Theor Chim Acta (1993) 85: 231–247

Theoretica Chimica Acta

Determining proton positions in an enzyme-inhibitor complex is a first step for theoretical mechanistic studies of aspartic proteinases

Amiram Goldblum, Meir Glick, and Anwar Rayan

Department of Pharmaceutical Chemistry, School of Pharmacy, Hebrew University of Jerusalem, Jerusalem, 91120, Israel

Received September 24, 1991/Accepted August 28, 1992

1 Introduction

The mechanism of peptide cleavage by aspartic proteinases (AP) received by now only little attention in theoretical studies [1]. Recently, due to the vast interest in medicinal aspects of the AP of HIV-1 [2] and of renin [3], this well known family of proteinases (that also includes pepsins, chymosin, some cathepsins and retroviral proteinases) is now studied by many theoretical research groups. Most of the efforts are directed to the development of efficient and selective inhibitors for members of this family, that could prevent or block disease conditions such as AIDS, hypertension, as well as some cancers and ulcers.

Most of the information about the structure of AP is based on X-ray structure determinations, both of the native enzymes and of their complexes with inhibitors [4]. Two aspartic acid moieties are found to be at the center of the active sites of AP in a nearly coplanar relationship. Electronic density at the midst of this aspartic pair has been attributed by most researchers to an oxygen of a tightly bound water molecule (Fig. 1). Other water molecules are found in the vicinity, and a water molecule is believed to be the nucleophile in a suggested

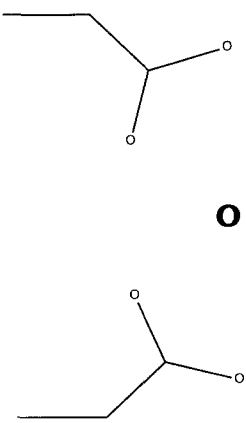


Fig. 1. Schematic arrangement in space of two aspartic side chains and oxygen of a water molecule, as found in native aspartic proteinases

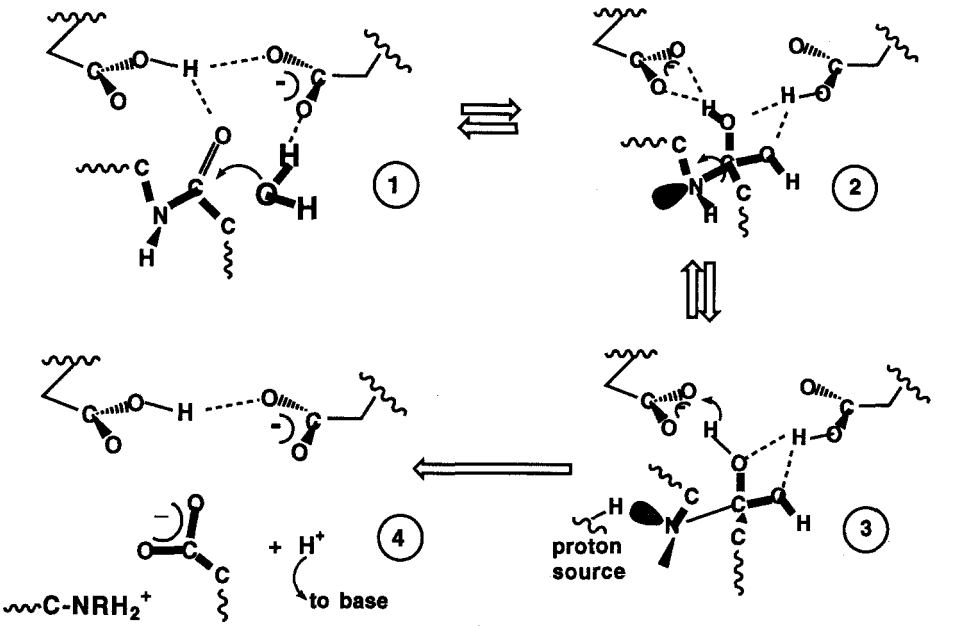


Fig. 2. One of several suggestions for the catalytic pathway of aspartic proteinases in the "general acid-general base" mechanism

"general acid-general base" (GAGB) mechanism. One possible pathway for such a catalytic peptide cleavage is presented in Fig. 2: polarization of the peptide carbonyl by a proton is succeeded by a nucleophilic attack of water (step 1) leading to an intermediate or a transition state (TS) gem-diol (step 2). Rotation about the C–N bond facilitates subsequent protonation of the nitrogen (step 3). This leads to C–N bond cleavage, with final dissocation of the products and restoration of the active site (step 4).

NMR studies [5a] have shown that a pepstatin ketone (statone) inhibitor binds to porcine pepsin in a tetrahedral hydrate form, thus supporting the GAGB mechanism. Cryoenzymological experiments have been employed to search for an expected "burst" of an amine product, if peptide cleavage is achieved by another mechanism, a direct nucleophilic attack [6] of an aspartate on the scissile bond, thus creating an anhydride intermediate [5b]. The lack of such a product is the main evidence against the nucleophilic mechanism (Fig. 3 depicts the first step). The monoanionic nature of the active site is strongly suggested by the pH dependence of pepsin catalytic activity and by the pKa values for the active site aspartates [5c].

In the widely accepted "general acid-general base" mechanism, it is still unclear which water molecule is involved in the attack on the carbonyl, how polarization of the carbonyl is achieved, and which residue (or water) is the proton donor to the nitrogen towards the final C–N bond cleavage. Most of the mechanistic refinements were based not on kinetics, but on "static" structures of AP-inhibitor complexes. Various structures for transition states and for intermediates were suggested [7]. Figure 4 presents two such ideas for the GAGB mechanism: In the upper part (path A), the proton bridging between the closest oxygens of the two aspartates polarizes the carbonyl, whose oxygen atom

