PROTEINS: Structure, Function, and Genetics 12:158–170 (1992)

Domain Flexibility in Aspartic Proteinases

Andrej Šali,[1] B. Veerapandian,[1] Jon B. Cooper,[1] David S. Moss,[1] Theo Hofmann,[2] and Tom L. Blundell[1]
[1]ICRF Unit of Structural Molecular Biology, Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, England; and [2]Department of Biochemistry, Kings College Circle, The University of Toronto, Toronto M5S 1A8, Canada

ABSTRACT **Comparison of the three-dimensional structures of native endothiapepsin (EC 3.4.23.6) and 15 endothiapepsin oligopeptide inhibitor complexes defined at high resolution by X-ray crystallography shows that endothiapepsin exists in two forms differing in the relative orientation of a domain comprising residues 190–302. There are relatively few interactions between the two parts of the enzyme; consequently, they can move as separate rigid bodies. A translational, librational, and screw analysis of the thermal parameters of endothiapepsin also supports a model in which the two parts can move relative to each other. In the comparison of different aspartic proteinases, the rms values are reduced by up to 47% when the two parts of the structure are superposed independently. This justifies description of the differences, including those between pepsinogen and pepsin (EC 3.4.34.1), as a rigid movement of one part relative to another although considerable distortions within the domains also occur. The consequence of the rigid body movement is a change in the shape of the active site cleft that is largest around the S_3 pocket. This is associated with a different position and conformation of the inhibitors that are bound to the two endothiapepsin forms. The relevance of these observations to a model of the hydrolysis by aspartic proteinases is briefly discussed.**

Key words: X-ray structure, TLS analysis, aspartic proteinases, inhibitor complexes, catalysis

INTRODUCTION

Aspartic proteinases are a class of proteolytic enzymes characterized by two essential aspartic residues and specific inhibition by the microbial oligopeptide pepstatin. High-resolution X-ray structures of three fungal (EC 3.4.23.6) and three mammalian (EC 3.4.23.1) aspartic proteinases are available: the fungal enzymes include endothiapepsin,[1,2] penicillopepsin,[3,4] and rhizopuspepsin[5,6]; the mammalian enzymes are pepsin (EC 3.4.23.1), which has been refined in the monoclinic[7,8] and hexagonal[9] crystal forms, chymosin[10,11] (EC 3.4.23.4), pepsinogen,[12,13] and human renin.[14] Coordinates of several complexes between aspartic proteinases and their oligopeptide inhibitors are also available. These include the structure of the complex of rhizopuspepsin with its inhibitor[15] and 15 structures of endothiapepsin complexes, which have been solved at Birkbeck College (Table I for names and references).

The fold of aspartic proteinases consists of two structurally similar lobes of about 160 residues.[16] Each of these two lobes contributes one aspartate residue to the active site center, which is buried in the middle of a long and deep active site cleft at the interface between the two lobes. The approximate diad axis of symmetry relating the two lobes passes through the active site center. The fold consists predominantly of β-strands, but it also contains several helices. A structural comparison of mammalian and fungal aspartic proteinases and different forms of the same enzyme shows that, although the same secondary structure elements are used in all cases, the spatial relationships between some of them are considerably different. See ref. 62 for the latest review of the structure and function of the aspartic proteinases.

Extensive data on specificity, steady- and pre-steady-state kinetics (ref. 24 reviews early work), isotope exchange experiments,[17–19] a new interpretation of transpeptidation reactions,[20,21] cryoenzymology,[22] and an analysis of the dependence of reaction on pH[23] led to the establishment of the noncovalent general acid–base model of catalysis[3,15,24–29] at the expense of the covalent acyl- or amino-enzyme model.[24,30,31] This was supported by the crystallographic studies which showed that the geometry of the rigid and buried active site center does not allow for the formation of a covalent inter-

Received July 4, 1990; revision accepted June 14, 1991.

Address reprint requests to Dr. Tom L. Blundell, Department of Crystallography, Birkbeck College, The University of London, Malet Street, London WC1E 7HX, England.

