron map for the deuterium atom. In contrast, however, when the hydroxyl oxygen is well-ordered and its deuterium is involved in a hydrogen bond, density representing the deuterium atom's location usually appears quite distinct. One example of a hydroxyl hydrogen not exchanging is Thr66. In addition to its interaction with HN 36 as an acceptor, it H-bonds as a donor to the buried carboxylate group (Oε1) of Asp60. As judged by the low temperature factors of the side-chain Oγ1 and Cγ2 (B = 6 and 8 Å^2, respectively), the side chain is highly ordered, yet at the expected location for the Dγ1 there is a small residual negative peak. This, along with the fact that the peptide amide HN of Asp36 is unexchanged, strongly suggests that the hydroxyl hydrogen has not fully exchanged out over the period the crystal was soaked in D_2O.

The finding that an interior hydroxyl is resistant to exchange has precedence in the neutron analysis of trypsin (Kossiakoff et al., 1990). In trypsin one buried serine, Ser54, was substantially protected from exchange even after a year of exchange. The hydroxyl is embedded in a β-sheet structure acting as a H-bonding conduit between two strands of the sheet.

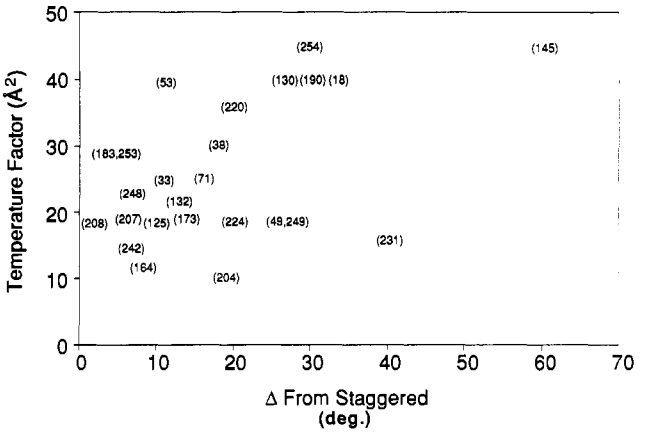

FIGURE 2: Plot showing correlation between the observed deviation of the hydroxyl rotors from a staggered conformation and the refined temperature factor of the hydroxyl deuterium. The locations of the individual residues are represented by their corresponding numbers in the plot.

(*a*) *Hydroxyl Group Orientation and H-Bonding Patterns.* The data base for defining hydroxyl rotor trends for subtilisin is not as complete as that which we have derived from trypsin using neutron solvent difference maps (Kossiakoff et al., 1990). However, by eliminating those groups that cannot be accurately defined because of thermal or positional disorder, the remaining groups can give an informative picture of the characteristics of these groups in subtilisin.

As was the case for the serine and threonine groups in trypsin, the trans conformation for the hydrogen is the highest populated conformer (61% in subtilisin, 46% in trypsin). The other two low-energy conformers (60° and 300°) are about equally populated in subtilisin. Figure 2 shows a plot correlating the observed displacement of the hydroxyl rotors from a low-energy conformer (60°, 180°, 300°) and the hydroxyl hydrogen's temperature factor. By and large, those groups with the largest deviation from minimum energy are those with high B factors. At the resolution of the data used in this study (2.0 Å), the accuracy is typically about 10–15°.

In the trypsin analysis it was found that serine hydroxyls preferentially act as H-bond acceptors and were rarely found as H-bond donors alone (Kossiakoff et al., 1990). The situation differs in subtilisin where, of the 24 serines that were involved in hydrogen bonds and could be evaluated by interpretable density for the hydroxyl deuterium, there are five instances that the hydroxyl is the H-bond acceptor (i.e., H bond through the oxygen), nine where it is exclusively the donor, and ten cases where it functions as both by forming two H-bonds.[4] In an α-helix, a common H-bonding motif observed when the hydroxyl functions exclusively as a donor is when the group replaces the main-chain nitrogen, usually at a point where the helix is appreciably deformed.

(*b*) *Comparison of Observed Conformations with Predicted Orientations.* Figure 3 shows, for a representative set of well-ordered hydroxyl groups, a comparison of the experimentally determined dihedral conformations with their predicted orientations based on calculated energy criteria. Assignment of hydroxyl deuterium positions was based strictly on observed density in the Fourier map; no nonbonded stereochemistry such as H-bonding geometry was taken into consideration. The energy profiles were calculated on the basis

[4] In trypsin the H-bonding donor/acceptor ratio for serine was 10 acceptors, 1 donor, and 10 in which the group functioned as both.

