Crystallographic Analysis of Transition-State Mimics Bound to Penicillopepsin: Phosphorus-Containing Peptide Analogues†

Marie E. Fraser,‡ Natalie C. J. Strynadka,‡ Paul A. Bartlett,§ John E. Hanson,§ and Michael N. G. James*,‡

Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7, and Department of Chemistry, University of California, Berkeley, California 94720

Received January 7, 1992; Revised Manuscript Received March 13, 1992

ABSTRACT: The molecular structures of three phosphorus-based peptide inhibitors of aspartyl proteinases complexed with penicillopepsin [**1**, Iva–L-Val–L-Val–Sta^{P}OEt [Iva = isovaleryl, Sta^P = the phosphinic acid analogue of statine [(*S*)-4-amino-(*S*)-3-hydroxy-6-methylheptanoic acid] (IvaVVSta^{P}OEt)]; **2**, Iva–L-Val–L-Val–L-Leu^P–(O)Phe–OMe [Leu^P = the phosphinic acid analogue of L-leucine; (O)Phe = L-3-phenyllactic acid; OMe = methyl ester] [Iva VVL^P(O)FOMe]; and **3**, Cbz–L-Ala–L-Ala–L-Leu^P–(O)-Phe–OMe (Cbz = benzyloxycarbonyl) [CbzAAL^P(O)FOMe]] have been determined by X-ray crystallography and refined to crystallographic agreement factors, R ($= \sum||F_o| - |F_c||/\sum|F_o|$), of 0.132, 0.131, and 0.134, respectively. These inhibitors were designed to be structural mimics of the tetrahedral transition-state intermediate encountered during aspartic proteinase catalysis. They are potent inhibitors of penicillopepsin with K_i values of **1**, 22 nM; **2**, 2.8 nM; and **3**, 1600 nM, respectively [Bartlett, P. A., Hanson, J. E., & Giannousis, P. P. (1990) *J. Org. Chem. 55*, 6268–6274]. All three of these phosphorus-based inhibitors bind virtually identically in the active site of penicillopepsin in a manner that closely approximates that expected for the transition state [James, M. N. G., Sielecki, A.R., Hayakawa, K., & Gelb, M. H. (1992) *Biochemistry 31*, 3872–3886]. The *pro-S* oxygen atom of the two phosphonate inhibitors and of the phosphinate group of the Sta^P inhibitor make very short contact distances (~2.4 Å) to the carboxyl oxygen atom, $O^{\delta 1}$, of Asp33 on penicillopepsin. We have interpreted this distance and the stereochemical environment of the carboxyl and phosphonate groups in terms of a hydrogen bond that most probably has a symmetric single-well potential energy function. The *pro-R* oxygen atom is the recipient of a hydrogen bond from the carboxyl group of Asp213. Thus, we are able to assign a neutral status to Asp213 and a partially negatively charged status to Asp33 with reasonable confidence. Similar very short hydrogen bonds involving the active site glutamic acid residues of thermolysin and carboxypeptidase A and the *pro-R* oxygen of bound phosphonate inhibitors have been reported [Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H., & Matthews, B. W. (1987) *Biochemistry 26*, 8542–8553; Kim, H., & Lipscomb, W. N. (1991) *Biochemistry 30*, 8171–8180]. The hydrogen-bonding scheme proposed for the phosphonate inhibitors with the active site of penicillopepsin at pH 4.4 in the crystals can be used to rationalize the pH dependence of the inhibition constant. These data indicate that there are pK_a values near 3.5 and near 5.5 which can be associated with the titration of two protons in the molecular complex. The binding of IvaVVL^P(O)FOMe and CbzAAL^P(O)FOMe to penicillopepsin has allowed, for the first time, experimental characterization of the extensive hydrophobic interactions between the P_1' residue and the S_1' binding site. This structure confirms our proposals for substrate binding to penicillopepsin [James, M. N. G., & Sielecki, A. R. (1985) *Biochemistry 24*, 3701–3713]. Detailed comparisons of the phosphonate inhibitor binding mode with that of a difluorostatone peptide mimic of the tetrahedral transition state intermediate show that these two inhibitors are shifted by ~0.6 Å relative to one another. In spite of this global shift both inhibitors exhibit equivalently high affinities for the active site of penicillopepsin. Comparison of the binding of IvaVVL^P-(O)FOMe and CbzAAL^P(O)FOMe suggests that the origin of the 570-fold difference in binding affinity may be entropic, due to a different number of ordered water molecules displaced in the two cases.

Penicillopepsin (EC 3.4.23.X) is a fungal aspartic proteinase having 323 amino acids in a single polypeptide chain. It was one of the first members of this family to have its three-dimensional structure elucidated by X-ray crystallographic methods (Hsu et al., 1977; James & Sielecki, 1983). This enzyme family is characterized by two similarly folded predominantly β-sheet domains, each of which contributes one aspartic acid side-chain carboxyl group to the active site of substrate hydrolysis (Tang et al., 1978; James & Sielecki, 1987; Davies, 1990; Sielecki et al., 1990).

The search for effective inhibitors of human renin, an aspartic proteinase involved in the regulation of the angiotensin system, has resulted in a large number of chemical compounds designed to mimic the tetrahedral transition-state intermediate encountered in the hydrolysis of peptide substrates (Boger, 1985; Szelke et al., 1983; Szelke, 1985). Many of the initial syntheses were based on statine, an unusual amino acid that has a tetrahedral carbon mimic of the transition-state intermediate (Marciniszyn et al., 1976; Marshall, 1976). Com-

† The work in Berkeley was supported by a grant from the National Institutes of Health (CA-22747) to P.A.B. The crystallographic work was supported by the Medical Research Council of Canada grant to the Group in Protein Structure and Function at the University of Alberta. M.F. was an AHFMR postdoctoral fellow; N.C.J.S. was an AHFMR doctoral student.

* Author to whom correspondence should be addressed.

‡ University of Alberta, Edmonton.

§ University of California, Berkeley.

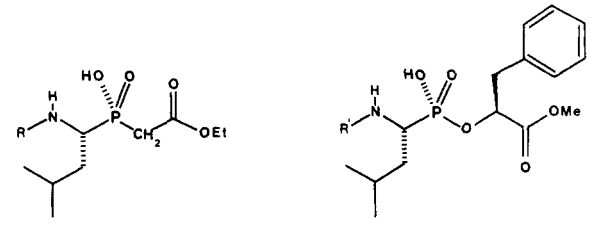
0006-2960/92/0431-5201$03.00/0

pounds containing statine [(3*S*,4*S*)-3-hydroxy-4-amino-6-methylheptanoic acid, Sta] have been bound to several of the fungal enzymes and crystallized. The crystal structures of these complexes have confirmed that the (*S*)-3-hydroxyl group of the statine residue is indeed bound between the two active-site aspartate carboxyl groups in a manner expected for a tetrahedral intermediate of a peptide substrate (James et al., 1982, 1985; James & Sielecki, 1985; Bott et al., 1982; Cooper et al., 1989). The intermolecular interactions observed for IvaValValStaOEt (Iva = isovaleryl) binding to penicillopepsin aided in the development of proposals for a possible catalytic pathway for the aspartic proteinases (James & Sielecki, 1985, 1987). Other proposals have also been advanced (Suguna et al., 1987b; Polgar, 1987; Pearl & Blundell, 1984; James et al., 1992).

The various mechanistic proposals agree that the aspartic proteinases catalyze peptide bond hydrolysis via noncovalent catalysis as first suggested by Fruton (1976). Like the zinc metalloproteinases, the aspartic proteinases catalyze the general-base-assisted direct addition of water to the carbonyl carbon atom of the amide linkage. The resulting *gem*-diol tetrahedral intermediate would be stabilized in the active site by hydrogen-bonding interactions to the carboxyl groups of Asp33 and Asp213 as was proposed for penicillopepsin (James & Sielecki, 1985, 1987). The structure of the molecular complex of penicillopepsin with a difluorinated keto analogue of statine revealed the details of these intimate interactions for the *gem*-diol, further supporting the proposals for a probable catalytic pathway (James et al., 1992). The only serious point of disagreement among the several mechanistic proposals has been the identification of the nucleophilic H_2O molecule in the active site (James & Sielecki, 1985, 1987; Suguna et al., 1987b; Pearl & Blundell, 1984). In the past, the Edmonton group has supported a proposal involving a water molecule that is bound to Asp32, whereas the other groups suggest that the nucleophile is the water molecule bound between both of the active site carboxyl groups (Asp32 and Asp215 in the pepsin numbering). It is now generally accepted that the nucleophilic H_2O molecule is the one bound between the two carboxyl groups of Asp32 and Asp215 (James et al., 1992).

Phosphorus-containing analogues of oligopeptide substrates, some occurring naturally [phosphoramidon (Suda et al., 1973; Komiyama et al., 1975)] and others produced synthetically (Holmquist, 1977; Nishino & Powers, 1979; Holmquist & Vallee, 1979; Bartlett & Marlowe, 1983, 1987) are exceedingly potent inhibitors of the zinc metalloproteinases, thermolysin (TLN), and carboxypeptidase A (CPA). It was proposed that phosphorous-containing inhibitors were likely transition-state analogues of the metalloproteinases (Weaver & Matthews, 1977). More importantly, it has been demonstrated experimentally that some of the phosphorus-containing inhibitors are indeed transition-state analogues of the equivalent peptide substrates (Bartlett & Marlowe, 1983; Hanson et al., 1989). Crystallographic studies of several of these inhibitors bound to TLN (Tronrud et al., 1987; Holden et al., 1987) and to CPA (Kim & Lipscomb, 1990, 1991) confirm that the tetrahedral phosphonyl (and phosphinyl) moieties are bound noncovalently in the enzymes' active sites in a manner consistent with that of a transition-state mimic. These metalloenzyme complexes are characterized by one short (~2.1 Å) Zn^{2+} ion to phosphonyl oxygen distance and a longer distance (~3.1 Å) from the second phosphonyl oxygen atom to the metal ion. The high-resolution refinement of diisopropylphosphoryl trypsin showed that phosphorous-containing inhibitors are also mimics of the tetrahedral tansition state in serine proteinase hydrolysis (Chambers & Stroud, 1977). Oligopeptide analogues incorporating an electrophilic phosphorus group are irreversible inhibitors of a variety of serine proteinases (Sampson & Bartlett, 1991). Crystallographic analysis of inhibitor binding to α-lytic proteinase shows that the phosphoryl group is covalently bound to the active site serine hydroxyl (Bone et al., 1991).

Because of the mechanistic parallels between the zinc metalloproteinases and the aspartic proteinases, several phosphorus-containing oligopeptides have been synthesized and tested as inhibitors of pepsin (Bartlett & Kezer, 1984), renin (Allen et al., 1989), penicillopepsin (Bartlett et al., 1990), and the HIV protease (Dreyer et al., 1989; Grobelny et al., 1990). These studies have shown that, indeed, they are good inhibitors of the aspartic proteinases. In order to characterize these compounds as transition-state mimics more fully, we have carried out the crystal structure analyses of three of the synthetic inhibitors complexed with fungal penicillopepsin.