Abbreviations: TLS analysis, translational, librational, and screw analysis; B_{iso}, isotropic temperature factor; rms, root-mean-square deviation. The nomenclature by Schechter and Berger is used to label ligand and enzyme binding sites (Schechter, I., Berger, A. On the size of the active site in proteases: I. Papain. Biochem. Biophys. Res. Commun. 27:157–162, 1967).

TABLE I. Endothiapepsin Complexes Used in This Analysis

Inhibitor	Brookhaven code	Type	Length	Resolution (Å)	*R*-factor	References
BW624	1ER1	Amino alcohol	P_1–P_3'	2.0	0.21	55
H142	1ER4	Hydroxyl	P_6–P_3'	2.1	0.19	33,34,56–58
H261	*	Hydroxyl	P_5–P_3'	2.6	0.19	34
H256	1ER6	Reduced	P_4–P_3'	2.0	0.20	56–58
CP-81,282	*	Fluorostatone	P_3–P_1'	2.0	0.17	59
L-363,564	1ER9	Statine	P_5–P_3'	2.2	0.18	38,56
BW625	*	Amino alcohol	P_1–P_3'	2.2	0.23	59
CP-69,799	1ER2	AHS†	P_3–P_3'	1.8	0.16	37
ACRIP	1ER0	Statine	P_6–P_3'	3.0	0.19	38
H189	1ER5	Statine	P_6–P_4'	1.8	0.15	59
H77	1ER8	Reduced	P_6–P_3'	2.0	0.20	59
CP-71,362	1ER3	Hydroxyl	P_3–P_3'	2.0	0.19	59
PD125967	2ER1	Hydroxyl	P_4–P_2'	2.0	0.15	59
H261	1ER7	Hydroxyl	P_5–P_3'	1.6	0.14	59
Pepstatin	2ER2	Statine	P_3–P_4'	2.0	0.16	59

*This complex is being deposited to Brookhaven Protein Databank.[60,61]
†AHS, azahomostatine.

mediate during catalysis.[3] The same conclusion was reached from the X-ray analyses of fungal enzymes complexed with oligopeptide inhibitors.[5,32,33] The mode of binding of the inhibitors allowed modeling of the transition state of the hydrolysis. These studies together with kinetic experiments played a major role in elucidation of some aspects of the aspartic proteinase mechanism and in the design of potent inhibitors (see ref. 34 for a review).

However, the crystallographic data were not sufficient to rationalize all experimental evidence from kinetic studies. For example, the structural details of transpeptidation reactions,[20,21,31] the reasons for the differences in hydrolysis of short and long substrates, where k_{cat} increases sharply but K_m remains constant,[24,35] and the precise nature of multistep inhibitor binding[36] are still not clear. A tentative interpretation of these processes involved conformational changes in the enzyme. The crystallographic structures known at the time gave no indications of a major conformational change in the enzyme that might take place during catalysis. The only exception was a movement of the "flap" (β-hairpin 73–82) that covers the active site center. The "flap" has to open to allow the access of the ligand into the active site cleft and occupied a slightly different position in the complex than in the unliganded state.[32]

In this paper, we extend the description of conformational changes associated with inhibitor binding to endothiapepsin[37] and with activation of pepsinogen to pepsin.[9] We show that a large part of the conformational variability in the aspartic proteinase family as a whole can be accounted for by a rigid body movement of the C-domain (190–302) relative to the rest of the molecule. We discuss the relative orientations of the two rigid bodies in terms of an interface structure and TLS analysis of thermal motions. We use 15 inhibitor complexes with endothiapepsin to describe the consequences of the rigid body movement for the shape of the active site cleft and for the conformation and position of the ligand in the cleft. Finally, we discuss implications of these observations on the conformational variability for the mechanism of aspartic proteinases.