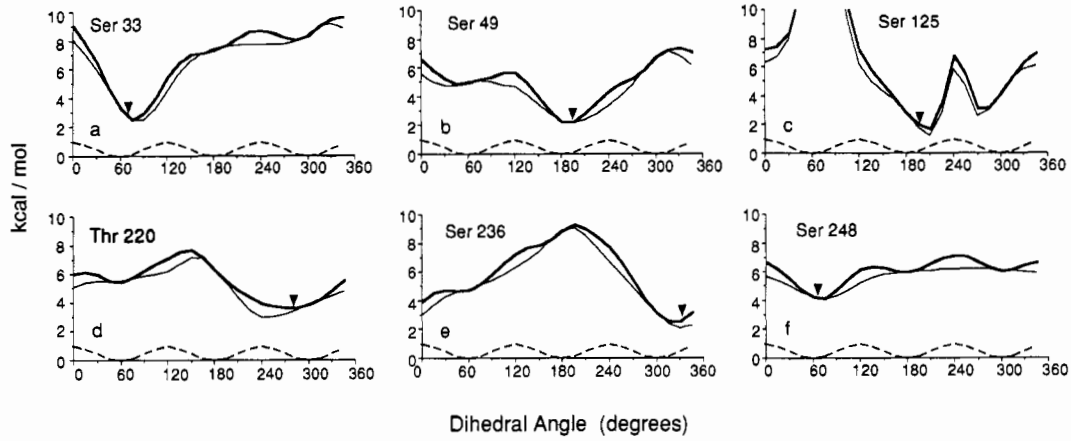

FIGURE 3: Energy profiles calculated for a representative set of hydroxyl rotors. The profiles were determined by rotating the hydroxyl hydrogen around the axis of rotation (Cβ–Oγ bond) in increments of 10° and calculating the energy in each position. The energy function was taken from Amber (Weiner et al., 1984) and is displayed as a total energy (heavy line) as well as its component parts—torsional (barrier to rotation; dashed line) and nonbonded (van der Waals and electrostatic; thin line). The position of the observed orientation is marked by an arrow.

of the combination of the torsional energy, i.e., barrier to rotation of the hydroxyl hydrogen, and nonbonded terms which included van der Waals attraction and repulsion effects and electrostatic interactions.[5] As seen in Figure 3, the dominant effects defining the overall characteristics of the energy profile are dictated by the nonbonded interactions. Except in cases of steric close contacts, i.e., Ser125 (Figure 3c) between 30° and 120°, electrostatic contributions are larger than the van der Waals energies. There is a strong correspondence between the energy minima defined by both torsional and nonbonded factors, and as was the case for trypsin, there is a striking overall correspondence between experiment and theory.

The hydroxyl rotor plots show that each hydroxyl rotor has its own characteristic energy profile—some have a broad minimum while others have quite deep and narrow energy wells. The shapes of the profiles are a good indicator of the character of the local forces influencing the hydroxyl group's conformation. For hydroxyl groups of Ser33, Ser49, Thr220, Ser236, and Ser248, which are observed to be displaced slightly out of the low-energy staggered conformation, it is tempting to assign significance to the observed perturbations because they mirror the calculated nonbonded energy minimum, which is similarly displaced from the torsional energy minimum. It is emphasized, however, that, at the resolution of this analysis, the error associated with defining the hydroxyl conformations is between 10° and 15°. Displacements of Ser49, Thr220, and Ser236 are probably significant. The hydroxyl of Ser49 participates in two H-bonds, as a donor to O 54 and as an acceptor to HN 51. This is a rigid network of H-bonds as revealed by the finding that HN 51 is highly protected from exchange. The nonbonded energies favor a dihedral angle of about 205° (Figure 3b, about 25° displaced from staggered); the observed angle of about 190° suggests a compromise between the nonbonded and torsional terms. Figure 3d shows that the nonbonded energy minimum for Thr220 is displaced toward the high-energy eclipsed conformer (240°), while the summation of the energy terms produces a broad minimum at about 270°, which is close to the observed value of 280°. The breadth of this minimum also correlates with the diffuse density of the hydroxyl deuterium (B = 36 Å^2) in the Fourier map.

The energy profile for Ser236 suggests a different partitioning of the relative contributions of the torsional and nonbonded energy terms than was indicated for Ser49 (Figure 3e). For Ser236 the observed dihedral displacement of 40° (to 340°) coincides more closely with the nonbonded minimum (330°) than the overall energy minimum of 320°. This suggests a limited influence by the torsional term. This finding, in light of the other examples that indicate measurable influence of the torsional parameter, complicates making general assumptions qualifying the relative importance of each of the terms to the overall mix that defines the hydroxyl's conformation. The discrepancy in the effect of the torsional parameter between 236 and 49, however, may be explained by electrostatic effects. Ser236 is in a helix, and its hydroxyl deuterium back-bonds onto the carbonyl oxygen of 233 (Figure 4). Unlike Ser49, the 236 hydroxyl acts only as an H-bond donor. This H-bond is shared with the peptide HN of 237, resulting in a steric close contact between Dγ 236 and HN 237 (2.1 Å). In the observed orientation Dγ 236 is 1.95 Å from O 233; modeled in the staggered position (300°), the Dγ–O distance is lengthened by about 0.15 Å. This lengthening would be expected to have an appreciable effect on the H-bond strength of the interaction and apparently is the determining factor of the geometry. Similar dominance of the effect of electrostatics over the torsional component in defining H-bond geometry was reported for trypsin (Kossiakoff et al., 1990).