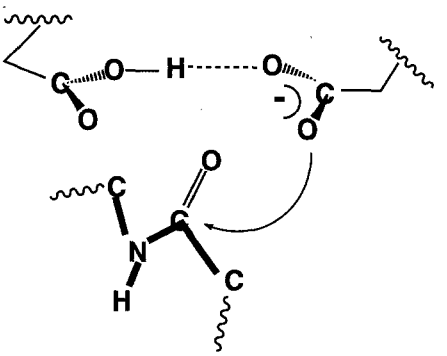


Fig. 3. The first step in the direct nucleophilic mechanism

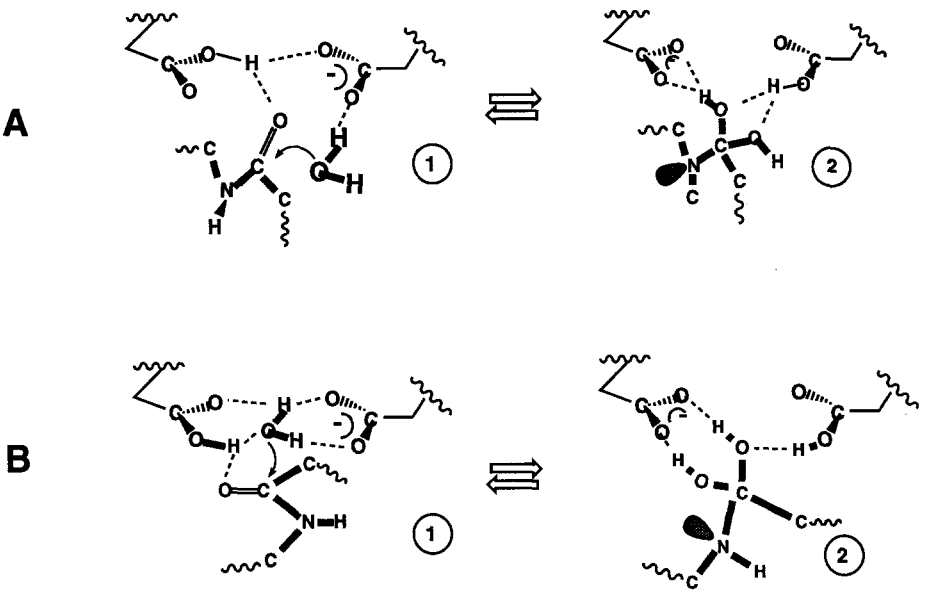


Fig. 4. Two alternatives (A and B) for binding and water attack (A1, B1) on a peptide substrate in the active site of aspartic proteinase and for the structure of an intermediate (A2, B2) gem-diol

replaced the oxygen of water at the midst of the active site. The nucleophilic attack is achieved by an external water molecule, leading to a gemdiol that is stabilized by hydrogen bonds with the aspartates. The second suggestion, path B, retains the internal water molecule while polarization is accomplished by a proton that resides on an "external" oxygen atom of the aspartates. The direction of nucleophilic attack is from the center of the active site, resulting in a gem-diol with different geometric and hydrogen bonding characteristics than those of A2 of Fig. 4. Each of the two pathways advances through one or more TS. The TS structure is believed to be especially important for understanding the mechanism and its characterization will hopefully improve the rational basis for designing "transition state analogs" that could be efficient inhibitors. But, as one can not follow an enzyme-substrate reaction by crystallography, indirect meth-

ods must be devised. For crystallography, a natural choice is to employ good inhibitors that would imitate intermediates or transition states along the pathway.

However, such an inhibitor that would mimic the TS of a substrate is difficult to design. Daniel Rich [8] indicates that inhibitors containing the irregular amino acid statine, from the universal inhibitor of AP, pepstatin [9] can be best characterized as "collected substrate-inhibitors". Inhibitors that could be most closely identified with a TS or a structurally related intermediate are those of a phosphonate analog of pepstatin [10] and the gem-diols obtained in water from α-difluoromethylene ketones [11]. Two recent crystal determinations of AP complexes with such ketone inhibitors show very similar structural characteristics [12]. A complex of difluorostatone with *penicillopepsin* has been solved at a resolution of 1.8 Å [12b]. The coordinates of a complex between a difluoromethylene peptide inhibitor and *endothiapepsin* have been solved at a resolution of 2.0 Å [12a]. In both structures, similar distance relations were found between the oxygen atoms of a gem-diol function of the inhibitors and the oxygens of the two aspartic side chains. They resemble the positions of heavy atoms in structure B2 of Fig. 4, for the substrate gem-diol form. The difluoromethylene function of these inhibitors replaces the N–H function of the substrates. While CF_2 has very different electronic character than N–H, the resemblance of the inhibitor to a substrate is found in the main chain direction away from the difluoromethylene function (Fig. 5, center).

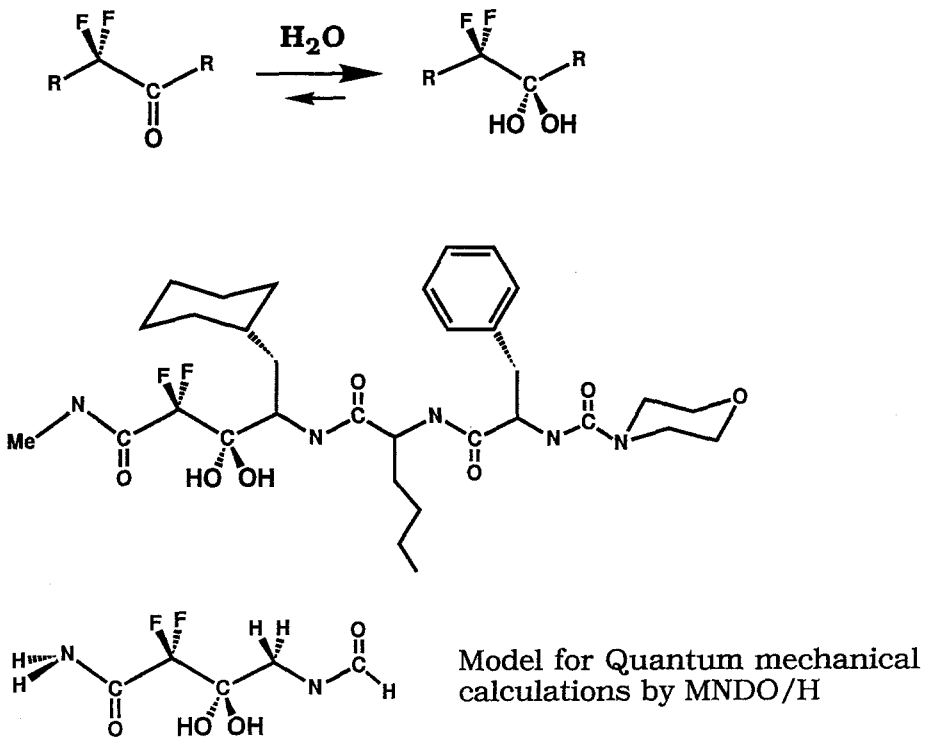


Fig. 5. A tripeptide difluoroketone inhibitor (*middle*) exists in water as the gem-diol (*upper*). The model chosen for this study is shown at the bottom

