Stein, O. J., Vanderhoek, J., & Stein, Y. (1976) *Biochim. Biophys. Acta 431*, 347.
Stein, O., Friedman, G., Chajek-Shaul, T., Halperin, G., Olivecrona, T., & Stein, Y. (1983) *Biochim. Biophys. Acta 750*, 306.
Stein, Y., Dabach, Y., Hollander, G., Halperin, G., & Stein, O. (1983) *Biochim. Biophys. Acta 752*, 98.
Tabas, I., & Tall, A. R. (1984) *J. Biol. Chem. 259*, 13897.
Tikkanen, M. J., Nikkila, E. A., Kuusi, T., & Sipinen, S. (1981) *Clin. Chim. Acta 115*, 63.
Tikkanen, M. J., Nikkila, E. A., Kuusi, T., & Sipinen, S. (1982) *J. Clin. Endocrinol. Metab. 54*, 1113.
Van't Hooft, F., van Gent, T., & van't Tol, A. (1981) *Biochem. J. 196*, 877.
Van Tol, A., van Gent, T., van't Hooft, F., & Vlaspolder, F. (1978) *Atherosclerosis (Shannon, Irel.) 29*, 439.
Van Tol, A., van Gent, T., & Jansen, H. (1980) *Biochem. Biophys. Res. Commun. 94*, 101.
Wehrli, F. W., & Wirthlin, J. (1978) *Interpretation of Carbon-13 NMR Spectra*, p 145, Heyden, London.

Stereochemical Analysis of Peptide Bond Hydrolysis Catalyzed by the Aspartic Proteinase Penicillopepsin†

Michael N. G. James* and Anita R. Sielecki

Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received August 17, 1984

ABSTRACT: The X-ray crystal structures of native penicillopepsin and of its complex with a synthetic analogue of the inhibitor pepstatin have been refined recently at 1.8-Å resolution. These highly refined structures permit a detailed examination of peptide hydrolysis in the aspartic proteinases. Complexes of penicillopepsin with substrate and catalytic intermediates were modeled, by using computer graphics, with minimal perturbation of the observed inhibitor complex. A thallium ion binding experiment shows that the position of solvent molecule O39, between Asp-33(32) and Asp-213(215) in the native structure, is favorable for cations, a fact that places constraints on possible mechanisms. A mechanism for hydrolysis is proposed in which (1) Asp-213(215) acts as an electrophile by protonating the carbonyl oxygen of the substrate, thereby polarizing the carbon–oxygen bond, (2) a water molecule bound to Asp-33(32) (O284 in the native structure) attacks the carbonyl carbon as the nucleophile in a general-base mechanism, (3) the newly pyramidal peptide nitrogen is protonated, either from the solvent after nitrogen inversion or by an internal proton transfer via Asp-213(215) from a hydroxyl of the tetrahedral carbon, and (4) the tetrahedral intermediate breaks down in a manner consistent with the stereoelectronic hypothesis. The models permit the rationalization of observed subsite preferences for substrates and may be useful in predicting subsite preferences of other aspartic proteinases.

An understanding of the hydrolytic mechanism of the aspartic proteinase family is of paramount importance, due to the roles that some of its members have in the control of a variety of biological processes (e.g., renin in hypertension, cathepsins D and E in protein turnover). Toward this end, the tertiary structures of three microbial aspartic proteinases have been determined (Hsu et al., 1977a,b; Subramanian et al., 1977a,b; Jenkins et al., 1977). The structure of porcine pepsin (Andreeva et al., 1984) is highly homologous to these enzymes. Inhibitor binding to *Rhizopus chinensis* pepsin (Bott et al., 1982) and to penicillopepsin (James et al., 1982) has also been analyzed by crystallographic techniques. On the basis of these structures, possible binding modes of substrates and catalytic mechanisms have been suggested (James, 1980; James et al., 1977, 1981; Blundell et al., 1980; Andreeva et al., 1981; Foltmann, 1981; Bott et al., 1982). Unfortunately, none of the proposals have been based on refined crystal structures and therefore require further examination in light of the more accurate atomic coordinates now available.

The crystal-structure refinement of penicillopepsin at 1.8-Å resolution has recently been completed (James & Sielecki, 1983). This refinement showed a solvent peak, O39, bound between the carboxyl groups of Asp-33(32)[1] and Asp-213(215) in the final electron density map. The size of this peak is consistent with a neutral water molecule or with an ion (NH_4^+ or H_3O^+). Further experiments reported in this paper indicate that this site is a favorable cationic binding site. This fact imposes important limitations on the hydrolytic pathway.

Additionally, the synthetic analogue of pepstatin Iva-Val-Val-Sta-OEt [Iva = isovaleryl; Sta-OEt = ethyl ester of statine [(4*S*,3*S*)-4-amino-3-hydroxy-6-methylheptanoic acid]] has been refined at 1.8-Å resolution (James et al., 1983). The results of these higher resolution structural studies have given rise to more concrete proposals for the productive mode of substrate binding to penicillopepsin (Hofmann et al., 1984).

The earlier extensive chemical and kinetic data on pepsin and synthetic substrates [see reviews by Clement (1973), Fruton (1970, 1976, 1977), Hofmann (1974), and Knowles

† This research was supported by grants from the Medical Research Council of Canada to the Group in Protein Structure and Function at Edmonton.

[1] The sequential numbering of penicillopepsin (James & Sielecki, 1983) is used throughout; the corresponding residue numbering of porcine pepsin (Tang et al., 1973) is in parentheses.

0006-2960/85/0424-3701$01.50/0 © 1985 American Chemical Society

(1970)] have allowed for the deduction of the following structure–function relationships. Two aspartic acid groups, Asp-33(32) and Asp-213(215), are directly involved in the covalency changes experienced by substrates. The pH dependence of the kinetic parameters for pepsin-catalyzed hydrolysis [reviewed by Clement (1973)] suggests pK_a values of $\simeq$1.2 and 4.7, although it is not established which of the above-mentioned carboxyl groups has the lower pK_a. On the basis of the geometrical disposition of the groups in penicillopepsin at pH 4.4, the proton resides predominantly on Asp-213(215) (James & Sielecki, 1983). The detailed analysis of k_{cat}/K_m data for many synthetic substrates with pepsin suggested an extended binding region for oligopeptide substrates (Fruton, 1970, 1976). The subsequent crystal-structure analyses of the fungal enzymes showed that aspartic proteinases were essentially all β-structures with two major similarly folded domains, the junction of which formed this extended binding cleft and active-site region. Enzyme flexibility (Hollands et al., 1969; Fruton, 1970) and conformational changes induced upon peptide binding to penicillopepsin (Wang et al., 1974) have important roles in catalysis. One of the many conformational changes possible when substrates (or inhibitors) bind has been observed crystallographically for the residues of the "flap", Ser-72(72)–Ser-82(81) (James et al., 1982, 1983). Catalysis should also be affected by the presence of a basic residue, arginine in pepsin (Knowles, 1970) and lysine in penicillopepsin (Grippon & Hofmann, 1981). The only common basic residue between these two enzymes is located at position 304(308), a partially buried residue in both structures. A previously suggested charge relay scheme (James et al., 1977) involving this buried basic residue, along with Asp-300(304), is nonetheless not operable in these enzymes in light of the newer structural results (James & Sielecki, 1983). We propose herein an alternative explanation for the role of this residue in the reported inactivation experiments. Chemical analysis of products in transpeptidation reactions (Wang & Hofmann, 1976; Newmark & Knowles, 1975) and the determination of the order of product release for a variety of substrates with pepsin have brought into question the involvement of a covalently attached acyl- or amino-enzyme intermediate in the aspartic proteinase catalytic pathway (Fruton, 1976, 1977). Fruton's deeply insightful suggestion of no covalent intermediates is supported by the X-ray crystal-structure studies (Hsu et al., 1977b; James et al., 1977). The hydrolytic mechanisms subsequently advanced for aspartic proteinases based on the crystallographic results have not involved covalent intermediates (James et al., 1977, 1981; Bott et al., 1982). More recently, cryoenzymatic studies of pepsin (Dunn & Fink, 1984) and of penicillopepsin (Hofmann & Fink, 1984) designed to trap either amino or acyl intermediates of good substrates of these enzymes on the hydrolytic pathway were also unsuccessful.

Proteolytic enzymes have three common features in their catalytic apparatus (Drenth, 1980; James, 1980): an electrophilic component to enhance the polarization of the carbonyl bond of the substrate, a nucleophilic component to attack the substrate's carbonyl carbon atom, and a proton donor to make the peptide nitrogen atom a good leaving group. It is generally accepted that amide hydrolysis proceeds through a series of intermediates, starting with a tetrahedral species that results from the nucleophilic attack on the carbonyl carbon atom of the scissile bond. For the serine and thiol proteinases, both of which involve a covalently attached tetrahedral intermediate, the nature of the nucleophile is well established. In the case of the aspartic proteinases several candidates have been proposed for the electrophile, the nucleophile, and the proton donor.

There is a consensus among those working with structural data that the electrophilic component of the aspartic proteinase mechanism is the proton shared by Asp-213(215) and Asp-33(32) in a hydrogen-bonded interaction (James et al., 1977, 1981; Bott et al., 1982). There is also agreement on the general-base mechanism associated with nucleophilic attack by a hydroxide ion on the carbonyl carbon of the substrate, with the general base being the carboxylate of Asp-33(32). The determination of the solvent structure carried out during the refinement of penicillopepsin has enabled us to identify the position of water molecule O284 as the most likely attacking nucleophile. The source of the proton for the leaving group nitrogen is less clear. It had been suggested that the proton donor was the phenolic hydroxyl group of Tyr-75(75) (James et al., 1977, 1981). Inhibitor binding studies with penicillopepsin (James et al., 1982) and with *Rhizopus* pepsin (Bott et al., 1982) clearly ruled out that possibility. Bott and co-workers (Bott et al., 1982; Bott & Davies, 1983) propose that the carbonyl oxygen of Gly-35(34) facilitates the transfer of protons from Asp-33(32) to the hydroxide and from the hydroxide to the amide nitrogen. Such a role for a carbonyl oxygen is chemically unsound, and the proposal has most likely resulted from inaccurate atomic coordinates. The *Rhizopus chinensis* pepsin structure and its complex with pepstatin have been determined at 2.5-Å resolution; the amino acid sequence has not been completed, and the refinement of the native molecule is still at an intermediate stage. As a result, the proposed substrate binding mode for the *Rhizopus* enzyme, while similar to our earlier proposals for penicillopepsin (James, 1980), differs in important details from the productive binding mode presented here. The shortcomings of their model of the tetrahedral intermediate are suggested by the presence of interatomic distances considerably shorter than accepted van der Waals contacts between atoms [Table I of Bott et al. (1982)]. In this paper we propose that the source for protonation of the leaving group nitrogen is the bulk solvent, for those proteinases exhibiting acidic pH optima, or the proton may be otherwise transferred via Asp-213(215). Moreover, in our analysis of substrate binding (Hofmann et al., 1984), based on highly refined structures at 1.8-Å resolution, we show that the carbonyl oxygen of Gly-35(34) is most probably involved with substrate binding as a hydrogen-bond acceptor.