1: R = Iva-Val-Val- (IvaVVSta^{P}OEt); **2**: R' = Iva-Val-Val- (IvaVVL^P(O)FOMe); **3**: R' = Cbz-Ala-Ala- (ZAAL^P(O)FOMe)

The measured values for the K_i's of these compounds with penicillopepsin are **1**, 22 nM; **2**, 2.8 nM; and **3**, 1600 nM (Bartlett et al., 1990). The present analysis provides for rationalization of these numbers on a firm structural basis. The refined crystallographic atomic coordinates of all three complexes have been deposited with the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Experimental Procedures

Synthesis of Inhibitors and Evaluation of Inhibition Constants. The syntheses of IvaVVSta^{P}OEt, IvaVVL^P(O)FOMe, and CbzAAL^P(O)FOMe along with their evaluation as inhibitors of penicillopepsin and pepsin have been described previously (Bartlett et al., 1990).

Crystallization and Data Collection Details. Samples of penicillopepsin were generously provided by Professor T. Hofmann, University of Toronto. The inhibitors were complexed to penicillopepsin in much the same way as was done previously with other Sta-based inhibitors (James et al., 1982), except that in order to dissolve the phosphastatine-containing inhibitor, a minimal volume of dimethylformamide was added to the methanol. Crystallization was achieved from solutions that were 10 mg/mL complex, 34–40% saturated $(NH_4)_2SO_4$, and 0.1 M $NaC_2H_3O_2$, pH 4.4, placed in hanging drop experiments. The crystals grew at 20 ± 2 °C to adequate sizes in 2–3 weeks; they exhibited a morphology typical of crystals of penicillopepsin that have been complexed with inhibitor molecules (James et al., 1982). The unit cell dimensions of the three complex crystals are given in Table I.

Intensity data for the crystal of the molecular complex of penicillopepsin with the phosphastatine-containing peptide (IvaVVSta^{P}OEt) were collected on an Enraf-Nonius CAD4 diffractometer. Only a single quadrant of data on one crystal was collected (*hkl* and *hk*$\bar{l}$). Additional reflections were collected to perform an intensity decay correction (Table I). The absorption corrections were made from measurement of

Table I: Crystal Data for Phosphonate Inhibitor Complexes with Penicillopepsin

	penicillopepsin complexed to		
	IvaVVSta^{P}OEt	IvaVVL^P(O)-FOMe	ZAAL^P(O)-FOMe
space group	*C*2	*C*2	*C*2
cell dimensions			
a (Å)	97.48	97.50	97.40
b (Å)	46.55	46.46	46.59
c (Å)	66.42	65.93	66.29
β (°)	116.14	115.2	115.7
data collection	CAD4 diffractometer	SDMS area detector	SDMS area detector
X-rays Cu K$\bar{\alpha}$	40 kV, 26 mA	40 kV, 40 mA	40 kV, 40 mA
resolution (Å)	45–1.8	60–1.7	60–1.7
total measurements	29 429	88 917	100 309
unique measurements	24 917	26 533	26 564
total possible (unique)	26 533	29 601	29 687
merging R^a	0.033	0.060	0.055
isomorphous R^b	0.275	0.179	0.242
decay range^c $I_t/I_{t=0}$	0.95–0.71		
max absorption correction	1.381		

$^a R_M = \sum(I_i - \bar{I})/\sum I_i$, where I_i are the intensities of the multiply measured reflections and $\bar{I}$ is their average value. $^b R_I = \sum||F_{complex}| - |F_{native}||/\sum|F_{complex}|$, where $|F_{complex}|$ and $|F_{native}|$ are the structure factor amplitudes measured from crystals of the complexed and native enzyme, respectively. ^{c} In addition to five standard reflections measured at regular intervals throughout the data collection, the initial 1500 reflections were measured a second time at the end of the data collection. This provided an estimate of the intensity decay as a function of time and resolution (Hendrickson, 1976). The ratios given above correspond to the values at the minimum and maximum resolution ranges, respectively.

Table II: Restrained-Parameter, Least-Squares Refinement Statistics for the Penicillopepsin–Phosphonate Inhibitor Complex Structures

parameters	IvaVVSta^P-OEt	IvaVVL^P-(O)FOMe	ZAAL^P-(O)FOMe
R factor (= $\sum||F_o| - |F_c||/\sum|F_o|$)^a	0.132	0.131	0.134
resolution range (Å)	8.0–1.8	8.0–1.7	8.0–1.7
no. of reflections [$\|F\| \geq 3\sigma(F)$]	22 145	24 760	25 614
R factor (all data) range of resolution	0.142	0.148	0.150
no. of protein atoms (+sugars)	2386	2390	2386
no. of inhibitor atoms	35	32	42
no. of solvent atoms^b	285	282	315
rms deviations from ideal values			
covalent bond distances (Å)	0.022	0.024	0.019
interbond angle distances (Å)	0.044	0.047	0.043
planar 1–4 distances (Å)	0.046	0.049	0.045
planar groups planarity (Å)	0.018	0.021	0.019
chiral centers (Å^3)	0.213	0.203	0.163
nonbonded contact distances			
single torsions (1–4)	0.250	0.246	0.244
multiple torsion	0.182	0.181	0.174
possible hydrogen bonds^c	0.204	0.226	0.227
planar peptide group torsion angles	1.7	3.9	3.7

$^a |F_o|$ and $|F_c|$ are the measured and calculated structure factor amplitudes, respectively. ^{b} In addition to refining the molecular complex of penicillopepsin and the three phosphorous-based inhibitors, several new features of the penicillopepsin molecule itself have become apparent. Two *O*-linked sugar residues were discovered: an α-D-mannose on Ser3 O$^\gamma$ and a tentatively assigned xylopyranose monosaccharide on Thr7 O$^{\gamma 1}$. In addition, a strongly bound solvent feature was reinterpreted as a sulfate anion, and an extensive revision of the strongly bound, highly ordered solvent structure was performed. Lastly, alternative side-chain conformations were detected for a number of residues. ^{c} No restraints were applied on possible hydrogen bonds when the distance between the atoms involved was larger than the sum of the corresponding van der Waals radii.

the variation in X-ray transmission observed for ψ scans on two reflections (North et al., 1968). The large value of the isomorphous agreement index, R_I, is due primarily to the lack of isomorphism with the native crystals (the *c*-axis of the crystals of the complex is larger by ~1.5% and the β-angle is increased by ~0.7° over those parameters in the native penicillopepsin crystals).

Intensity data from one crystal of each of the two phosphonate inhibitors [IvaVVL^P(O)FOMe and CbzAAL^P(O)-FOMe] bound to penicillopepsin were collected on the SDMS twin area detector system (Hamlin, 1985) presently installed at the University of Alberta (Table I). The general data collection strategy was similar to that already described (Xuong et al., 1985). Data processing, including local scaling, Lorentz and polarization corrections, and merging to a unique data set, was carried out with the program package supplied (Howard et al., 1985). Redundancy factors of ~3.3 and 4.0 resulted after the symmetry equivalent reflections were averaged. The overall merging *R* value for each of these two data sets was ~6%; the weak data near the limits of resolution (1.7 Å) merged less well (~20%). The data sets collected for all three complexes are of similar quality. The phosphastatine complex data set is of only slightly lower resolution (1.8 Å) than those of the two phosphonate complexes (1.7 Å).

Initial coordinates for the atoms of the three inhibitors were derived from difference electron density maps that were computed with coefficients $|F_o| - |F_c|$ and calculated phases, α_c, from the refined complexes of penicillopepsin with other statine-containing inhibitors (James et al., 1982, 1985, 1992). The atoms of the other inhibitors and the solvent molecules in the active site were omitted in the phase calculations. Molecular models for the three inhibitors and of waters in the active site were readily derived from these difference electron density maps. Least-squares refinement was initiated for the three complexes. Restrained-parameter least-squares (Hendrickson, 1985; Hendrickson & Konnert, 1980) was used exclusively for the phosphastatine-containing complex (IvaVVSta^{P}OEt). The initial refinement stages for the two phosphonate-based inhibitors were done with the program XPLOR (Brünger et al., 1987). The strategy of a heating stage at 2000 K followed by annealing at 300 K was essentially the same as described in other papers from that laboratory (Brünger, 1988, and references cited therein). Parameters for the groups that are not included with XPLOR were chosen using data from the Cambridge Structural Database (Allen et al., 1983). These molecular dynamics steps with the additional energy restraint of the crystallographic data reduced the initial *R* factors of ~0.36 to *R* factors of ~0.22 and 0.20 for each complex. At this point restrained-parameter least-squares refinement using the Konnert–Hendrickson programs was initiated.

Restraints on bond lengths, bond angle lengths, planar groups, and chiral volumes were included throughout the three refinements. The weights (target values) that were applied to these parameters in the final stages reflected the root-mean-square (rms) deviations from ideal values [0.020, 0.045, and 0.020 Å and 0.180 Å^3, respectively (Table II)]. Restraints on the nonbonded contacts were weak, and the restraints applied to the thermal parameters were relaxed during the final refinement stages. No restraints on side-chain conformational angles or possible hydrogen bonds with lengths greater than the sum of the van der Waals radii of the atoms involved were applied at any stage. The solvent structure was analyzed for each complex in a manner similar to that described by Fujinaga et al. (1985).

Results and Discussion

The least-squares refinement process has produced excellent agreement between the calculated and observed structure factor amplitudes (Table II). The relatively low *R* factors (~0.13) have been achieved while maintaining close agreement of the molecular structure to that expected from the

Table III: Contact Distances for Possible Hydrogen Bonding between the Phosphonyl Inhibitors and Penicillopepsin

	atoms involved			distances (Å)		
	inhibitor[a]		penicillopepsin	···Sta^{P}OEt	IvaVV···	CbzAA···
P_3	Val (Ala) NH	···	Thr217 $O^{\gamma 1}$	2.81	2.88	2.94
	Val (Ala) CO	···	Thr217 NH	3.07	3.03	2.99
P_2	Val (Ala) NH	···	Asp77 $O^{\delta 1}$	2.88	2.93	2.79
	Val (Ala) CO	···	Asp77 NH	3.09	3.00	3.00
		···	Gly76 NH	3.27	3.27	3.19
P_1	Sta^P (Leu^P) NH	···	Gly215 CO	3.15	3.22	3.19
	Sta^P (Leu^P) PO	···	Asp33 $O^{\delta 1}$	(3.67	3.71	3.70)[b]
			Asp33 $O^{\delta 2}$	3.05	3.21	3.17
			Asp213 $O^{\delta 1}$	2.56	2.60	2.65
			Asp213 $O^{\delta 2}$	3.08	3.02	2.96
	Sta^P (Leu^P) POH	···	Asp33 $O^{\delta 1}$	2.39	2.40	2.54
			Asp33 $O^{\delta 2}$	(3.22	3.33	3.44)[b]
P_1'	Sta^P (Leu^P) CO	···	Gly76 NH	2.82	2.89	2.93

[a] The nomenclature of Schechter and Berger (1967) is used to describe the interactions of enzyme and inhibitors. P_n and P_n' denote residues on either side of the scissile bond, by definition between P_2 and P_1'. S_n and S_n' denote the corresponding binding pockets in the enzyme. [b] The stereochemistry of these interactions is inconsistent with hydrogen bonding, but the distances are included here for completeness.

amino acid templates deduced from small molecule crystal structures. In addition to the bound inhibitor molecules, these structural analyses have identified two *O*-linked sugar residues covalently attached to penicillopepsin. The solvent structures in each complex have been carefully assessed, and the final description of the solvent structure includes approximately 300 highly ordered solvent molecules. The solvent molecules were refined as oxygen atoms; only their positions and atomic *B* factors were included as refineable parameters.