The relatively low X-ray resolution in protein crystallography does not allow direct positioning of protons in proteins, and only the "heavy atom" coordinates are given. Without the proton positions, the map of hydrogen bonding interactions can not be fully appreciated and its contribution to stabilization for enzyme-ligand and solvent interactions can not be well understood. Proton positions should be assigned for the whole protein, but three of them are of special interest for the active site-inhibitor interactions: one proton should be positioned on one of four oxygens of the aspartates (in the assumed monoanionic state), and two others – on the oxygens of the gem-diol. Each of the alternative proton positions in the active site can lead to a different rotamer for one or both of the inhibitor's hydroxyl groups. The best approach to the "real" structures is to characterize the different options by computing their energies. This is achieved by the construction of models that are naturally limited according to the method employed for energy calculations. One of the crystallographic groups [12a] employed *ab initio* calculations for a small model, resulting in the same preferred option for H-bonding relations as the one based on their intuition and experience.

We employed coordinates of the *endothiapepsin* complex [12a] as a basis for quantum mechanical and molecular mechanics modeling. MNDO/H, an improved approach to the geometric and energy description of H-bonded structures [13], was used to study relatively small models for the various options. Previous such studies with small models gave encouraging results for AP acidity and ligand binding [14]. Force field calculations for the small model were then compared to the MNDO/H results and a larger model was constructed, of a full enzyme (AP of HIV-1), including a water environment. Minima for the various options found by quantum mechanics for the original inhibitor model, in its (heavy atoms) X-ray atomic positions, were employed for successive transformations of the gem-diol inhibitor towards a substrate gem-diol derivative.

With this approach, we hope to obtain answers to the following questions: 1. What is the most favorable hydrogen bonding scheme for the difluoroketone inhibitor in the monoanionic active site? 2. Does the α-difluoro function affect the preferred protonation of the active site and the resulting H-bonding scheme? 3. What is the best option for active site binding of a substrate gem-diol derivative and is it different than the best option of the inhibitor? 4. Does the binding of the difluoroketone inhibitor alter the preferred protonation state of the native enzyme? The problems of reaction pathway for AP-substrate interaction and the steps of the catalysis require a separate study, since very different structures than the ones used for the present study must be modeled, without direct experimental guidance. However, our results supply starting points for considering the alternative reaction pathways.

2 Methods

Coordinates for the complex of *endothiapepsin* with the difluoroketone tripeptide inhibitor (Fig. 5, center) served as the basis for construction of the model. The inhibitor is present in its gem-diol form due to the balance of difluoroketone reactions with water, shown at the upper part of Fig. 5. The heavy atoms (C,N,O) of eight residues (Asp32-Thr33-Gly34-Ser35; Asp215-Thr216-Gly217-Thr218) as well as that portion of the inhibitor that is closer to the active site were kept at their crystallographic positions and hydrogen atoms were added

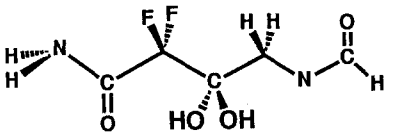
with the INSIGHT package of BIOSYM [15]. These eight residues have many hydrogen bonds between themselves and much of the active site's well-known rigidity and stabilization is conferred by their peptide bonds and side chains. The model of the inhibitor (Fig. 5, lowest) retains the crucial gem-diol and the difluoro function, as well as two amides that can supply additional hydrogen bonding stabilization with the active site residues. The methylenecyclohexyl function of the inhibitor was replaced by a hydrogen, as it projects to the side that is remote from the active site aspartates.

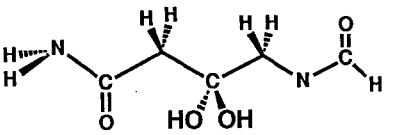
Prior to the addition of protons, all threonines were transformed into serines and the N- and C-terminals for each partial sequence were eliminated. This has been previously shown to be justified [14], since the crucial –OH function of the Thr is retained in these transformations, and its hydrogen bonding pattern is unscathed. The N- and C-terminals of each partial sequence do not contribute any H-bonding interactions with the eight active site residues, and their absence has a minor, non-qualitative effect, on the computed results, while their omission shortens the computation time.

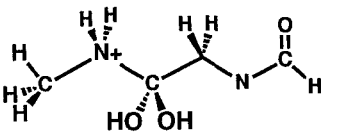
Protonation of the model was followed by geometry optimization of the protons' positions. The active site was assumed to be in the mono anionic state due to the pH (4.5) of crystallization [12a] and the known pKa values. The options for a single proton residing alternatively on the four oxygens of the aspartic moieties were considered. For each of those, a full conformation analysis was done for the two hydroxyls of the model inhibitor with the standard force field of DISCOVER [16]. All the minima were then employed as starting geometries for semiempirical geometry optimizations with MNDO/H. In the optimizations, the heavy atoms of the enzyme's residues and of the inhibitor model were kept fixed in the first stage of calculations, as our goal was to correctly position the protons for the experimental coordinates. Subsequently, the full model inhibitor was optimized only for the two lowest energy structures, which were much more stable than the others prior to full optimization. We did not attempt to search for different positionings of the inhibitor model towards the active site model: such a potential surface search is extremely complex.

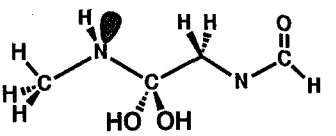
To compare the protons' positions in the active site of the complex to those in the native structure, we employed the coordinates of *endothiapepsin* (4APE) from the protein data bank [17]. A comparison between the active site structure of this native enzyme with the active site of the complex (from which the inhibitor was excluded) should reveal the effects of the inhibitor, if any, on the structure of the active site.