MATERIALS AND METHODS

Structures

Structures analyzed in this paper include 15 endothiapepsin complexes (Table I) and eight forms of five different aspartic proteinases (Table II). Endothiapepsin-inhibitor complexes can be crystallized in two different unit cells (Fig. 4): one isomorphous[1] to native endothiapepsin and one not.[38]

Definition of Rigid Bodies

A difference distance matrix is frequently used to represent conformational differences between two related molecules.[39] It is defined as the absolute difference between the two C_α distance plots of the molecules compared. A difference distance matrix was used to define rigid bodies for native and inhibitor bound endothiapepsin. The same division was later justified for other comparisons by corresponding difference distance plots, least-squares superpositions, and inspection on a graphics terminal. Where different enzymes were compared the alignment used was obtained by the program COMPARER.[40]

Definition of Rigid Body Movement

Rigid body movement is best described in terms of a screw transformation. A translation along and a rotation around a line in space are specified, such that one rigid body moves from its relative position in the first molecule to its relative position in the second molecule. The parameters of screw transformation were obtained from two least-squares superpositions: the first one was the superposition of the

TABLE II. Aspartic Proteinases Used in This Analysis

Protein	Code	Resolution (Å)	*R*-factor	Reference
Endothiapepsin	4APE*	2.1	0.18	1,2
Endothiapepsin–CP-69,799	1ER2*	1.8	0.16	37
Penicillopepsin	2APP*	1.8	0.14	3,4
Rhizopuspepsin	2APR*	1.8	0.14	5,6
Rhizopuspepsin-reduced bond inhibitor	3APR*	1.8	0.15	15
Porcine pepsin	MON†	2.0	0.18	9
Pepsinogen	PGEN‡	1.7	0.21	13
Calf chymosin	CHY‡	2.2	0.20	10

*These structures were deposited to Brookhaven Protein Data Bank. The same code is used here.
†Porcine pepsin refined in a monoclinic unit cell (MON) is essentially identical to hexagonal pepsin[9] that was deposited to Brookhaven Protein Data Bank with the code 2PEP.
‡These structures appear in the April 1990 release of the Brookhaven Protein Data Bank. The Brookhaven codes for chymosin and pepsinogen are 3CMS and 1PSG, respectively.

reference rigid bodies and the second one was the superposition of the moving rigid bodies starting from the orientation obtained in the first step. The rotation matrix and the translation vector acquired in the second step were then expressed in terms of the screw parameters.[41] Each least-squares superposition started with a list of equivalent C_α atoms from the alignment by COMPARER[40] but that list was then iteratively reduced until all C_α atoms in the list were at a distance of $\leq$ 3.0 Å from their equivalents in the other molecule. This ensured that local distortions in loop regions did not influence the estimate of the global rigid body movement.

TLS Analysis

In the most general harmonic model of thermal rigid body motion in molecules, the motion is described by T, L, and S tensors.[42] Kinematically, it is possible to express this general motion as a sum of six simple independent motions: three translations and three screw motions.[42] A simplifying assumption was made here that motion of a rigid body can be described by one screw oscillation only such that the displacement vector is

$$\mathbf{u} = k[t\mathbf{n} + (A - A_R)\mathbf{n}]. \quad (1)$$

The position of each atom in the rigid body changes as parameter k assumes values between $-\infty$ and $+\infty$ with probability $p(k)dk$, where $p(k)$ is any probability density function with mean equal to 0. Parameter t determines the pitch of the screw; it is the translation corresponding to 1 radian rotation. Column vector $\mathbf{n}$ is a unit vector on the screw axis. Matrix A contains Cartesian coordinates of the moving point (x,y,z):

$$A = \begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix}.$$

Matrix A_R is an equivalent matrix consisting of coordinates of any point on the screw axis.

The screw motion defined in Eq. (1) results in the following T, L, and S tensors

$$\begin{aligned} T &= \sigma^2 MNM^T \\ L &= \sigma^2 N \\ S &= \sigma^2 NM^T \end{aligned} \quad (2)$$

where $M = (tI - A_R)$ and $N = \mathbf{n}\mathbf{n}^T$. σ is standard deviation of $p(k)$ and measures a magnitude of the motion. Note that the ability of a certain screw motion to account kinematically for given T, L, and S tensors does not prove that this screw motion is the actual trajectory of the rigid body in the crystal lattice.