Additional restrictions on the stereochemical course of the formation and breakdown of the tetrahedral intermediate in peptide bond hydrolysis can be realized in terms of the stereoelectronic hypothesis (Deslongchamps, 1975; Bizzozero & Dutler, 1981). The possible orientation of lone pair electrons on heteroatoms antiperiplanar to bonds undergoing formation or cleavage on a tetrahedral carbon atom governs the course of the reaction. It is equivalent to the principle of least nuclear motion (Hine, 1977) and provides, in many cases, a useful rule of thumb for the predictions of products in acyl hydrolysis. The limitations of this hypothesis have been indicated for the case of glycoside hydrolysis (Hosie et al., 1984).

In this paper we analyze the stereochemical course of the hydrolytic reaction of peptide bonds catalyzed by aspartic proteinases. Important similarities and differences to the hydrolytic pathway of serine proteinases are also discussed. The analysis of this pathway provides a rationale for the observed subsite preference for certain amino acids in the pepsin hydrolysis of proteins (Powers et al., 1977) and can be used to predict the cognitive interaction of aspartic proteinase inhibitors.

Materials and Methods

The purification of penicillopepsin has been described (Sodek & Hofmann, 1970). Crystals of native penicillopepsin are grown from $(NH_4)_2SO_4$ buffered with $NaC_2H_3O_2$ to pH 4.4 and subsequently transferred to 2.5 M Li_2SO_4–0.1 M $NaC_2H_3O_2$ solution, pH 4.4 (Hsu et al., 1977a). In order to test the hypothesis that solvent site O39 in the active site of the native enzyme is occupied by a positively charged ion, we soaked a crystal for 24 h in a solution of mother liquor containing 38 mM thallium acetate. Intensity data to a resolution of 2.5 Å were collected on an Enraf-Nonius CAD-4 diffractometer. Crystallographic details are similar to those described for crystals of the native enzyme (James & Sielecki, 1983) or for the pepstatin analogue inhibited enzyme crystals (James et al., 1982). The difference electron density map was computed with coefficients $||F_{Tl}| - |F_N||$ and phases α_N, where $|F_{Tl}|$ and $|F_N|$ are the structure-factor amplitudes of the thallium acetate soaked crystal and the native penicillopepsin crystal, respectively; α_N's are the calculated phase angles from the refined native penicillopepsin structure at 1.8-Å resolution and at an agreement factor $R = 0.136$ [see James & Sielecki (1983)].[2]

All molecular modeling was done on an MMS-X interactive computer graphics system (Barry et al., 1976). The refined atomic coordinates of penicillopepsin and those of penicillopepsin in complex with the pepstatin analogue inhibitor formed the fixed reference frame of the display. These coordinates were not varied during the process of analyzing the stereochemical fit of the animated tetrapeptide substrate Ac-Ala-Lys-Tyr-Ala-NH_2 to the active site of penicillopepsin. The model for the substrate was constructed from a dictionary of standard amino acid components (Sielecki et al., 1979). The scissile bond is the Lys–Tyr bond [see Hofmann et al. (1984)]. All model manipulations were done with the program M3, devised and implemented on the MMS-X by Colin Broughton (Sielecki et al., 1982). Only single-bond rotations, ϕ and ψ for the main chain and χ_i for the side chains, were allowed to vary. A continuous check on too short, nonbonded contacts from atoms of the tetrapeptide to the atoms of penicillopepsin facilitated the elimination of unallowed regions of conformational space.

The model of the tetrahedral intermediate was constructed by allowing the carbonyl carbon atom of the P_1-lysyl residue[3] and the nitrogen atom of the P_1'-tyrosyl residue to adopt sp^3 pyramidal character following the principles enunciated by Bürgi et al. (1973). The apices of the developing pyramids on the C and N atoms were directed in opposite directions from one another relative to the C–N bond direction. The C^α–C and N–C^α bond lengths for the P_1-lysyl and P_1'-tyrosyl residues were fixed at 1.47 and 1.51 Å, respectively. The peptide C–N bond was increased to 1.45 Å, close to the accepted C–N single-bond length. Likewise, the two carbon to oxygen bonds of the tetrahedral intermediate were given lengths of 1.40 Å, close to the expected sp^3 carbon to hydroxyl oxygen bond length. Thus, in the tetrahedral intermediate the carbon atom has a *gem*-diol configuration and the nitrogen atom has a pyramidal character. Staggered conformations about all freely rotatable single bonds were chosen initially. This model was positioned by global translations and rotations to fit the refined coordinates of the acyl portion of the pepstatin analogue (James et al., 1983). Subsequently, free rotations were used to adjust the tyrosylalanylamide portion of the tetrahedral intermediate to fit as closely as possible to the corresponding portion derived for productive substrate binding (Hofmann et al., 1984). Hydrogen-bonding and noncovalent van der Waals interactions between enzyme and model substrate were maintained during this process.

Results and Discussion

The present analysis of the stereochemical course of aspartic proteinase catalyzed hydrolysis of peptide bonds is based on the 1.8-Å resolution refined crystal structures of penicillopepsin, $R = 0.136$ (James & Sielecki, 1983), and of penicillopepsin complexed with the inhibitor Iva-Val-Val-Sta-OEt, $R = 0.131$ (James et al., 1982, 1983). Not only is it important to define the native enzyme structure as accurately as possible, but it is also mandatory to have refined coordinates for inhibitors or substrate analogues so that nonbonded interactions and intermolecular hydrogen bonds can be described with confidence. By way of example, the pepstatin analogue exhibited a root mean square (rms) coordinate difference of 0.90 Å, maximum difference 3.36 Å, between initial map-fitted and final refined positions, after 21 cycles of least-squares refinement (James et al., 1983). The difference electron density map from which the initial inhibitor coordinates were deduced was of exceptionally good quality (James et al., 1982), and yet large atomic movements resulted from the refinement procedure.

Both structures were refined by using the algorithm of Hendrickson & Konnert (1980). In this process stereochemical restraints on molecular geometry are imposed since the ratio of observations to parameters refined is very small for protein crystal structures. In order to judge the quality of the molecular model that results from this least-squares refinement procedure, it is important to compare the root mean square deviations of the geometry of the refined coordinates to that observed in the accurate, small molecule structure determinations of amino acids and peptides. Such a comparison has been summarized in Table 1 of James et al. (1983) for penicillopepsin and its complex. From the data reported therein, it is evident that the deviations from "ideal" geometry in both refined structures are well within acceptable tolerances, allowing for a high level of confidence in the description of intra- and intermolecular nonbonded interactions.

The molecule of penicillopepsin exhibits varying degrees of flexibility as reflected in the values of the refined isotropic temperature factors B. An extensive and intricate hydrogen-bonded network involving main-chain and side-chain atoms in the vicinity of the catalytic aspartyl residues confers relative rigidity to the active-site region as evidenced by the very low values of the associated isotropic B factors (James & Sielecki, 1983). On the other hand, the residues from Ser-72(72) to Ser-82(81), the flap region, exhibit some of the largest isotropic B factors in the molecule. This region undergoes a major conformational change when Iva-Val-Val-Sta-OEt binds (James et al., 1982).

The environments and conformations of Asp-33(32) and Asp-213(215) are extraordinarily symmetric (James & Sielecki, 1983). The polypeptide strands Asn-31(30)–Ser-36(35) and Leu-121(120)–Gly-123(122) are closely related by an approximate 2-fold axis to strands Ile-211(213)–Thr-216(218) and Ile-297(301)–Gly-299(303). A least-squares superposition

[2] The R factor is defined as $\sum||F_o| - |F_c||/\sum|F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes, respectively.

[3] The nomenclature of Schechter & Berger (1967) is used to describe the interactions of enzyme and substrate. P_n and P_n' denote residues of the substrate on either side of the scissile bond, with n increasing toward the N- and C-terminal directions, respectively. S_n and S_n' denote regions on the enzyme to which the amino acid side chains P_n and P_n' bind. Portions of the S_n and S_n' binding subsites are formed only after conformational changes in penicillopepsin occur.

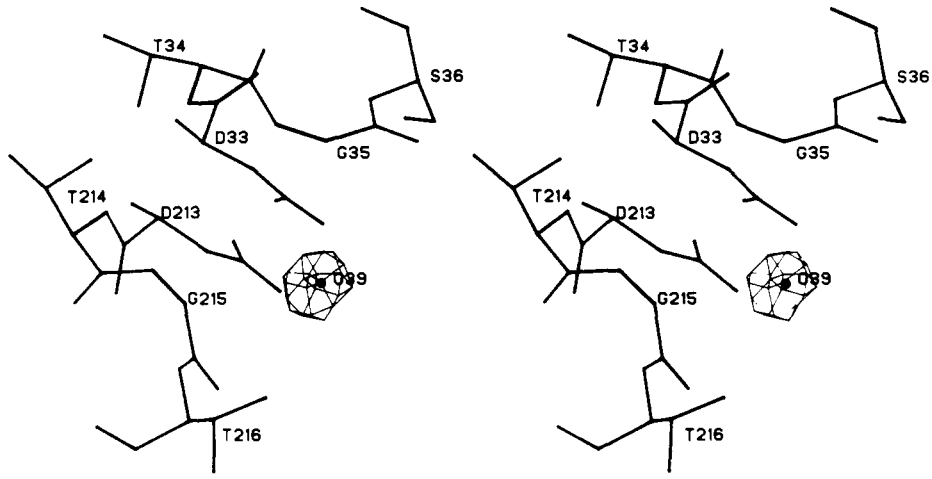


FIGURE 1: Stereoscopic view of the active site of penicillopepsin with the difference electron density map $||F_{Tl}| - |F_N||$ and α_N superimposed (see Materials and Methods for details). The positive density, contoured at 0.1 e Å^{-3}, indicates that the Tl^+ binding site coincides with the position of solvent peak O39 in the native enzyme structure. No conformational changes in penicillopepsin occur on Tl^+ binding.

procedure involving the 46 common atom pairs in these regions results in an rms difference of 0.49 Å. However, in spite of this high degree of internal molecular symmetry at the active site, the productive binding mode of the substrate is of necessity asymmetric (Hofmann et al., 1984). It is this asymmetric association that allows for some of the deductions regarding mechanism in the following discussion.

Charged State of the Active Site. Penicillopepsin crystals are grown from an $(NH_4)_2SO_4$ mother liquor buffered by 0.1 M $NaC_2H_3O_2$ to pH 4.4. Pepsin exhibits pK_a values of 1.2 and 4.7 for the carboxyl groups at the active site (Cornish-Bowden & Knowles, 1969). These data suggest that the net charge at the active site of penicillopepsin is −1. However, the identification of the lower pK_a with either Asp-33(32) or Asp-213(215) is not firm. Neither can the chemical nature of solvent peak O39, found between these two carboxyl groups in the native enzyme structure, be established from the crystallographic results (James & Sielecki, 1983). It could be a water molecule, H_2O, an ammonium ion, NH_4^+, or a hydronium ion, H_3O^+, on the basis of the size of the associated electron density peak, occupancy (0.9), and *B* factor (17.4 Å^2).