The tetrahedral character of each of the phosphorus-based mimics of the transition-state intermediate was clearly revealed in the initial and subsequent electron-density maps (Figure 1). In each of the three complexes, this region was one of relatively rigid geometry as the *B* factors of the atoms involved were all <10 Å^2. The coordinate accuracy in these regions of low *B* factor [estimated from Luzzati (1952) plots or from Cruickshank's method (Cruickshank, 1949, 1967)] is of the order of 0.06 Å for each complex. Thus, there is a high degree of confidence in the nonbonded distances between atoms in the vicinity of the active site as a result of the high-resolution refinements.

The two phosphonyl groups, Leu^P(O)FOMe, of the two phosphonate-based inhibitors and the phosphinyl group of the Sta^{P}OEt containing peptide all bind in the active site of penicillopepsin in essentially an identical fashion (Figure 2a,b). This is in spite of the chemical differences between the phosphinate and the phosphonate groups. Of course, the different C-terminal groups, L-β-phenyllactate methyl ester and the ethyl ester moiety of the phosphastatine, are bound differently. There is an alternative binding mode for the terminal CH_3 of the ethyl group of the ester in the phosphastatine inhibitor, whereas the methyl ester of the (O)Phe is more ordered in the crystals of the phosphonate complexes. The mean temperature factor ($\bar{B}$) for the atoms of the methyl esters are 12.0 and 15.0 Å^2 for IvaVVL^P(O)FOMe and CbzAAL^P(O)FOMe, respectively.

Hydrogen-Bonding Scheme at the Active Site. The hydrogen-bond distances (Figure 2, Table III) between oxygen atoms of the carboxyl groups of Asp33, Asp213, and the phosphonyl and phosphinyl groups of the three inhibitors define the proton positions (Figure 3). This is unusual for an X-ray crystal structure determination of a large protein. Hydrogen atom positions are seen reliably with X-rays only for those structures done at the highest attainable resolution [e.g., crambin (Whitlow & Teeter, 1985) and avian pancreatic polypeptide (Glover et al., 1983)] because their contributions to the X-ray scattering is small. Accurate hydrogen atom positions can best be determined by neutron diffraction (Mason et al., 1984; Teeter & Kossiakoff, 1984). The hydrogen positions in the structures of these three complexes can be inferred reliably because of the short nonbonded distances between the oxygen atoms of the active site (Figure 2).

The distance of 2.4–2.5 Å between the *pro-S* oxygen atoms of the phosphonate inhibitors and the $O^{\delta 1}$ atom of Asp33 is indicative of a very short hydrogen bond between the carboxyl group and the phosphonate (Figure 3). Such features have been well characterized in many small molecule crystal structures of dicarboxylic acids, e.g., the maleate monoanion (Darlow & Cochran, 1961; Ellison & Levy, 1965; James & Williams, 1974; James & Matsushima, 1976) and many other acid salts of carboxylic acids (Speakman, 1972). The average O···O distance in these structures is close to 2.44 Å, and, for many, the O···H···O bond can be represented by a symmetrical single-well potential energy curve with the proton position located midway between the two oxygen atoms. The very much longer distance of ~3.2 Å from the *pro-R* oxygen of the phosphonate to the $O^{\delta 2}$ atom of Asp33 confirms that there is only one proton between the phosphonate and the carboxyl group of Asp33. In Figure 3b,c we show the likely hydrogen bonding at a pH between ~3.5 and ~4.5 (Bartlett et al., 1990). In this interpretation, we have chosen to represent a double-well potential with the proton localized either on the carboxyl group (Figure 3b) or on the phosphonyl group (Figure 3c). From the results of this X-ray diffraction experiment, it is not possible to distinguish between this situation and one with a symmetrical single-well potential energy. The *pro-S* oxygen atom of the phosphonyl group lies on the least-squares plane of the carboxyl group of Asp33 (within 0.15 Å), whereas the *pro-R* oxygen atom is more than 1.0 Å from the same carboxyl plane (Figure 2).

The *pro-R* oxygen atom of the phosphonyl group (and phosphinyl group) is a hydrogen-bond acceptor. The carboxyl group of Asp213 donates a proton to this oxygen atom. The resultant hydrogen bond is longer than the one between the carboxyl group of Asp33 and the phosphonate *pro-S* oxygen atom (2.6 vs 2.4 Å). As well, we can place the proton on $O^{\delta 1}$ of Asp213 rather than on $O^{\delta 2}$ (Table III, Figure 3) because the *pro-R* oxygen to $O^{\delta 1}$ distance is consistently ~0.4 Å shorter than the *pro-R* oxygen to Asp213 $O^{\delta 2}$ distance (~3.1 Å) for the three inhibitors. The shorter distance and the stereochemical disposition of the heavier atoms make it much more favorable to allocate the Asp213 carboxyl proton position as shown in the first three panels of Figure 3.

The *pro-R* oxygen atom is in the plane of the carboxyl group of Asp213 (within 0.1 Å). On the other hand, the $O^{\delta 2}$ atom of Asp33 is more than 1.0 Å out of the plane of the Asp213 carboxyl group. Thus, it is very unlikely that there is a hydrogen bond between the two aspartate carboxyl groups as has been suggested for the native enzyme (James & Sielecki, 1983). The average distance from Asp33 $O^{\delta 2}$ to Asp213 $O^{\delta 1}$ of ~3.1 Å seen in these three complexes is an acceptable van der Waals contact distance between oxygen atoms and is ~0.3 Å longer than the distance observed in the native penicillopepsin molecule (2.8 Å). Recent interpretations of the hydrogen bonding in the native enzyme (James et al., 1992) suggest that there is no hydrogen bond between Asp33 and Asp213 in the native penicillopepsin either.

Panels b and c of Figure 3 show that the net negative charge in the active site when the three phosphorus-based inhibitors

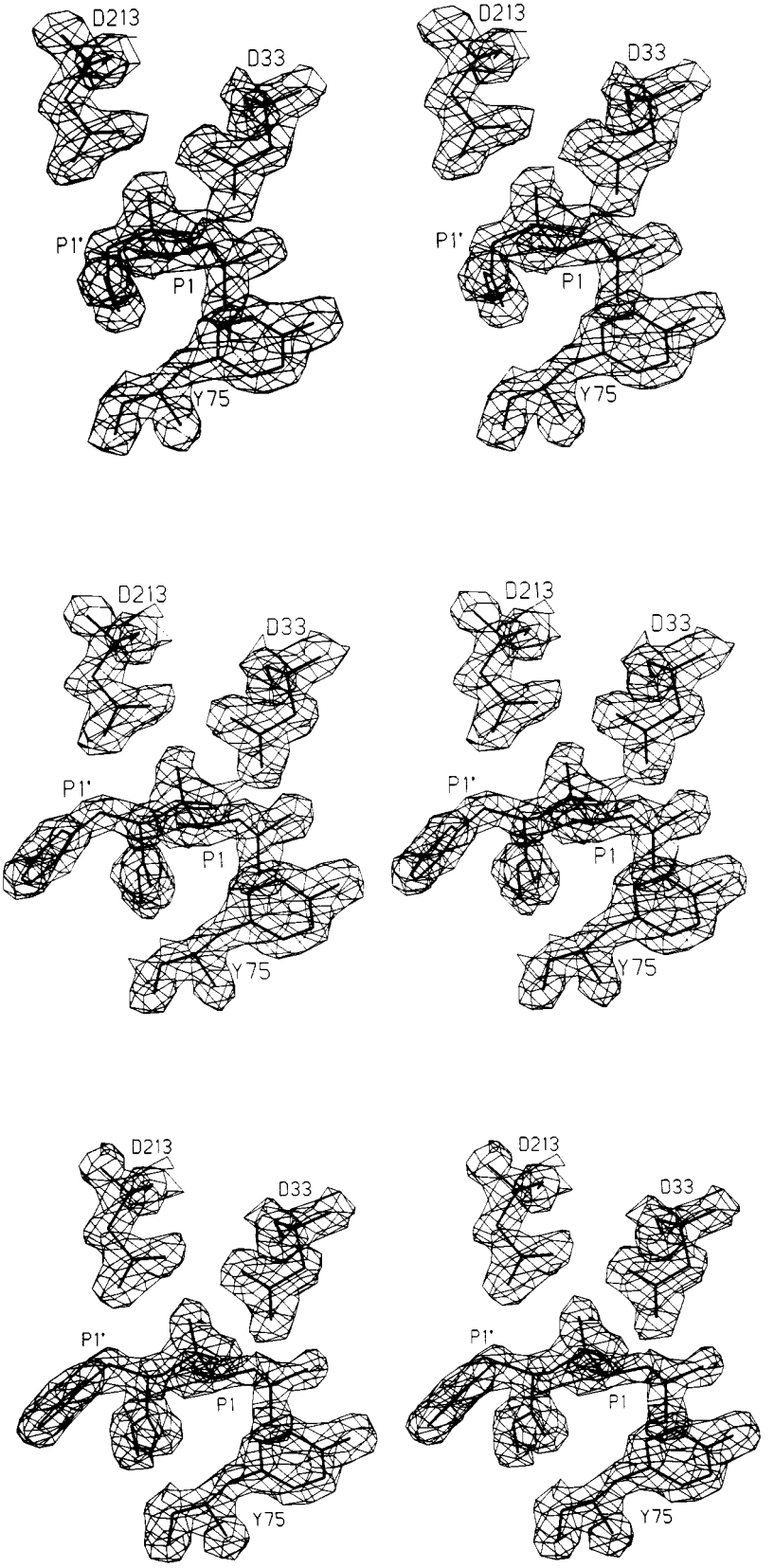


FIGURE 1: Electron density distributions of the final $(2|F_o| - |F_c|)$ α_c maps contoured in the immediate vicinity of the catalytic aspartic acid residues, Asp33 and Asp213. Tyr75, on the flap, is also depicted in the closed conformation that is stabilized by the binding of the inhibitors in the active site cleft. The contour level is 0.60 eÅ^{-3}. Data in the range 8.0–1.7 Å were used in the calculation of these maps. (a, top) Phosphostatine (IvaVVSta^{P}OEt) complex. (b, middle) Phosphonate [IvaVVL^P(O)FOMe] complex. (c, bottom) Phosphonate [CbzAAL^P(O)FOMe] complex.

are bound to penicillopepsin is localized either on the phosphonyl moiety or on the carboxyl group of Asp33. Our interpretation of the observed distances and stereochemistry of the groups involved has the carboxyl group of Asp213 always protonated and therefore neutral (uncharged). Both carboxyl groups of Asp33 and Asp213 receive two additional hydrogen bonds from surrounding hydrogen-bond donating groups of penicillopepsin (Figure 2). The hydrogen-bonding distances

