Sreebny, L. M., and Meyer, J., Ed., (1964), Salivary Glands and Their Secretions, New York, N. Y., The Macmillan Co.

Takeda, T., DeBusk, J., and Grollman, A. (1969), *Amer. J. Physiol. 216*, 1194.

Taylor, J. M., Cohen, S., and Mitchell, W. M. (1970), *Proc. Nat. Acad. Sci. U. S. 67*, 164.

Waldhausl, W. K., Lucas, C. P., Conn, J. W., Lutz, J. H., and Cohen, E. L. (1970), *Biochim. Biophys. Acta 221*, 536.

Werle, E., Trautschold, I., Krammer, K., and Schmal, A. (1968), *Hoppe Seyler's Z. Physiol. Chem. 349*, 1441.

Werle, E., Vogel, R., and Goldel, L. T. (1957), *Arch. Exp. Pathol. Pharmakol. 230*, 236.

Whitaker, J. R. (1963), *Anal. Chem. 35*, 1950.

Zacharius, R. M., Zell, T. E., Morrison, J. H., and Woodlock, J. J. (1969), *Anal. Biochem. 30*, 148.

Subtilisin; a Stereochemical Mechanism Involving Transition-State Stabilization†

Jon D. Robertus,‡ Joseph Kraut,* Richard A. Alden, and Jens J. Birktoft

ABSTRACT: The difference-Fourier method was used to determine the binding geometry for various polypeptides to subtilisin BPN′. The principal finding was that polypeptides corresponding to the acylating portion of good substrates bind in a fashion very similar to that seen previously for polypeptide chloromethyl ketone derivatives. The binding of these virtual substrate polypeptides was assumed to represent a model for the Michaelis complex between the enzyme and actual polypeptide substrates. The carbonyl oxygen of the susceptible bond is near two potential hydrogen-bond donating groups, the backbone NH of Ser-221 and Nδ2 of Asn-155, but the distance is too large for the bonds to be formed in the Michaelis complex. The focus of these two hydrogen bonds is termed the "oxyanion hole." In addition the amido nitrogen of the substrate specificity residue P_1 (Schechter, I., and Berger, A. (1967), *Biochem. Biophys. Res. Commun. 27*, 157) is poised near to, but again too far from, the carbonyl oxygen of Ser-125 for a hydrogen bond, termed the "S_1–P_1 hydrogen bond," to be formed. Model-building experiments readily led to models for the two other likely intermediates, a tetrahedral addition compound and the acyl-enzyme. In the tetrahedral addition compound construction of a covalent linkage between the Ser-221 Oγ and the carbonyl carbon of the susceptible bond results in distortion of that carbon to a tetrahedral conformation, and closer contact between the substrate specificity residue P_1 and the enzyme. Consequently, the S_1–P_1 hydrogen bond is formed in the addition compound, and the oxygen of the specificity residue, now carrying a formal negative charge, is stabilized in the oxyanion hole by formation of two additional hydrogen bonds. These three hydrogen bonds, plus a fourth between the leaving group and His-64, stabilize the tetrahedral intermediate and thereby lower the free energy of activation for rate-limiting transition states. Furthermore, model building also indicates that when the tetrahedral intermediate collapses to the acyl-enzyme, steric constraints imposed on a specific substrate by the enzyme prevent the carbonyl oxygen of the resulting ester linkage from being maintained in the oxyanion hole. Finally, the stereochemical similarity between subtilisin and chymotrypsin previously described for the charge relay system and the binding site is shown to extend as well to the oxyanion hole, apparently so important for catalysis. This provides further evidence that the two enzymes have converged in an evolutionary sense to the same mechanism of action.

In the previous paper of this series we described the binding to subtilisin BPN′ of four chloromethyl ketone analogs of good polypeptide substrates, as seen by the difference-Fourier method (Robertus *et al.*, 1972). All four inhibitors contained an L-phenylalanine residue at P_1. The polypeptide portion in each case was hydrogen bonded to an extended segment, Ser-125–Leu-126–Gly-127, of the enzyme backbone to form an anti-parallel β structure, with the P_1 phenylalanine side chain fitting snugly into a hydrophobic crevice. One wall of this crevice consisted of the extended backbone segment 125–127.

This geometry represented a plausible model for binding of the P_1–P_3 residues, that is, the acylating portion of polypeptide substrates. Evidence for the validity of this model included the convincing manner in which it explained various aspects of the enzyme's specificity at P_1–P_3 (Morihara *et al.*, 1970) and the virtual identity of the model with that proposed independently for substrate binding to γ-chymotrypsin on the basis of similar experiments by Segal *et al.* (1971).

Several basic questions remained to be answered, however, regarding subtilisin's mechanism of action. What was the effect on the observed binding due to covalent attachment of the inhibitors to His-64 *via* a methylene carbonyl linkage? Would a Michaelis complex with a polypeptide substrate be appreciably different? Is the binding mode the same when P_1 is other than phenylalanine? How would the leaving portion (P_1', P_2', etc.) be bound in the Michaelis complex? What is the

† From the Department of Chemistry, University of California at San Diego, La Jolla, California 92037. *Received July 5, 1972.* This work was supported by research grants from the National Institutes of Health (GM 10928 and GM 16717) and the National Science Foundation (GB 30828X and GB 23054) and by a Public Health Service Research Career Development award to R. A. A. from the National Institute of General Medical Sciences (GM 15401).

‡ Present address: M. R. C. Laboratory of Molecular Biology, Cambridge, England.

significance of the quite different binding observed for Bz-Arg[1] (Wright *et al.*, 1972)? What is a reasonable stereochemical picture for the subsequent catalytic hydrolysis of a bound polypeptide substrate?

In order to explore these questions further we have studied the binding of several polypeptides to subtilisin BPN′ in crystals of the active enzyme, as there is reason to assume that properly chosen polypeptides will exhibit a substrate-like binding geometry. Specifically it has been shown that chymotrypsin catalyzes ^{18}O exchange between *N*-acetyl-L-tryptophan and solvent water (Bender and Kemp, 1957). The obvious interpretation of this phenomenon is that such hydrolysis products of good substrates are bound in a productive mode for some fraction of the time, that is, they act as "virtual substrates." Furthermore, the binding of *N*-formyl-L-tryptophan to α-chymotrypsin as seen crystallographically by Steitz *et al.* (1969) was compatible with this concept since the carboxyl group of the presumed virtual substrate was close to the catalytic site Ser-195 and in a stereochemically plausible orientation for interaction with the charge-relay system. Since we now know that subtilisin and chymotrypsin possess almost identical catalytic and substrate binding sites, we expected that properly selected polypeptides should also bind like virtual substrates in the case of subtilisin.

Before going on to the details, we shall briefly outline our conclusions and their experimental basis. (1) The principal binding mode for peptides that are presumed to be good virtual substrates is almost identical with that found for the polypeptide chloromethyl ketone inhibitors previously described. An important exception is that the hydrogen bond between the amido nitrogen of the specificity residue P_1 and the carbonyl oxygen of Ser-125 is not formed. This hydrogen bond, and its analog in chymotrypsin will henceforth be referred to as the S_1–P_1 hydrogen bond. (2) Other areas on the enzyme surface in the vicinity of the catalytic site are found to possess affinity for portions of the polypeptides tested. We have assumed these "sticky" areas are the S_1', S_2', etc., binding sites. (3) On the basis of (1) and (2) we have built a hypothetical model for the Michaelis complex between subtilisin BPN′ and a good polypeptide substrate. (4) Given the stereochemical constraints imposed by this model we were able to arrive at a plausible geometry for likely intermediates in the catalytic mechanism. (5) The most significant feature of the proposed mechanism is that one of these intermediates, the tetrahedral addition compound, is stabilized by interaction with the enzyme, thereby also stabilizing structurally similar rate-controlling transition states. (6) It is likely that the same conclusions apply as well to chymotrypsin and related serine proteases.

Experimental Section

Materials and X-Ray Methods. All polypeptides for this study (ZGGT, ZAP, ZGGL, and ZGGR) were obtained from Cyclo Chemical Co., Los Angeles. Subtilisin BPN′ was purchased from the Enzyme Development Corp., New York, and purified on a carboxymethylcellulose column as previously described (Robertus *et al.*, 1972). The eluent from that column was concentrated by making the solution 65% in ammonium sulfate, and the resulting protein precipitate was redissolved in 0.05 M sodium acetate (pH 5.9), and desalted on a Sephadex G-25 column equilibrated with the same solvent. The protein was then crystallized in the usual manner (Wright *et al.*, 1969).