To try to elucidate the distribution of charges on the active site, we attempted the soaking experiment with thallium acetate because of the similarity of its charge and ionic radius, 1.47 Å, to that of an ammonium ion, 1.43 Å. Tl^+ has 80 electrons and should be easily recognizable on a difference electron density map.

The crystal of penicillopepsin soaked in thallium acetate had unit cell dimensions of a = 97.37 (9) Å, b = 46.66 (5) Å, c = 65.39 (7) Å, and β = 115.3 (2)° and was, within experimental error, isomorphous with crystals of the native enzyme. The agreement index between the structure-factor amplitudes of the Tl^+ soaked crystal and the native crystal is 0.07 ($=\sum||F_{Tl}| - |F_N||/\sum|F_N|$). The difference electron density map at 2.5-Å resolution (5780 terms) shows no gradients associated with atomic positions of the enzyme. This indicates that no detectable conformational movement of atoms of penicillopepsin took place upon Tl^+ binding. There are only two significant features in the map, one of them a positive peak coincident with the refined position of O39 (Figure 1). This peak, representing the Tl^+ replacement of O39, is 0.15 e Å^{-3} in height, whereas the second peak in the map has approximately twice that height, 0.33 e Å^{-3}. This latter binding site is at an intermolecular contact zone, and the Tl^+ ion replaces another strongly bound solvent molecule, O15.

In addition, solvent site O39 is coincident with the site of both Hg^{2+} and uranyl ion binding in the multiple isomorphous phasing process (Hsu et al., 1977a). Therefore, this site is certainly favorable for binding a positively charged ionic species, and the solvent detected during refinement of the native molecule could most likely be an NH_4^+ ion that was not displaced by the transfer of the crystals to the Li_2SO_4 solution (Li^+ has too small an ionic radius, 0.65 Å, to effectively replace NH_4^+). Should this site indeed be occupied by an hydronium ion in the aspartic proteinases during the initial stages of the hydrolytic process, no matter how fleetingly, it could represent the electrophilic component of the catalytic mechanism.

A possible hydrogen-bonding scheme for the groups that comprise the active site of penicillopepsin is presented in Figure 2. It should be emphasized that the polar hydrogen atom positions are tentatively assigned on the basis of the refined positions of the non-hydrogen atoms, with accepted bond distances and angles determined for similar groups in small molecule crystal-structure analyses. At the present resolution of 1.8 Å the location of the hydrogen atoms cannot be determined from the experimental data. The representation in Figure 2a assumes that O39 is an ammonium ion and that the negative charge in the active site is localized on Asp-33(32). The chemical and stereochemical arguments that justify this assumption are discussed in James & Sielecki [1983, section 4f(i)]. The alternative hydrogen-bonding scheme in which the proton resides on the carboxyl group of Asp-33(32) and the negative charge is associated with Asp-213(215) cannot be excluded on the basis of the crystallographic evidence for the native enzyme alone. Due to the geometric disposition of the different components, the former alternative favors O284 (see Figure 2a) as the attacking nucleophile in a general-base mechanism, whereas the latter would assign this function to O61, a solvent molecule hydrogen bonded to Asp-213(215). We do not consider O39 for the role of the nucleophile because of the arguments favoring this being the site of a positively charged ionic species, at least in native penicillopepsin.

It is also possible to propose a hydrogen-bonding scheme for the active site of penicillopepsin in the presence of the bound pepstatin analogue, and one possibility is presented in Figure 2b. In this case, the refined structure of the complex indicates that minor conformational changes involving Asp-33(32) have taken place, with the result that the weak interaction between Ser-36(35) O^γ and Asp-33(32) O^{δ_1} (3.43 Å in the native enzyme) becomes a strong hydrogen bond in the complex (2.74 Å). The 3-hydroxyl oxygen atom of statine[4]

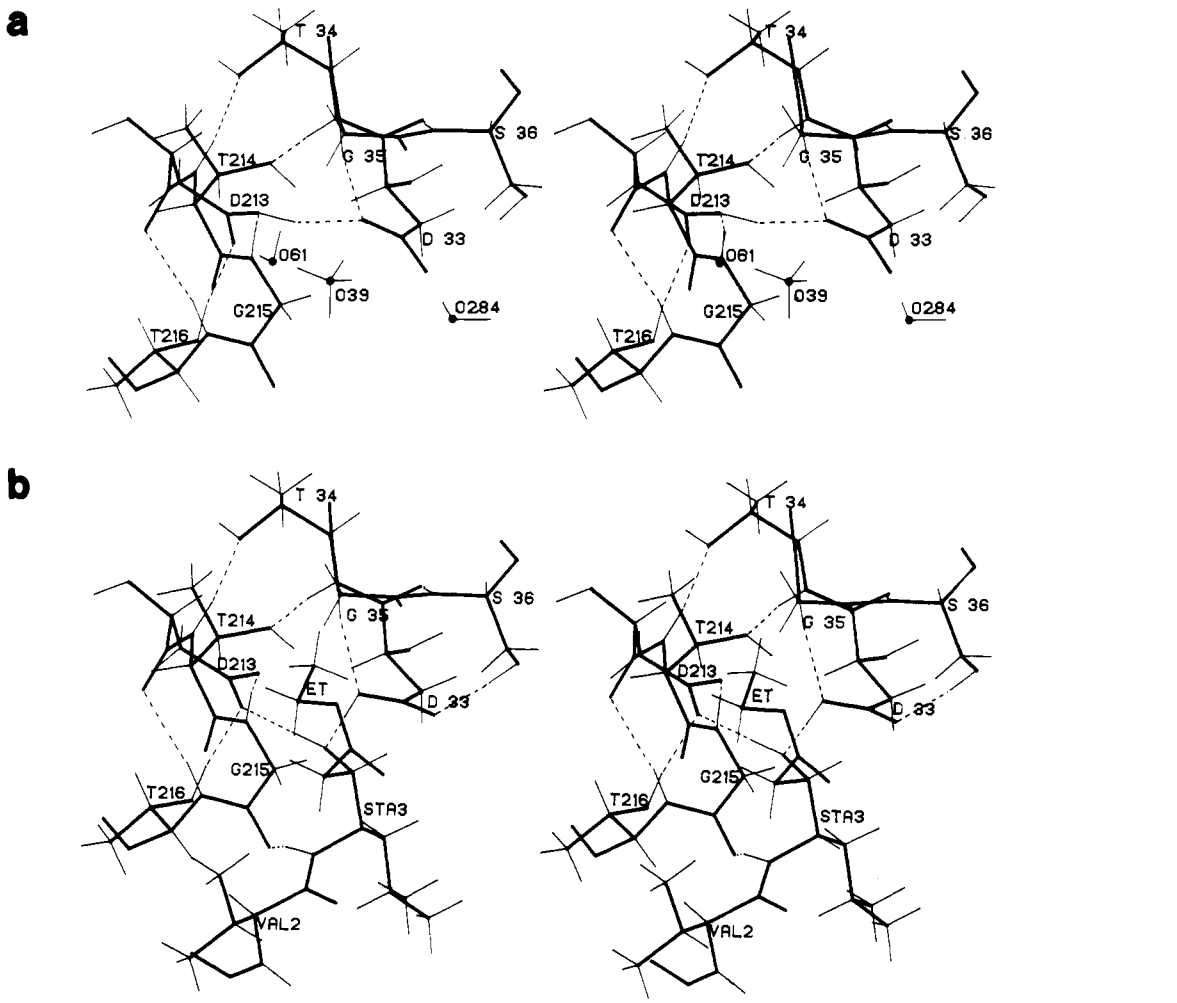


FIGURE 2: (a) Possible hydrogen atom positions for hydrogen-bonding interactions in native penicillopepsin in the region of the active site. Solvent site O39 is represented here as an ammonium ion, the other solvent sites being water molecules. The aspartic acid with the lower pK_a is assumed to be Asp-33(32) [see James & Sielecki (1983)]. Hydrogen bonding is indicated by dashed lines drawn from the hydrogen to the acceptor atom. (b) A possible hydrogen-bonding scheme for the pepstatin analogue at the active site of penicillopepsin. Only the P_1 Sta and P_2 Val residues of Iva-Val-Val-Sta-OEt are represented. In the presence of the inhibitor, we conclude that the net negative charge at the active site is on Asp-213(215). This is supported by the values of hydrogen-bond distances (James et al., 1983) between acceptor and donor atoms. The small conformational change of Asp-33(32) upon inhibitor binding allows hydrogen bonding to Ser-36.

lies within 0.25 Å from the site that solvent O39 occupies in the native enzyme. The hydrogen bonding proposed here implies that the net negative charge is on Asp-213(215). Note that solvent molecules O39, O61, and O284 of the native structure are among the 11 ordered solvent molecules displaced by inhibitor binding and the concomitant conformational change of the "flap". Both O39 and the 3-OH oxygen atom lie in the carboxyl plane of Asp-213(215). The rotation of the carboxyl group of Asp-33(32) about the C^{β}–C^{γ} bond upon binding of the inhibitor allows for an acceptable contact distance between the tetrahedral carbon atom C3 of the statine residue and $O^{\delta 1}$ of Asp-33(32) (3.48 Å). It is also likely that the tetrahedral carbon atom C3, with its attached hydrogen, is the cause for displacing solvent molecule O284, a situation that would not arise in the binding of a substrate.

Our main interest is, of course, to elucidate the distribution of charges in the active site in the presence of a substrate in the productive binding mode. Two of the alternatives are portrayed schematically in Figure 3, with the direction of approach of the substrate that each distribution would imply.

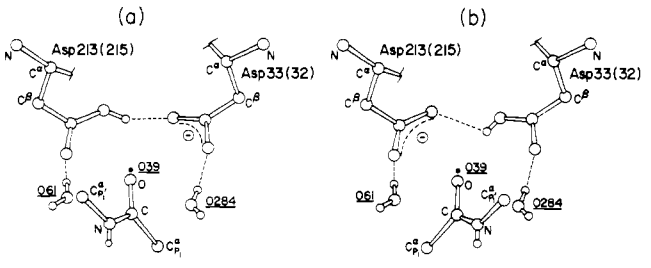
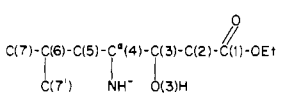


FIGURE 3: Two alternatives for the charged state of the active-site aspartic acid residues in the presence of bound substrate. Hydrogen-bonding interactions are shown as dashed lines. The position of O39 in native penicillopepsin is shown as a filled circle. In both possibilities of substrate binding in (a) and (b) O39 would be displaced by the carbonyl oxygen of the scissile peptide bond of the substrate. (a) The negative charge of the active site is shown localized on the carboxyl group of Asp-33(32). The substrate binding for this possibility has been deduced from the bound position of Iva-Val-Val-Sta-OEt (Figure 7). This charge distribution favors O284 as the attacking nucleophile as solvents O39 and O61 are displaced by substrate. (b) In this alternative Asp-213(215) is the general base with attacking nucleophile O61. This possibility is not favored since a different mode of enzyme–substrate binding would be required (as shown) and such a mode cannot use the binding interactions deduced experimentally for residues in P_1–S_1 and P_1'–S_1' subsites for the pepstatin analogue.