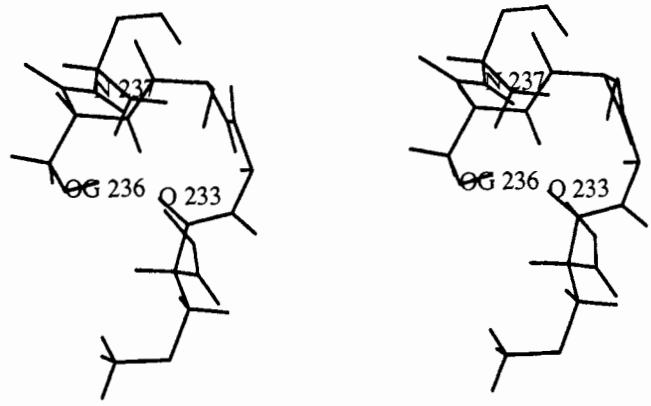

FIGURE 4: The hydroxyl orientation of Ser 236 relative to the peptide amide of residue 237.

The high degree of complementarity between the calculated energy minimum and the observed rotor conformation noted above and also reported for the analysis of trypsin (Kossiakoff et al., 1990), even where the groups are perturbed out of the staggered orientation, indicates that energy calculations based

[5] The Amber force field was used in these calculations (dielectric constant = 3) (Weiner et al., 1984).

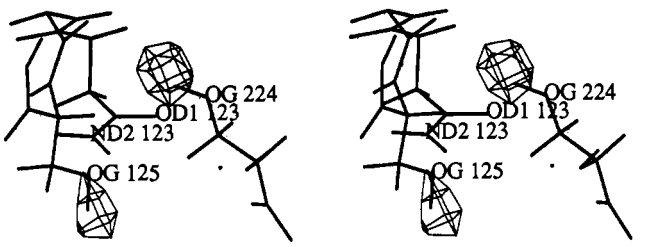

FIGURE 5: Difference map calculated with the hydroxyl deuteriums of 224 and 125 removed from the phasing model.

on local bonded and nonbonded effects are a very good predictor of actual rotor conformations. However, there is a particularly puzzling example in subtilisin where the predicted energy criteria, which accurately predicted a majority of hydroxyl conformations, are disobeyed. The Ser224 Oγ is in a buried position about 3.1 Å from the Oδ1 of Asn123. The orientation of the Asn123 side chain is directly assigned on the basis of the well-defined density for the Nδ2 deuteriums. In the absence of knowledge of the hydroxyl deuterium position, on the basis of stereochemical factors, it would be assumed to point toward the Oδ1 forming a H-bond. The neutron density defining the hydroxyl deuterium position shows unambiguously that this is not the case; the deuterium actually points away from the Oδ1 (Figure 5). Even if it were pointing directly at the Oδ1, it would have poor H-bonding geometry.[6] However, considering the close distance between the two oxygens, on the basis of electrostatic arguments alone, it would still provide a buffer between the two electronegative atoms.

At the resolution of the analysis and expected accuracy of the phasing model it is certain that there are some few artifacts in the map. Normally, an outlier from the norm such as the hydroxyl of 224 would be questioned and perhaps eliminated from the data base unless there were significant supporting data. To establish a level of confidence for the assigned hydroxyl orientation of 224, several findings were considered. The refined position and temperature factor for the deuterium (B = 14 Å^2) indicated it to be one of the five best ordered hydroxyls in the structure (Figure 2). The rotor conformation is very near a dihedral energy minimum (197°), and it is unlikely that a noise peak would coincidentally fall at an energy minimum. Figure 5 shows a section of the difference density map where both the hydroxyl deuteriums of Ser224 and Ser125 have been removed from the phasing model. The corresponding peaks came back at about the same strength at their original positions. (Ser125 is also a well-ordered hydroxyl, B = 16 Å, and H-bonds to the Nδ2 of 123.) In a difference map where the 224 hydroxyl deuterium has been placed to H-bond to the Oδ1 of 123, a negative peak appears at that position accompanied by a positive peak at the originally assigned position. Taken together, these observations support the original assignment of the hydroxyl rotor.

Hydrogen–Deuterium Exchange of Peptide Amide Groups. The observed exchange pattern of the peptide amide groups can be used to evaluate relationships between exchangeable sites and structural and chemical properties of the molecule. To date, neutron analyses of the H/D characteristics of six proteins have been reported (Kossiakoff, 1982, 1984; Wlodawer & Sjolin, 1983; Mason et al., 1984; Raghavan & Schoenborn, 1984; Teeter & Kossiakoff, 1984; Wlodawer et al., 1984). Taken together, they indicate a strong correlation of exchange rates with the location of the sites within regions of the protein's secondary structure elements. One shortcoming of these studies, however, is that, with the exception of trypsin (23 kDa), the proteins were relatively small, ranging from 8 to 15 kDa, and consequently were limited in the variety and size of their secondary structure types and had relatively small solvent-inaccessible cores. Thus, within any one of these protein systems it was not possible to evaluate how different structural features as a group impacted the relative degrees of protection afforded peptide sites.

Several features of the structure of subtilisin make it an attractive system for study of the H/D properties. It is a protein of 275 amino acids (27 kDa) with a considerable hydrophobic core. It has extensive regions of α-helix and β-sheet, as well as typical types of loops and turns, and thus provides in one system a more substantial data base from which correlations can be drawn. It has no disulfide bridges yet is stable toward denaturants. A question of interest was whether there is an identifiable characteristic in the exchange pattern that might provide insight into the source of this stability.