Several active subtilisin crystals were then transferred to a 2.1 M ammonium sulfate solution (pH 5.9), containing 0.05 M sodium acetate and a quantity of the polypeptide to be tested for binding to the enzyme. Usually three crystals were transferred to a 0.5-ml aliquot containing 5 mg of the polypeptide. In all cases this was more than sufficient polypeptide to saturate the solution. The crystals were allowed to soak at 4° for 10–20 days.

Crystals were mounted in the usual fashion and data collected on a Hilger and Watts four-circle automatic diffractometer. A single crystal was used to measure the intensity of all unique *hkl* and *hkl̄* reflections to the desired Bragg resolution for each enzyme–polypeptide complex. Data were collected to 2.5-Å resolution for ZGGT and ZGGL, to 2.8-Å resolution for ZAP, and to 3.0-Å resolution for ZGGR. Unit cell parameters were calculated from diffractometer settings and were the same as for the active and PMS·BPN′ crystals within 1% in all cases. Difference-Fourier maps were calculated using $F_{deriv} - F_{nat}$ as coefficients and multiple-isomorphous-replacement phases from PMS·BPN′ (Wright *et al.*, 1969). In this notation F_{deriv} is the reflection amplitude for a given enzyme–polypeptide complex and F_{nat} the reflection amplitude for the native enzyme crystal. All maps were calculated on a 1 Å × 0.7 Å × 1 Å grid and plotted on a 2-cm/Å scale. The root-mean-square difference density, σ, was calculated over all grid points in the unit cell. The range of σ was 0.03–0.04 e/Å^3. Contours on the ZGGT and ZAP maps were drawn at $\pm 3\sigma$ and at subsequent intervals of $\pm 1.5\sigma$; ZGGL was contoured at $\pm 2\sigma$ and subsequent intervals of $\pm \sigma$; and ZGGR at $\pm 3\sigma$ with intervals of $\pm \sigma$. Interpretation of these maps was facilitated by use of an optical comparator (Richards, 1968).

A Fourier map using coefficients $2F_{ZGGL} - F_{PMS\cdot BPN'}$ and multiple-isomorphous-replacement phases from PMS·BPN′ was also calculated. Such a map tends to show the electron density for the enzyme–peptide complex while minimizing the effects of using PMS·BPN′ phases. This map was plotted on a 1-cm/Å scale and positive contours were drawn at intervals of 0.1 e/Å^3. The map was transferred to transparent sheets by means of a Xerox copier and then attached to Plexiglass sheets for vertical spacing.

To estimate the degree of occupancy for the complexes, the entire positive region corresponding to the bound polypeptide in each map was integrated. This result was compared with the integrated density corresponding to the completely occupied ZGGP chloromethyl ketone derivative previously described (Robertus *et al.*, 1972), which served as a normalization standard.

Interpretation of Difference Maps. Each of the four electron-density difference maps corresponding to the four bound polypeptides (ZGGT, ZAP, ZGGL, and ZGGR) contained as its most prominent feature a ribbon of continuous density running essentially parallel to the extended backbone segment 125–127 in the enzyme molecule. It will be recalled that this segment constituted the polypeptide binding site described in our previous paper on binding of polypeptide chloromethyl ketones. The maximum difference density in this ribbon ranged from 6σ to 11σ.

[1] Abbreviations used are: AAGP, acetyl-L-alanylglycyl-L-phenylalanyl; Bz-Arg, *N*-benzoyl-L-arginine; PMS·BPN′, phenylmethanesulfonyl-subtilisin BPN′; Z, benzyloxycarbonyl; ZAGP, benzyloxycarbonyl-L-alanylglycyl-L-phenylalanyl; ZAP, benzyloxycarbonyl-L-alanyl-L-phenylalanine; ZGGL, benzyloxycarbonylglycylglycyl-L-leucine; ZGGP, benzyloxycarbonylglycylglycyl-L-phenylalanyl; ZGGR, benzyloxycarbonylglycylglycyl-L-arginine; ZGGT, benzyloxycarbonylglycylglycyl-L-tyrosine.

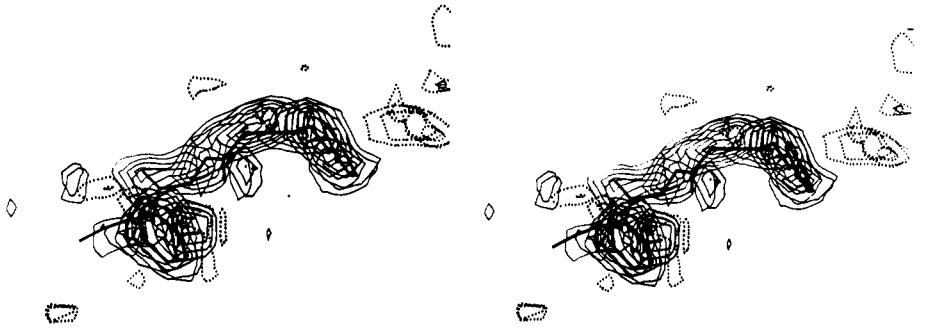

FIGURE 1: Stereoscopic pair showing electron density difference at 2.5-Å resolution between ZGGT·subtilitisn BPN′ complex and native enzyme. Positive and negative contousr, represented by solid and dotted lines, respectively, are drawn at ± 0.08 e/Å^3 ($\pm 3\sigma$) and at subsequent intervals of ± 0.04 e/Å^3 ($\pm 1.5\sigma$). Superimposed is an outline of the ZGGT peptide, with the Z group at the upper right, the carboxyl terminus at the lower left, and the P_1 tyrosine side chain projecting toward the viewer.

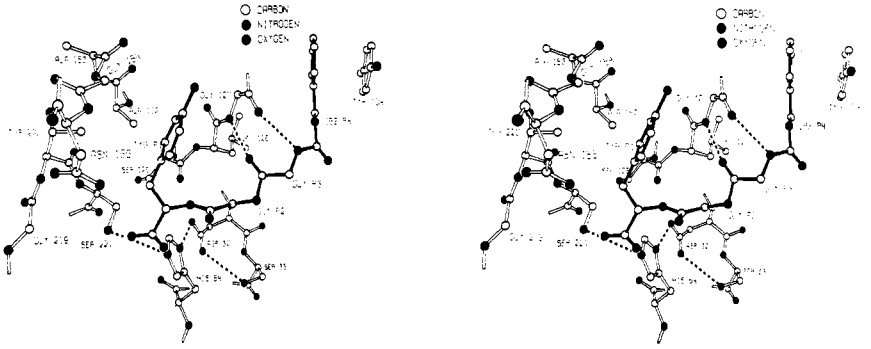

FIGURE 2: Active-site region of subtilisin BPN′ with bound peptide ZGGT. Bonds in the peptide are shown in black for clarity. Cbz denotes the benzyloxycarbonyl group.

A second prominent feature common to the ZGGT, ZGGL, and ZGGR maps but not seen in the ZAP map was a large peak of 4σ–10σ residing in the S_1 crevice occupied by the P_1 phenylalanine side chain of the chloromethyl ketone derivatives.

From these general considerations it is already evident that the model for polypeptide binding arrived at as a result of the chloromethyl ketone study will not be greatly modified by this investigation.

Another feature common to all four difference maps was a cluster of two peaks and a hole of amplitude 4σ to 6σ surrounding the side chain of Tyr-217. This was interpreted as movement of the side chain by 1 or 2 Å in response to additional binding, the precise nature of which could not be defined. However, it seems fairly certain that the effect is real, and not an artifact of the isomorphous replacement phases, as this feature was not observed in the chloromethyl ketone maps or in the subtilisin Novo *minus* subtilisin BPN′ map (Robertus *et al.*, 1971) for which the same set of phases had been used. This "sticky" area will be identified with the S_1' binding site on the leaving group side.

Similarly, a second cluster of prominent peaks and holes of amplitude 3σ to 5σ is observed near the side chain of Phe-189 in all three tripeptide difference maps, but not in the ZAP map. Again, this feature will be identified as the S_2' binding site following the same line of argument as applied to the S_1' site, even though the precise nature of the interaction cannot be defined from these maps alone.

We shall now present a more detailed description of those portions of each of the four difference maps containing the polypeptide binding sites.

BENZYLOXYCARBONYLGLYCYLGLYCYL-L-TYROSINE. This was the clearest of all four difference maps and was readily interpreted. Figure 1 shows a model of the ZGGT polypeptide superimposed upon a slab containing the most prominent feature in the map. Integration of all positive density in that feature showed that the polypeptide binding sites were about 60% occupied.