The isotropic temperature factor B_{iso} of any atom can be calculated from the T, L, and S tensors using the following relations[43,44]:

$$\begin{aligned} U &= T + ALA^T + AS + S^TA^T \\ B_{iso} &= 8\pi^2 \tfrac{1}{3}\text{trace}(U). \end{aligned} \quad (3)$$

The least-squares refinement program RESTRAIN[43,45] was used to calculate T, L, and S tensors for the two rigid bodies of the endothiapepsin–CP-69,799 complex. The RESTRAIN program was first modified to allow discontinuous rigid bodies consisting of any type of atoms. Main chain atoms only were included in the calculation of the tensors. The initial tensors for the start of refinement were calculated from Eqs. (2) using screw motion parameters as determined from the comparison of the native and CP-69,799 bound endothiapepsin (Table III). Ten cycles of refining T, L, and S tensors simultaneously with positions of all atoms and isotropic temperature factors for atoms other than those in protein main chain resulted in the refined T, L, and S tensors. Finally, the overdetermined system of Eqs. (2) was solved by Newton–Raphson and linear least-squares methods to obtain a single screw motion best accounting for the refined T, L, and S tensors. This screw motion is described by a position of the screw in space, A_R, orientation in space, $\mathbf{n}$, pitch of the screw, $2\pi t$, and standard deviation of rotation, σ. If we assume a Gaussian distribution for rotation, then rotational displacement from an equilibrium position will be within $\pm\sigma$ for

66.7% of time. The details of this new method for analysis of dynamic rigid body motion based on X-ray data will be published elsewhere (ref. 46 and A. Šali and D.S. Moss, in preparation).

Principal Components Analysis of Rigid Body Orientations

Principal components analysis is a standard clustering technique[47] that projects a set of points from a higher to a lower dimensional space in such a way that the first axis in the lower dimensions accounts for the largest possible fraction of variance among the original data points. The second axis accounts for the largest possible fraction of the remaining variance, etc. If it happens that only a small number of these principal axes account for a large fraction of the original variance then considerable simplification in the representation of original relationship between the points is possible by plotting them in the lower dimensional space.

The relative orientation of two rigid bodies can be described by six independent parameters [cf. Eq. (1)]. These parameters specify a point in a six-dimensional space. However, because they depend on an arbitrary choice of the reference structure, such a simple description of a rigid body orientation is biased. Bias can be removed by using, in turn, each structure as a reference when describing relative orientation in one protein and then joining individual parameter sets. In our analysis of a seven membered family, seven sets of six coordinates were joined in a 42-dimensional vector specifying an unbiased relative orientation of the two rigid bodies in one protein. These seven 42-dimensional vectors can then be subjected to a standard principal components analysis[47] to reduce the number of dimensions still accounting for a large fraction of the original variance. Before the eigenvalues/eigenvectors problem was solved, the vectors were normalized, so that the mean and variance of the rotational and translational components of the seven vectors were 0 and 1 respectively.[47]

Miscellaneous

Mainchain hydrogen bonds were found by program DSSP[48] using an energy cutoff of −1.0 kcal/mol. Hydrogen bonds involving side chains were defined using a simple 3.5 Å cutoff for the distance between any protein donor and acceptor atoms. This cutoff was relaxed to 3.6 Å when water molecules were included in the analysis. Hydrophobic contacts were obtained as in ref. 40. Solvent contact areas of residues were calculated by the algorithm of Richmond and Richards.[49] The interface between the two rigid bodies in endothiapepsin was defined as a union of residues obtained by three criteria: (1) Residues with at least one atom closer than 4.5 Å to any atom from the other rigid body. (2) Residues that lose more than 2 Å^2 of their solvent contact area when the two artificially separated rigid bodies form the folded structure. (3) Residues that are hydrogen bonded via water molecules to any residue in the other rigid body.