[4] The atom numbering scheme for the statine residue is



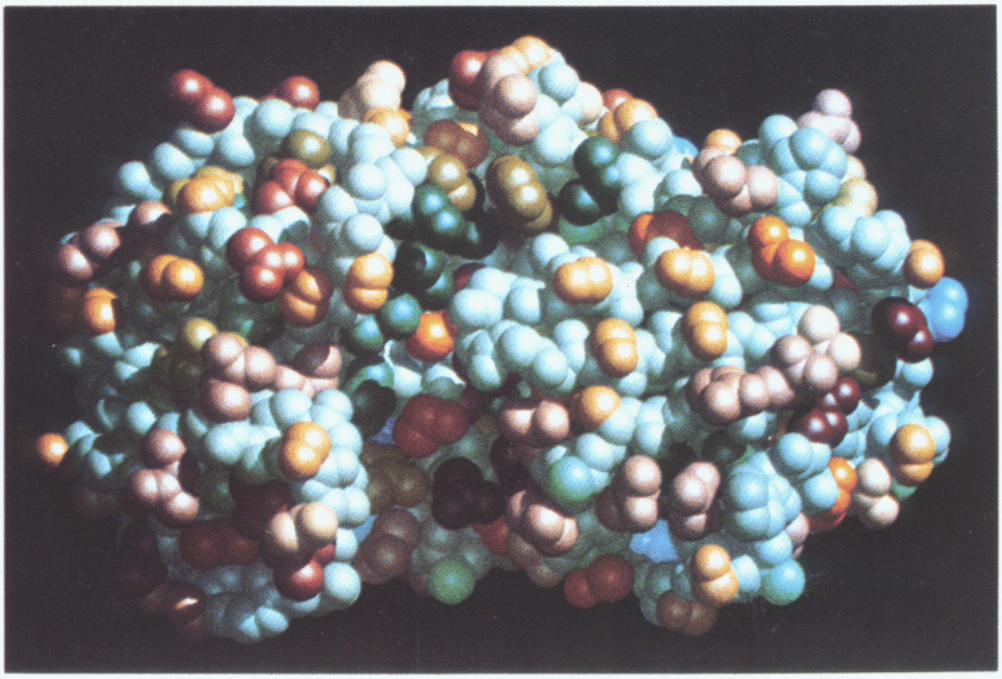


FIGURE 4: van der Waals surface representation of the penicillopepsin molecule. The color coding for this figure and for Figures 5 and 6 is as follows: increasingly deeper shades of green, Ala, Val, Leu, and Ile; shades of brown, Tyr, Phe, and Trp; shades of pink, Asn and Gln; shades of orange, Ser and Thr; shades of red, Asp and Glu; shades of blue, His, Lys, and Arg; light gray, Gly and the polypeptide chain; gray, Pro. This view shows the substrate binding cleft running diagonally across the molecular surface and the "flap" (residues 72–82) projecting across the active site. We are indebted to David Bacon for the development of the computer program for the representation of molecular surfaces that was used to produce this figure and Figures 5 and 6.

In the ensuing discussion we shall show that the scissile bond of the substrate interacts with the catalytic aspartyl groups in an asymmetric fashion [see also Hofmann et al. (1984)]. In the modeling of this binding, the side chain of the P_1' residue would displace water O61, so that in the E–S complex the role of the nucleophile could only be ascribed to O284, suggesting the distribution of charges depicted in Figure 3a. We will discuss this further in a subsequent section.

Substrate Binding Site in Penicillopepsin. The crystal-structure analysis of the bound inhibitor Iva-Val-Val-Sta-OEt provides definitive data on binding sites S_4–S_1' in penicillopepsin (James et al., 1982, 1983). A series of elegant experiments with porcine pepsin demonstrated that the kinetic effect of increasing the chain length of synthetic substrates was to enhance k_{cat} with little effect on K_M (Fruton, 1976). These studies were consistent with a binding region for pepsin sufficiently long to accommodate seven or eight amino acid residues. Similar studies with penicillopepsin have also shown kinetically the presence of an extended binding cleft (Hofmann & Hodges, 1982) that penetrates deeply into the enzyme and runs diagonally across the molecule in the view of Figure 4. The "flap", an antiparallel β-loop [Ser-72(72)–Ser-82(81)], projects over the cleft forming a channel into which a substrate binds. Thus, substrates or inhibitors are enclosed between the active site and the flap. The enzyme–substrate interactions involving residues of the flap are mainly to residues P_1, P_2, and P_3 (James et al., 1983).

The hydrophobic nature of a portion of the substrate binding cleft in penicillopepsin can be appreciated in the close-up view of Figure 5. The two catalytically important aspartic acid residues are centrally located (for clarity, the residues of the flap have been removed in Figures 5 and 6). The approximate 2-fold relationship between the N- and C-terminal domains (Tang et al., 1978; James & Sielecki, 1983) is evident in this view. The hydrophobic patch lies to the left of, and above, Asp-213(215) and comprises the phenolic ring of Tyr-274-(275), Leu-284(284), Leu-218(220), Leu-220(222), Ile-293-(297), Phe-295(299), Ile-297(301), Ile-211(213), Phe-190-(189), and Ile-129(128). These residues are part of the S_4, S_2, and S_1' binding subsites.

Iva-Val-Val-Sta-OEt binds at the active site of penicillopepsin in the manner of a good substrate (Figures 6 and 7), and the 3-carbon of statine, with its associated hydroxyl group, mimics the tetrahedral intermediate as proposed by Marshall (1976) and Marciniszyn et al. (1976). The interactions of the isovaleryl group, the side chain of Val-2 (P_2), and the ethyl group of the ester with the hydrophobic residues mentioned above are clear. The P_1 Sta side chain (equivalent to the isobutyl group of a leucyl side chain) interacts mainly with Tyr-75(75) and Leu-121(120) (see Figure 7). Hydrogen bonding between penicillopepsin and the peptide backbone of the acyl portion of the substrate is critical to the proper positioning of the scissile bond relative to the aspartyl residues. This is entirely analogous to the binding of the acyl portion of oligopeptide substrates or products to the serine proteinases such as SGPA (James et al., 1980). These latter enzymes also exhibit major kinetic effects on k_{cat} with increasing chain length of the substrates (Bauer et al., 1976).

Productive Binding Mode for Substrates. In the discussion of the catalytic pathway for aspartic proteinases that follows, we favor the view that a covalently bound tetrahedral intermediate is not mechanistically or kinetically important. The suggestion was made originally by Fruton (1976) after many fruitless efforts by a large number of investigators to trap a covalently attached intermediate. In our earlier proposals (James et al., 1977, 1981) we adopted this same idea. Our position is unchanged and has some experimental support from the isotope-exchange studies of Sharon et al. (1962) and Antonov et al. (1978, 1981).

A productive binding mode for the tetrapeptide Ac-Ala-Lys-Tyr-Ala-NH_2 was derived (Hofmann et al., 1984). This model is presented in Figure 8; the overall binding mode is

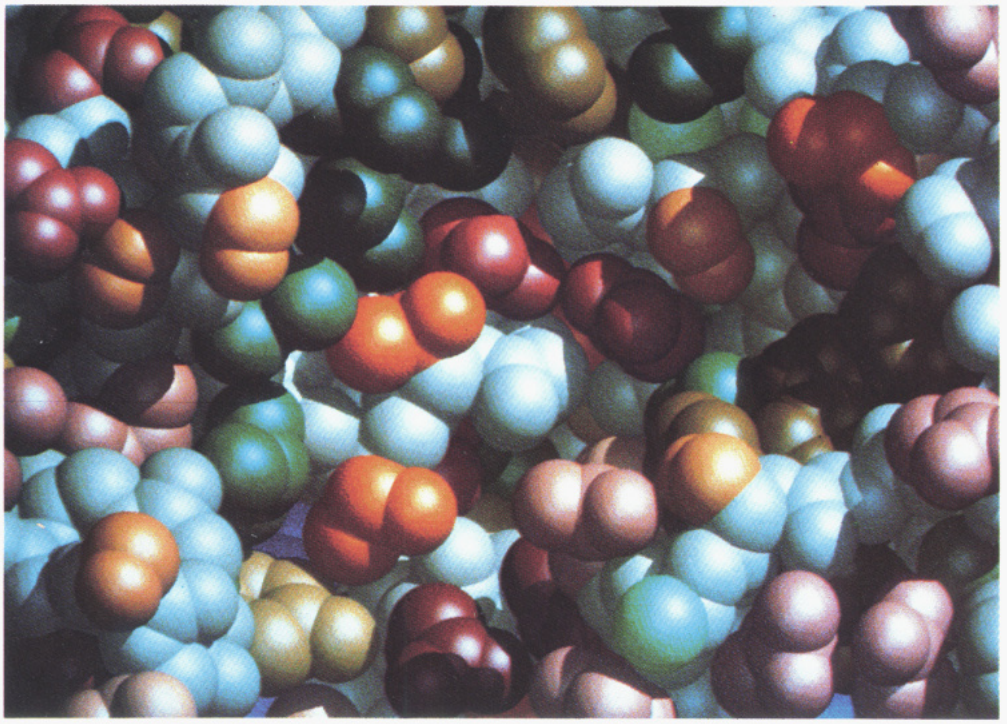


FIGURE 5: Close-up of the penicillopepsin active site with the "flap" removed for clarity. The two aspartic acid residues Asp-33(32) and Asp-213(215) are centrally located. The approximate molecular dyad relating the N- and C-terminal domains passes between these two side chains. The carboxyl groups of Asp-33(32) and Asp-213(215) share a proton in a hydrogen-bonded interaction. To the left of Asp-213(215) is the hydrophobic patch of residues that forms the binding sites S_4, S_2, and S_1'. The approximate dyad does not reproduce this hydrophobic patch on the right of Asp-33(32).