The binding geometry, as seen in Figure 2, closely resembles that found for tripeptide chloromethyl ketones, with ZGGT bound in anti-parallel β conformation to extended backbone segment Ser-125–Leu-126–Gly-127 and the P_1 tyrosine side chain inserted in the S_1 hydrophobic crevice. The P_1 carboxyl group displaces solvent molecule W-202 and so the added density is weak at this location. However, steric constraints permit the carboxyl group to be unambiguously positioned adjacent to the Ser-221–His-64 portion of the charge relay system.

It is important to note several significant differences between the binding geometry seen here and that of the covalent inhibitors. In general terms, the entire P_1 residue is moved about 1 Å out from binding site S_1, undoubtedly owing to the absence of a covalent bond to His-64. The P_1 tyrosine side chain is about 1 Å less deeply embedded in the S_1 crevice than was the corresponding P_1 phenylalanine side chain of the chloromethyl ketone inhibitors, and unlike the latter the plane of the aromatic ring is tilted about 30° away from parallel with the external wall of the crevice composed of the backbone segment 125–127. The difference map also indicates that the P_1 tyrosine hydroxyl group is hydrogen bonded to a solvent molecule, W-310, lying in the back of the crevice. A peak of 3σ at the W-310 site further suggests

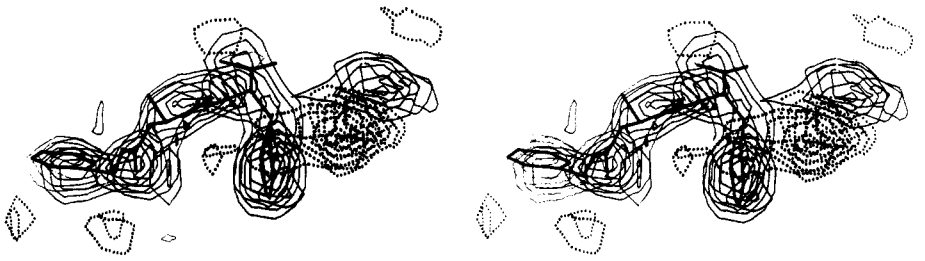

FIGURE 3: Stereoscopic pair showing electron density difference at 2.8-Å resolution between ZAP·subtilisin BPN′ complex and native enzyme. Contours are drawn at ± 0.08 e/Å^3 ($\pm 3\sigma$) and at subsequent intervals of ± 0.04 e/Å^3 ($\pm 1.5\sigma$). Superimposed is an outline of the ZAP peptide, with the Z ring at the left, the carboxyl terminus at the top center and the phenylalanine side chain below it and slightly to the right. Note that the sense of the peptide chain is opposite to that of ZGGT in Figure 1. The large negative and positive peaks at the right represent movement of the Tyr-104 side chain.

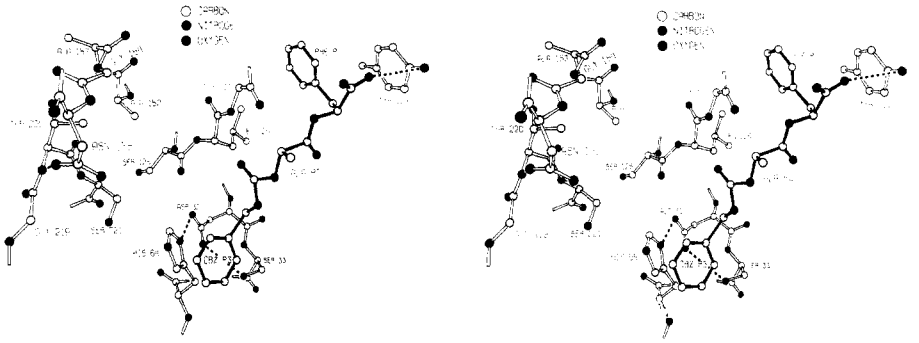

FIGURE 4: Binding of ZAP peptide near the active-site region of subtilisin BPN′.

that it either becomes more completely occupied or that its apparent temperature factor is lowered, or both, upon binding of the tyrosine side chain.

The most important distinction between the binding of ZGGT and that of the covalent inhibitors is the absence of a P_1–S_1 hydrogen bond between the P_1 amido group and the carbonyl oxygen of Ser-125. That this hydrogen bond fails to form in the present case is clear from the fact that the distance between the atoms involved is 4 Å. Indeed there is a 5σ hole in the difference map lying between the amido and carbonyl groups, corresponding to displacement of solvent molecule W-201. This hole is not present in the chloromethyl ketone derivative maps because it is filled in by positive density corresponding to the P_1 amido group.

Interaction of the remainder of the polypeptide, P_2–P_4, with the enzyme is essentially identical at the present resolution with that described for the ZGGP and ZAGP chloromethyl ketone derivatives and need not be further discussed here.

BENZYLOXYCARBONYL-L-ALANYL-L-PHENYLALANINE. This difference map was also clear and relatively easy to interpret. Figure 3 shows a model of the ZAP polypeptide superimposed upon the most prominent region of difference density. A very striking and unexpected result is that the chain runs in the opposite direction from that just described for ZGGT and previously observed for polypeptide chloromethyl ketone derivatives. Integration of the positive region corresponding to the ZAP polypeptide yielded an approximate occupancy of 65%.

The binding geometry is shown in Figure 4. In this case the P_1 phenylalanine side chain occupies the site S_4, where previously only the Z group of benzyloxycarbonyl tripeptide derivatives were found. In three previously described difference maps, ZGGT and two tripeptide chloromethyl ketone derivatives (Robertus *et al.*, 1972), binding the Z group at S_4 was accompanied by some movement of the Tyr-104 side chain, but here for the first time this movement is quite clearly defined. In fact the Tyr-104 side chain is evidently anchored at its new position by a hydrogen bond between the polypeptide carboxyl group and the tyrosine hydroxyl group. Moreover, the P_1 phenylalanine side chain lies about 1 Å deeper into the S_4 crevice than did any of the Z groups.

Although the bound ZAP polypeptide chain is now parallel to backbone segment 125–127, instead of anti-parallel, the two chains appear to be poorly aligned for optimal formation of the expected parallel β-structure hydrogen-bond system.

The Z group occupies a region "normally" occupied by the P_2 side chain, as seen in the ZAP chloromethyl ketone derivative and as inferred from the orientation of the P_2 glycine residue in other derivatives. The S_1 crevice is not occupied in the ZAP·subtilisin BPN′ complex.

BENZYLOXYCARBONYLGLYCYLGLYCYL-L-LEUCINE. The difference map for the complex between this polypeptide and subtilisin BPN′ was considerably more complicated and difficult to interpret unambiguously than the previous two. The most likely interpretation is that the polypeptide binds in two predominant modes at the S_1–S_4 sites to the extent of about 30% occupancy for each mode. In the difference map the prominent and extended feature thus interpreted reaches a density of 7.5σ in the S_4 crevice. This corresponds to 0.3 e/Å^3, or about the same maximum density as that for the well-defined ZGGT complex.

The first of the ZGGL binding modes at P_1–P_4 is essentially the same as that seen for ZGGT, except of course that a leucine side chain replaces a tyrosine side chain in the S_1 crevice. The maximum density associated with this branch of the fea-

ture referred to above is 4σ, and occurs in the S_1 crevice.

The second of the two binding modes proposed for ZGGL has its P_1 leucine residue twisted in such a way that the side chain is down in the vicinity of His-64, approximately where the Z group of ZAP is in Figure 4. The maximum density associated with this branch is also 4σ. Otherwise, the geometry of this second binding mode is probably much the same as the first, and, in fact, P_3 glycine and the Z group appear to coincide for the two.

There is also considerable binding in the vicinity of a surface depression formed by the side chains of Tyr-217, Met-222, and His-67 as evidenced by the peak of height 6σ. A second peak of height 5σ is located in another surface depression under the side chain of Phe-189. Unfortunately these features cannot be unambiguously interpreted to represent a bound ZGGL molecule. As mentioned earlier, similarly uninterpretable, though less pronounced, features also appear in exactly the same locations for the other two tripeptides ZGGT and ZGGR. Most probably these areas on the enzyme surface are "sticky spots" that tend to bind only portions of substrate molecules, leaving the rest to rotate freely.

BENZYLOXYCARBONYLGLYCYLGLYCYL-L-ARGININE. The complex of subtilisin BPN′ with this peptide produced the least satisfactory difference map, but a reasonable interpretation could be made when it was considered in conjunction with the maps previously described. In general terms the ZGGR difference map resembled the ZGGL map except that the second binding mode was relatively stronger than the first. The maximum density associated with the region corresponding to binding mode "one" was 5σ while that for binding mode "two" was 6σ, at what is presumed to be the guanidinium group of the P_1 side chain. As in the case of binding mode "two" for ZGGL, the P_1 residue appears to be twisted down into close proximity with His-64.