Changing the two fluorine atoms to hydrogens, in all the minima found for the difluoro gem-diol model, allows us to draw conclusions about the effect of the fluorines on the relative energies of the various forms. The transformations of the difluoro inhibitor through a few steps, to a hydrated substrate model, were achieved by "computer mutations" of the relevant atoms, as shown in Fig. 6. The transformations do not include the substrate model, as indicated in the introduction. It is shown only for reference. For each of the models, the six minima found for the active site-difluoroketone inhibitor interactions were employed as starting points for optimization. For each of these models, full geometry optimization of the transformed inhibitor was applied, including the protons of the active site model, but excluding the heavy atoms of the enzyme. In a recent study of active site mutations [18] we have shown that such restrictions on the active site heavy atoms are necessary in quantum mechanical









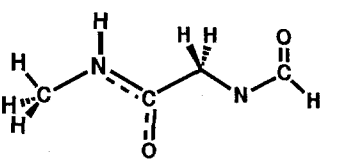


Fig. 6. Transformations of the inhibitor model (*upper structure*) were done in steps through the methylene analog to a protonated model of a hydrated substrate (*middle*)

models of AP, because the restraining effect on the movement of active site atoms by their neighbors is not included.

To test the effect of the enzyme and the environment on the results of the smaller models, the force field of DISCOVER was applied to calculations of the active site protonation options in the native AP of HIV-1 (3HVP in the protein data bank), to which the full inhibitor from the complex with *endothiapepsin* was introduced by employing its spatial relations to active site residues in the crystal. This enzyme is much smaller than *endothiapepsin*, having 198 residues in a highly symmetric dimeric structure for the native enzyme. Thus it is more easily amenable for the required calculations. The main four options for active site protonation and their associated gem-diol conformations (obtained from MNDO/H results) were compared in the full enzyme. Geometry optimization was done for all hydrogens, than for hydrogens + full inhibitor. To test the effect of water on the results, a 5 Å water layer (1151 water molecules) was added to two of the four options (DA with an "outer" proton and CA with an "inner" proton), and all the solvation shell was geometry optimized. The resulting hydration shells were then "switched" between the two structures, in order to eliminate any bias that could be due to the different water organization from two

optimizations, each affected by a different enzyme force field due to the altered proton and charge positions. Results of this part of the study must be regarded with more caution, since the native AP of HIV-1 has an "open flap", a sequence of some 10 residues from each monomer that change their conformation in enzyme-inhibitor complexes (and possibly with substrates as well), approaching closer to the active site while squeezing water molecules away. Optimization of the enzyme coordinates, especially with many water molecules present in the active site, can not lead to such a conformational change. Thus, results of this additional study, with thousands of polar and non-polar atoms around the active site, should best be regarded as balancing the other calculations which were done in vacuo.

All quantum mechanical computations were done on a VAX 9000-210 at the computer center of the Hebrew University of Jerusalem. DISCOVER calculations of the enzyme-complex and the environment were run on a Silicon Graphics 4D/280 machine.

3 Results and discussion

From the various minima of the molecular mechanics force field, six distinct minima were identified by the quantum mechanical optimizations (Fig. 7). They are shown schematically in Fig. 8. For two (B and C) of the four options for active site protonation, two separate minima were identified for each. The energies of the models (active site + difluoro inhibitor) in these six minima and of the two structures (A and D) with a fully optimized inhibitor are depicted in Fig. 9, on a relative energy scale, where DA(opt) serves as the zero reference. In all the alternative binding modes, the extended conformation of the inhibitor is assisted by hydrogen bonds from two of its N–H to the carbonyls of Gly-34 and Gly-217. The differences between the binding modes are mainly with respect to the position of the single proton in the active site and the concomitant conformations of the gem-diol hydroxyls. The structural difference between BA and BA1 (see Figs. 7 and 8) is in both the rotation angle of one OH of the inhibitor and the rotation angle of OH from Asp-32. CA and CA1 differ structurally only in the rotation angle of one OH of the inhibitor model.

Structure DA, with residue 215 protonated on the "outer" oxygen (the one more remote from the second aspartate), is found to have the lowest energy of all the alternatives. The inhibitor is bound in an asymmetric manner with respect to an apparent symmetry of the aspartate carboxyl groups (a C_2-axis can be imagined in the plane of the carboxylates, between the two aspartic acids). The relative stability of structure DA, is 16.5 kcal/mol with respect to the next most stable structure AA. This exact difference has been found also by *ab-initio* computations with the STO-4G basis set [12a]. All other structures are less stable than DA by more than 20 kcal/mol. Our molecular mechanics small model is for a somewhat different system, including the two aspartates and the full tripeptide inhibitor of Fig. 5. Again, optimization was only applied to proton positions and to rotations of the gem-diol hydroxyls. The results indicate a large preference, of at least 14.5 kcal/mol, for structure DA over the others (Fig. 10). In both these and the quantum mechanical computations, it is not surprising to find different energies for structures that seem to be symmetrically oriented – B and D on one hand, A and C on the other: The inhibitor is bound in a non-symmetric manner.