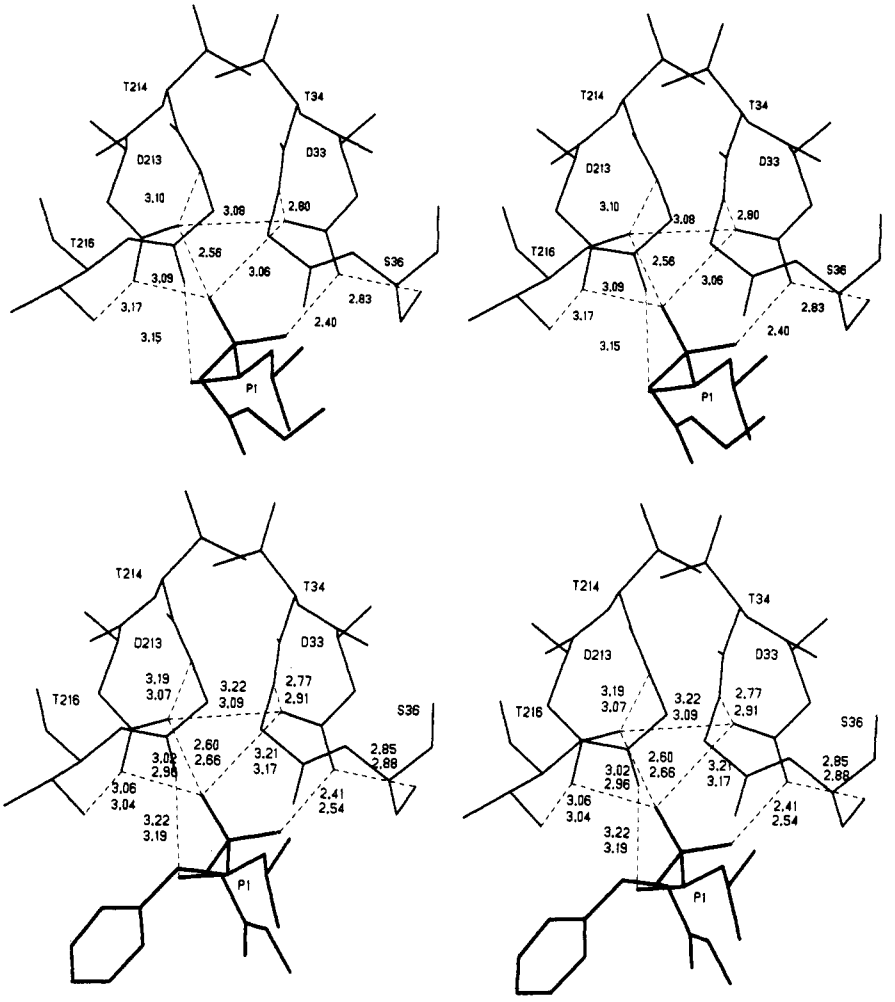


FIGURE 2: Stereoview of the hydrogen bonding and the distances between oxygen atoms of the active site catalytic residues and bound inhibitor. Bonds connecting the protein atoms are shown as thin solid lines, inhibitor bonds as thick solid lines. Hydrogen bonds are shown as dashed lines with the appropriate distance adjacent. (a, top) IvaVVSta^{P}OEt complex. (b, bottom) IvaVVL^{P}(O)FOMe complex; in each instance the second value (given below) is that of the corresponding distance in the CbzAALP(O)FOMe complex.

from these donor groups to the carboxyl oxygen atoms are consistent with a neutral Asp213 and a partial negative charge on the carboxylate of Asp33. The hydrogen-bond distances involving Asp33 are ~0.3 Å shorter (and therefore presumably stronger) than the equivalent distances involving Asp213 (Table III, Figure 2). For example, both the distances Ser36 O^{γ}···Asp33 $O^{\delta 1}$ and Gly35 NH···Asp33 $O^{\delta 2}$ are shorter by 0.3 Å than the distances Thr216 $O^{\gamma 1}$···Asp213 $O^{\delta 2}$ and Gly215 NH···Asp213 $O^{\delta 1}$. In Figure 2a,b it can be seen that the peptide dipole of the Thr34 to Gly35 peptide bond is directed exclusively at Asp33 $O^{\delta 2}$. On the other hand, the corresponding peptide dipole in the Thr214 to Gly215 peptide bond points almost equally toward both carboxyl groups so that the Gly215 NH approximates a bifurcated hydrogen bond. Clearly, the partially negatively charged character of the Asp33 carboxylate in the presence of the phosphonate inhibitors has disrupted the approximate 2-fold symmetry of the native penicillopepsin active site (James & Sielecki, 1983).

Short Hydrogen Bonds in Other Systems. Short hydrogen bonds between phosphonyl oxygen atoms of bound inhibitors and oxygen atoms of carboxyl groups have been reported for thermolysin (TLN) (Tronrud et al., 1987; Holden et al., 1987) and carboxypeptidase A (CPA) (Kim & Lipscomb, 1990, 1991). These short bonds range in length from 2.2 to 2.5 Å with an average of 2.4 Å. This distance is similar to the average value of 2.44 Å for the three complexes reported here (Table III) and all of these distances are ~0.3–0.4 Å shorter than normal O–H···O hydrogen bonds observed in a variety of other compounds (Donohue, 1968; Hamilton & Ibers, 1968). Holden et al. (1987) have interpreted the charge distribution on the phosphonyl group bound in the active site of TLN so that the negative charge is located on the *pro-S* oxygen atom that is closer (2.2 Å) to the zinc ion. In their interpretation this oxygen would be equivalent to the carbonyl oxygen of a peptide substrate and would develop a net negative charge as the general-base-assisted nucleophilic attack on the carbonyl carbon atom takes place. Electrostatic stabilization of this developing charge would come both from the Zn^{2+} ion and from the hydrogen bond made by the protonated $N^{\epsilon 2}$ of His231 in TLN. The hydrogen-bonding distances observed for the phosphonate inhibitors suggests an alternative scheme shown in Figure 3d,e, however. In this scheme, the very short O···H···O distance of ~2.4 Å between the Glu143 carboxyl and the *pro-R* phosphonyl oxygen atom would correspond to a symmetrical single-well potential energy function as we have proposed for the phosphonate inhibitors of the aspartic proteinases. The 2.7–2.9 Å hydrogen-bond distance from the other phosphonyl oxygen to His231 supports this as it is not considered to be a short distance.

The hydrogen-bonding distances in the CPA–phosphonate complexes are equivalent to those of the TLN–phosphonate complexes, so the above discussion should apply equally to those structures (Kim & Lipscomb, 1990, 1991). On the other hand, no short hydrogen-bonding distances were observed in

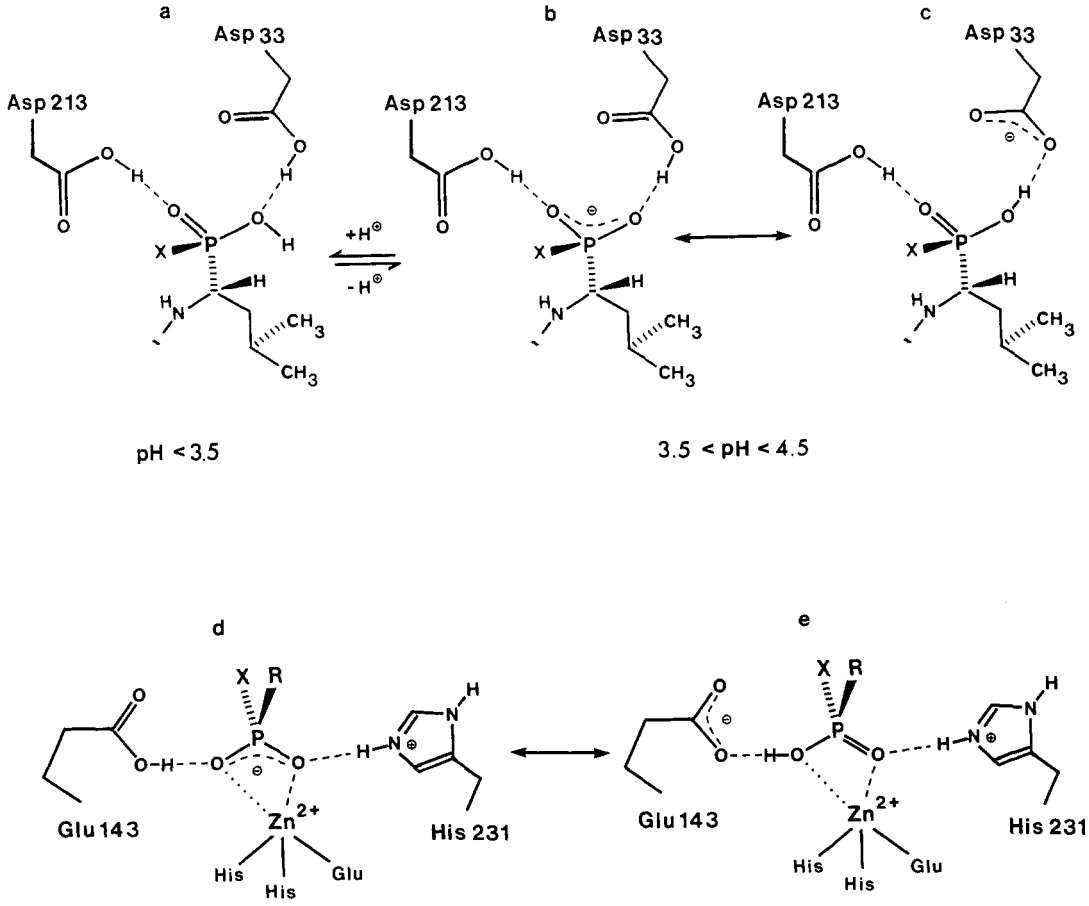


FIGURE 3: Proposed proton positions in the active sites of penicillopepsin (a–c) and thermolysin (d and e) in the presence of phosphorus-based peptide analogue inhibitors. The proton distribution is based upon the hydrogen-bonding distances and nonbonded contact distances shown in Figure 2 for penicillopepsin and reported previously for TLN (Holden et al., 1987) and CPA (Kim & Lipscomb, 1991). Panels b and c show the tautomeric forms for the short (2.4 Å) hydrogen bond between Asp33 and the phosphonate *pro-S* oxygen. It is not possible to distinguish this proton distribution from one that would have a central symmetric position for the proton with a single-well potential energy profile (Speakman, 1972). Panel a shows the proton distribution at lower pH values in which the active site is uncharged. The pH dependence of the K_i suggests that the pK_{a1} is near 3.5 (Bartlett et al., 1990). Panels d and e show the tautomeric forms expected for phosphonate inhibitors bound to the active sites of the metalloenzymes TLN and CPA. Very short hydrogen bonds are reported between the active site general base glutamate residues and the *pro-R* oxygen of the phosphonate inhibitors. As with the aspartic proteinases and the phosphonate inhibitors, it is not possible to distinguish between the situation shown in panels d and e and a symmetric single-well potential for the proton.

the structure of the irreversible phosphonate complex with α-lytic protease (Bone et al., 1991). The hydrogen bonds from Gly193 NH and from Ser195 NH to the *pro-S* phosphinyl oxygen atom that occupies the oxyanion-binding site are of normal expected lengths, 2.6 and 2.8 Å, respectively.