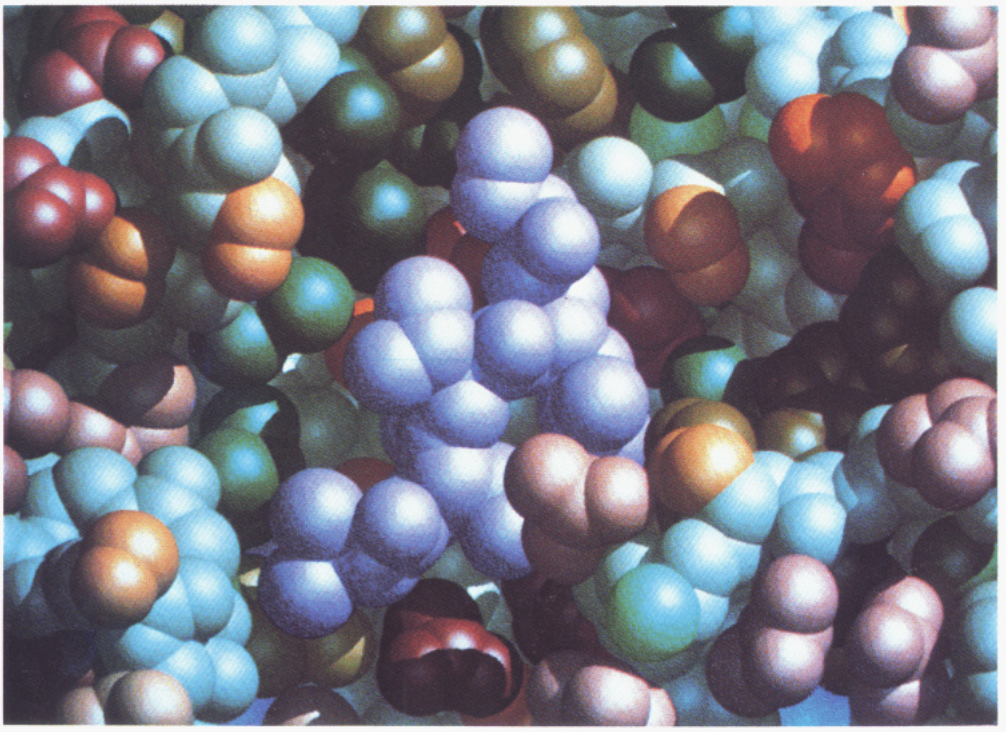


FIGURE 6: Same view as that shown in Figure 5 but with the pepstatin analogue Iva-Val-Val-Sta-OEt represented in purple. The hydrophobic interactions involving the residues of penicillopepsin and Iva (P_4), Val-2 (P_2), and the ethyl group of the ester (P_1') are visible in this view. The statine side chain fits into the S_1 binding pocket formed by Leu-121(120), Phe-112(111), and Tyr-75(75) (not shown).

very similar to that predicted at the 2.8-Å resolution level (James, 1980). However, the relative positioning and hydrogen bonding to the enzyme have been modified since the model has now been built on the basis of the experimentally observed binding and refined positions of the inhibitor Iva-Val-Val-Sta-OEt (James et al., 1982, 1983). The sequence assigned to the tetrapeptide was based on the synthetic substrate Ac-Ala-Ala-Lys-(*p*-NO_2)Phe-Ala-Ala-NH_2 used by Hofmann & Hodges (1982) for extensive kinetic studies.

Two hydrogen-bonded interactions from the peptide NH's of P_1 lysine and P_2' alanine to the main-chain carbonyl oxygen atoms of Gly-215(217) and Gly-35(34), respectively, are worth

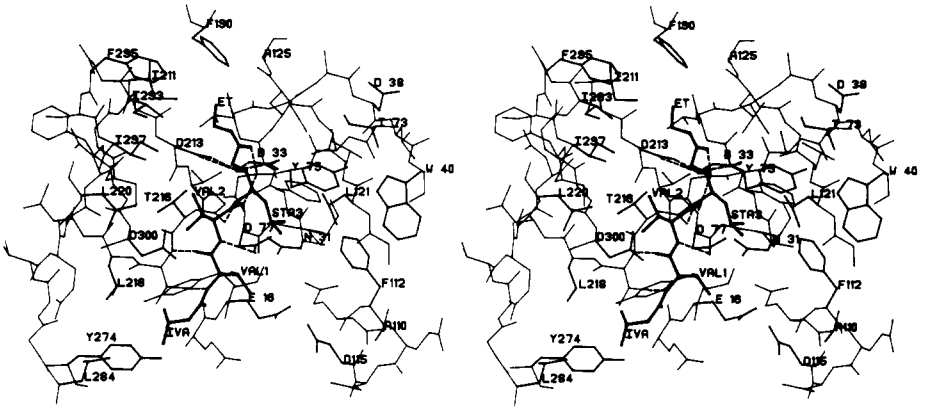


FIGURE 7: Stereoview of the hydrogen-bonding interactions of the main chain of Iva-Val-Val-Sta-OEt to penicillopepsin. The pepstatin analogue and important residues on the enzyme are shown with a thicker line. Hydrogen bonds are deduced from the refined atomic coordinates of the complex and are depicted by dashed lines. A close-up of this view showing the interactions of the P_1 statine residue is given in Figure 2b.

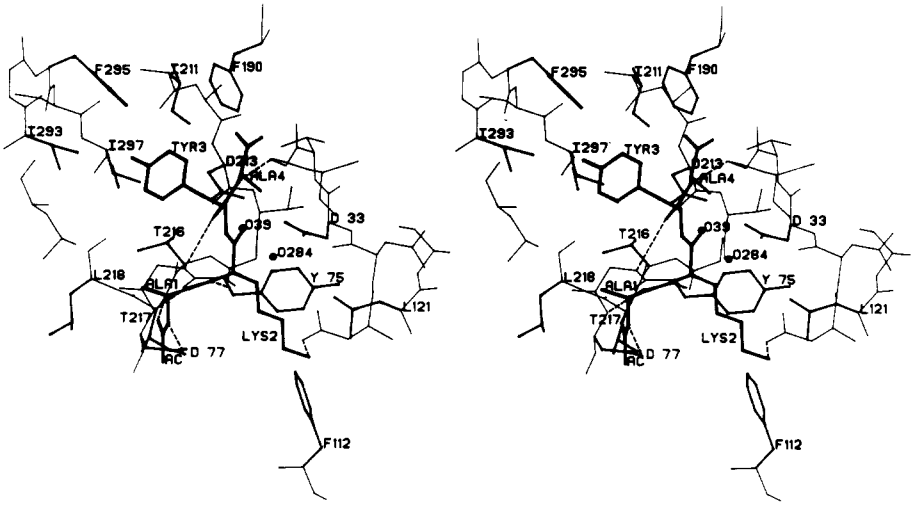


FIGURE 8: Hypothetical model of the proposed productive binding mode of a substrate Ac-Ala-Lys-Tyr-Ala-NH_2 to penicillopepsin. This model was constructed from the observed binding mode for the pepstatin analogue (Hofmann et al., 1984). Hydrogen-bonding interactions involving the Ac-Ala-Lys portion of the substrate are the same as those described for the inhibitor structure. Gly-76 NH to Tyr-3 O and Ala-4 NH to Gly-35(34) O are proposed as a result of the molecular modeling. The positions of O39 and O284 (the nucleophilic water) are also shown in this representation. Obviously, O39 would be displaced by the substrate.

noting. In the present binding mode, these hydrogen bonds are rather long ($\simeq$3.2 Å) but could be shortened as the substrate is converted to a tetrahedral intermediate. The situation would then be exactly analogous to that occurring in substrate binding to the serine proteinases. The importance of forming a strong hydrogen bond from the P_1 NH of peptide substrates to the C=O of Ser-214 in serine proteinases, thus assisting in the formation and stabilization of the tetrahedral intermediate, was recognized early on (Robertus et al., 1972) and has since been observed crystallographically in complexes of bacterial serine proteinases (James et al., 1980; Fujinaga et al., 1982; Read et al., 1983) and complexes of trypsin (Huber & Bode, 1978).

The binding mode shown in Figure 8 suggests the electrophilic and nucleophilic groups in the catalytic mechanism. The earlier proposals from this laboratory (James et al., 1977; James, 1980) suggested that the electrophile (the component enhancing the polarization of the peptide carbonyl–oxygen bond) is the proton shared by Asp-33(32) and Asp-213(215). Also, a general-base mechanism implicated a water molecule bound to Asp-33(32) as the nucleophilic component. In line with our previous proposals, we maintain these groups as electrophile and nucleophile. The positions of solvent sites O39 (displaced by substrate binding) and O284 are also shown in Figure 8. The carbonyl oxygen of the model-built substrate lies within 0.6 Å of site O39, very favorably positioned to accept a proton from Asp-213(215) and thus initiate the formation of the tetrahedral intermediate. Kinetic evidence is strongly in favor of such an electrophilic character for the hydrolytic process (Fruton, 1976).

Water O284 in the native enzyme is ideally positioned for nucleophilic attack on the substrate carbonyl (Figure 3a). It is activated by the proximity of the carboxylate of Asp-33(32), and the nucleophile is the generated OH^- ion. The ideal position for a nucleophile relative to a carbonyl carbon has been analyzed (Bürgi et al., 1973). Site O284 lies within 0.2 Å of such an ideal position (Figure 8), 2.8 Å from the carbonyl carbon atom of P_1 lysine. A line connecting these two atoms would be roughly perpendicular to the scissile peptide plane. This plane contains solvent atom O39. In addition, the vector between the carbonyl carbon atom and the carbonyl oxygen atom points midway between the two carboxyl groups of Asp-33(32) and Asp-213(215) (Figure 8).

Changes to the ϕ,ψ conformational angles of P_1 lysine and P_1' tyrosine can achieve a slightly different binding mode for the substrate (Figure 9). In this mode, the peptide plane lies midway between the positions of O39 and O284, with the carbonyl carbon to oxygen atom vector directed approximately at the carboxyl group of Asp-33(32). The role of the nucleophile could then be ascribed to O39, interpreted as a bound

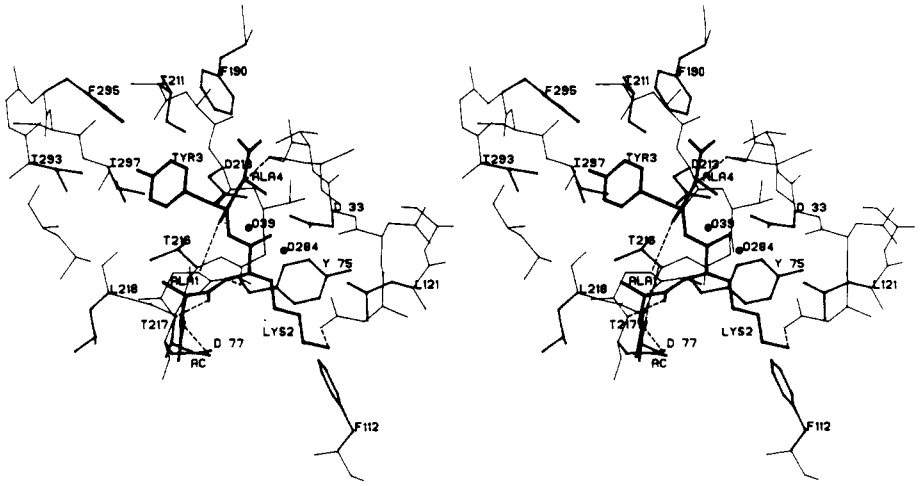


FIGURE 9: An alternate position for the substrate scissile bond in the active site of penicillopepsin. In this orientation the carbonyl oxygen atom of the peptide bond points at the carboxyl group of Asp-33(32). There would be no electrophilic component in this model. Also, as discussed in the text, it would be unlikely that the role of nucleophile could be ascribed to solvent O39.