The ZGGR difference map, like that for ZGGL, also contained significant peaks of density at the sticky spots mentioned earlier, in the vicinity of Tyr-217 and Phe-189. Their peak heights are respectively 6σ and 5σ.

Binding of Virtual Substrates. The results just described confirm the binding geometry for the acylating portion, P_1–P_4, of polypeptide substrates of subtilisin BPN′ proposed earlier on the basis of studies carried out with covalently bound polypeptide chloromethyl ketone inhibitors (Robertus *et al.*, 1972). That is, the binding geometry seen for ZGGT (Figure 2), a polypeptide which is likely to represent a good virtual substrate, differs little from that seen in the case of the covalently bound tripeptide inhibitors ZGGP and ZAGP, where the P_1 residue is linked *via* a methylene carbonyl bridge to $N\epsilon_2$ of His-64. Furthermore the results reported here from binding studies with the free tripeptides ZGGL and ZGGR are consistent with the idea that this same binding geometry occurs to a significant extent even in the case of these much less specific virtual substrates. Henceforth we shall identify this binding geometry as the "productive mode." Our observation that a tyrosine residue at P_1 causes the polypeptide to bind almost exclusively in the productive mode, whereas leucine or arginine at P_1 are much less effective in this respect, is in accord with a large but amorphous body of kinetic data on hydrolysis of polypeptides, amides, and amino acid esters (Barel and Glazer, 1968; Morihara and Tsuzuki, 1969; Morihara *et al.*, 1970, 1971; Markland and Smith, 1971). These data clearly show that substrates with tyrosine at P_1 are hydrolyzed rapidly, those with leucine at P_1 are hydrolyzed much less rapidly, and substrates with arginine at P_1 are hydrolyzed rather slowly.

However, it must be reiterated that slight but significant differences exist between the presumed productive binding geometry of polypeptide virtual substrates described here and the binding of polypeptide chloromethyl ketones described earlier. In the former case the P_1 residue is more loosely held and, in particular, the distance (4 Å) between the P_1 amido group and the backbone carbonyl oxygen of Ser-125 is too long for hydrogen bonding. We propose that this looser binding geometry represents the Michaelis complex for a virtual substrate. Furthermore we propose that this looser binding is also characteristic of the Michaelis complex for true specific substrates, *i.e.*, amides, esters, and polypeptides.

Michaelis Complex for Polypeptides. Another important question to be dealt with is the location of subsites S_1', S_2', etc., and the geometry of the Michaelis complex on the leaving-group side. Kinetic data for subtilisin BPN′ presented by Morihara *et al.* (1970) show that there is a significant decrease in k_{cat}/K_m when D-alanine is substituted for L-alanine at either P_1' or P_2'. Thus the existence of such subsites, interacting with polypeptide substrates, is highly probable. There is, however, an obvious but important difficulty in directly mapping these S_1' and S_2' subsites, namely that no clear-cut specificity (other than the D,L stereospecificity just referred to) has been discovered for these subsites. Thus it would be difficult to choose a polypeptide which might be expected to bind to those subsites in the natural, productive mode with sufficient stability for X-ray crystallographic study.

Given this situation, how then may one obtain further insight into the binding geometry characteristic of an entire polypeptide substrate? In this communication we report our conclusions based upon model building experiments taking as their point of departure the model for virtual substrate binding just described. These experiments suggest a very plausible model for the Michaelis complex of a polypeptide, and more importantly suggest a convincing model for catalytic intermediates which are stabilized by specific interaction with the enzyme. These models can be incorporated into a mechanism of action for subtilisin BPN′, which is also applicable to serine proteases in general, and which is consistent with the available kinetic data for specific substrates.

We turn now to a detailed description of this proposed Michaelis complex for a polypeptide substrate. Such a hypothetical complex between the enzyme and a representative polypeptide substrate

$$-\text{Phe-Ala-Ala-Tyr}\overset{\downarrow}{-}\text{Leu-Leu}-$$

is shown in Figure 5. The substrate is in position to be cleaved at the bond indicated by the arrow. Thus residue P_1 is tyrosine, residue P_4 is phenylalanine, and residues P_1' and P_2' are both leucine. Reasons for this particular choice of hypothetical substrate to illustrate various features of our model will become apparent as these features are described each in its turn.

Positioning of residues P_1–P_3 is disposed of rather easily. As mentioned earlier, we simply assume that the geometry seen for the ZGGT virtual substrate complex in Figure 2 is also valid for a true polypeptide substrate bearing amino acid residues beyond the tyrosine at P_1. Further, the glycine residues are replaced by alanines at P_2 and P_3 to indicate side-chain orientations at these sites, and these side chains are in fact just where they were found in the covalent inhibitors ZAP, ZAGP, and AAGP.

As mentioned in our discussion of these chloromethyl ketone inhibitors, the Z group bears a strong steric resem-

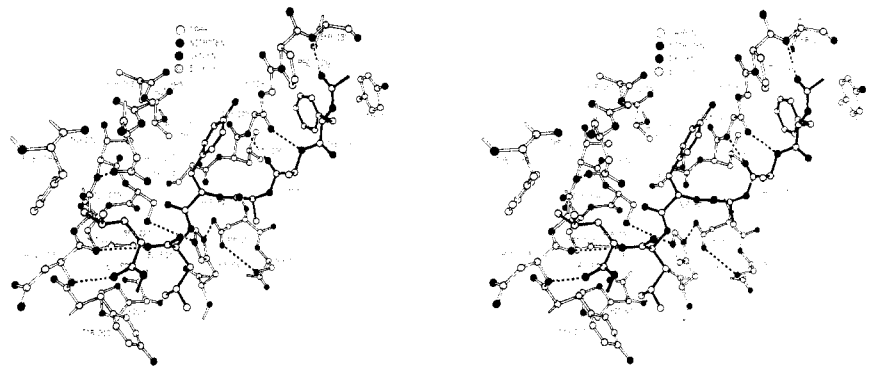

FIGURE 5: Hypothetical Michaelis complex between the peptide substrate -Phe-Ala-Ala-Tyr-Leu-Leu- and the binding channel region of subtilisin BPN'.

blance to a phenylalanine residue, and in building our hypothetical Michaelis complex we have converted the Z group of ZGGT into a phenylalanine at P_4. This is readily accomplished with little disturbance of the binding. The acylating group can then be extended somewhat toward the amino terminus, with the P_4 backbone running along the segment Gly-128–Pro-129. The P_5 carbonyl group may also form a hydrogen bond to the backbone NH of Ser-130, but the substrate chain appears to leave the enzyme surface beyond this point.

Next, we consider the probable location of residues P_1' and P_2' in the Michaelis complex. To begin with it is reasonable to suppose that the Michaelis complex for a polypeptide amide, for example, ZGGT amide, would be essentially the same as for ZGGT and could be represented by simply replacing one of the polypeptide carboxyl oxygens with an NH_2 group. Following Steitz *et al.* (1969) in their interpretation of *N*-formyl-L-tryptophan binding to α-chymotrypsin we conclude that an NH_2 group would replace the oxygen nearer to the active-site His-64, that is, the lower one in Figure 2, since it is now generally accepted that the leaving group is protonated by the active-site histidine during catalysis by serine proteases (Hess, 1971). In addition to this rather compelling rationale, there is another reason for believing that we have made the correct choice in positioning the leaving amine. Had we placed it at the site of the other carboxyl oxygen it would then be impossible to extend the leaving group in a stereochemically sensible fashion by even a single residue. A polypeptide continuing in that direction runs immediately into a pocket in the enzyme surface formed by the side chain of Asn-155 and the backbone of Thr-220 and Ser-221.

Having arrived in this way at a model for the productive Michaelis complex between the enzyme and polypeptide amides, we can now inquire as to what happens when the amide is extended to become an L-amino acid residue. Further model building reveals that the constraints of standard bond lengths and angles within the substrate plus the local topography of the enzyme strongly determine the placement of this P_1' residue. The resulting model has the P_1' backbone in van der Waals contact with both the side chain of His-64 and with the side chain of Met-222, while the carbonyl group points into the surrounding solvent, away from the body of the enzyme. The P_1' residue has been shown in Figure 5 as a leucine, rather arbitrarily chosen because its side chain is of medium size. It is important to note that the side chain of the Leu P_1' has fallen naturally within the sticky spot near the side chain of Tyr-217, justifying our identification of this region as subsite S_1'.