pH Dependence of Inhibition. The hydrogen-bonding scheme discussed above for the aspartic proteinases, and shown diagrammatically in Figure 3b,c, is fully consistent with there being a net negative charge at the active site of penicillopepsin at pH 4.4 (the pH of the crystallization mother liquor). This charge is probably shared equally by the phosphonyl and phosphinyl groups of the inhibitors and the carboxylate of Asp33 on the enzyme. The pK_a values of the phosphinate and the phosphonate inhibitors are 3.0 and 1.5, respectively, so the inhibitors would be negatively charged in solutions of pH 4.0 and above (Grobelny et al., 1989). The pH dependence of the K_i values for IvaVVSta^{P}OEt and CbzAAL^P(O)FOMe upon binding to penicillopepsin was shown to be consistent with protonation of the active site upon binding (Bartlett et al., 1990). That this is the case is certainly supported by the structural results presented here. Indeed, the behavior of K_i upon dropping the pH of the buffer to 3.5 suggests that a neutral active site (Figure 3a) resulting from protonation of the *pro-S* oxygen atom, the most solvent-accessible oxygen atom in the active site region, leads to ~10-fold tighter binding of these inhibitors [K_i = 0.19 μM for CbzAAL^P(O)FOMe (Bartlett et al., 1990)]. On the other hand, increasing the buffer pH to 5.5 results in a marked decrease in inhibitor binding affinity [the K_i is increased 40-fold from its value of 2.7 μM at pH 4.5, (Bartlett et al., 1990)]. This is consistent with the removal of the proton shared by the *pro-S* phosphonyl oxygen atom and the Asp33 carboxyl group. The affinity of the inhibitor would clearly be reduced if two negative charges must be buried in close proximity. Since the two pK_a's of the complex are not known directly from titration, they can only be inferred from our structural results done at pH 4.4. The pK_{a1} must be less than 4.4, as we assume that the very short hydrogen bond, 2.4 Å, is a result of the equal sharing of the single proton by two negatively charged groups (Figure 3). Similar situations are common in proteins where the carboxyl groups of aspartic or glutamic acid residues are in close proximity (Sawyer & James, 1982). The lower pH value of 3.5 leading to increased inhibitor affinity for the penicillopepsin active site suggests that the first pK_a of the system is close to this value, ~3.5. The pH at which the second proton (shared by Asp33 and the phosphonate in Figure 3) would be removed is likely to be ~1 pH unit above the solution pH in the present study. Thus, we would expect the second pK_a to be in the region of pH 5.5. The fact that CbzAAL^P(O)FOMe is only a weak inhibitor (K_i = 107 μM) at pH 5.5 (Bartlett et al.,

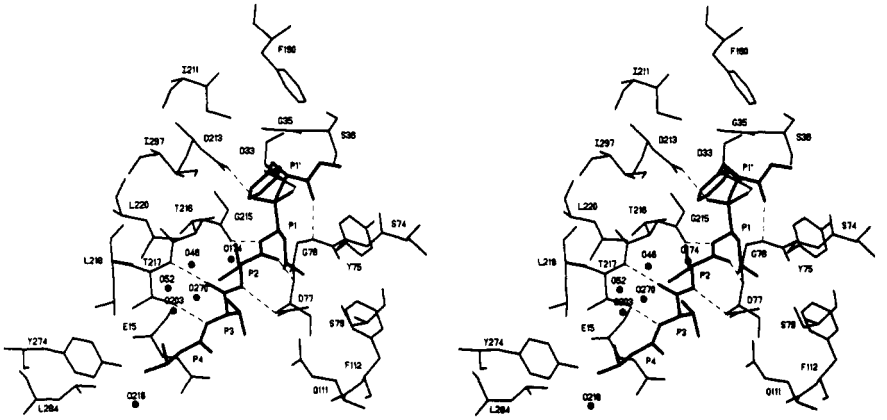


FIGURE 4: Global stereoview of IvaVVL^P(O)FOMe bound to penicillopepsin. Bonds in the inhibitor are shown as thick solid lines with inhibitor residues P_1', P_1, P_2, P_3, and P_4 labeled appropriately. Water molecules are depicted as filled circles. Hydrogen bonds are shown as dashed lines.

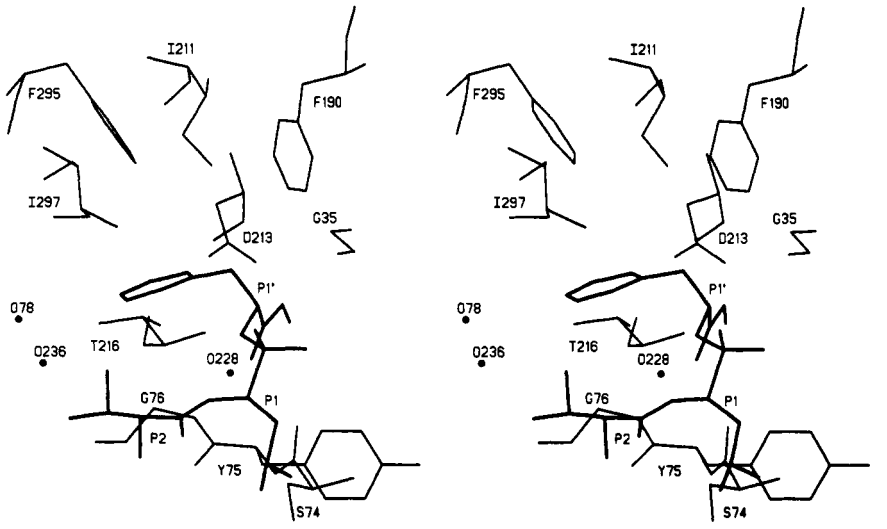


FIGURE 5: Close-up view of the S_1'-binding site in penicillopepsin. The P_1' phenylalanine side chain of the phosphonate inhibitor (depicted by thick solid lines) makes hydrophobic contacts with Gly76, Phe190, Ile211, Phe295, and Ile297. The side chain of Ile297 undergoes a 60° conformational change (χ^2 changes from ~140° in the native enzyme to ~80° in the complex).

1990) suggests that the endpoint for the titration of the second proton is in this vicinity.

Backbone Hydrogen Bonding. As with the other inhibitors of penicillopepsin that have been studied crystallographically, the phosphorus-based peptide analogues studied here are highly complementary to the active site of the enzyme (Figure 4). There are extensive regions of nonpolar contact between the inhibitors' P_4, P_2, P_1, and P_1' residues and the binding subsites S_4, S_2, S_1, and S_1' on penicillopepsin. In addition to the hydrophobic interactions, there are polar interactions that involve main-chain backbone hydrogen bonding between inhibitor and enzyme (Table III). These hydrogen bonds have been detailed in most of the other papers that describe the fungal aspartic proteinase/inhibitor interactions (Davies, 1990; Suguna et al., 1987a,b; James et al., 1992; Cooper et al., 1987, 1989; Foundling et al., 1987; Sali et al., 1989).

S_1'–P_1' Interactions in Penicillopepsin. IvaVVL^P(O)FOMe and CbzAAL^P(O)FOMe are the only inhibitors bound to penicillopepsin and studied crystallographically that have a residue extending toward the carboxyl end of the scissile bond mimic [i.e., (O)PheOMe]. Thus, it is now possible to evaluate the molecular modeling that was done for the proposals for the catalytic pathway in which a P_1' Tyr residue was considered (James & Sielecki, 1985). The conformational angles that were proposed for the P_1' Tyr residue were approximately ϕ, −120°; ψ, +150°; χ_1, +90°, χ_2, +90°. The equivalent values observed here for the P_1' (O)Phe residue ar_ _19°, +173°, +58°, and +104° (Table IV), indicating that the molecular modeling described a possible substrate-binding mode to penicillopepsin reasonably well. The 30° discrepancy in the value of ψ is most likely the result of our proposal that the P_2' NH forms a hydrogen-bonded interaction with the C=O of Gly35 [Figures 8, 11, and 12 of James and Sielecki (1985)]. Such an interaction is not possible for (O)PheOMe because the inhibitor atom equivalent to the P_2' NH atom is the ester oxygen of the methyl ester; it lies approximately 3.7 Å from the C=O of Gly35.

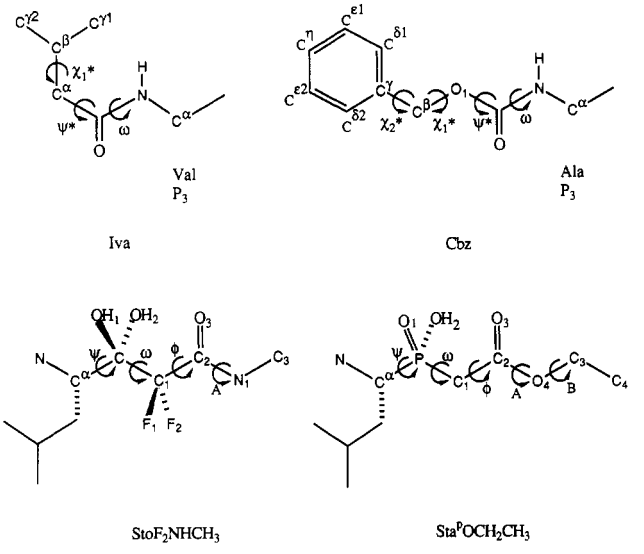
The residues that line the S_1'-binding site in penicillopepsin are predominantly hydrophobic in character (Figure 5). Residues that provide atoms to form the S_1'-binding region in penicillopepsin are Gly35, Gly76, Phe190, Ile211, Asp213, Ile293, Phe295, and Ile297. The P_1' Phe in the two phosphonate inhibitors makes contacts that are ≤4.2 Å with most of these residues. Only minor conformational changes in the positions of the side chains of Ile293 and of Ile297 from their positions in native penicillopepsin result upon inhibitor binding to S_1'.

Comparisons of Inhibitor Binding to Penicillopepsin. In order to make comparisons of the binding modes of these several inhibitors in the active site of penicillopepsin, the backbone NC$^\alpha$CO atoms of all the residues in penicillopepsin were fitted to each other pairwise by a least-squares procedure

Table IV: Conformational Angles (deg) of Phosphorous-Based Peptide Analogue Inhibitors Bound to Penicillopepsin

conformational angles[a]	···Sta^{P}OEt	IvaVV···	CbzAA···	StoF$_2$[a]
P_4				
ψ^*	−148	−140	−157	−160
ω	−176	−174	177	−177
χ_1^*	66	59	−115	61
χ_2^*			139	
P_3				
ϕ	−116	−123	−80	−110
ψ	171	162	164	166
ω	176	179	178	−179
χ_1	−62	−68		−57
P_2				
ϕ	−134	−132	−135	−138
ψ	91	99	86	101
ω	−179	−180	−178	−178
χ_1	−174	−176		−175
P_1				
ϕ	−113	−123	−115	−124
ψ^*	64	73	71	71
ω	83	147	149	76
χ_1	−54	−48	−47	−53
χ_2	177	159	175	176
P_1'				
ϕ	96	−113	−119	97
A	−163	−178	−158	−178
ψ		174	−173	
χ_1		63	58	
χ_2		96	104	
B	134			

[a] Definition of conformational angles. Iva, $\omega = C^\alpha$ (Iva)–C–N–C^α (Val); $\psi^* = C^\beta$–C^α–C–N; $\chi_1^* = C^{\gamma 1}$–C^β–C^α–C. Cbz, ω = O1 (Cbz)–C–N–C^α (Ala); $\psi^* = C^\beta$–O1–C–N; $\chi_1^* = C^\gamma$–C^β–O1–C; $\chi_2^* = C^{\delta 1}$–C^γ–C^β–O1. The atom numbering is shown in the diagram below:



(Table V). The resulting orientation matrix was then used to place one of the inhibitors being compared into the reference frame of the other inhibitor, and the coordinate differences were taken. The resulting rms differences are given in the lower triangular portion of the comparison matrix in Table V. In the comparisons of the several complexed penicillopepsin molecules with the native uncomplexed form, the nine residues that constitute the flap (Ile73–Ala81) were left out of the least-squares fitting. For the comparisons involving only the complexed forms of the enzyme, all 1293 main-chain atoms were used.