To obtain a model of the exchange pattern of a protein, it is essential to be able to discriminate accurately between those sites that are fully exchanged or fully protected and those that are partially exchanged. The presence of partially exchanged groups indicates that they have a degree of protection, which provides additional information in the context of the exchange pattern for the fully exchanged and unexchanged groups. However, even estimation of the partially exchanged groups is problematic. For example, if the site is 60% exchanged, the scattering contribution from the D is about 4.0 F (0.6 × 6.7 F), and from H, −1.5 F (0.4 × −3.8 F), leading to a reduction to about 40% (2.5 F) that of a full deuterium. The expected peak in a neutron Fourier map, even with very high resolution data, would be close to the noise level of the map. Nevertheless, even though exchange ratios between 15 and 70% D can only be qualitatively assigned, the limit of the partitioning of exchange characteristics into three principal categories, exchanged (100–70%), partially exchanged (70–15%), and unexchanged (15–0%), still provides a fruitful data base from which structure–protein dynamics correlations can be drawn.

Exchange Patterns in Secondary Structure Units. In subtilisin 55% of the peptide amide hydrogens were found to be nearly fully exchanged, 15% were unexchanged, and the remaining 30% were classified as being partially exchanged. The ratio of partially exchanged groups is comparatively high, presumably because there apparently is a significant number of moderately protected groups in the molecule. The crystal was soaked in D_2O for a relatively short period of time (1 month) before data collection was started, while complete data collection spanned 10 months. Under these experimental conditions, for the slowly exchanging sites, the technique involves a time-averaged measurement of a dynamic process; groups that were only marginally exchanged at the initiation of the data collection could be significantly exchanged over the period for full data acquisition.

A vast majority of the unexchanged groups were found in β-sheet and α-helix secondary structures. The core of the protein is composed of an extended seven-stranded parallel β-sheet unit (designated β_p), which is sandwiched between four lengths of α-helix. As judged by solvent accessibility calculations,[7] the main-chain atoms of the residues contained in the sheet are, with the exception of strand I and several other isolated cases, well shielded from direct contact with solvent.

[6] The H-bond angle defined by C(123)–O(123)–Dγ(224) is 125°, with the Dγ aligned perpendicular to the plane of the 123 side chain (Oδ1–C–Nδ2).

[7] Solvent accessibility calculations were carried out by using the method of Lee and Richards (1971) with a water probe size of 1.4 Å.

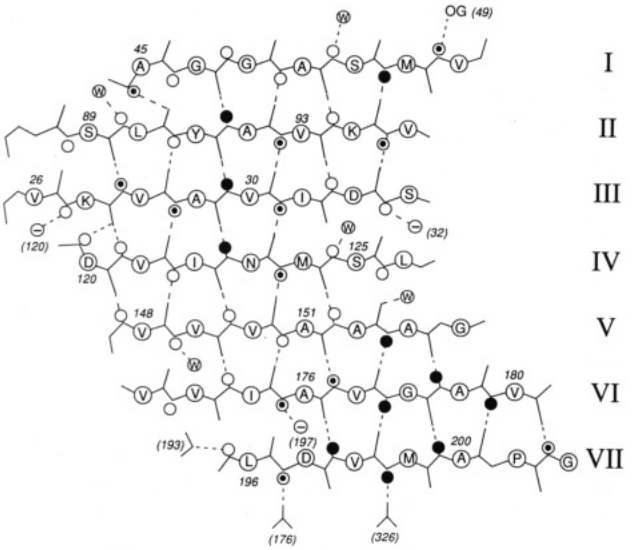

FIGURE 6: Schematic representation of H/D exchange at peptide amide sites in the β_p structure of subtilisin. Key: O, full exchange; ⊙, partial exchange; ●, unexchanged; ⊖, carboxylate group; circled "W", water molecule. Amino acid types are designated by their single-letter codes.

The solvent accessibilities of the helix units are variable and are usually characteristically amphipathic.

β-Sheet Structures. The amino acid composition of the β-sheet is unusual, being extremely rich in alanine and valine. The high population of valine (18%) is characteristic of β-sheet composition, but about 22% of the sheet (10 out of 45 total) is composed of alanine, which is not a good β-sheet promoting group. The β-strands are between 6 and 7 residues in length; four of the strands are preceded by 3_{10}-type turns. Within sections of strands V and VI, 6 of the 8 residues are alanine or glycine. Of the 45 residues contained within the β_p sheet structure, 11 are fully or nearly fully unexchanged (Figure 6). The largest concentration of protected groups is in strands VI (G178, A179, V180) and VII (V198, M199, A200). These sites form a distinct patch near the edge of the sheet. The other unexchanged groups are M50 (strand I), A92 (in II), V30 (in III), N123 (in IV), and A153 (in V).

Of the two interior strands, which might be expected to be the most resistant to exchange, strand IV has one unexchanged and one partially protected group, while V is fully exchanged except for A153, which is unexchanged. The exchange behavior of strand V might be explained by local H-bonding irregularities. The peptide nitrogen of V149 is at the interface of a pocket of three partially buried waters and, rather than H-bonding to a main-chain carbonyl oxygen, H-bonds directly to one of the interior waters. In addition, at the C-terminal end of the strand, the carbonyl oxygen of A152 does not H-bond to the expected partner in the adjacent strand, 126 HN, but rather to another water. Thus, the regular interstrand H-bonding network is not conserved for this strand.