The P_2' residue, again shown in Figure 5 as Leu, can be positioned along the extended chain, Asn-218–Gly-219 of the enzyme backbone. The P_2' amido group forms a hydrogen bond to the backbone carbonyl oxygen of Asn-218 while the P_2' carbonyl oxygen is hydrogen bonded to the Asn-218 backbone amido group, resulting in a short segment of anti-parallel β structure. In exactly the same manner as for the P_1'–S_1' interaction, the P_2' leucine side chain lies within the S_2' sticky spot under the Phe-189 side chain. The existence of a favorable interaction as proposed for the S_2' site appears to be consistent with the available kinetic data. For example, Morihara *et al.* (1970) have shown that extension of

$$\text{Z-Gly-Leu}\overset{\downarrow}{\text{-}}\text{Gly to Z-Gly-Leu}\overset{\downarrow}{\text{-}}\text{Gly-Gly}$$

causes the rate of hydrolysis by subtilisin to increase 380-fold.

The substrate does not appear from our model-building experiments to make specific contacts with the enzyme at sites more remote than P_2' and we have shown it leaving the surface of the protein molecule beyond this point.

In passing, it is interesting to note that if Gly-219 were any residue other than glycine, its side chain would interfere with binding of a polypeptide substrate at S_1' and S_2'. This would account for the observed invariance of Gly-219 (Dayhoff, 1969) in the subtilisin family of proteases.

To summarize, the binding site of subtilisin BPN' is a rather pronounced surface channel capable of accommodating at least six amino acid residues of a polypeptide substrate. The acylating or specificity side of the substrate, P_1–P_4, lines up along the extended enzyme backbone segment, Ser-125–Ser-130, in an anti-parallel β-sheet arrangement, with the substrate side chains being held as well at three subsites, S_1, S_2, and S_4. The leaving portion P_1'–P_2' appears to be held less tightly as it runs along the backbone segment Asn-218–Thr-220 and specific complementary contacts between enzyme and substrate side chains appear to be less extensive than on the P_1–P_4 side.

Two features of our Michaelis complex model deserve emphasis at this point, as they will be invoked again later in our discussion of the enzymic mechanism. They are (1) the S_1–P_1 hydrogen bond is not present in the Michaelis complex, although the groups involved are poised to form such a bond; and (2) solvent molecule W-202 has been displaced in the Michaelis complex. In the native enzyme W-202 is located at the entrance to a hole lined by the side chain of Asn-155, the α carbon of Gly-219, the backbone amido group and side-chain Oγ of active-site Ser-221, Nε2 of active-site His-64,

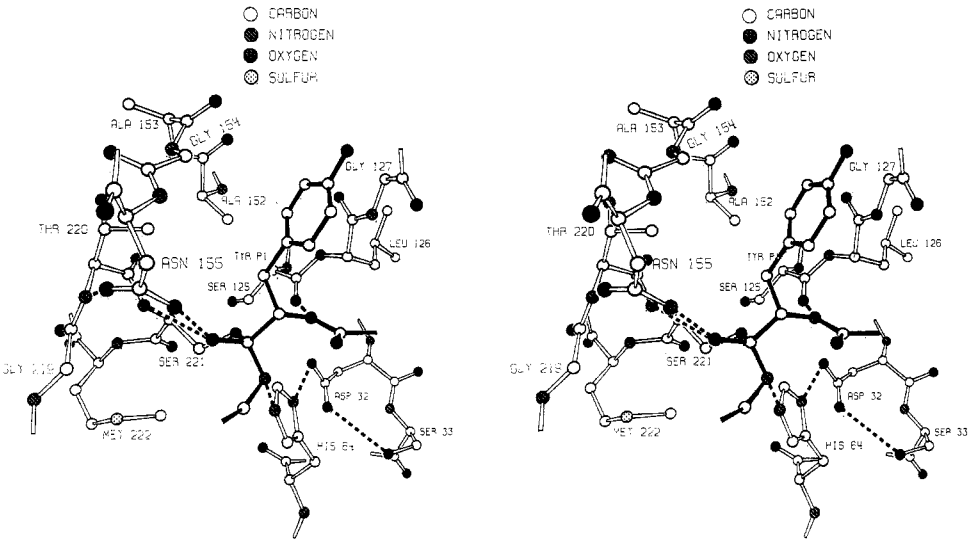

FIGURE 6: Model of tetrahedral addition compound for a good substrate. All features not essential to the discussion in the text are eliminated for simplicity. Removal of leaving group would give rise to acyl intermediate shown in Figure 7. Note that the backbone NH of Ser-221, Nδ2 of Asn-155, and Sδ of Met-222 are arranged in a very rough tetrahedron around the P_1 oxyanion.

and Sδ of Met-222. It appears to be a water molecule hydrogen bonded to Oγ of Ser-221, but it is probably too far from the other neighboring potential hydrogen-bonding groups actually to form acceptable hydrogen bonds with them. We suspect that W-202 and the hole play an important role in the enzymic mechanism.

Considerable support for our model of a Michaelis complex is provided by comparison between the substrate conformation in our model and the conformation of the recognition site in basic pancreatic trypsin inhibitor (Huber *et al.*, 1971). Such a comparison is warranted in view of the great similarity between the catalytic and binding sites of subtilisin and chymotrypsin (Robertus *et al.*, 1972), and in view of the observation that several proteinase inhibitors from avian egg white (Feeney, 1971) will complex both enzymes with concomitant cleavage of the same specific inhibitor peptide bond.

A Kendrew–Watson model of the basic pancreatic trypsin inhibitor was constructed from atomic coordinates kindly provided by Dr. R. Huber. The recognition site was found to fit readily into the subtilisin binding channel, making all of the specific complementary contacts with the P_1–P_2' region already described for our model of the Michaelis complex. Indeed, inhibitor segment Cys-14–Lys-15–Ala-16 has almost identically the same conformation as the P_1'–P_2 segment of the polypeptide substrate in our model. The susceptible Lys-15–Ala-16 bond of the inhibitor can be oriented in the same way as the P_1'–P_1 bond in our model, with the Lys-15 side chain fitting into the S_1 specificity crevice.

Further to quantify the comparison, coordinates for the backbone chain and β-carbon atoms of the P_1'–P_2 residues of the inhibitor were compared to those of the corresponding segment of the model polypeptide by a least-squares rotation procedure. They were found to agree surprisingly well, with a root-mean-square deviation of 0.2 Å, and a maximum deviation of 0.4 Å.

It is also gratifying to note that these model-building experiments provide an explanation for the inability of basic pancreatic trypsin inhibitor to complex with subtilisin despite the similarity between the binding sites of subtilisin and the chymotryspin family of serine proteases. When the recognition site Cys-14–Lys-15–Ala-16 of the basic pancreatic trypsin inhibitor model is seated into the subtilisin binding channel, the backbone and side chains of inhibitor residues Cys-38–Arg-39 overlap completely with an external loop of the enzyme formed by residues Gly-97 through Gly-102. An equivalent interfering loop does not exist in the chymotrypsin family.

Before proceeding we should dispose of questions raised by the apparently anomalous binding geometry found here for ZAP and by Wright *et al.* (1972) for Bz-Arg, and by what appear to be multiple binding modes for ZGGL and ZGGR. We believe all of these are examples of nonproductive binding. From the result obtained with ZAP, we tentatively conclude that the affinity of the S_4 crevice for a phenylalanine side chain is very strong, and in this case caused the predominant binding mode to be abortive. In fact we would suggest that the S_4 crevice may play an important role in determining the cleavage point for polypeptide substrates. It was for this reason we chose to illustrate a hypothetical Michaelis complex with a polypeptide having a phenylalanine at P_4. Bz-Arg, on the other hand, now appears to be a case where predominant binding has occurred at the leaving group site, with the benzoyl group bound at S_2' and the arginine residue at S_1'. Finally the "second" binding mode already described for ZGGL and ZGGR must also represent extensive nonproductive binding, if our proposed model for a Michaelis complex is correct. The important conclusion to be drawn from all of this is that nonproductive binding appears to occur to a significant extent for many synthetic substrates, and should be taken into account when interpreting kinetic data.

Models for Catalytic Intermediates and a Mechanism. Given the detailed model of a Michaelis complex between subtilisin and a polypeptide substrate just described, we can now attempt to arrive at a plausible stereochemical basis for subsequent steps in the catalytic mechanism. In the following discussion we shall assume that subtilisin and the chymotrypsin family of serine proteases share a common transacylation mechanism and that experimental data bearing on this mechanism for the chymotrypsin family of enzymes are also valid for subtilisin.