The rms differences between the several complexed forms of penicillopepsin with the unbound native structure range from 0.20 to 0.26 Å. The largest conformational differences are due to differences in position of the flap residues, by up to 4.1 Å at the C^α of Gly76 (James et al., 1982). This changed

Table V: Comparison of the Various Molecular Complexes of Inhibitors with Penicillopepsin, Main-Chain Atoms[a] rms (Å)

	PP1[b]	PP2[b]	PP3[b]	StoF$_2$[b]	StaF$_2$[b]
native	0.255	0.198	0.226	0.259	0.256
PP1		0.201	0.147	0.110	0.086
PP2	0.150 (30)		0.158	0.167	0.204
PP3	0.289 (23)	0.290 (32)		0.126	0.155
StoF$_2$	0.259 (30)	0.217 (30)	0.322 (23)		0.094
StaF$_2$	0.216 (29)	0.160 (29)	0.317 (22)	0.108 (35)	

[a] The upper right triangular portion of this matrix has the rms values that result from least-squares fitting of the 1293 main-chain atoms of penicillopepsin from each complex in a pairwise fashion. For those comparisons that involve the native structure, the nine residues (36 atoms) that comprise the flap (Ile73–Ala81) have been omitted from the calculation. The atoms of the flap undergo a very large conformational change (up to 4.1 Å at C^α of Gly76) when the complexes are formed (James et al., 1982). The lower triangular portion of the matrix gives the rms differences that result from the best fit of the common atoms of the inhibitors when compared in a pairwise fashion among themselves. The numbers in parentheses are the number of atoms in the least-squares fitting. [b] PP1 = IvaVVSta^{P}OEt; PP2 = IvaVVL^P(O)FOMe; PP3 = CbzAAL^P(O)FOMe; StoF$_2$ = IvaVVStoF$_2$NHMe; StaF$_2$ = IvaVVStaF$_2$NHMe.

conformation is due to the favorable intermolecular interactions that form from residues on the flap, Tyr75, Gly76, and Asp77 in particular, to the atoms of the inhibitor. The flap–inhibitor interactions account for approximately 20% of the total solvent-accessible surface area of the inhibitors that is buried when the molecular complexes form. There are other regions of the penicillopepsin molecule that exhibit consistent but much smaller conformational differences when the inhibitors bind. These regions border the entrance to the substrate binding cleft and constitute the following sets of contiguous residues: Gln50–His54 (this segment of chain lies adjacent to a short helical segment, Ser109–Gln114, but does not make direct contacts with the inhibitors), Ile108–Gln114 (helical region), Gly235–Gly244 (this region packs against the disordered segment of Pro276–Thr282), Pro276–Thr282 [disordered or poorly ordered in all aspartic proteinase structures (James & Sielecki, 1983, 1987)], and Ser291–Phe295 (a small flap-like region that constitutes part of the S_1'-binding pocket).

Comparison of the penicillopepsin molecule in the IvaVVL^P(O)FOMe complex with that in the CbzAAL^P(O)FOMe complex shows that the two differ only in minor respects (rms difference ~0.16 Å for 1293 main-chain atoms). However, the comparison of the 32 common atoms of the two inhibitors when their positions are overlapped as described above reveals a number that is almost twice as large as this (0.29 Å, Table V). The inter- and intramolecular contacts that are most likely behind this larger difference can be appreciated in Figure 6. From the NH of P_3 Val (Ala) to the main-chain C^α of P_1' (O)Phe, the fit is extremely good, 0.17 Å rms, certainly within the coordinate accuracy. The largest departures are found at the N-terminus, where the chemical differences between the carbobenzoxyl and the isovaleryl groups predominate (Figure 6), and at the P_1' phenyl ring. The differences in position of the P_1' phenyl ring are due to the difference in the P_2 side chains of the two inhibitors. In IvaVVL^P(O)FOMe, the larger volume of the valyl residue in comparison to alanine in CbzAAL^P(O)FOMe causes the P_1' phenyl ring to be shifted by an rms distance of 0.64 Å. Not only is the P_1' phenyl ring shifted but also the atoms of the methyl ester are shifted (Figure 6). These altered conformations are primaily the result of an ~5° opening of the bond angle O–C^α–C^β (of the L-β-phenyllactate) in the IvaVVL^P-(O)FOMe inhibitor relative to the equivalent angle in CbzAAL^P(O)FOMe. The atoms of the methyl ester are

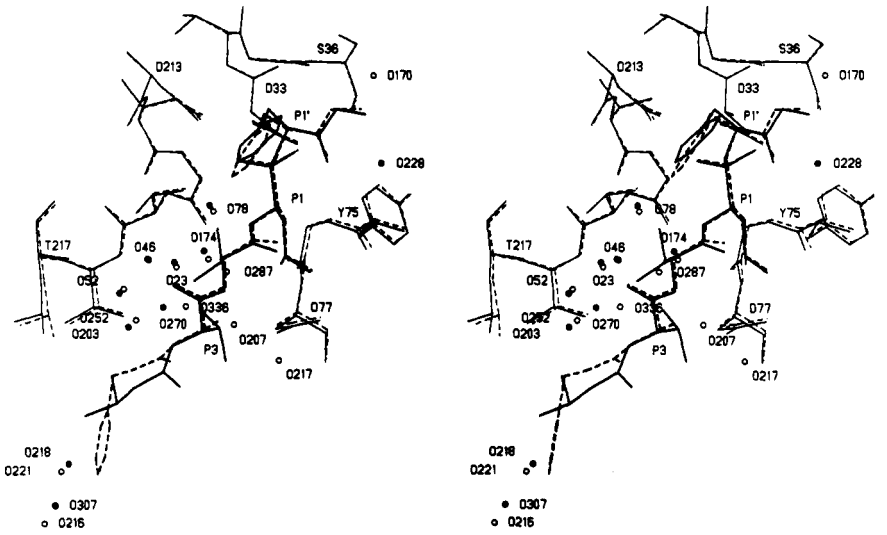


FIGURE 6: Stereoview showing the superposition of the two phosphonate inhibitor–penicillopepsin complexes (Table V). IvaVVL^P(O)FOMe and residues in the active site are represented by solid lines; the solvent molecules of this complex are represented by filled circles. The equivalent penicillopepsin residues and CbzAAL^P(O)FOMe in that complex are represented by dashed lines, the solvent molecules by open circles. Major conformational differences are seen at the Iva and Cbz blocking groups on the N-termini and in the P_1' Phe residues. Solvent molecules O207, O217, and O287 are present in the CbzAAL^P(O)FOMe complex but not in the IvaVVL^P(O)FOMe due to the differences in the P_3 and P_2 side chains. Other solvents in the vicinity of the P_4 and P_2 side chains are seen to shift positions in the presence of the other inhibitor.

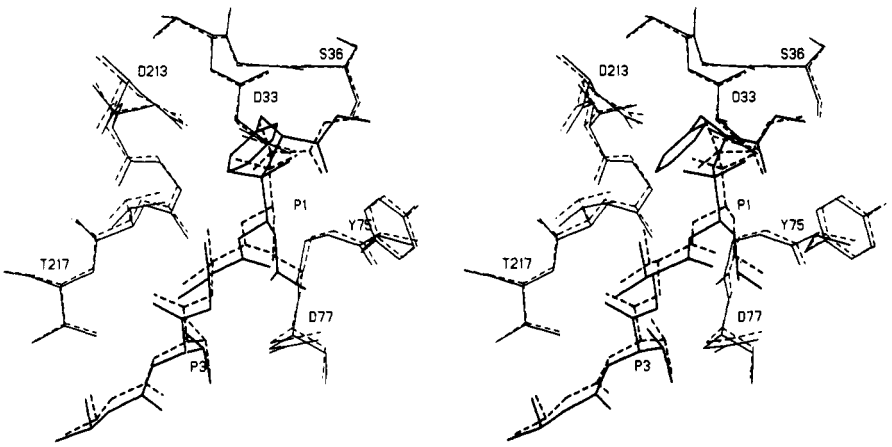


FIGURE 7: Stereoview of the superposition of the residues in the penicillopepsin active site in the complexes wth IvaVVL^P(O)FOMe (solid lines) and IvaVVStoF$_2$NHMe (StoF$_2$, difluorostatone) (dashed lines). The main-chain atoms of the two penicillopepsin molecules were superimposed as outlined in text. The inhibitors are shifted globally relative to the active site carboxyl groups by ~0.6 Å, but the hydrogen-bonding interactions are largely preserved in both complexes.

shifted as a result of an ~5° reduction in the bond angle O–C$^\alpha$–C in the IvaVVL^P(O)FOMe inhibitor relative to the CbzAAL^P(O)FOMe inhibitor. Thus, the bond angle C$^\beta$–C$^\alpha$–C remains unchanged in both inhibitors.

Figure 6 also reveals other differences in the binding of these two inhibitors to penicillopepsin. Three well-ordered solvent molecules (B factors <30 Å^2) are bound in the vicinity of the P_2 Ala and P_3 Ala residues in the CbzAAL^P(O)FOMe inhibitor complex. These solvent molecules are not present in the IvaVVL^P(O)FOMe complex crystals. Three other solvent molecules in the proximity of the P_2 Val side chain are shifted further from the inhibitor atoms than they are in the complex with CbzAAL^P(O)FOMe.

These molecular differences, relatively distant from the phosphonate groups, must be directly related to the 570-fold difference in K_i ($\Delta\Delta G$ = −3.7 kcal/mol) for the two inhibitors (Bartlett et al., 1990). On the one hand, the larger hydrophobic volume of the P_4 Cbz group over that of the P_4 Iva group could be expected to enhance the affinity for the S4 site. This potentially favorable enthalpic component of the binding free energy seems to be more than offset by the slight difference in the positions of the P_1' phenyl rings and by the displacement of the additional three water molecules by the IvaVVL^P(O)FOMe inhibitor. It is likely that the latter entropic contribution to the inhibitor affinity is the larger component contributing to the differences in binding free energy.