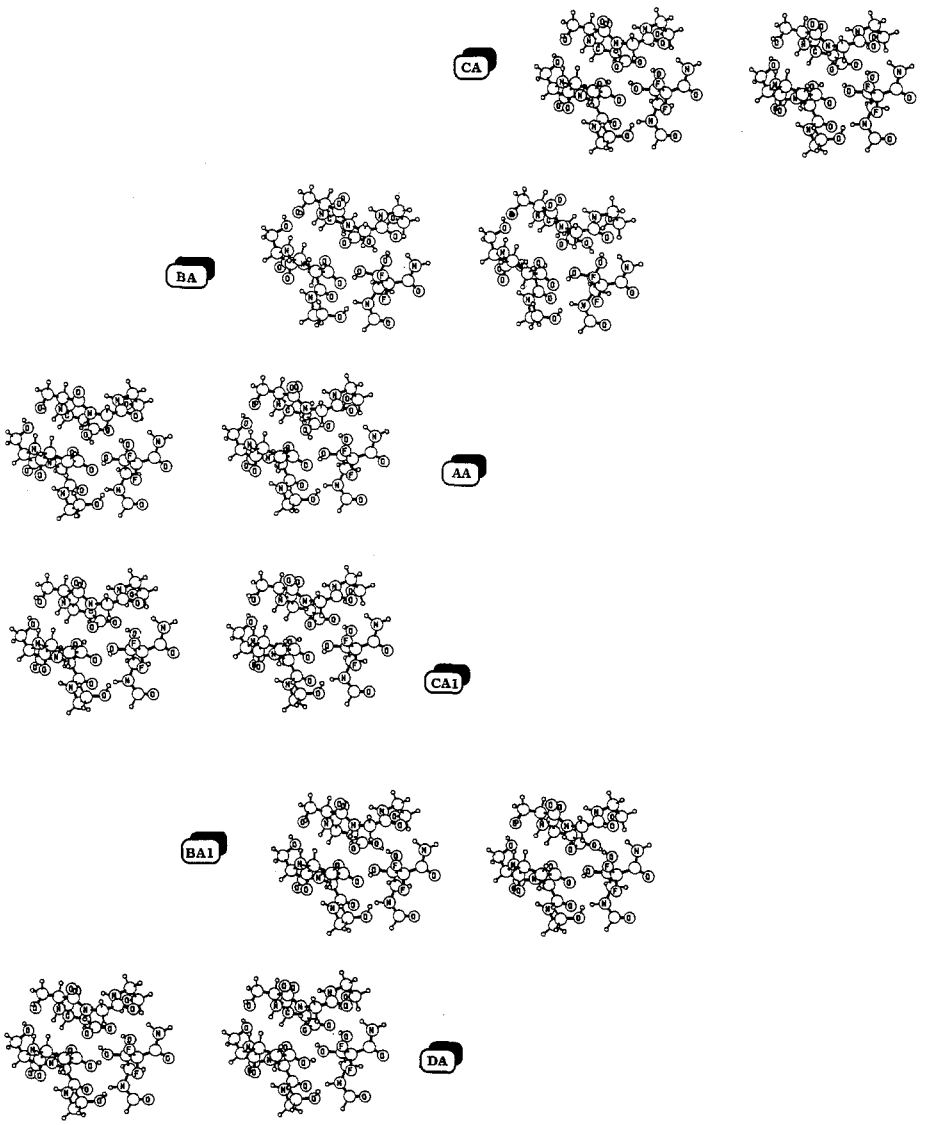
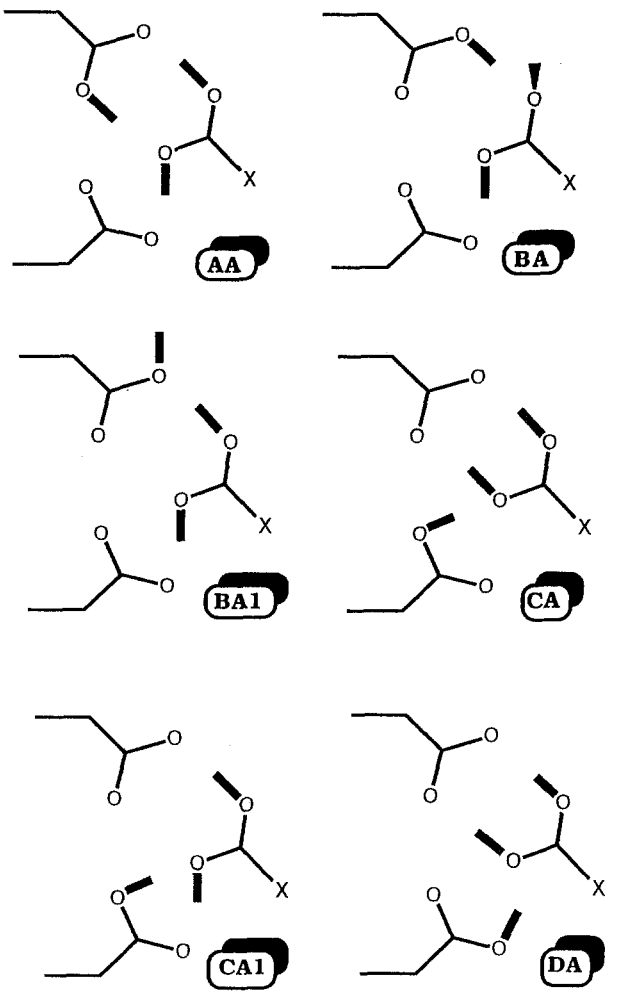


Fig. 7. Stereo plots of the six minima found for the active site-inhibitor complex, by shifting protons and hydrogen bonds and optimizing their positions, with no changes in the positions of heavy atoms

The effects of a full enzyme and some of its water environment on the four different options (A, B, C, and D) were tested on the AP of HIV-1, after docking the tripeptide inhibitor into this active site according to its position in *endothiapepsin*. Figure 11 shows the energies for three sets of calculations in this large system: Optimization of all hydrogens only, optimization of hydrogens and of the whole inhibitor, and optimization of the water structure (only for C and D). Each step was started from the final point of the previous one. It is pleasing to find that the differences between the most stable structure D and others in the



Inhibitor model: $HCO\text{-}NH\text{-}CH_2\text{-}C(OH)_2\text{-}X\text{-}R$

For $X = CF_2, CH_2$: $R = CO\text{-}NH_2$

For $X = NH_2^+$, NH: $R = CH_3$

Fig. 8. Schemes for the relations between the inhibitor's gem-diol and the mono-anionic active site of the aspartates. The schemes portray the minima shown in Fig. 7. The transformations of inhibitor to other structures were based on these six minima

small quantum- and molecular-mechanics models, are retained in this comparatively gigantic model. Thus, we conclude that form D is the best option for the binding of this inhibitor, and it is unlikely that better computational methods would reverse this result.

What is the source of the greater stabilization of this structure compared to the other options? Table 1 lists the hydrogen bonding interactions from the quantum mechanical results of these two structures, AA and DA. We included in this table some longer distances (> 3.6 Å) in order to compare them to shorter H-bonds that are symmetrically related to them. In general, the hydrogen bonds between the inhibitor and the active site are shorter than those within the active site. The active site hydrogen bonds are in the upper part of the table and

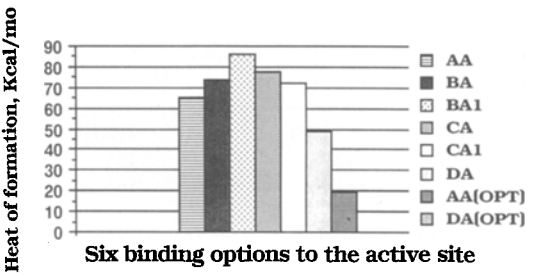


Fig. 9. Energies of the six binding modes of the inhibitor model are shown relative to the energy of the fully optimized (all inhibitor atoms + active site hydrogens) structure DA (0.0 kcal/mol on this scale). On the right side, the fully optimized AA structure is ca. 19 kcal/mol less stable. All other structures are compared on the basis of partial optimization (hydrogens only)