Within the β_p structure there are 12 partially protected amide peptide groups. These are primarily found contiguous to the protected groups, indicating a gradual demarcation between fully exchanged and fully protected regions of the structure. There appears to be little correlation between H-bond length and angle and whether the group is partially or fully unexchanged. In fact, some of the longest H-bonds in the sheet [between strands VI and VII, O 176–HN 197 (3.1 Å), O 176–HN 198 (3.2 Å), HN 178–O 198 (3.0 Å), O 178–HN 200 (3.0 Å), HN 180–O 200 (3.0 Å), O 180–HN 202 (3.1 Å)] are between segments of chain that are highly protected.

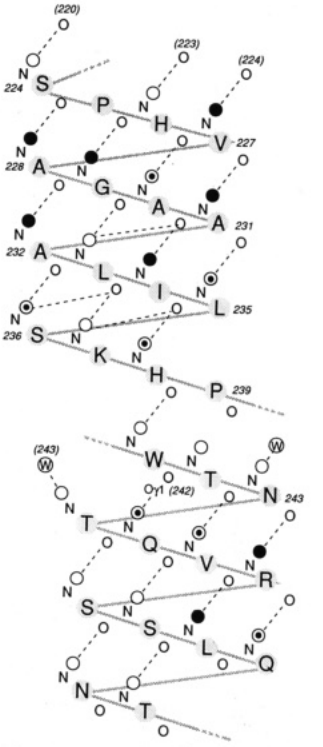

FIGURE 7: Schematic representation of H/D exchange in the α-helix structures 220–239 and 242–253. Key is the same as described in Figure 6. Peptide nitrogens are designated by N, oxygens by O.

Compared to trypsin, which has extensive regions of antiparallel β-sheet, the β_p structure in subtilisin is considerably more susceptible to exchange. In the case of exchange in trypsin's antiparallel sheets, essentially all the groups contained in the central sections of the sheet are highly protected; the only sites that exchange are at the edges of the sheet (Kossiakoff, 1982, 1984). The β_p structure in subtilisin does not appear to possess the same pattern of cooperativity since regions along strands within the central region of the sheet exchange more rapidly than some at the edges.

α-Helix Structures. There are five α-helix structures in subtilisin. These helices differ considerably in length, accessibility, and H-bonding regularity. The first helix in the sequence, 13–21, is more accurately described as a mixture of interspersed α and 3_{10} turns. As such, it is not surprising that the H-bonding network does not effectively protect the peptide groups from exchange. The only group that is reasonably well protected is L12, which is involved in a 3_{10} turn with 9 O.

The most highly protected helix contains residues 224–236, a significant number of which are buried in the protein interior. Three consecutive residues, V227, A228, and G229, are unexchanged. The next residue (A230) is partially exchanged followed by two unexchanged groups (A231, A232) (Figure 7). The H-bonding and exchange pattern abruptly changes at L233, which is fully exchanged. The regularity of the helix geometry is perturbed here, with the HN of 233 forming H-bonds with both 229 O (2.9 Å) and 230 O (3.0 Å). A more optimal helix geometry follows, and the exchange pattern reverts back to protection for I234.

Exchange of N237, which is flanked on either side by partially protected groups, marks another H-bonding deformation in the helix where the peptide nitrogen forms a bifurcated interaction with 233 O (3.3 Å) and 234 O (3.1 Å). The helix is terminated at P239, and after a short loop to T242 another helix is initiated. This is a short helix of approximately two turns and is more solvent accessible than the preceding

one. There are two protected groups (R247, L250) and three partially protected groups (Q245, V246, Q251) (Figure 7). Within this group only V246 and L250 are solvent inaccessible as measured by coordinates of the static structure. The hydrophilic nature of the other protected side-chain types indicates that accessibility is not the principal factor in the exchange characteristics of this helix. As is the case for the 224–236 helix, the unexchanged groups are sequestered within a single patch on the helix surface.

In contrast to the situations discussed above, two other helices consisting of residues 104–117 and 132–147 are more fully exchanged. Whereas the usual situation is one where an unexchanged peptide is flanked by groups of similar character, within the helix 104–117, the only unexchanged group, Q112, is adjacent to fully exchanged sites. Thus, since the accessibilities of the groups are similar, Q112 may be a point in the helix that is not appreciably affected by breathing fluctuations in the protein.

Residues 64–85 form an unusual discontinuous helix (Richardson, 1981). After the initial two turns (residues 64–74) the helix is interrupted by a loop containing 8 residues, which forms the primary Ca^{2+} binding site of the protein. This loop is followed by the continuation of the helix by addition of a third turn comprised of residues 83–85. Within the first two turns of this helix only the peptide amides of residues T66, V68, and G70 are highly protected from exchange; in the final turn, two peptide amides (G83 and A85) are significantly protected.

An aspect of the exchange findings for the helices worth noting is the number of glycines in the helices that are fully or partially protected. Glycine is generally considered a helix "breaker" and is not often found in helices (Chou & Fasman, 1974). Consequently, it might be expected that because of fewer conformational restrictions, the glycine sites in the helix would be more apt to display local conformational flexibility leading to increased exchange rates. Such a situation clearly does not exist; of the five glycines located in helices, three are unexchanged and two are partially protected.