Comparisons with Difluorostatine and Difluorostatone Inhibitors. Since the difluorostatone-containing peptide inhibitor, IvaVVStoF$_2$NHMe, contains a hydrated ketone as a mimic of the tetrahedral transition-state intermediate (James et al., 1992), it is of interest to compare that structure bound to penicillopepsin to the bound conformation of IvaVVL^P-(O)FOMe (Figure 7). The latter inhibitor is shifted by ~0.6 Å toward the N-terminus of the inhibitor-binding site of penicillopepsin, a difference that is reflected in the conformations of the two aspartic acid residues in the two complexes. Hydrogen-bonding interactions that involve the oxygen atoms on the tetrahedral mimics are most likely the most important determinants of these differences. Since the IvaVV portions of the two inhibitors are identical, one would not expect differences to arise from these regions of molecules. The hydrated ketone forms a hydrogen bond (~2.65 Å) from each of the tetrahedral oxygen atoms to each of the carboxyl oxygen atoms of Asp33. This is different from the single very short hydrogen

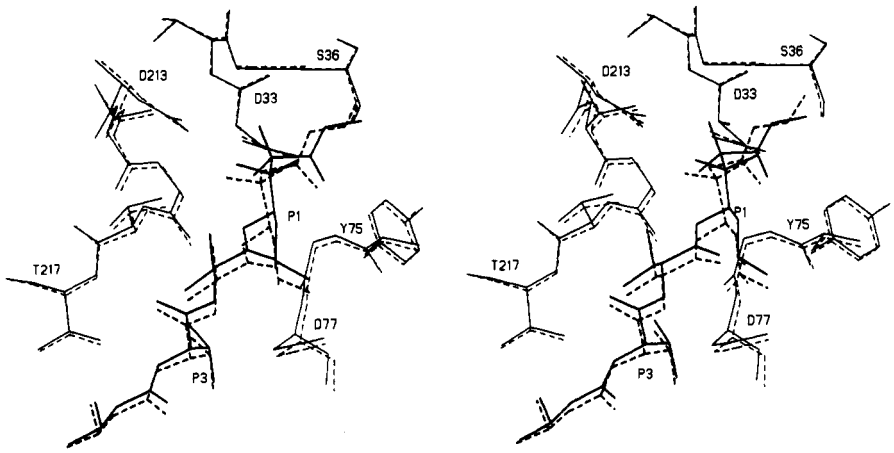


FIGURE 8: Stereoview of the phosphastatine inhibitor IvaVVSta^{P}OEt (dashed lines) superimposed on the difluorostatone inhibitor IvaVVStoF_2NHMe (solid lines) bound in the active site of penicillopepsin. The shift in position of the two inhibitors is ~0.6 Å with the phosphorus-based inhibitor positioned further towards the S_n binding sites. The superposition of IvaVVSta^{P}OEt and IvaVVStaF_2NHMe is very similar to that depicted here.

bond (2.4 Å) from the *pro-S* phosphonate oxygen to Asp33 $O^{\delta 1}$. We propose that Asp213 is protonated in both of these inhibitor complexes, but the refined distances and the stereochemistry of the two complexes suggest that it is the alternate carboxyl oxygen that is protonated in each complex. That is to say, Asp213 $O^{\delta 1}$ is the proton donor to the *pro-R* phosphonyl oxygen atom in the phosphonate inhibitor (Figures 2 and 3), whereas Asp213 $O^{\delta 2}$ is the proton donor to the hydrated ketone oxygen atom in the StoF_2 containing complex [Figure 8 of James et al. (1992)].

In spite of the 0.6-Å shift in global position of these two inhibitors relative to one another in the binding cleft of penicillopepsin, each inhibitor is able to form the identical hydrogen-bonding interactions to side-chain and main-chain atoms that line the binding cleft in penicillopepsin. In addition, if one compares the bound conformations of the two inhibitors, it is clear that both have very similar overall conformations (Table IV). The rms difference in position for the 30 atoms deemed to be common between the two is 0.217 Å (Table V). Most of the difference lies in the two termini where the Iva group has a difference of 20° in ψ^* (mainly affecting $C^{\gamma 1}$ and $C^{\gamma 2}$ and the carbonyl oxygen). The phosphonate and the hydrated ketone are remarkably similar in conformation up to the ω angle which is ~150° in IvaVVL^P(O)FOMe compared to the 76° in the StoF_2-containing inhibitor (James et al., 1992). Thereafter, the chemical differences dictate the major departure in conformation. In spite of these differences, the amide carbonyl oxygen atom of StoF_2NHCH_3 and the ester carbonyl oxygen atom of FOMe both form good hydrogen-bonding interactions with the NH of Gly76 on the flap (Figure 7). It would be of interest to see what would be the effect on the affinity of incorporating a C-terminal amide into the L^P(O)FOMe inhibitor.

The phosphastatine-containing inhibitor, IvaVVSta^{P}OEt, forms very strong hydrogen-bonding interactions with the two carboxyl groups of the penicillopepsin active site (Figures 2 and 3). These interactions are very similar to those described for the phosphonate-containing inhibitors. Thus, the phosphastatine moiety, by virtue of the tetrahedral phosphinyl group, has essentially the same intermolecular interactions with penicillopepsin. This is not the case when one compares the StaF_2NHMe inhibitor with the Sta^{P}OEt analogue. The former has only the (*S*)-hydroxyl group on the tetrahedral carbon atom, which allows for a much closer approach to Asp33 than occurs with Sta^{P}OEt. In fact, both of the difluorinated Sta- and Sto-containing inhibitors approach the carboxyl group of Asp33 more closely than does the Sta^{P}OEt containing inhibitor (Figure 8). So, in spite of the similar tetrahedral character of the three statine analogues, the Sta^{P}OEt binds to the active site more in a manner that resembles the other two phosphonate inhibitors. This must be due to the chemical similarity of the three phosphorus-containing inhibitors, in spite of the fact that the phosphastatine has a methylene (–CH_2–) group bridging the phosphorus atom and the ethyl ester group. There is a global difference between the binding position of IvaVVSta^{P}OEt and that of IvaVVStaF_2NHMe (~0.5–0.6 Å shift, Figure 8). That the shift is global is corroborated by the very close conformational similarity of these two inhibitors (rms difference is 0.16 Å, Table V).

Table VI: Solvent-Accessible Surface Area (Å^2)[a]

	uncomplexed			complexed				
inhibitor	h/phil	h/pho	total	h/phil	h/pho	total	Δ^b	(%)
IvaVVSta^P-OEt	192	606	798	26	138	163	635	80
IvaVVL^P-(O)FOMe	192	673	866	26	124	150	716	83
CbzAAL^P-(O)FOMe	207	673	881	22	152	174	707	80
IvaVVSta-F_2NHMe	212	574	785	26	129	155	630	81
IvaVVSto-F_2NHMe	230	556	786	26	124	150	636	80

[a] Calculated according to the algorithm of Richards (1977) using a program developed by Connolly (1983a,b). h/phil and h/pho stand for the hydrophilic and hydrophobic surface areas, respectively. [b] Δ is the difference in the total solvent-accessible surface areas calculated for the uncomplexed and complexed forms of the inhibitors.

Solvent-Accessible Surfaces. The burial of hydrophobic surfaces is a major determinant in estimating the strength of affinity between inhibitors and enzymes or between interacting subunits of multisubunit proteins (Chothia, 1974, 1976; Chothia & Janin, 1975). The aspartic proteinases are characterized by a strong preference for hydrophobic amino acids in oligopeptide substrates (Powers et al., 1977). As shown in Table VI, approximately 80% of the total solvent-accessible surface area of the several inhibitors that have been bound to penicillopepsin is buried when the complex is formed. Approximately 20% of the total buried surface area is due to the interaction of residues of the flap with the various inhibitors (Figure 9). The remaining 150 Å^2 of accessible surface area on IvaVVL^P(O)FOMe is predominantly hydrophobic in character and comprises the hydrocarbon components of P_4

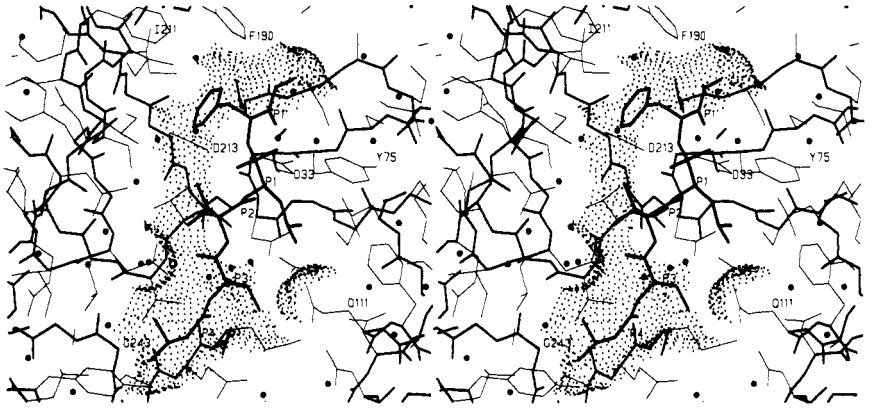


FIGURE 9: Stereoview of the van der Waals surface remaining accessible to solvent for the inhibitor IvaVVL^P(O)FOMe when bound to penicillopepsin. The dotted surface represents that portion of the total surface area of the inhibitor that remains accessible to solvent in the presence of penicillopepsin (Table VI). The inhibitor is shown as thick solid lines, water molecules as filled circles, the side chains of penicillopepsin are the thinnest lines present.

Iva, P_2 Val, P_1' Phe, and the methyl group of the C-terminal methyl ester (Figure 9). Seven water molecules are in contact with the remaining solvent accessible surface but only one of them is within hydrogen-bonding distance of a hydrophilic atom on the inhibitor (Wat O52, Figure 4).

Transition-State Mimicry. Are the phosphonate inhibitors reliable transition-state mimics of the tetrahedral intermediate of good peptide substrates of the aspartic proteinases? We approach this question from a structural viewpoint. Although statine and difluorostatine have tetrahedral carbon atoms that mimic the tetrahedral nature of the transition-state intermediate, there is only one oxygen in these chemical groups that can interact with the active site aspartates. The microorganisms that synthesize the statine amino acid analogue have made best use of their chemical mimicry by choosing the (3*S*)-isomer that would interact optimally with both of the aspartate carboxyl groups. Rich et al. (1985) have synthesized inhibitors containing the (*R*)-statine isomer and shown that they have an order of magnitude weaker binding affinity to pepsin than the (3*S*)-isomers, as would be expected, since the (3*R*)-isomer can interact only with Asp33.

The hydrated difluorostatone-containing inhibitor is an excellent mimic of the tetrahedral transition-state intermediate (James et al., 1992). The *gem*-diol interacts with both aspartate carboxyl groups in the manner expected for the tetrahedral intermediate (James & Sielecki, 1985, 1987). However, the difluorostatone analogue has two additional main-chain atoms, $-CF_2-(C{=}O)-$, that would not be present in the tetrahedral intermediate of a normal substrate. In addition, the conformation of this group differs very much from that of an oligopeptide inhibitor or substrate with a P_1' residue whose side chain would fit into the S_1'-binding site [cf Leu^P(O)FOMe ω of 148° vs the StoF$_2$NHMe ω angle of 76°]. Thus, the ω angle of the phosphonates resembles more closely the antiperiplanar conformation of ~180° expected for a normal substrate. From this point of view, the phosphonates are more faithful mimics of the transition-state intermediate.