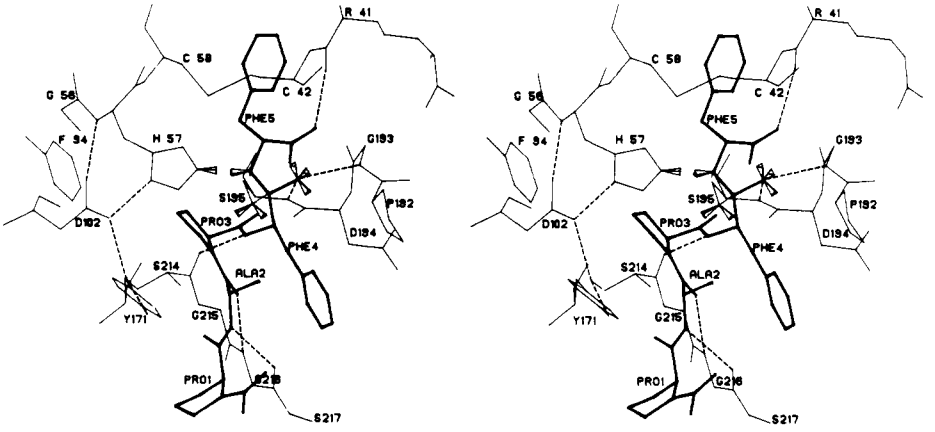


FIGURE 10: A tetrahedral intermediate in the active site of the serine protease SGPA. Lone pair electrons on the two oxygen atoms of the tetrahedral carbon are oriented antiperiplanar to the C–N bond. Protonation of the nitrogen atom of the P_1' residue makes the nitrogen a good leaving group. This proton is transferred from Ser-195 to the imidazole of His-57 and hence to the nitrogen atom of the leaving group.

water molecule and not a positively charged ion. In addition, in such a substrate binding mode, water O284 would be displaced by the substrate.

Two major arguments can be made against this latter proposal. First, there is no clearly defined electrophilic component for the polarization of the carbonyl carbon–oxygen bond. Second, it is unlikely that O39 would be favored as a negatively charged OH^- group in light of its environment (see above discussion regarding the Tl^+ ion binding). Additionally, if O39 were the nucleophile, the direction of nucleophilic attack by the generated hydroxide ion is unfavorably oriented relative to the peptide plane of the scissile bond.

Formation of the Tetrahedral Intermediate. The stereoelectronic hypothesis has been useful in the analysis of the hydrolytic mechanism of esters and amides (Deslongchamps, 1975). This proposal suggests that antiperiplanar lone pair electrons on one or more heteroatoms of a tetrahedral intermediate favor the departure of the leaving group. The hydrolytic mechanism of the serine proteinases has been analyzed following this hypothesis (Bizzozero & Dutler, 1981). The stereochemistry at the active sites of the serine proteinases is fully consistent with the most stable conformer of the tetrahedral intermediate. Formation and breakdown of this intermediate would have antiperiplanar lone pair electrons on the two other heteroatoms relative to (a) the bond formation from O^γ of Ser-195 to the carbonyl carbon atom of the substrate and (b) cleavage of the C–N bond after protonation of the P_1' nitrogen atom to make a good leaving group (Figure 10). The importance of the oxyanion binding site for this enzyme family has recently been verified experimentally (Asbóth & Polgár, 1983). The direction of nucleophilic attack by the alkoxide ion of Ser-195 has been firmly established.

We propose that the formation of the tetrahedral intermediate in the aspartic proteinase catalytic pathway is initiated by the protonation of the carbonyl oxygen of the substrate. This would disrupt the delocalization of the π electrons in the peptide bond. The subsequent nucleophilic attack by an OH^- ion, assisted by Asp-33(32), would result in the pyramidalization of the carbonyl carbon and the imino nitrogen atoms (Figure 11). Lone pair electrons on the protonated carbonyl oxygen and the P_1' nitrogen are shown antiperiplanar to the direction of nucleophilic attack by the OH^- of the O284 solvent site (Figure 11). The electrophilic protonation of the carbonyl oxygen in this mechanism is analogous to the effect of the oxyanion binding site of the serine proteinases.

Several favorable electrostatic interactions could help to provide stabilization of this proposed tetrahedral intermediate. In the conformer represented in Figure 11, the two hydroxyl groups (a *gem*-diol on the carbonyl carbon of the substrate) are shown hydrogen bonded to the protonated carboxyl group of Asp-33(32). [The proton on the carboxyl group of Asp-33(32) has been omitted from this diagram and that of Figure 12 for reasons of clarity.] Asp-33(32) is a hydrogen-bond donor to the hydroxyl group that arises from O284. The other

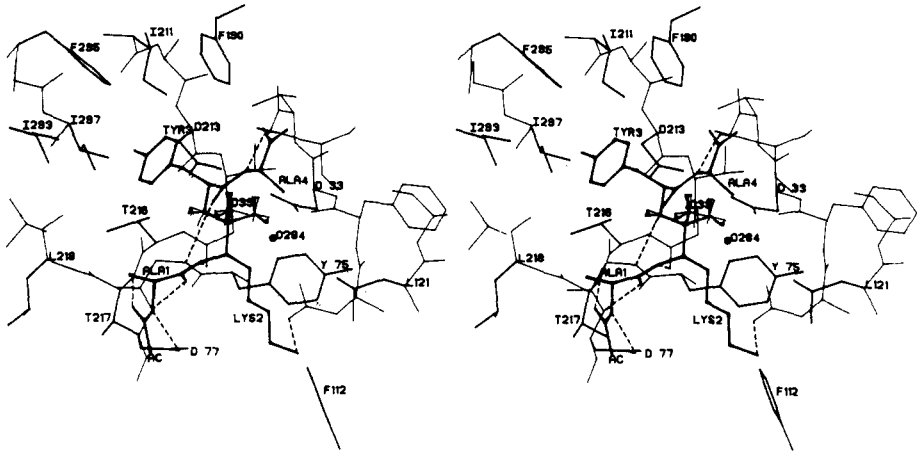


FIGURE 11: Model of a tetrahedral intermediate for penicillopepsin resulting from protonation of the carbonyl oxygen of the substrate, followed by nucleophilic attack by O284 as an OH^- ion. The base assisting in this attack is assumed to be the carboxylate of Asp-33(32). Lone pair orbitals on the carbonyl oxygen and the tetrahedral sp^3 nitrogen are shown antiperiplanar to the direction of approach of O284 in accordance with the principle of the stereoelectronic effect (Deslongchamps, 1975). The positions of O39 and O284 are indicated relative to the tetrahedral intermediate only for reference.

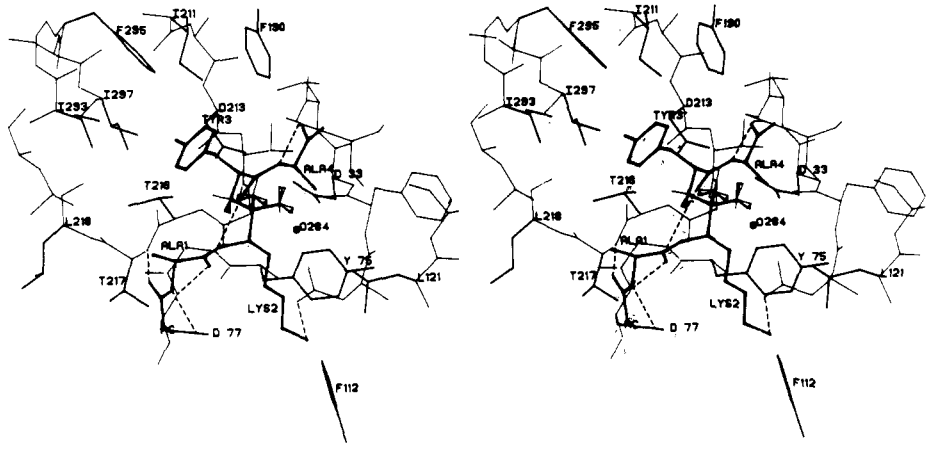


FIGURE 12: The second tetrahedral intermediate has lone pair electrons on the two OH groups of the tetrahedral carbon atom oriented antiperiplanar to the C–N bond. The sp^3 nitrogen has undergone inversion of configuration followed by rotation of $\simeq 60°$ in order that the lone pair electrons point toward solution for protonation. The alternative to this inversion is discussed in the text.

oxygen of the carboxyl group is a hydrogen-bond acceptor from the hydroxyl group that arises from protonation of the carbonyl oxygen of the scissile peptide bond of the substrate. This interaction resembles, but clearly is not electronically, a carboxylic acid type dimer.

The other electrostatic interaction discernible from this proposed tetrahedral intermediate involves the lone pair electrons on the pyramidal nitrogen atom of the scissile bond (Figure 11). In native penicillopepsin and in the Iva-Val-Val-Sta-OEt complex, $O^{\gamma 1}$–H of Thr-216(218) forms a strong hydrogen bond to $O^{\delta 1}$ of Asp-213(215) (see Figure 2). In the tetrahedral intermediate of Figure 11 this proton could form a bifurcated hydrogen bond to both Asp-213(215) and the lone pair electrons on the nitrogen of the substrate.

Cleavage of the C–N Bond of the Tetrahedral Intermediate. A necessary prerequisite to cleavage of the C–N bond is the protonation of the pyramidal nitrogen atom, thus creating a favorable leaving group (Drenth, 1980). There are two alternatives for proton donation to the nitrogen atom of the leaving group. In one of them, inversion of configuration at the nitrogen, followed by a rotation of $\sim 60°$ about the C–N bond, would leave the lone pair electrons on the pyramidal nitrogen directed toward bulk solvent, where protonation could easily take place for those aspartic proteinases having a strongly acidic pH optimum (e.g., pepsin and other gastric proteinases). This may or may not be the case for kidney renin that has a pH optimum of $\simeq 7$ (Inagami et al., 1977); a proposal involving similar inversion of configuration at the nitrogen atom has been made for the serine proteinases in order to facilitate protonation of the leaving group nitrogen atom by the imidazolium side chain of His-57 (Bizzozero & Dutler, 1981). If inversion of configuration of the nitrogen atom is associated with the protonation event (Figure 12), then stabilization of the resulting tetrahedral intermediate could involve an altered hydrogen-bonding interaction with $O^{\gamma 1}$ of Thr-216(218). This alteration involves re-formation of the strong hydrogen bond from Thr-216(218) $O^{\gamma 1}$ to Asp-213(215) $O^{\delta 1}$ and $O^{\gamma 1}$ of Thr-216(218) accepting a hydrogen bond from the pyramidal N–H of the scissile bond nitrogen. No such intermediate stabilization could be provided for those aspartic proteinases lacking Thr or Ser at position 216(218) (i.e., human kidney renin).