Other Structural Factors Affecting H/D Exchange. Although the majority of protected groups are located in regions of regular secondary structure, there are a significant number of cases where side-chain H-bonding or steric factors affect rates of exchange of peptide amide groups. Below, the nature of these effects is discussed and categorized by whether they are principally due to H-bonding, sterics, or both.

(*a*) *Side-Chain H-Bonding.* A majority of the peptide amides that are H-bonded to side-chain groups are fully exchanged. This implies, as expected by the nature of their conformational degrees of freedom, that side chains, even when H-bonded, are more conformationally variable than regions of the folded main chain. There are, however, instances where side-chain H-bonding impedes exchange. These are principally in situations where the side chain is highly restrained due to secondary H-bonding interactions.

There are three instances in subtilisin where exchange is impeded by a H-bond to the hydroxyl oxygen of either a serine [HN 51–Oγ 49 (3.0 Å), HN 179–Oγ 190 (2.9 Å)] or a threonine [HN 36–Oγ1 66 (3.0 Å)]. In each instance the groups are buried, and for Oγ 49 and Oγ1 66 the hydroxyls are involved in a second H-bond as a donor. Even though the amide peptide is protected from exchange, the hydroxyls of 49 and 190 are exchanged, showing that exchange of a hydroxyl is chemically and sterically a lower energy process than exchange of a backbone hydrogen.

(*b*) *β-Turns.* A common secondary structural element in proteins is the β (or 3_{10}) turn, which provides for reversal in chain direction. The turn involves four contiguous residues that are stabilized by a backbone H-bond between the carbonyl O of residue n and the peptide nitrogen of residue $n + 3$. However, because this type of structure is somewhat rigid and is usually accompanied by some additional H-bonds involving other residues in the turn, inherently the $n + 3$ peptide HN is afforded a degree of protection against exchange.

In subtilisin there are 33 β-turn type structures. (Note that these structures are defined here simply as involving the O_n–N_{n+3} H bond and not a specific set of ϕ,ψ main-chain torsion angles.) Characteristic bond lengths and angles defining geometries of β-turns have been compiled by several investigators (Richardson, 1981; Baker & Hubbard, 1984; Wilmot & Thornton, 1988). They found that in general these geometries were characterized by slightly longer H-bond lengths (≅0.1 Å) and less optimum H-bonding angles than those associated with α-helices and β-sheets. Of the 33 $n + 3$ peptide amides contained within these turns, one is fully protected and another six are partially protected from exchange. For the fully protected group (L42) and three of the six partially protected groups (Q54, L196, V227) there is an additional structural factor that presumably affects exchange rates. In these cases, proline is the $n + 1$ residue, a feature that apparently incorporates a higher degree of rigidity in the turn. Although this data base for protected β-turn groups is small, the fact that all four of the turns having a proline in the $n + 1$ position are unexchanged to an extent suggests that this feature plays a significant role in their protection.

(*c*) *β-Hairpin Turns.* In addition to the β_p structure, subtilisin has a β-hairpin loop (residues 205–217) with three sets of antiparallel H-bonding interactions. The peptide amide groups of the two interior H-bonds (209:213, 207:215) display a degree of protection from exchange. This is not a function of H-bonding structure alone since the N 213–O 209 H-bonding distance is long (3.2 Å); at the end of the hairpin (205:217) the peptide amides are fully exchanged even though the H-bond distances are shorter [N 205–O 217 (2.8 Å), N 217–O 205 (3.0 Å)].

Active Site, the Catalytic Machinery. The mechanism of the serine protease class of enzymes has been an intensely debated issue in the enzyme mechanism field. This family of enzymes is characterized by three conserved residues at the active site, a histidine (64),[8] an aspartic acid (32), and a serine (221). The principal question was whether His64 was the actual chemical base in the hydrolysis reaction or whether, because of local environmental effects, the histidine acts as an intermediary through which Asp32 functions as the base. This question was unequivocally resolved by the neutron analysis of a transition-state complex of trypsin, showing the histidine to be the chemical base (Kossiakoff & Spencer, 1980; Kossiakoff & Spencer, 1981), indicating that the side-chain groups retain their inherent chemical characteristics during catalysis. In the NMR study of α-lytic protease a similar finding was found for the uncomplexed form of a serine protease enzyme (Bachovchin & Roberts, 1978).

This subtilisin analysis differs in several respects from that of trypsin. First, subtilisin has a different evolutionary lineage than trypsin, and consequently, even though there is a quite similar stereochemical arrangement of the Asp–His–Ser, the neighboring environment is significantly different. Second, the subtilisin structure studied here is that of the "free" en-

[8] In the trypsin family of enzymes the active site numbering is His57, Asp102, and Ser195.

zyme, so that the conformations of the active site groups are those unaltered by a bound substrate.