RESULTS

Fold of the Aspartic Proteinases

A conventional starting point for the description of the architecture of the aspartic proteinases is a division of the fold into two structurally similar lobes.[16] However, as the interface between these two lobes involves a well defined six stranded β-pleated sheet, we may alternatively divide the molecule into three parts (Fig. 1). In this description, there are two mainly β-sheet domains—the N- and C-domains—that provide the active site cleft and are also related by a diad axis of symmetry. These are packed on top of the third part, a central motif involving the interlobe β-sheet and a helix at each end. The C-domain is more clearly separated from the rest of the molecule than the N-domain (Fig. 1). Superposition of the C_α backbones of seven forms of five different aspartic proteinases establishes that by far the largest differences across the family as well as between different forms of the same enzyme involve the C-domain (Fig. 1).

The diad axis of symmetry implies similarity between the conformation and orientation of the N- and C-domains. However, there are many important differences between these two domains that may explain why the conformational changes described later do not occur symmetrically with respect to the aspartic proteinase fold as a whole. First, the neighboring edge strands of the central motif (a, −2–7) and the N-domain (b, 13–20) are much closer than the equivalent strands in the C-domain (a′, 179–184, b′, 190–199). The average distance between juxtaposed C_α atoms of residues 3–7 and 17–13 is 5.5 Å whereas the equivalent distance in the C-domain is 6.6 Å. This results in the gap between the central motif and the C-domain that does not occur in the N-domain. Second, the conformational differences between the two domains include the presence of the short 3_{10} helix (269–274) in the C-domain that interfaces to the central motif via the α-helix 304–310. There is no equivalent of the helix 269–274 in the N-domain. Additionally, while the N-domain can be viewed as composed of two β-sheets 1_N and 2_N,[2] the sheet 2_C of the C-domain has a definite discontinuity at strands 243–246 and 289–285. These two strands are not hydrogen bonded to each other, contrary to the equivalent strands 80–87 and 108–99 in the N-domain. Third, a major difference between the two domains is the presence of the long β-hairpin "flap" (71–82) and α-helix 108–114 in the N-domain. When the ligand binds, the "flap" contributes to S_2, S_1, and S_2' pockets whereas the helix contributes to binding the P_3 residue. The "flap" is just a short turn (241–244) in the C-domain and the

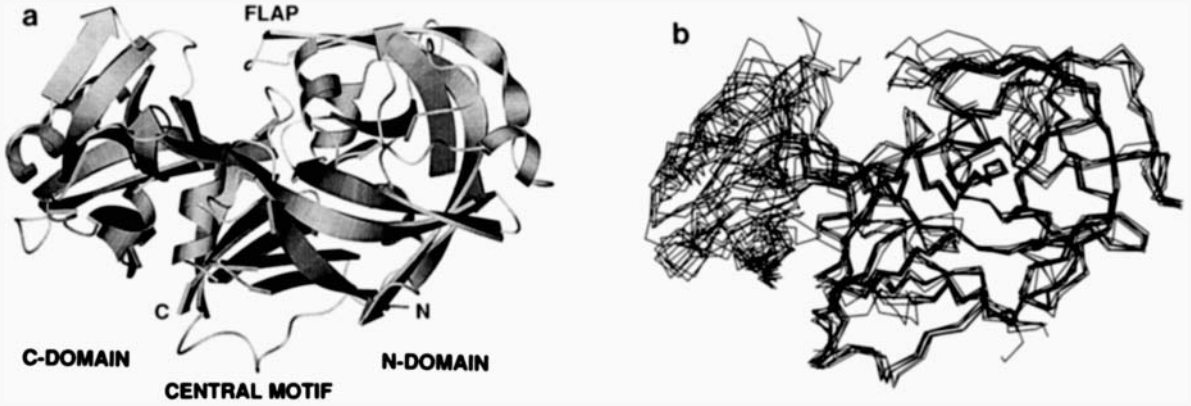

Fig. 1. Structural organization of aspartic proteinases. **(a)** Secondary structure ribbon diagram for endothiapepsin. **(b)** Aspartic proteinases listed in Table II superposed on their central motifs and N-domains; for clarity, the amino-terminal part of pepsinogen up to the residue 12 (pepsin numbering) is not shown.

helix does not exist at all breaking dramatically the symmetry of the binding pockets.