The alternative to inversion of configuration of the nitrogen atom involves transfer of the proton from the innermost hydroxyl group of the tetrahedral intermediate to the nitrogen atom. This transfer could occur via the carboxylate of Asp-213(215) in a two-step process. In this case, it would not be necessary for the nitrogen to undergo inversion of configuration. However, a propitious conformation of the tetrahedral intermediate would require a syn conformation for the hydroxyl proton and the nitrogen lone pair. The energetic differences between these two alternatives are difficult to assess,

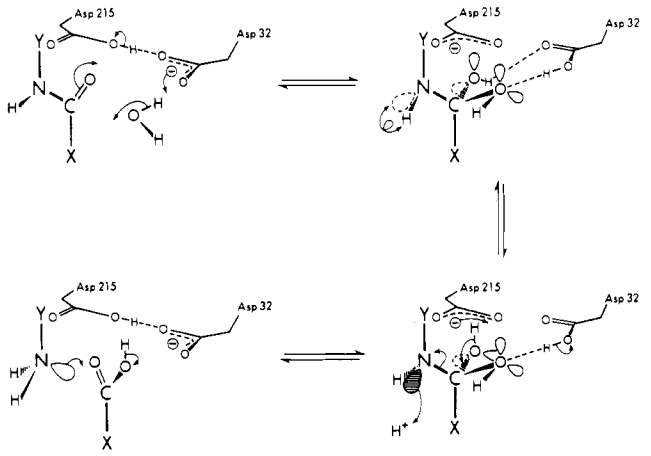


FIGURE 13: A summary of the proposed hydrolytic pathway of the aspartic proteinases.

but one or the other represents the only available possibility for proton donation to the leaving group nitrogen in this proposed binding mode.

Breakdown of the tetrahedral intermediate into a free amino group and a free protonated carboxyl group would be facilitated by the lone pair orbitals on the two oxygen atoms oriented antiperiplanar to the C–N bond as depicted in Figure 12. This conformer is easily attainable from that of the intermediate depicted in Figure 11 by single-bond rotations that should imply small rotational energy barriers.

The hydrolytic pathway proposed is summarized in Figure 13. In each of the steps we chose to display only one of the alternatives that has been discussed above. In this pathway, the electrophile is the proton shared by the two catalytic aspartic acid residues; the nucleophile is the OH^- ion resulting from proton transfer from the bound water molecule to the ionized aspartyl 33(32); the proton is donated to the leaving group nitrogen from bulk solvent.

Predictions for the Mode of Inhibitor Binding. Knowledge of the pepstatin analogue mode of binding to penicillopepsin can be used to predict interactions between the enzyme and other inhibitors. It was possible to deduce that a lysyl side chain in P_1 could form an ion pair with the carboxyl carboxylate pair Glu-16(13), Asp-115(114), suggesting an enhanced K_i for such an inhibitor (Hofmann et al., 1984). In order to verify this prediction, the analogue of statine 4,8-diamino-3-hydroxyoctanoic acid (DAMOA) was synthetically incorporated into Iva-Val-Val-DAMOA-OEt (Rich et al., 1983). In agreement with our prediction, the K_i of this inhibitor was lower by a factor of 100 over that for Iva-Val-Val-Sta-OEt. Thus, we feel confident that our structural results can be used effectively to predict binding modes of other inhibitors of aspartic proteinases.

Structures based on the reduced peptide bond $-C^{\alpha}-CH_2-NH_2^+-C^{\alpha}-$ or on the hydroxy isostere $-C^{\alpha}-CH(OH)-CH_2-C^{\alpha}-$ also make potent inhibitors of aspartic proteinases (Szelke et al., 1983). Incorporation of the hydroxy isostere into a peptide produces the most potent modification yet synthesized for a renin inhibitor. It should bind to aspartic proteinases in a manner exactly analogous to that observed for the statine-type interaction. The added advantage that the hydroxy isostere has over the statine-based inhibitors is that the isosteric hydroxy–peptide bond has no "extra" atoms relative to those of a tetrahedral intermediate.

The proposed binding mode of an inhibitor based on the reduced peptide to penicillopepsin is shown in Figure 14. We have again used Ac-Ala-Lys-Tyr-Ala-NH_2 as the model peptide, with the reduced peptide bond between Lys-Tyr. In this configuration the strongest interaction that the inhibitor makes with the enzyme is probably the "internal" ion-pair attraction between the positively charged immonium ion and the negatively charged active-site carboxyl groups of Asp-33(32) and Asp-213(215). Thus, the nitrogen atom of the reduced peptide would be directed toward the two aspartic acid residues. The lack of a carbonyl oxygen atom in the reduced peptide allows the inhibitor to adopt a conformation that would be unallowed for a normal peptide (Hofmann et al., 1984).

In addition to providing strong predictive potential on substrate or inhibitor binding, our analysis of binding modes and the proposed hydrolytic pathway for aspartic proteinases allows for a rational explanation of the dislike pepsin displays for prolyl residues at position P_2, P_1, or P_2' and for β-branched amino acids like Thr, Val, and Ile in position P_1 (Powers et al., 1977).

Proline is unable to act as a hydrogen-bond donor due to the cyclic attachment of its side chain to the α-imino nitrogen atom. The importance of hydrogen-bond interactions from the substrate main-chain nitrogen atoms in positions P_1 and P_2' is emphasized in our proposed substrate binding mode [Figure 8 and as discussed in Hofmann et al. (1984)]. A

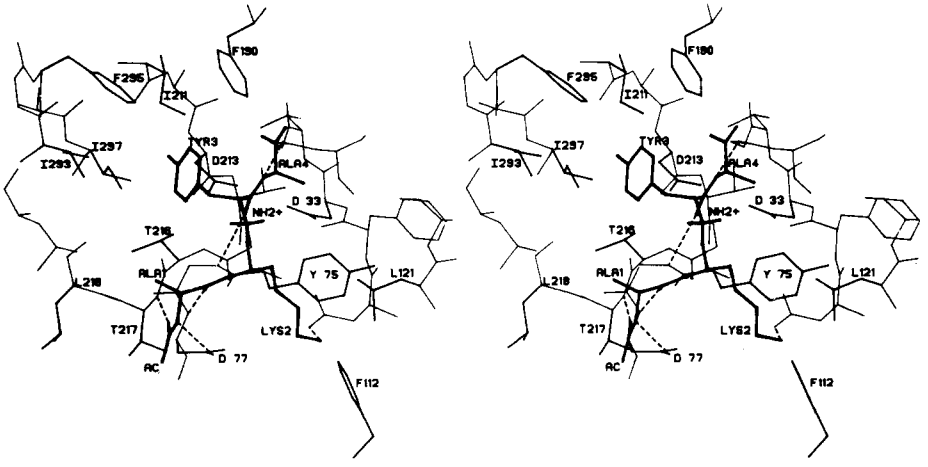


FIGURE 14: Proposed binding of an inhibitor with a reduced peptide bond to the active site of penicillopepsin. The inhibitor is Ac-Ala-Lys-Tyr-Ala-NH_2, with the Lys–Tyr bond as the reduced peptide $-CH_2-NH_2^+-$ (Szelke et al., 1983). In this model, the positively charged nitrogen atom points in toward the two carboxyl groups of Asp-33(32) and Asp-213(215). Such a conformation is possible for this inhibitor but not for a normal substrate due to seriously short nonbonded contacts that occur between adjacent peptide carbonyl oxygen atoms (Hofmann et al., 1984).

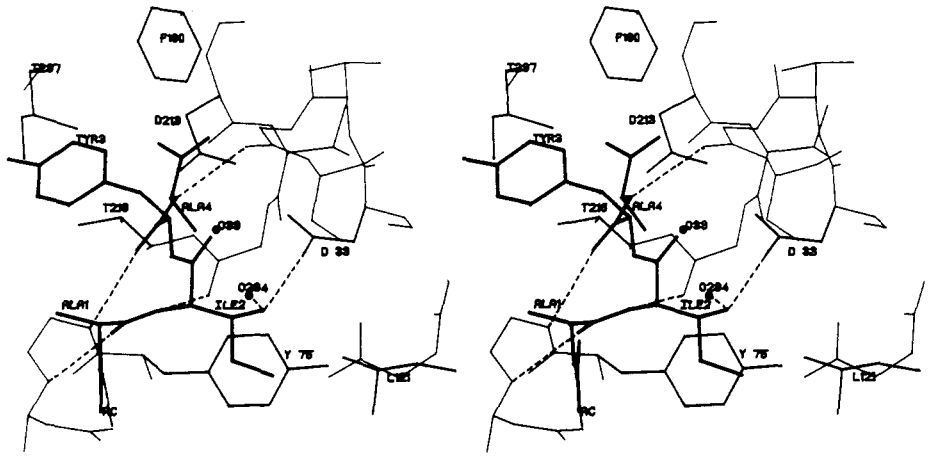


FIGURE 15: A model of a substrate with a β-branched amino acid in the P_1 position bound to the active site of penicillopepsin. An isoleucine residue was fitted to the P_1 lysine side chain of a good substrate. Too close contacts of $C^{\gamma 2}$ with Asp-33(32) prohibit good binding interactions and may explain, in part, why peptides with Ile, Thr, or Val residues in the P_1 position are poor substrates for porcine pepsin (Powers et al., 1977).

proline residue in either P_1 or P_2' precludes these important substrate binding interactions. A prolyl residue in subsite P_2 of a substrate also rules out the possibility of a hydrogen bond to Thr-77 $O^{\gamma 1}$ of pepsin (analogous to the hydrogen bond to the side chain of Asp-77 in penicillopepsin). The restriction for the peptide ϕ angle for prolyl residues to values close to $-70°$ will also have an important influence on the possible conformations of the substrate and the observed enzyme subsite preferences.

The β-branched amino acids threonine, valine, and isoleucine display cleavage probabilities significantly lower than the mean when located in subsite P_1 (Powers et al., 1977). Figure 15 shows why this is so. The lysyl side chain of the "good" penicillopepsin substrate of Figure 8 has been replaced by an isoleucyl side chain with χ_1, χ_2 angles at the conformationally preferred values of $-60°$, $+165°$ (James & Sielecki, 1983). Whereas the C^{β}, C^{γ}, and C^{δ} carbon atoms superimpose almost exactly with the corresponding methylene carbon atoms of the lysyl side chain, the additional $C^{\gamma 2}$ methyl group interferes with the side chain of Asp-33(32) and would also displace the nucleophilic water molecule O284, thus precluding the formation of the tetrahedral intermediate. Any other conformation of the β-branched residues, corresponding to less frequently observed side-chain angles, fails to relieve the unfavorable close contacts equivalent to those displayed in Figure 15 for an isoleucine P_1 residue. It is harder to rationalize, though, why subsites S_3 and S_1 have such a strong dislike for the basic residues Arg, Lys, and His in pepsin.