The mother liquor in which the crystal was soaked before data collection had a pH of 6.1 (corrected for deuterium isotope effect). Although it is difficult to assign exactly the pH in the crystal itself, the pH is probably not perturbed more than 0.3 pH unit, and thus it was surprising to find that the imidazole of His64 was in a neutral state (Figure 1). This contradicts what would be expected from activity titrations (Russell et al., 1987) and NMR assignments (Jordan et al., 1985; Bycroft & Fersht, 1988) which indicate the pK_a to be close to 7. It was also clear from the neutron density maps that the pK_a of His64 is lower than that of His17 and His238. This order of apparent pK_a's is also in disagreement with a set of NMR assignments (Bycroft & Fersht, 1988). Because of the proximity of hydrophilic groups of neighboring molecules in the crystal lattice, it is possible that the protonation of the histidines might be affected by a perturbed charge environment compared to that experienced in the fully solvated case. In other BPN′ subtilisin structures that were solved from different crystal forms at pH's below 6.5 there was evidence for a bound sulfate adjacent to the His64 imidazole, indicating that the histidine was protonated (Matthews et al., 1977). In the crystal form analyzed here there is no evidence in the electron density difference maps for a similar binding sulfate anion.

In the free enzyme, the imidazole makes only a H-bond to the carboxylate of Asp32 (Oδ1–Nδ1, 2.9 Å), since the distance between the Ser221 Oγ and His64 Nε2 is too long (3.4 Å) for a hydrogen bond interaction. It is not surprising, therefore, that the 221 Dγ (deuterium of Oγ) adopts multiple conformations and has no identifiable density features. This finding indicates that an assumption made by Smith et al. (1989), based on a solid-state NMR analysis of α-lytic protease, is not applicable to all other serine proteases. These investigators suggested that the reason why crystallographic analyses of native serine proteases showed no His–Ser hydrogen bond was that in crystals, because of the high salt content, the active site histidine has an aberrantly high pK_a, making it protonated at the pH of the analyses. They further state that they observe a H-bond only when the imidazole is unprotonated and the serine hydroxyl is the donor group. Aside from the fact that it is hard to rationalize on chemical grounds why a hydroxyl hydrogen bond could not be equally satisfied whatever the protonation state of the imidazole, the neutron results showing the imidazole Nε2 to be unprotonated and the hydroxyl oxygen still 3.3 Å away demonstrate that, at least in the case of subtilisin, their assumption is not operative.

Substrate and inhibitor studies have shown that the side chain of Asn155 plays an important role in stabilizing the transition state by forming a H-bond to the oxyanion oxygen (Matthews et al., 1975; McPhalen & James, 1988; McPhalen et al., 1985; Bode et al., 1986, 1987). It has been reported by McPhalen and James that in subtilisin–protein inhibitor structures the Asn155 side chain makes two additional H-bonds, through its Oδ1 to the NH of Thr220 and Oγ1 of the threonine hydroxyl (McPhalen & James, 1988). A similar H-bonding configuration is seen in several boronic acid inhibitor complexes of the enzyme (Butcher and Kossiakoff, unpublished results). In M222Q the Oδ1–NH 220 H-bond is formed, but the H-bond to the hydroxyl is not. The distance Oγ1 220 to Nδ2 155 is 3.4 Å and Oγ1 to Oδ1 is 3.3 Å. In addition, the hydroxyl deuterium is not pointing toward either atom, but rather at the Cγ. The hydroxyl rotor profile of 220 (Figure 2) shows the minimum-energy well to be quite broad and slightly rotated out of a staggered conformation. This corresponds well with the experimental finding also showing the rotor to be perturbed 20° and the density being diffuse. Taken together with the inhibitor results, it is suggestive that, on binding of substrate, there are some small adjustments in the active site that facilitate the 155–220 side-chain H-bonds being formed.

Cation Binding Sites. Subtilisin BPN′ contains two cation binding sites (Drenth et al., 1972). The cation in the first site is octahedrally coordinated with average ligand bond lengths of about 2.4 Å. These stereochemical features along with the fact that the ion refines to the equivalent of 20 e^- makes it almost certain that this site is occupied by Ca^{2+}. The bond distances (2.9 Å) and the geometry (pentagonal bipyramid) of the ligands forming the second site, however, are more characteristic of a monovalent site as is the refined occupancy of the atom (Bott et al., 1988). Thus it has been presumed to be a Na^+ (or K^+) ion or a water (Drenth et al., 1972; Bott et al., 1988).

The scattering of neutrons by Ca (4.7 F), K (3.7 F), and Na (3.5 F) is small, even less than for C or O. It certainly is not possible to differentiate among these atoms on the basis of scattering magnitude. However, the question of whether the second site is a monovalent ion, a water, or a mix of the two can be addressed uniquely by neutron diffraction. The density in the Fourier map at both sites 1 and 2 is only slightly above the noise level. If the second site were a water (D_2O) or even a mixture with water being at least a 30% component, the density would be significantly higher. Consequently, it can be said with some confidence that the site does not contain a water, but rather a monovalent cation.

ACKNOWLEDGMENTS

We thank Dr. B. P. Schoenborn and his colleagues at Brookhaven National Laboratory who maintain the crystallography station at the HFBR. Dr. Robert McDowell performed the energy calculations for the hydroxyl rotors. We also thank Dr. R. Bott for use of coordinates prior to publication and M. Randal for help with the crystallization.

Registry No. Ser, 56-45-1; Thr, 72-19-5; His, 71-00-1; subtilisin, 9014-01-1.

REFERENCES

Bachovchin, W. W., & Roberts, J. D. (1978) *J. Am. Chem. Soc. 97*, 8041–8047.

Bacon, G. E. (1975) in *Neutron Diffraction*, Clarendon Press, Oxford.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.