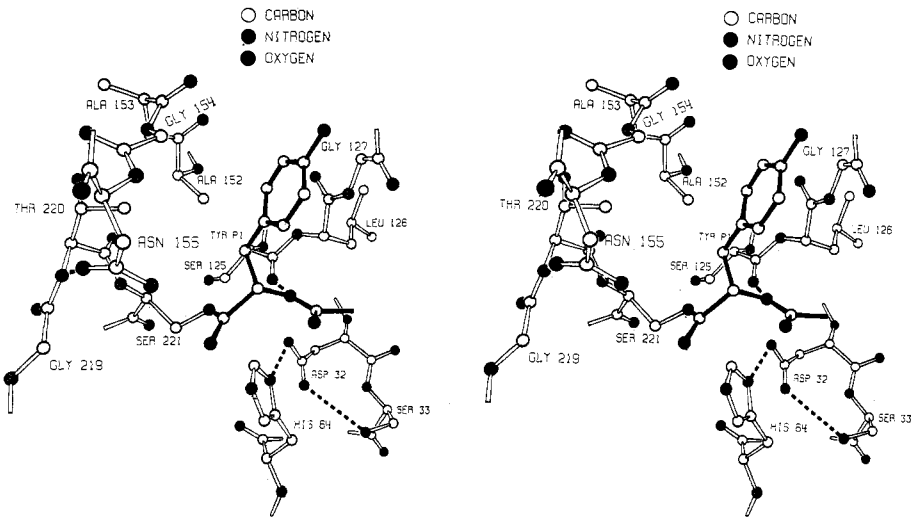

FIGURE 7: Model of an acyl intermediate for a good substrate.

Such an assumption would appear to be amply justified by all available kinetic and structural evidence (Markland and Smith, 1971; Kraut, 1971; Robertus *et al.*, 1972). We should emphasize that this identity applies only to the transacylation mechanism, and not to all aspects of their specificity where obvious differences do occur between the two classes of proteases.

As a first step toward investigating possible mechanisms we break the susceptible bond between P_1 and P_1' and construct a covalent ester linkage to Oγ of Ser-221 to form an acyl intermediate. Evidence for the existence of an acyl intermediate in the course of hydrolysis of ester substrates is abundant, and there is no reason to suppose such an intermediate does not also occur in the case of amide substrates (Hess, 1971). The resulting model is shown in Figure 7. Construction of this model requires rotation about the Cα–Cβ bond of Ser-221 by approximately 120°. Note that the ester group in this model has been constructed in the conventional planar conformation. To make this bond it is necessary to move the entire P_1 residue about 1 Å further in toward the enzyme surface, accompanied by deeper penetration of the P_1 tyrosine side-chain ring into the S_1 specificity crevice. The resulting structure is almost identical with that observed for the polypeptide chloromethyl ketone derivatives (Robertus *et al.*, 1972). Since the P_1 amido group moves to within about 3 Å of the carbonyl oxygen of Ser-125, presumably the S_1–P_1 hydrogen bond is now formed. This hydrogen bond was observed in the chloromethyl ketone inhibitor derivatives (Segal *et al.*, 1971; Robertus *et al.*, 1972) of γ-chymotrypsin and subtilisin and was inferred to play a role in substrate interaction with α-chymotrypsin by Steitz *et al.* (1969) and by Henderson (1970).

We consider nonexistence of the S_1–P_1 hydrogen bond in the Michaelis complex and its subsequent formation in the acyl-enzyme to be an important feature of the enzymic mechanism.

Several observations can now be made concerning the acyl intermediate. (1) The carbonyl oxygen of the ester linkage is constrained to be in the "down" position (Figure 7) to retain favorable interactions between the rest of residue P_1 and the enzyme, especially that between the P_1 side chain and the S_1 specificity crevice. (2) The orientation of the ester linkage and its location with respect to other important groups very closely resembles the corresponding linkage found by Henderson (1970) in indoleacryloyl-α-chymotrypsin. Incidentally, a satisfactory hypothetical model for indoleacryloyl-subtilisin was readily constructed from Henderson's coordinates for the indoleacryloyl group. The relationship between the latter and the enzyme's catalytic and binding residues was found to be almost identical for subtilisin and chymotrypsin. Further it should be noted that carbamyl-α-chymotrypsin also contains a similarly oriented ester linkage (Robillard *et al.*, 1972). In both cases, the ester carbonyl group is pointed "down," as in Figure 7, where it can hydrogen bond to an external water molecule. It should be mentioned that in the two structures, indoleacryloyl- and carbamyl-α-chymotrypsin, differently located water molecules are involved in this interaction, and that our model of the acyl intermediate for subtilisin is not sufficiently reliable to distinguish which of the alternatives it most closely resembles; but this detail is unimportant to the present discussion. (3) Following arguments presented by Henderson (1970) and by Robillard *et al.* (1972) for α-chymotrypsin, it is likely that the ester carbonyl group must be rotated "upward," from its position in Figure 7 in order to undergo hydrolysis. Briefly, their arguments are that the carbonyl oxygen of the acyl group in this conformation blocks access of an attacking activated water molecule to the carbonyl carbon; that the carbonyl group must be rotated about 45° upward to position the p orbital of the carbonyl carbon for maximum overlap with an orbital of an attacking water molecule; and that the carbonyl oxygen in the proposed upward position would be stabilized by hydrogen bonds from the backbone chain NH groups of Ser-195 and Gly-193 in the case of chymotrypsin. (4) The backbone NH of active-site Ser-221 in subtilisin BPN′ lies in the same location with respect to other critical groups as does the corresponding backbone NH of α-chymotrypsin's active-site Ser-195. In fact when the least-squares rotation of corresponding groups in α-chymotrypsin and subtilisin (Robertus *et al.*, 1972) is repeated, as will be described later in more detail, this time including the active-site serine backbone NH groups, the deviation between these two groups is only 0.3 Å. Furthermore, a second potential hydrogen-bond donor group, analogous to α-chymotrypsin's Gly-193 backbone NH, exists in subtilisin BPN′, namely, the side-chain NH_2 of Asn-155. Although these two analogous groups are *not* in closely corresponding locations, the side-chain NH_2 of Asn-155 in our subtilisin model lies in a very

favorable position to donate a hydrogen bond to the carbonyl oxygen atom in the proposed upward position. Since it is unlikely that this additional stereochemical correspondence between subtilisin and chymotrypsin is merely coincidental, we are convinced of the validity of point 3 and that it probably applies to serine proteases generally. We shall henceforth refer to this site as the oxyanion hole. The rationale underlying this nomenclature will develop as we proceed. An interesting footnote to this portion of our discussion is that Sδ of Met-222, the backbone NH of Ser-221, and the side-chain NH_2 of Asn-155 appear to be arranged in a roughly tetrahedral manner about the location of the carbonyl oxygen atom when in its proposed upward orientation. In chymotrypsin the position of Sγ of Cys-42 corresponds to Sδ of Met-222 in subtilisin. The possible functional role of the sulfur atom in the oxyanion hole is not obvious. That it does play some role, however, is suggested by its presence in the same relative location in both subtilisin and chymotrypsin, and by its evolutionary invariance in both families of serine proteases (Dayhoff, 1969; Johnson and Smillie, 1971). (5) We come now to the most significant point of this discussion. When we attempt to construct a model of the planar acyl intermediate with its carbonyl group in the upward orientation we encounter a serious difficulty: the conformation of substrate residue P_1 must be severely distorted from that seen in the Michaelis complex. Specifically, if the P_1 side chain is to remain more or less within the S_1 specificity crevice, and if the P_1–S_1 hydrogen bond is to be retained, it is necessary to rotate the Cα–N bond of P_1 by about 100°, the Cα–Cβ bond by about 90°, and the Cβ–Cγ bond by about 45°. Furthermore, as a result of these bizarre contortions the Cβ atom is forced out of the S_1 crevice, a conformation that has never been observed crystallographically for any of the bound virtual substrates or covalent inhibitors in either subtilisin or chymotrypsin.

In view of these considerations, we decided to reject a structure for the planar acyl intermediate with its carbonyl group pointing up. However, the arguments presented in points 3 and 4 are still quite cogent. Therefore we attempted to find an alternative model for the next step in deacylation in which the carbonyl oxygen can be positioned in the oxyanion hole, while retaining all the favorable interactions between a good substrate and the enzyme.

There is one obvious possibilty. Recent evidence points to the existence of a tetrahedral intermediate in the acylation of α-chymotrypsin by a variety of substrates (Caplow, 1969; Frankfater and Kézdy, 1971; Ferscht and Requena, 1971; Ferscht, 1972), suggesting that we try to fit such a structure into our subtilisin model. The result was satisfactory in that we could place the "carbonyl" oxygen in the oxyanion hole while maintaining the other specific substrate–enzyme interactions. Thus, we conclude that the enzyme's structure is in fact complementary to and stabilizes the proposed tetrahedral intermediate.