Finally, the role of a basic group in the catalytic hydrolysis of peptides by the aspartic proteinases can be rationalized. Phenylglyoxal partially inactivates pepsin (Kitson & Knowles, 1971). The basic group involved in this inactivation is most likely Arg-308 (Lys-304 in penicillopepsin). Rather than the disruption of the complicated charge relay mechanism previously suggested (James et al., 1977), we propose the occurrence of a major conformational change in a portion of the C-terminal domain as a result of reaction with phenylglyoxal. Lys-304 (Arg-308) is partially buried in penicillopepsin but is accessible to the bulk solvent via a solvent channel. Penetration of the bulky reagent into this channel could disrupt the surrounding region and lead to an altered domain structure. This domain movement would directly affect the binding subsites S_4, S_2, and S_1' leading to weakened substrate binding interactions and thus partial enzyme inactivation.

The highly refined crystal structures of penicillopepsin and of its complex with a pepstatin analogue at 1.8-Å resolution have provided the framework for a detailed analysis of the stereochemistry of peptide bond hydrolysis. Many of the features of a tentative hydrolytic pathway proposed earlier, based on the 2.8-Å resolution, multiple isomorphous replacement solution of the structure of penicillopepsin, are retained in the present proposal. We have made new suggestions for the nature of the proton donor. The stereoelectronic hypothesis has assisted in the interpretation of proposed intermediates on the catalytic pathway. Many of the features of the serine proteinase mechanism of peptide bond hydrolysis are common to the aspartic proteinases. The present work has also allowed us to rationalize some of the observed subsite preferences and dislikes for substrates binding to pepsin. We expect this model to be useful for further predictions of subsite specificities for other aspartic proteinases.

REFERENCES

Andreeva, N. S., Gustchina, A. E., & Fedorov, A. A. (1981) *Dokl. Akad. Nauk USSR 259*, 1264.

Andreeva, N. S., Zdanov, A. S., Gustchina, A. E., & Fedorov, A. A. (1984) *J. Biol. Chem. 259*, 11353–11365.

Antonov, V. K., Ginodman, L. M., Kapitannikov, Y. V., Barshevskaya, T. N., Gurova, A. G., & Rumsh, L. D. (1978) *FEBS Lett. 88*, 87–90.

Antonov, V. K., Ginodman, L. M., Rumsh, L. D., Kapitannikov, Y. V., Barshevskaya, T. N., Yavashev, L. P., Gurova, A. G., & Volkova, L. I. (1981) *Eur. J. Biochem. 117*, 195–200.

Asbóth, B., & Polgár, L. (1983) *Biochemistry 22*, 117–122.

Barry, C. D., Molnar, C. E., & Rosenberger, F. U. (1976) Technical Memorandum No. 229, Computer Systems Laboratory, Washington University, St. Louis, MO.

Bauer, C.-A., Thompson, R. C., & Blout, E. R. (1976) *Biochemistry 15*, 1296–1299.

Bizzozero, S. A., & Dutler, H. (1981) *Bioorg. Chem. 10*, 46–62.

Blundell, T. L., Jones, H. B., Khan, G., Taylor, G., Sewell, B. T., Pearl, L. H., & Wood, S. P. (1980) *Proc. FEBS Meet. 60*, 281–288.

Bott, R. R., & Davies, D. R. (1983) *Pept.: Struct. Funct., Proc. Am. Pept. Symp., 8th*, 531–540.

Bott, R., Subramanian, E., & Davis, D. R. (1982) *Biochemistry 21*, 6956–6962.

Bürgi, H. B., Dunitz, J. D., & Shefter, E. (1973) *J. Am. Chem. Soc. 95*, 5065–5067.

Clement, G. E. (1973) *Prog. Bioorg. Chem. 2*, 177–238.

Cornish-Bowden, A. J., & Knowles, J. R. (1969) *Biochem. J. 113*, 353–362.
Deslongchamps, P. (1975) *Tetrahedron 31*, 2463–2490.
Drenth, J. (1980) *Recl. Trav. Chim. Pays-Bas 99*, 185–190.
Dunn, B. M., & Fink, A. L. (1984) *Biochemistry 23*, 5241–5247.
Foltmann, B. (1981) *Essays Biochem. 17*, 52–84.
Fruton, J. S. (1970) *Adv. Enzymol. Relat. Areas Mol. Biol. 33*, 401–443.
Fruton, J. S. (1976) *Adv. Enzymol. Relat. Areas Mol. Biol. 44*, 1–36.
Fruton, J. S. (1977) *Proc. Am. Philos. Soc. 121*, 309–315.
Fujinaga, M., Read, R. J., Sielecki, A., Ardelt, W., Laskowski, M., Jr., & James, M. N. G. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4868–4872.
Gripon, J.-C., & Hofmann, T. (1981) *Biochem. J. 193*, 55–65.
Hendrickson, W. A., & Konnert, J. H. (1980) *Comput. Crystallogr., Lect. Int. Winter Sch.*, 13.01–13.23.
Hine, J. (1977) *Adv. Phys. Org. Chem. 15*, 1–61.
Hofmann, T. (1974) *Adv. Chem. Ser. No. 136*, 146–185.
Hofmann, T., & Hodges, R. S. (1982) *Biochem. J. 203*, 603–610.
Hofmann, T., & Fink, A. L. (1984) *Biochemistry 23*, 5247–5256.
Hofmann, T., Hodges, R. S., & James, M. N. G. (1984) *Biochemistry 23*, 635–643.
Hollands, T. R., Voynick, I. M., & Fruton, J. S. (1969) *Biochemistry 8*, 575–588.
Hosie, L., Marshall, P. J., & Sinnott, M. (1984) *J. Chem. Soc., Perkin Trans. 2*, 1121–1131.
Hsu, I.-N., Delbaere, L. T. J., James, M. N. G., & Hofmann, T. (1977a) *Nature (London) 266*, 140–145.
Hsu, I.-N., Delbaere, L. T. J., James, M. N. G., & Hofmann, T. (1977b) in *Acid Proteases, Structure, Function and Biology* (Tang, J., Ed.) pp 61–81, Plenum Press, New York.
Huber, R., & Bode, W. (1978) *Acc. Chem. Res. 11*, 114–121.
Inagami, T., Murakami, K., Misono, K., Workman, R. J., Cohen, S., & Suketa, Y. (1977) in *Acid Proteases, Structure, Function and Biology* (Tang, J., Ed.) pp 225–248, Plenum Press, New York.
James, M. N. G. (1980) *Can. J. Biochem. 58*, 251–271.
James, M. N. G., & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299–361.
James, M. N. G., Hsu, I.-N., & Delbaere, L. T. J. (1977) *Nature (London) 267*, 808–813.
James, M. N. G., Sielecki, A., Brayer, G. D., Delbaere, L. T. J., & Bauer, C.-A. (1980) *J. Mol. Biol. 144*, 43–88.
James, M. N. G., Hsu, I.-N., Hofmann, T., & Sielecki, A. (1981) in *Structural Studies on Molecules of Biological Interest* (Dodson, G., Glusker, J. P., & Sayre, D., Eds.) pp 350–389, Clarendon Press, Oxford.
James, M. N. G., Sielecki, A., Salituro, F., Rich, D. H., & Hofmann, T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6137–6141.
James, M. N. G., Sielecki, A. R., & Moult, J. (1983) *Pept.: Struct. Funct., Proc. Am. Pept. Symp., 8th*, 521–530.
Jenkins, J. A., Tickle, I., Sewell, T., Ungaretti, L., Wollmer, A., & Blundell, T. (1977) in *Acid Proteases, Structure, Function and Biology* (Tang, J., Ed.) pp 43–60, Plenum Press, New York.
Kitson, T. M., & Knowles, J. R. (1971) *FEBS Lett. 16*, 337–338.
Knowles, J. R. (1970) *Philos. Trans. R. Soc. London, Ser. B 257*, 135–146.
Marciniszyn, J., Jr., Hartsuck, J. A., & Tang, J. J. N. (1976) *J. Biol. Chem. 251*, 7088–7094.
Marshall, G. R. (1976) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 35*, 2494–2501.
Newmark, A., & Knowles, J. R. (1975) *J. Am. Chem. Soc. 97*, 3557–3559.
Powers, J. C., Harley, A. D., & Myers, D. V. (1977) in *Acid Proteases, Structure, Function and Biology* (Tang, J., Ed.) pp 141–157, Plenum Press, New York.
Read, R. J., Fujinaga, M., Sielecki, A. R., & James, M. N. G. (1983) *Biochemistry 22*, 4420–4433.
Rich, D. H., Salituro, F. G., & Holladay, M. W. (1983) *Pept.: Struct. Funct., Proc. Am. Pept. Symp., 8th*, 511–520.
Robertus, J. D., Kraut, J., Alden, R. A., & Birktoft, J. J. (1972) *Biochemistry 11*, 4293–4303.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.
Sharon, N., Grisaro, V., & Neumann, H. (1962) *Arch. Biochem. Biophys. 97*, 219–221.
Sielecki, A. R., Hendrickson, W. A., Broughton, C. G., Delbaere, L. T. J., Brayer, G. D., & James, M. N. G. (1979) *J. Mol. Biol. 134*, 781–804.
Sielecki, A. R., James, M. N. G., & Broughton, C. G. (1982) in *Computational Crystallography* (Sayre, D., Ed.) pp 409–419, Clarendon Press, Oxford.
Sodek, J., & Hofmann, T. (1970) *Can. J. Biochem. 48*, 425–431.
Subramanian, E., Liu, M., Swan, I. D. A., & Davies, D. R. (1977a) in *Acid Proteases, Structure, Function and Biology* (Tang, J., Ed.) pp 33–60, Plenum Press, New York.
Subramanian, E., Swan, I. D. A., Liu, M., Davies, D. R., Jenkins, J. A., Tickle, I. J., & Blundell, T. L. (1977b) *Proc. Natl. Acad. Sci. U.S.A. 74*, 556–559.
Szelke, M., Jones, D. M., Atrash, B., & Hallett, A. (1983) *Pept.: Struct. Funct., Proc. Am. Pept. Symp., 8th*, 579–582.
Tang, J., Sepulveda, P., Marciniszyn, J., Chen, K. C. S., Huang, W.-Y., Too, N., Liu, D., & Lanier, J. P. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3437–3439.
Tang, J., James, M. N. G., Hsu, I.-N., Jenkins, J. A., & Blundell, T. L. (1978) *Nature (London) 271*, 618–621.
Wang, T.-T., & Hofmann, T. (1976) *Biochem. J. 153*, 691–699.
Wang, T.-T., Dorrington, K. J., & Hofmann, T. (1974) *Biochem. Biophys. Res. Commun. 57*, 865–869.