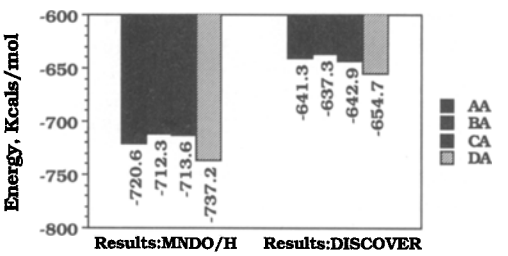


Fig. 10. Energies for 4 binding modes in quantum-mechanical (heats of formation) and molecular mechanics calculations

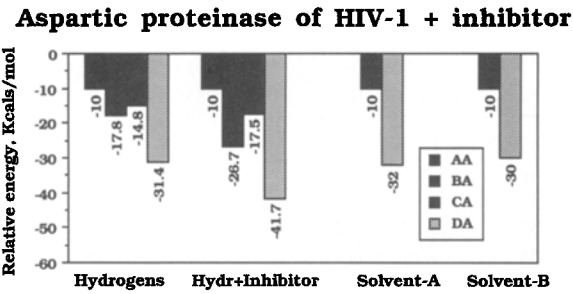


Fig. 11. The 4 binding modes in the field of a full aspartic proteinase with optimization of hydrogens and hydrogens + inhibitor. Solvent structure (optimized) effects are given for CA and DA only

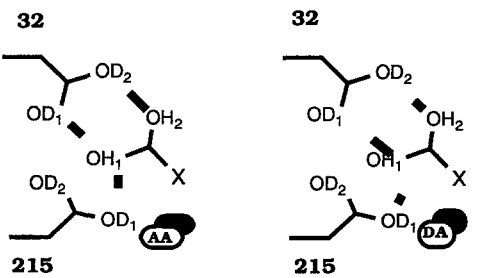
those between the inhibitor and the active site are in the lower part. In the upper part, H-bonds are compared in pairs according to the expected symmetric arrangement of the residues. The first pair in the table is the one called "fireman's grip" [19], where two threonines, one from each partial sequence (33 and 216) increase the rigidity of the active site by hydrogen bonds between their N–H to an OH of the other Thr. The overall number of hydrogen bonding interactions is larger for the complex with structure DA, and those H-bonds can justify the greater stability of this state. The excess hydrogen bonds of DA over AA is most evident in the lower part of the table, for the inhibitor-active site interactions. The inhibitor is, however, a model, and thus it is important to test if this preference remains upon full optimization of the inhibitor in both structures. Also, it is interesting to find out if the energy difference between DA and AA depends on the active site-inhibitor interaction alone, or whether some of it arises from the deposition of active site residues in the two alternative states AA and DA, even when no inhibitor is present.

Following the initial optimization of hydrogen positions in the complex, the structure and position of the model inhibitor in the active site was fully

Table 1. Hydrogen bonds in the active site-inhibitor complexes AA and DA

Acceptor Y		Donor X		Distance X···Y	From complex AA		DA	
					H···Y	ϑ(XHY)	H···Y	ϑ(XHY)
OG1	216	N	33	3.117	2.260	142.0	2.340	133.3
OG1	33	N	216	3.301	2.319	164.1	2.318	164.0
OD1	32	N	34	3.249	2.450	130.9	2.273	161.7
OD2	215	N	217	2.957	1.982	157.5	2.043	148.3
OD2	215	N	34	3.307	2.443	143.1	2.753	114.8
OD1	32	N	217	3.682	3.308	103.2	3.203	110.4
OD2	32	OG	35	2.664	1.744	159.0	1.721	162.9
OD1	215	OG	218	2.899	2.176	131.7	2.363	115.4
O	32	N	35	3.652	2.685	162.8	2.663	170.0
O	215	N	218	3.251	2.304	155.7	2.335	150.4
OD2	215	OD1	32	3.344	2.757	118.7	—[a]	—[a]
OD1	32	OH1	1NH	2.899	—[a]	—[a]	1.940	172.7
OD1	215	OH1	1NH	2.575	1.612	161.2	—[a]	—[a]
OD2	32	OH1	1NH	3.452	—[a]	—[a]	2.751	130.1
OD2	32	OH2	1NH	2.580	1.620	168.0	1.583	171.4
O	34	N1	1NH	2.667	1.642	172.4	1.638	173.3
OH1	INH	OD1	215	2.575	—[a]	—[a]	1.602	163.6
OH1	INH	OD1	32	2.900	1.938	164.2	—[a]	—[a]
F1	INH	OD2	215	3.184	—[a]	—[a]	2.617	116.1
O	217	N	INH	3.157	2.297	142.4	2.274	145.4
OG	218	N	INH	2.901	2.341	114.0	2.369	111.9

[a] Missing values indicate that an H-bond does not exist



optimized (in addition to the active site's protons) for the two more stable structures DA and AA. This computer experiment alters the positions of atoms with respect to the X-ray coordinates, and can not be regarded as reflecting any real trend of the system. The fully optimized structures are much lower in energy than their predecessors. This is mainly a result of the difference between the bond lengths in the refined X-ray structure (that were kept constant in the initial optimization) and those in MNDO/H, where energy dependence on bond length

is very sensitive due to the exponential parametrization of the core-core repulsion integrals [20]. However, the difference between the fully optimized DA and AA (18.5 kcal/mol) is larger by only 2 kcal/mol with respect to the structure with fixed heavy atoms.

Our quantum mechanical calculations do not include any effect of the enzyme's environment, except for the, explicitly included, eight residues. Various methods were proposed for including the environment effect on reactions such as those of enzyme active sites [21]. On the other hand, the well-known rigidity of the active sites of AP has been recently used in algorithms that search for fit between crystallographic coordinates of HIV-PR and coordinates of small molecules [22]. However, the asymmetric binding of the inhibitor to a nearly symmetric active site could induce some changes in positions of both heavy atoms and hydrogens with respect to the native structure. It could thus lead to altered preferences of protonation of active site aspartates in the complex. This can be examined by theory if the proton positions on the active site aspartates are compared for the native structure and for the complex, by excluding the inhibitor's coordinates from the latter. Optimization of all the protons in the field of the stationary nuclei from the two sources of X-ray coordinates reveals a striking difference between the active site of the native enzyme and that of the complex. In the first, the most stable states have a hydrogen bond between the two inner oxygens of the aspartic acids. In the active site of the enzyme-inhibitor complex, stability is much higher for protons on the two outer oxygens. These results are shown in Fig. 12. Structures AA and CA are lower than BA and DA in the native structure, but higher in the complex. In the coordinates of the native structure, relatively small energy differences are found between these two pairs of locally symmetric structures: for AA and CA with "inner" hydrogen bonds the difference is 1.7 kcal/mol, and for BA and DA with an "outer" proton the difference is 0.8 kcal/mol. In the complex, these differences are increased to 3.5 kcal/mol for both. Table 2 lists the various hydrogen bonds of the active sites. In the complex, most hydrogen bonds within the active site are longer (Y–X distances) than in the native enzyme. The active site of the complex has somewhat expanded. By undergoing such changes, it can probably improve the interactions with the inhibitor.