Definition of Rigid Body in Endothiapepsin

A difference distance plot (Fig. 2) shows the conformational changes in endothiapepsin when the oligopeptide inhibitor CP-69,799 binds into its active site cleft. If one neglects local distortions such as those involving the "flap," this plot shows that there are only two parts of the structure that retain their internal conformation but change their spatial relationship when the inhibitor binds. Thus the two rigid bodies in endothiapepsin with respect to the inhibitor binding are defined. The first rigid body comprises residues −2 to 189 and 303 to 326 and the second rigid body consists of residues 190 to 302. This division is in complete accord with the tripartite description of the aspartic proteinases fold: the first rigid body comprises both the central motif and the N-domain and the second rigid body corresponds to the C-domain.

Definition of the Rigid Body Movement in Endothiapepsin

The description of the rigid body movement in terms of the screw motion (Fig. 3) shows that the conformational change involved a 4° rotation around and a negligible translation along the screw axis that passes approximately through the active site center (Table III, the first line). The rms value for the two C-domains is 1.47 Å when only the central motifs and N-domains are used in the least-squares superposition (Table III). This is decreased to 0.46 Å (31% of the original value) when the two C-domains are separately superposed (Table III). This large reduction justifies description of the conformational transition as a rigid body movement as opposed to distortion within the C-domain.

The movement corresponds to changes in the relative positions of the C_α atoms of up to 2.5 Å. Changes in the active site cleft become appreciable (≈ 1.5 Å) at the S_3 pocket formed in part by the α-helix 108–114; the cleft opens when the inhibitor binds. After the "flap" β-hairpin, this helix undergoes the next largest decrease in thermal mobility of all regular secondary structure elements when the

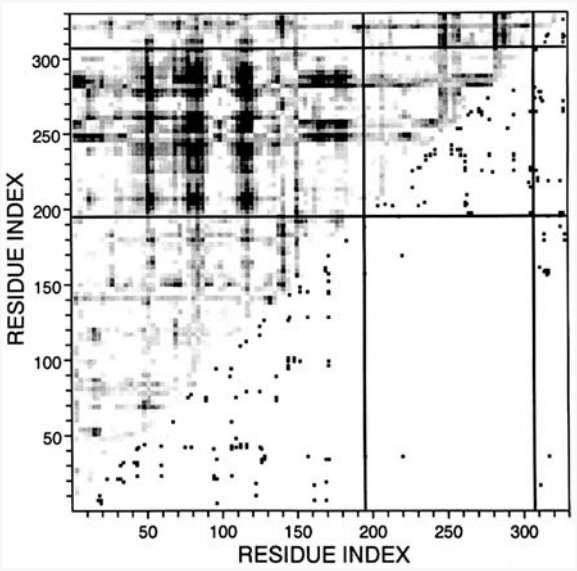

Fig. 2. Definition of rigid bodies in endothiapepsin. The upper half of this plot contains one-half of the symmetrical difference distance matrix obtained from the comparison of native endothiapepsin and the endothiapepsin molecule from the CP-69,799 complex. Shades of gray indicate differences in distances with white and black for values smaller and larger than 0.2 and 1.5 Å, respectively. One cell contains an average for the 2 × 2 residue by residue mask. The black lines divide the molecule into three parts that retain their internal conformation. The first rigid body is formed from the first and last part of the chain. The middle part of the chain is the second rigid body. The lower half contains a black square at position *(i,j)* if residues *i* and *j* are in a hydrophobic contact[40] in native endothiapepsin.