Since the structure of a transition-state complex cannot be determined by crystallographic techiques, it is risky to gauge the fidelity of a transition-state analogue on this basis. Transition-state similarity can be demonstrated in favorable cases by correlation of the K_i values for a series of related inhibitors with the K_m/k_{cat} values of the corresponding substrates (Bartlett & Marlowe, 1983; Hanson et al., 1989); however, such an analysis has not yet been reported for an aspartic proteinase. In the absence of such kinetic evidence, it is hazardous to designate a particular structure as a model of the transition state; the comparison between the phosphorus analogues and the difluorostatone derivative is a case in point. Both types of inhibitors position a mimic of the tetrahedral intermediate in the vicinity of the catalytic aspartate carboxyl groups, yet they are shifted relative to each other by ca. 0.6 Å. Does this difference arise from a repulsion between the negatively charged phosphorus oxyanions and the cataytic aspartates, thus pushing these compounds toward the amino terminus of the binding cleft, or does it arise from hydrogen bonding between the C-terminal CONHMe moiety of the StoF$_2$ inhibitor and the carbonyl of Gly35 which, because of the shorter chain length of this inhibitor, pulls it toward the other end of the cleft?

What is notable is that the difference in orientation of these inhibitors in the active site is so great, in view of their high affinity and their overall structural similarity. This observation contrasts with the generally held view that active site interactions are very sensitive to orientation and that small perturbations in structure and conformation can have a large impact on binding affinity.

ACKNOWLEDGMENTS

We thank Professor T. Hofmann for the generous gifts of penicillopepsin without which this work would have not been possible.

REFERENCES

Allen, F. H., Kennard, O., & Taylor, R. (1983) *Acc. Chem. Res. 16*, 146–153.

Allen, M. C., Fuhrer, W., Tuck, B., Wade, R., & Wood, J. M. (1989) *J. Med. Chem. 32*, 1652–1661.

Bartlett, P. A., & Marlowe, C. K. (1983) *Biochemistry 22*, 4618–4624.

Bartlett, P. A., & Kezer, W. B. (1984) *J. Am. Chem. Soc. 106*, 4282–4283.

Bartlett, P. A., & Marlowe, C. K. (1987) *Biochemistry 26*, 8553–8561.

Bartlett, P. A., Hanson, J. E., & Giannousis, P. P. (1990) *J. Org. Chem. 85*, 6268–6274.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasmui, M. (1977) *J. Mol. Biol. 112*, 535–542.

Boger, J. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 401–420, Walter de Gruyter, New York.

Bone, R., Sampson, N. S., Bartlett, P. A., & Agard, D. A. (1991) *Biochemistry 30*, 2263–2272.

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

Brünger, A. T. (1988) XPLOR *Manual*, Version 1.5. Yale University, New Haven, CT.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

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

Chothia, C. (1974) *Nature 248*, 338–339.

Chothia, C. (1976) *J. Mol. Biol. 105*, 1–14.

Chothia, C., & Janin, J. (1975) *Nature 256*, 705–708.

Connolly, M. L. (1983a) *J. Appl. Crystallogr. 16*, 548–558.

Connolly, M. L. (1983b) *Science 221*, 709–713.

Cooper, J. B., Foundling, S. I., Hemmings, A. M., Blundell, T. L., Jones, D. M., Hallett, A., & Szelke, M. (1987) *Eur. J. Biochem. 169*, 215–221.

Cooper, J. B., Foundling, S. I., Blundell, T. L., Boger, J., Jupp, R. A., & Kay, J. (1989) *Biochemistry 28*, 8596–8603.

Cruickshank, D. W. J. (1949) *Acta Crystallogr. 2*, 65–82.

Cruickshank, D. W. J. (1967) in *International Tables for X-ray Crystallography* (Kasper, J. S., & Lonsdale, K., Eds.) Vol. II, pp 318–340, Kynoch Press, Birmingham, England.

Darlow, S. F., & Cochran, W. (1961) *Acta Crystallogr. 14*, 1250–1257.

Davies, D. R. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 189–215.

Donohue, J. (1968) in *Structural Chemistry and Molecular Biology* (Rich, A., & Davidson, N., Eds.) pp 443–465, W. H. Freeman & Co., San Francisco, CA.

Dreyer, G. B., Metcalf, B. W., Tomaszek, T. A. J., Carr, T. J., Chandler, A. C. I., Hyland, L., Fakhoury, S. A., Magaard, V. W., Moore, M. L., Strickler, J. E. et al. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9752–9756.

Ellison, R. D., & Levy, M. A. (1965) *Acta Crystallogr. 19*, 260–268.

Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., Sibanda, B. L., Hemmings, A., Wood, S. P., Blundell, T. L., Valler, M. J., et al. (1987) *Nature 327*, 349–352.

Fruton, J. S. (1976) *Adv. Enzymol. Relat. Areas Mol. Biol. 44*, 1–36.

Fujinaga, M., Delbaere, L. T. J., Brayer, G. D., & James, M. N. G. (1985) *J. Mol. Biol. 183*, 479–502.

Glover, I., Haneef, I., Pitts, J., Wood, S., Moss, D., Tickle, I., & Blundell, T. L. (1983) *Biopolymers 22*, 293–302.

Grobelny, D., Goli, U. B., & Galardy, R. E. (1989) *Biochemistry 28*, 4948–4951.

Grobelny, D., Wondrak, E. M., Galardy, R. E., & Oroszlan, S. (1990) *Biochem. Biophys. Res. Commun. 169*, 1111–1116.

Hamilton, W. C., & Ibers, J. A. (1968) in *Hydrogen Bonding in Solids*, pp 181–182, W. A. Benjamin, Inc., New York.

Hamlin, R. (1985) *Methods Enzymol. 114*, 416–452.

Hanson, J. E., Kaplan, A. P., & Bartlett, P. A. (1989) *Biochemistry 28*, 6294–6305.

Hendrickson, W. A. (1976) *J. Mol. Biol. 106*, 889–893.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Biomolecular Structure, Function, Conformation and Evolution* (Srinivasan, R., Ed.) Vol. 1, pp 43–57, Pergamon Press, Oxford.

Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H., & Matthews, B. W. (1987) *Biochemistry 26*, 8642–8653.

Holmquist, B. (1977) *Biochemistry 16*, 4591–4594.

Holmquist, B., & Vallee, B. L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6216–6220.

Howard, A. J., Nielsen, C., & Xuong, N. H. (1985) *Methods Enzymol. 114*, 452–472.

Hsu, I-N., Delbaere, L. T. J., James, M. N. G., & Hofmann, T. (1977) *Nature 266*, 140–145.

James, M. N. G., & Williams, G. J. B. (1974) *Acta Crystallogr. B30*, 1249–1257.

James, M. N. G., & Matsushima, M. (1976) *Acta Crystallogr. B32*, 1708–1713.

James, M. N. G., & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299–361.

James, M. N. G., & Sielecki, A. R. (1985) *Biochemistry 24*, 3701–3713.

James, M. N. G., & Sielecki, A. R. (1987) in *Biological Macromolecules and Assemblies* (Jurnak, F. A., & McPherson, A., Eds.) Vol. 3, pp 413–482, John Wiley & Sons, New York.

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., & Hofmann, T. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 163–177, Walter de Gruyter, Berlin.

James, M. N. G., Sielecki, A. R., Hayakawa, K., & Gelb, M. H. (1992) *Biochemistry 31*, 3872–3886.

Kim, H., & Lipscomb, W. N. (1990) *Biochemistry 29*, 5546–5555.

Kim, H., & Lipscomb, W. N. (1991) *Biochemistry 30*, 8171–8180.

Komiyama, T., Suda, H., Aoyagi, T., Takeuchi, T., & Umezawa, S. (1975) *Arch. Biochem. Biophys. 171*, 727–731.

Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.

Marciniszyn, J. J., Hartsuck, J. A., & Tang, J. (1976) *J. Biol. Chem. 251*, 7088–7094.

Marshall, G. R. (1976) *Fed. Proc. 35*, 2494–2501.

Mason, S. A., Bentley, G. A., & McIntyre, G. J. (1984) in *Neutrons in Biology* (Schoenborn, B., Ed.) pp 323–334, Plenum Press, New York.

Nishino, N., & Powers, J. C. (1979) *J. Biol. Chem. 255*, 3482–3486.

North, A. C. T., Phillips, D. C., & Matthews, F. S. (1968) *Acta Crystallogr. A24*, 351–359.

Pearl, L., & Blundell, T. (1984) *FEBS Lett. 174*, 96–101.

Polgar, L. (1987) *FEBS Lett. 219*, 1–4.

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.

Rich, D. H., Bernatowicz, M. S., Agarwal, N. S., Kawai, M., Salituro, F. G., & Schmidt, P. G. (1985) *Biochemistry 24*, 3165–3173.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.

Sali, A., Veerapandian, B., Cooper, J. B., Foundling, S. I., Hoover, D. J., & Blundell, T. L. (1989) *EMBO J. 8*, 2179–2188.

Sampson, N. S., & Bartlett, P. A. (1991) *Biochemistry 30*, 2255–2263.

Sawyer, L., & James, M. N. G. (1982) *Nature 295*, 79–80.

Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.

Sielecki, A. R., Fedorov, A. A., Boodhoo, A., Andreeva, N. S., & James, M. N. G. (1990) *J. Mol. Biol. 214*, 143–170.

Speakman, J. C. (1972) *Struct. Bonding (Berlin) 12*, 141–199.

Suda, H., Aoyagi, T., Takeuchi, T., & Umezawa, H. (1973) *J. Antibiot. 26*, 621–623.

Suguna, K., Bott, R. R., Padlan, E. A., Subramanian, E., Sheriff, S., Cohen, G. H., & Davies, D. R. (1987a) *J. Mol. Biol. 196*, 877–900.

Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D., & Davies, D. R. (1987b) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7009–7013.

Szelke, M. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 421–441, Walter de Gruyter, Berlin.

Szelke, M., Jones, D. M., Atrash, B., & Hallett, A. (1983) in *Peptides, Structure and Function* (Hruby, V. J., & Rich, D. H., Eds.) pp 579–582, Pierce Chemical Co., Rockford, IL.

Tang, J., James, M. N. G., Hsu, I.-N., Jenkins, J. A., & Blundell, T. L. (1978) *Nature 271*, 618–621.

Teeter, M. M., & Kossiakoff, A. A. (1984) in *Neutrons in Biology* (Schoenborn, B., Ed.) pp 335–348, Plenum Press, New York.

Tronrud, D. E., Holden, H. M., & Matthews, B. W. (1987) *Science 235*, 571–574.

Weaver, L. H., & Matthews, B. W. (1977) *J. Mol. Biol. 114*, 119–132.

Whitlow, M., & Teeter, M. M. (1985) *J. Biomol. Struct. Dyn. 2*, 831–841.

Xuong, N. H., Sullivan, D., Neilsen, C., Hamlin, R., & Anderson, D. (1985) *J. Appl. Crystallogr. 18*, 342–350.