Thus, the inhibitor induces some structural modifications in the neighboring active site residues, so that the nearly symmetric structures found in the native enzyme (DA and BA on the one hand, AA and CA on the other) no longer exist. The positions of all the active site atoms were compared for the native enzyme and the complex: In structure DA, the difference is RMS = 0.290 for the "heavy atoms" and 0.462 for hydrogens. In structure AA, the RMS

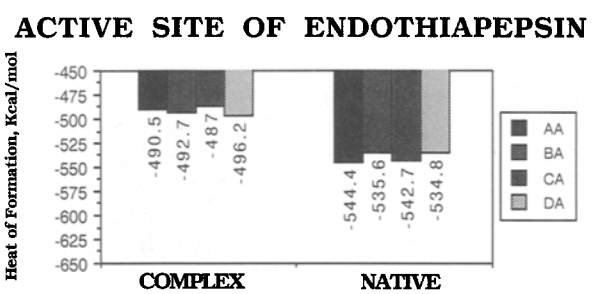


Fig. 12. Comparison of the heats of formation for the same set of active site residue from two crystal structure determinations

Table 2. Hydrogen bonds in the active sites of the native enzyme and of the complex

Acceptor		Donor		NATIVE: AA			DA		COMPLEX: AA			DA	
Y		X		Dist. Y—X	Dist. H···Y	ϑ(XHY)	Dist. H-Y	ϑ(XHY)	Dist. Y–X	Dist. H···Y	ϑ(XHY)	Dist. H···Y	ϑ(XHY)
OG1	216	N	33	3.095	2.148	155.5	2.154	154.8	3.117	2.265	141.3	2.370	130.4
OG1	33	N	216	3.152	2.171	163.3	2.169	163.1	3.301	2.320	164.0	2.318	163.8
OD1	32	N	34	3.321	2.428	146.9	2.358	157.6	3.249	2.518	128.7	2.252	166.6
OD2	215	N	217	2.845	1.923	147.4	2.032	135.4	2.957	1.977	158.4	2.058	146.4
OD2	215	N	34	3.464	2.739	128.9	2.865	118.3	3.307	2.425	145.3	2.808	110.8
OD1	32	N	217	3.447	3.072	102.6	2.976	109.2	3.682	3.308	103.2	3.203	110.4
OD3	32	OG	35	2.755	1.854	155.3	1.803	165.8	2.664	1.739	159.8	1.709	164.2
OD1	215	OG	218	2.844	1.892	170.0	1.915	165.1	2.899	2.117	138.1	2.349	116.6
O	32	N	35	3.373	2.410	159.7	2.376	170.3	3.652	2.678	164.8	2.657	172.9
O	215	N	218	3.194	2.192	171.9	2.208	165.1	3.251	2.296	157.3	2.344	149.2
OD2	215	OD	32	2.925	2.364	116.6	—a	—a	3.344	2.756	120.1	—a	—a

[a]Missing values indicate that an H-bond does not exist

is similar for "heavy atoms", and 0.449 for hydrogens. In the complex, the distance between the two "inner" oxygens of the aspartates is 3.34 Å, while in the native enzyme it is only 2.93 Å. This can be the source of lower stability for an "inner" hydrogen bond in the complex. The RMS for the heavy atoms of the modeled eight residues is, however, very small and reflects only minor movement of active site atoms in the complex, compared to the native structure.

Does the presence of fluorine atoms of the inhibitor affect the preference of conformation DA over the others? In structure DA, the outer proton on Asp-215 is at hydrogen bonding distance from one of the two fluorine atoms, in a bifurcated arrangement, as predicted by Veerapandian et al. [12a]. These fluorine atoms were transformed into hydrogens in the six minima found for the difluoro inhibitor. The results indicate that, for $-CH_2-$, the difference between DA and AA is somewhat reduced, to ca. 14.5 kcal/mol. This implies that the preferred binding mode, DA, is not an outcome of hydrogen bonding to the $-CF_2-$ group.

To test the possible binding mode of an intermediate that is closely related to the structure of the inhibitor, we transformed $-CF_2-$ to $-NH_2+-$, and the C-terminal amide was changed to a methyl group (representing a Cα). For this structure, the difference between DA and AA is smaller than for $-CF_2-$ and $-CH_2-$, while both are, in this case, much more stable with respect to CA and BA. The larger difference between DA and AA is found again when a proton is eliminated from the charged amine and a neutral hydrated peptide model is modeled. In the protonated species $-NH_2+-$, state DA is relatively destabilized due to the proximity of a proton on the outer oxygen of Asp-215 to a pair of positively charged protons on the nitrogen. The neutral species is a model for a hydrated substrate. Results for the three most stable stress in each model of this study are shown in Fig. 13. NH1 and NH2 represent the two options for a neutral amine prior to its protonation.

If binding of the substrate's gem-diol intermediate is similar to that of the difluoro inhibitor, we may conclude that such an intermediate should have the same binding preferences as those of the inhibitor, i.e., it should be bound as structure DA. Such a hydration intermediate can be produced by an attack of a water molecule that supplies one of two oxygens that comprise the gem-diol structure, while the other oxygen is that of the carbonyl. Two main pathways to this diol can be suggested, as depicted in Fig. 14. The first, in the upper right

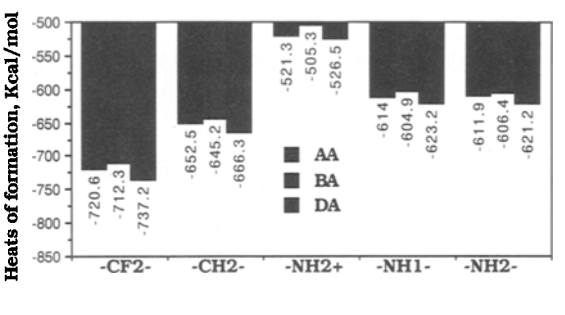


Fig. 13. Heats of formation of three binding modes for each of the calculated models