All of the foregoing enables us to propose the following stereochemical scenario for hydrolysis of a good polypeptide substrate by subtilisin BPN′. (1) A Michaelis complex forms with the geometry shown in Figure 5 and already described in detail. It is pertinent to recall that W-202 is displaced in this step. In the uncomplexed enzyme W-202 is at the molecular surface, blocking the entrance to the oxyanion hole; hydrogen bonded to Oγ of Ser-221 and close to Nε2 of His-64. Also the S_1–P_1 hydrogen bond is poised but not formed in the Michaelis complex; the carbonyl carbon of the susceptible peptide bond is favorably oriented for attack by Oγ of Ser-221, and the overall geometry is close to that proposed by Wang (1968, 1970) for optimal proton transfer to the amino terminus of the leaving group. Various features of the Michaelis complex are schematically illustrated in Figure 8a. (2) The next intermediate for which we can present a structural model is the tetrahedral addition compound. It is illustrated in Figure 6. For transition to take place from the Michaelis complex to the tetrahedral addition compound, several events must occur, although we cannot say from the present data in what order. Indeed these events may be concerted, and the following enumeration is not meant to imply any necessary temporal sequence. (a) The charge relay system (Blow *et al.*, 1969; Alden *et al.*, 1970) is activated, possibly assisted by desolvation of the side chain of Ser-221 when W-202 is removed, resulting in transfer of a proton from Oγ of Ser-221 to Nε2 of His-64. This step is shown schematically in Figure 8b. (b) The Ser-221 side chain swings upward about 120° while attacking the carbonyl carbon of the susceptible peptide bond. (c) The carbonyl group distorts to a tetrahedral conformation, rotating the carbonyl oxygen upward and outward about 50° and into the oxyanion hole. Note that at this stage an initial formal negative charge on buried Asp-32 of the charge relay system now resides on the "carbonyl" oxygen. This structure is stabilized by hydrogen bonds from the backbone NH group of Ser-221 and the side-chain NH_2 group of Asn-155. For these reasons we have called this site the oxyanion hole. It is important to note, however, that the negative charge is only a formal one, and in fact kinetic evidence suggests there is little charge development during hydrolysis of specific substrates by chymotrypsin (Jencks, 1971). (d) A hydrogen bond is formed between Nε2 of His-64 and the leaving NH group; possibly complete proton transfer occurs at this stage. (e) The P_1 residue moves about 1 Å toward the enzyme surface, resulting in deeper penetration of the S_1 specificity crevice by the P_1 side chain and formation of the S_1–P_1 hydrogen bond. The tetrahedral addition compound resulting from events b–e is schematically illustrated by Figure 8c. Since the degree of solvation in the tetrahedral intermediate is presumably the same as for the Michaelis complex, all of the energy gained from hydrogen-bond formation can be applied to the stabilization of the tetrahedral intermediate. (3) The third distinct intermediate for which a reasonable structure can be proposed is the acyl-enzyme. A model of this structure is shown in Figure 7, and a schemtic diagram is given in Figure 8d. To arrive at this intermediate from the tetrahedral addition compound, it is necessary to (a) complete the process of proton transfer to the leaving group, if this had not already been achieved at stage 2d; (b) break the susceptible C–N bond; (c) and to collapse the carbonyl carbon of the acyl-enzyme back to the planar conformation as the leaving group diffuses away. Note that Figure 7 depicts the carbonyl oxygen in the downward orientation, for reasons already discussed at length. As implied by these considerations, it would be sterically conceivable to retain the oxygen in the oxyanion hole only if the carbonyl carbon is tetrahedrally hybridized. However, such a situation is energetically so unfavorable as to make its occurrence very unlikely. (4) Deacylation occurs by replacing the leaving amino group with a water molecule and reversing the entire foregoing sequence of events. The entering water molecule would be oriented so as to donate a hydrogen bond to Nε2 of His-64, and in this sense would be activated by the charge relay system. It is quite possible for the side chain of His-64 to move slightly in this step, for optimum orientation. We know from the structure of PMS·BPN′ and from the polypeptide chloromethyl ketone

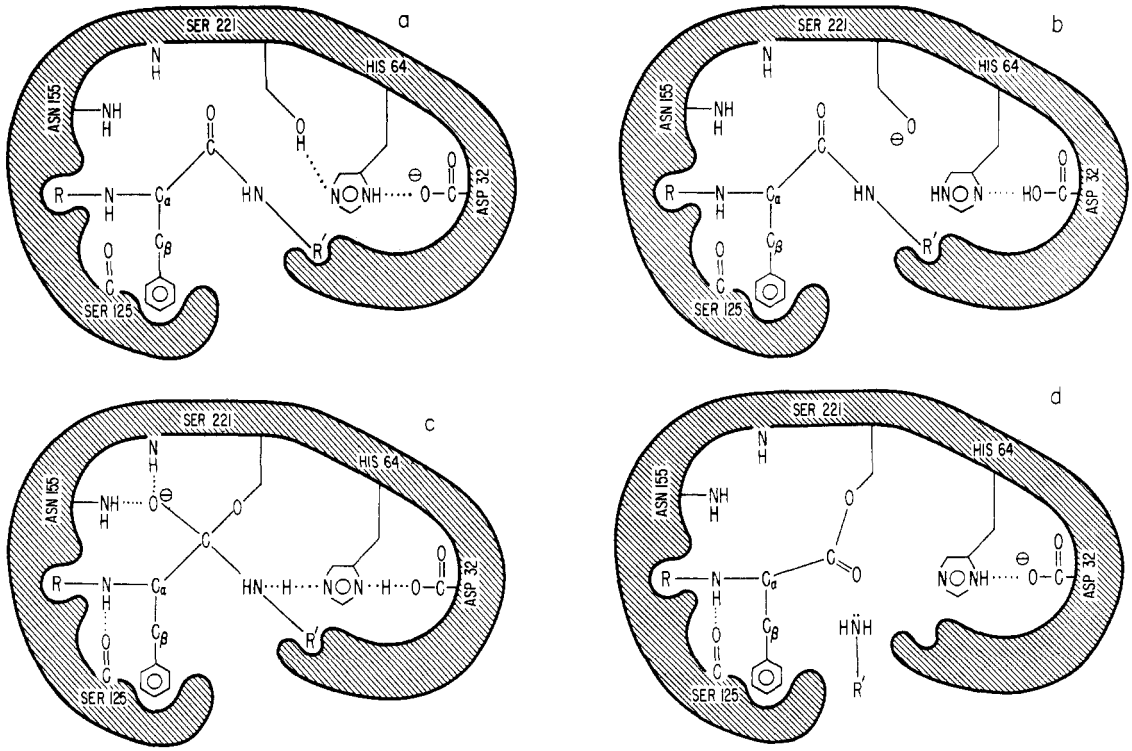

FIGURE 8: A schematic representation of the mechanism of action of subtilisin BPN′ for a good polypeptide substrate. Substrate residues P_2–P_4 are represented by R, residues P_1' and P_2' by R′. (a) Michaelis complex. (b) Activated charge relay system. (c) Tetrahedral addition product. (d) Acyl intermediate, with leaving group diffusing away.

derivative structures that the His-64 side chain is free to swivel to some extent.

It will be evident that the stereochemical mechanism proposed here for subtilisin is very similar to that proposed for chymotrypsin (Henderson, 1970; Henderson *et al.*, 1971). However we believe the differences in the two proposed schemes are very important because our scheme is consistent with transition-state theory as applied to enzymes (Wolfenden, 1972) whereas the chymotrypsin scheme would appear to be inconsistent with that theory.

The differences between the scheme presented here for subtilisin and the current one for chymotrypsin are (1) the Michaelis complex for subtilisin is *not stabilized* by forming the S_1–P_1 hydrogen bond, nor by having the carbonyl oxygen of the susceptible bond inserted into the oxyanion hole; and (2) the acyl-enzyme is *not stabilized* by having the ester carbonyl group oriented upward with its carbonyl oxygen in the oxyanion hole.

Although these differences may at first glance appear to be minor, we believe they are critical. The reason is that in the subtilisin scheme we have provided a stereochemical mechanism for stabilizing transition states *with respect to* both the Michaelis complex and the acyl-enzyme. This stabilization is achieved *via* the same favorable interactions occurring in the tetrahedral addition intermediate (Figures 6 and 8c), which must be very close in geometry to rate-determining transition states (Hammond, 1955). In other words, our stereochemical model predicts that the free energy of activation separating the Michaelis complex from the transition state is decreased by formation of at least three hydrogen bonds in the latter. Although it is very difficult to calculate with any precision the magnitude of such an effect on rate constants, it would evidently suffice to increase k_2 (the acylation rate constant) by about 10^8–10^9! Clearly similar considerations apply as well to the transition states between acyl-enzyme and products. In any case, according to Wolfenden (1972) "during enzyme reactions the substrate is bound to the enzyme very much more tightly than in the Michaelis complex" and "if the transition state were not more tightly bound, *catalysis would not occur*."