Bode, W., Papamokos, E., Musil, D., Seemueller, U., & Fritz, H. (1986) *EMBO J. 5*, 813–818.

Bode, W., Papamokos, E., & Musil, D. (1987) *Eur. J. Biochem. 166*, 673–692.

Bott, R., Ultsch, M., Kossiakoff, A., Graycar, T., Katz, B., & Power, S. (1988) *J. Biol. Chem. 263*, 7895–7906.

Bycroft, M., & Fersht, A. R. (1988) *Biochemistry 27*, 7390–7394.

Chambers, J. L., & Stroud, R. M. (1977) *Acta Crystallogr., Sect. B B33*, 1824–1837.

Chou, P. Y., & Fasman, G. D. (1974) *Biochemistry 13*, 211–222.

Drenth, J., Hol, W. G., Jansonius, J. N., & Kolkock, R. (1972) *Cold Spring Harbor Symp. Quant. Biol. 36*, 107–116.

Freer, S. T., Alden, R. A., Carter, C. W., & Kraut, J. (1975) *J. Biol. Chem. 250*, 46–54.

Hermans, J., & McQueen, J. E. (1974) *Acta Crystallogr., Sect. A A30*, 730–739.

Jordan, F., Polgar, L., & Tous, G. (1985) *Biochemistry 24*, 7711–7717.
Kossiakoff, A. A. (1982) *Nature 296*, 713–721.
Kossiakoff, A. A. (1983) *Annu. Rev. Biophys. Bioeng. 12*, 159–182.
Kossiakoff, A. A. (1984) in *Neutrons in Biology*, *Basic Life Sciences*, Vol. 27, pp 281–304, Plenum, New York.
Kossiakoff, A. A., & Spencer, S. A. (1980) *Nature 288*, 414–416.
Kossiakoff, A. A., & Spencer, S. A. (1981) *Biochemistry 20*, 6462–6474.
Kossiakoff, A. A., Shpungin, J., & Sintchak, M. D. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4468–4472.
Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.
Markland, F. S., & Smith, E. L. (1971) in *The Enzymes* (Boyer, P. D., Ed.) Vol. 3, pp 516–608, Academic Press, New York.
Mason, S. A., Bentley, G. A., & McIntyre, G. S. (1984) in *Neutrons in Biology*, *Basic Life Sciences*, Vol. 27, pp 323–334, Plenum, New York.
Matthews, D. A., Alden, R. A., Birktoff, J. J., Freer, S. T., & Kraut, J. (1975) *J. Biol. Chem. 250*, 7120–7126.
Matthews, D. A., Alden, R. A., Birktoft, J. J., Freer, S. T., & Kraut, J. (1977) *J. Biol. Chem. 252*, 8875–8883.
McPhalen, C. A., & James, M. N. G. (1988) *Biochemistry 27*, 6582–6598.
McPhalen, C. A., Svendsen, I., Jonassen, I., & James, M. N. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7242–7246.
Pantolino, M. W., Witlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L., & Bryan, P. N. (1989) *Biochemistry 28*, 7205–7213.
Poland, D., & Scheraga, H. A. (1967) *Biochemistry 6*, 3791–3820.
Polgar, L., & Bender, M. (1969) *Proc. Natl. Acad. Sci. U.S.A. 64*, 1335–1342.
Raghavan, W. V. & Schoenborn, B. P. (1984) in *Neutrons in Biology*, *Basic Life Sciences*, Vol. 27, pp 247–259, Plenum, New York.
Rao, S. N., Singh, U. C., Bash, P. A., & Kollman, P. A. (1987) *Nature 328*, 551–554.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Russell, A. J., Thomas, P. G., & Fersht, A. R. (1987) *J. Mol. Biol. 193*, 803–813.
Salemme, F. R., & Weatherford, D. W. (1981) *J. Mol. Biol. 146*, 101–117.
Schoenborn, B. P. (1984) in *Neutrons in Biology*, *Basic Life Sciences*, Vol. 27, pp 261–280, Plenum, New York.
Smith, S. O., Farr-Jones, S., Griffin, R. G., & Bachovchin, W. W. (1989) *Science 244*, 961–964.
Spencer, S. A., & Kossiakoff, A. A. (1980) *J. Appl. Crystallogr. 13*, 563–571.
Teeter, M. M., & Kossiakoff, A. A. (1984) in *Neutrons in Biology*, *Basic Life Sciences*, Vol. 27, pp 335–348, Plenum, New York.
Thaller, C., Eichele, G., Weaver, L. H., Wilson, E., Karlsson, R., & Jansonius, J. (1985) *Methods Enzymol. 14*, 132–136.
Weiner, S. J., Kollman, P. A., Case, D., Singh, U. L., Chio, C., Alagona, G., Profeta, P. S., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765–784.
Wells, J. A., & Estell, D. A. (1988) *Trends Biochem. Sci. 13*, 291–297.
Wilmot, C. M., & Thornton, J. M. (1988) *J. Mol. Biol. 203*, 221–232.
Wlodawer, A., & Sjolin, L. (1983) *Biochemistry 22*, 2720–2728.
Wlodawer, A., Walter, J., Huber, R., & Sjolin, L. (1984) *J. Mol. Biol.*, 301–329.
Wright, C. S., Alden, R. A., & Kraut, J. (1969) *Nature 221*, 235–242.