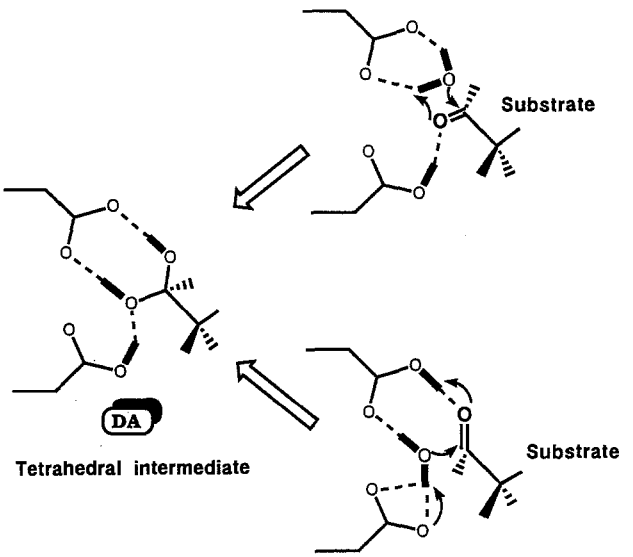


Fig. 14. The most stable structure of a hydrated substrate (tetrahedral intermediate) can be produced mainly by the two alternative pathways shown on the right. In these schemes, Asp-32 is the "upper" in each pair

part, has the substrate's carbonyl polarized by a proton on the outer oxygen of Asp-215, while a water molecule is polarized by anionic Asp-32. An attack on the carbonyl by water includes, in this case, a proton transfer from water to the carbonyl. In the second case, the carbonyl is polarized by a proton on the outer oxygen of Asp-32, and the attacking water transfers a proton to the outer oxygen of Asp-215. Other potential surfaces towards structure DA are more complex in that they require more proton shuttles or a larger change in the position of the inhibitor in the active site. Before such reaction pathways are computed, a larger potential surface for substrate binding to the active site must be searched.

Acknowledgments. This work was supported by an E.D. Bergmann grant for applied science, of the Hebrew University of Jerusalem and by grant no. 90-00311/1 of the United States–Israel Binational Science Foundation. We are grateful to Dr. T. Blundell for sending us the preprint and supplying the coordinates of the complex, for suggesting useful ideas and for discussing the results. We also thank Dr. M. James for sending us the preprint of his group's manuscript on the structure of the penicillopepsin-difluorostatone and difluorostatine complexes.

References

1. Hadzi D, Hodoscek M, Harb V, Turk D (1987) J Mol Struct (Theochem) 158:241; (1988) J Mol Struct (Theochem) 181:71
2. Huff JR, (1991) J Med Chem 34:2305
3. Greenlee WJ (1990) Medic Res Rev 10:173
4. McQuade TJ, Tomasselli AG, Liu L, Karacostas V, Moss B, Sawyer TK, Heinrikson RL, Tarpley WG (1990) Science 247:454; Roberts NA, Martin JA, Kinchington D, Broadhurst AV, Craig JC, Duncan IB, Galpin SA, Handa BK, Kay J, Krohn A, Lambert RW, Merrett JH, Mills JS, Parkes KEB, Redshaw S, Ritchie AJ, Taylor DL, Thomas GJ, Machin PJ (1990) Science 248:358
5. a) Rich DH, Bernatowicz MS, Schmidt PG (1982) J Am Chem Soc 104:3535; b) Hofmann T, Dunn BM, Fink AL (1984) Biochemistry 23:5247; c) Richards AD, Roberts R, Dunn BM, Graves MC, Kay J (1989) FEBS Lett 247:113

6. Antonov VK, Ginodman LM, Kapitannikov YV, Barshevskaya TN, Gurova AG, Rumsh LD (1978) FEBS Lett 88:87; Somayaji V, Keillor J, Brown RS (1988) J Am Chem Soc 110:2625
7. James MNG, Sielecki AR (1985) Biochemistry 24:3701; Suguna K, Bott RR, Padlan EA, Subramanian E, Sheriff S, Cohen GH, Davies DR (1987) J Mol Biol 196:877; Cooper JB, Foundling SI, Blundell TL, Boger J, Jupp RA, Kay J (1989) Biochemistry 28:8596; Miller M, Jaskolski M, Mohana Rao JK, Leis J, Wlodawer A (1989) Nature 337:576
8. Rich DH (1985) J Med Chem 28:263
9. Umezawa H, Aoyagi T, Morishima H, Matsuzaki H, Hamada M, Takeuchi T (1970) J Antibiot 23:259
10. Bartlett PA, Kezer WB (1984) J Am Chem Soc 106:4282
11. Thaisirvongs S, Pals DT, Kati WM, Turner SR, Thomasco LM (1985) J Med Chem 28:1553
12. a) Veerapandian B, Cooper JB, Sali A, Blundell T, Hoover DJ, Rosati RL, Dominy BW, Damon DB, in press (1992); b) James MNG, Sielecki AR, Hayakawa K, Gelb MH, in press (1992)
13. Goldblum A (1987) J Comput Chem 8:835
14. Goldblum A (1988) Biochemistry 27:1653; Goldblum A (1989) FEBS Lett 261:241; Goldblum A (1988) Biochem Biophys Res Comm 157:450
15. Biosym Technologies, 10065 Barnes Canyon Road, San Diego, CA 92121
16. Hagler AT, Osguthorpe DJ, Dauber-Osguthorpe P, Hemple JC (1985) Science 227:1309
17. Bernstein FC, Koetzle TF, Williams GJB, Meyer Jr EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) J Mol Biol 112:535
18. Arad G, Goldblum A, Chorev M, Shtorck A, Hughes S, Oroszlan S, Kotler M (1992) Submitted
19. Pearl L, Blundell TL (1984) FEBS Lett 174:96
20. Dewar MJS, Thiel W (1977) J Am Chem Soc 99:4899
21. a) Warshell A, Levitt M (1976) J Mol Biol 103:227; b) SJ Weiner, Chandra U, Kollman PA (1985) J Am Chem Soc 107:2219; c) Theodor B, Van Duijnen PT (1980) Theor Chim Acta (Berlin) 55:307
22. DesJarlais RL, Siebel GL, Kuntz ID, Furth PS, Alvarez JC, Ortiz De Monellano PR, DeCamp DL, Babe LM, Craik CS (1990) Proc Natl Acad Sci USA 87:6644