To be sure, Henderson (1970) did state that the tetrahedral intermediate might be stabilized by hydrogen bonds to the carbonyl oxygen, but in Henderson *et al.* (1971) essentially the same interactions with the enzyme are depicted for the tetrahedral intermediate as for the Michaelis complex and the acyl-enzyme. In any case we believe our scheme actually applies to the chymotrypsin family of serine proteases as well as to subtilisin. One important reason is that we were unable to construct a stereochemically plausible model of an acyl-chymotrypsin with a planar carbonyl group pointing upward so that its oxygen lay in the oxyanion hole, as depicted in Figure 4, III of Henderson *et al.* (1971). In these model building experiments, we used α-chymotrypsin coordinates from Birktoft and Blow (1972) and maintained constraints on a hypothetical polypeptide substrate similar to those described by Segal *et al.* (1971). As was the case for subtilisin, however, a tetrahedral intermediate with its "carbonyl" oxygen in the oxyanion hole could be constructed readily, although it should be noted that the backbone NH of Gly-193 did not seem to be quite optimally oriented for hydrogen bonding in our model.

Further support for the correctness of our stereochemical mechanism comes from recent kinetic data on chymotrypsin. Ingles and Knowles (1968) have presented evidence that can be interpreted as showing that the presence of an acylamido group at P_1 is important for catalysis but not for initial Michaelis complex formation. In fact, in agreement with this notion is the observation of Steitz *et al.* (1969) that the S_1–P_1 hydro-

gen bond is *not formed* in the complex of *N*-formyl-L-tryptophan with α-chymotrypsin. Further, Henderson (1970) notes that the indolyl ring of indoleacryloyl-α-chymotrypsin lies 0.5–1.0 Å deeper into the S_1 specificity crevice than does the corresponding ring in the complex of *N*-formyl-L-tryptophan with chymotrypsin. All of this is nicely consistent with our subtilisin mechanism in which the S_1–P_1 hydrogen bond is poised but not formed in the Michaelis complex.

Finally, kinetic data implicating the oxyanion hole in the chymotrypsin mechanism have been reported by Williams (1970). He interpreted substituent effects on hydrolysis rates of aryl hippurates as "evidence for electrophilic assistance at the carbonyl oxygen during the formation of a tetrahedral intermediate."

To summarize, then, we believe that the same arguments leading to our stereochemical mechanism for subtilisin also apply to chymotrypsin. Indeed, the possibility of deriving identical stereochemical mechanisms for these two otherwise unrelated enzymes constitutes strong support for the essential correctness of that mechanism.

In our previous publication (Robertus *et al.*, 1972) we compared the geometries of subtilisin's and chymotrypsin's catalytic and binding sites by computing a least-squares rotation to match the coordinates of 27 corresponding atoms. The mean deviation was 0.8 Å. We are now aware of a third common structural feature, namely, the oxyanion hole. To extend this comparison the least-squares calculation was repeated, this time including in addition the backbone NH of Ser-221, Nδ2 of Asn-155, and Sδ of Met-222 for subtilisin, on the one hand, and the corresponding NH of Ser-195, the backbone NH of Gly-193, and Sγ of Cys-42 for chymotrypsin, on the other hand. The result, not surprisingly, showed that the backbone NH groups of the active serine residues and the corresponding sulfur atoms in the two enzymes are in essentially the same relative locations as the rest of the active site groups, with deviations of 0.3 and 1.1 Å, respectively. The agreement between Nδ2 of Asn-155 (subtilisin) and the backbone NH of Gly-193 (chymotrypsin) however was less satisfactory, as they showed a discrepancy of 2.8 Å. Nevertheless it should be emphasized that there is little doubt Nδ2 of Asn-155 in subtilisin is in a good location for hydrogen-bond donation to the oxyanion. Furthermore, the coordinates of this atom used for the above comparison (Alden *et al.*, 1971) are relatively uncertain and refinement currently in progress suggests that a much closer correspondence between the two equivalent groups will ultimately emerge. It is also worth pointing out that a good hydrogen bond can be formed between the Oδ1 of Asn-155 and the backbone NH of Thr-220, holding this key side chain firmly in position.

Acknowledgment

We are grateful to Professor Robert Fahey for helpful discussions concerning likely reaction mechanisms and transition states.

References

Alden, R. A., Birktoft, J. J., Kraut, J., Robertus, J. D., and Wright, C. S. (1971), *Biochem. Biophys. Res. Commun. 45*, 337.

Alden, R. A., Wright, C. S., and Kraut, J. (1970), *Phil. Trans. Roy. Soc. London, Ser. B 257*, 119.

Barel, A. O., and Glazer, A. N. (1968), *J. Biol. Chem. 243*, 1344.

Bender, M. L., and Kemp, K. C. (1957), *J. Amer. Chem. Soc. 79*, 116.

Birktoft, J. J., and Blow, D. M. (1972), *J. Mol. Biol. 68*, 187.

Blow, D. M., Birktoft, J. J., and Hartley, B. S. (1969), *Nature (London) 221*, 337.

Caplow, M. (1969), *J. Amer. Chem. Soc. 91*, 3639.

Dayhoff, M. O. (1969), Atlas of Protein Sequence and Structure, Vol. 4, Silver Spring, Md., National Biomedical Research Foundation, p 49.

Feeney, R. E. (1971), *in* Proceedings of the International Research Conference on Proteinase Inhibition, Fritz, H., and Tschesche, H., Ed., Berlin, Walter de Gruyter, p 189.

Ferscht, A. R. (1972), *J. Amer. Chem. Soc. 94*, 293.

Ferscht, A. R., and Requena, Y. (1971), *J. Amer. Chem. Soc. 93*, 7079.

Frankfater, A., and Kézdy, F. J. (1971), *J. Amer. Chem. Soc. 93*, 4039.

Hammond, G. S. (1955), *J. Amer. Chem. Soc. 77*, 334.

Henderson, R. (1970), *J. Mol. Biol. 54*, 341.

Henderson, R., Wright, C. S., Hess, G. P., and Blow, D. M. (1971), *Cold Spring Harbor Symp. Quant. Biol. 36*, 63.

Hess, G. P. (1971), *Enzymes 3*, 213.

Huber, R., Kukla, D., Rühlmann, A., and Steigemann, W. (1971), *Cold Spring Harbor Symp. Quant. Biol. 36*, 141.

Ingles, D. W., and Knowles, J. R. (1968), *Biochem. J. 108*, 561.

Jencks, W. P. (1971), *Cold Spring Harbor Symp. Quant. Biol. 36*, 1.

Johnson, P., and Smillie, L. B. (1971), *Can. J. Biochem. 49*, 548.

Kraut, J. (1971), *Enzymes 3*, 547.

Markland, F. S., Jr., and Smith, E. L. (1971), *Enzymes 3*, 561.

Morihara, K., Oka, T., and Tsuzuki, H. (1970), *Arch. Biochem. Biophys. 138*, 515.

Morihara, K., and Tsuzuki, H. (1969), *Arch. Biochem. Biophys. 129*, 620.

Morihara, K., Tsuzuki, H., and Oka, T. (1971), *Biochem. Biophys. Res. Commun. 42*, 1000.

Richards, F. M. (1968), *J. Mol. Biol. 37*, 225.

Robertus, J. D., Alden, R. A., Birktoft, J. J., Kraut, J., Powers, J. C., and Wilcox, P. E. (1972), *Biochemistry 11*, 2439.

Robertus, J. D., Alden, R. A., and Kraut, J. (1971), *Biochem. Biophys. Res. Commun. 42*, 334.

Robillard, G. T., Powers, J. C., and Wilcox, P. E. (1972), *Biochemistry 11*, 1733.

Segal, D. M., Powers, J. C., Cohen, G. H., Davies, D. R., and Wilcox, P. E. (1971), *Biochemistry 10*, 3728.

Steitz, T. A., Henderson, R., and Blow, D. M. (1969), *J. Mol. Biol. 46*, 337.

Wang, J. H. (1968), *Science 161*, 328.

Wang, J. H. (1970), *Proc. Nat. Acad. Sci. U. S. 66*, 874.

Williams, A. (1970), *Biochemistry 9*, 3383.

Wolfenden, R. (1972), *Accounts of Chem. Res. 5*, 10.

Wright, C. S., Alden, R. A., and Kraut, J. (1969), *Nature (London) 221*, 235.

Wright, C. S., Alden, R. A., and Kraut, J. (1972), *J. Mol. Biol. 66*, 283.