Fig. 3. Geometry of the rigid body movement in endothiapepsin–inhibitor complexes. The movement of the second rigid body (C-domain) relative to the first one (central motif and N-domain) is described in terms of the screw motion. Thick line, native endothiapepsin; thin line, the second rigid body of the complex between endothiapepsin and H261. The orientation of this inhibitor complex is obtained from the superposition of its first rigid body with the first rigid body of the native endothiapepsin. The H261 complex is shown as a representative of the predominantly nonisomorphous group of inhibitor complexes that includes CP-69,799 inhibitor (Fig. 4).

inhibitor binds.[37] The average isotropic temperature factor of the helix falls from 37 to 16 Å^2. There is a further local movement of this helix of $\approx$ 0.8 Å in roughly the same direction as the global rigid body movement. This indicates that this could be the contact point on the enzyme used to trigger the gross conformational change by the ligand in the active site cleft. However, the reasons for a decrease in the average isotropic temperature factor and a positional change of this helix cannot be attributed unequivocally to the ligand binding because some of the helix residues are in contact with a neighbouring molecule in the unit cell of the inhibitor complex but not in that of the native enzyme.

Figure 4 shows the magnitude of rigid body changes for all 15 endothiapepsin–inhibitor complexes. In all cases, the orientation of the screw axis is the same as for the example described above. It is apparent that two forms of endothiapepsin exist. The first one with a very small rotation and translation is very close to the native endothiapepsin; the second one with $\approx$ 4° rotation and a very small translation of $\approx$ 0.3 Å is the form discussed above. Several conclusions can be easily made. First, there is no correlation between the chemical type of the P_1–P_1' bond of the inhibitor and the relative orientation of the rigid bodies, except that the statine inhibitors tend to associate with lower displacement from native conformation within each group of the complexes. Second, there is a clear but not an exclusive association of the rigid body movement with the nonisomorphous crystallographic unit cell. Third, the P_3 and/or P_2 residues of the ligand seem to be a necessary but not a sufficient condition for the rigid body movement to occur, since the absence of the P_3 and P_2 residues is the only outstanding feature of the BW625 inhibitor that crystallizes in the nonisomorphous unit cell and exhibits very small rigid body movement.

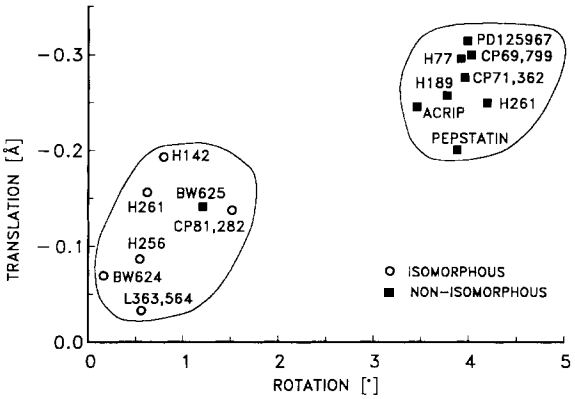

Fig. 4. Magnitude of the rigid body movement in 15 endothiapepsin–inhibitor complexes. A rotation and a translation component are specified for all 15 inhibitor complexes. In addition, the type of crystallographic unit cell is indicated by an empty circle for an isomorphous and a filled square for a nonisomorphous unit cell.

Comparison of Aspartic Proteinases

Figure 1 shows that most of the structural variability in the aspartic proteinase family involves the C-domain which corresponds to the second rigid body in endothiapepsin inhibitor binding. Therefore, the rigid body movement analysis including difference distance plots and screw movement fitting was extended from the endothiapepsin complexes to the whole family. A detailed description of some of the comparisons is presented in Figure 5 and Table III. Difference distance plots (data not shown) and a large reduction in the rms deviation between the second rigid bodies when the rigid body movement is allowed (Table III) confirm that the rigid body movement is also a convenient way of accounting for a large part of the differences within the aspartic proteinase family as a whole.

Principal components analysis (Fig. 5) of rigid body orientations in the aspartic proteinase family separates mammalian members from fungal ones. Within the mammalian group, the zymogen of pepsin is separated from pepsin and chymosin. Within

TABLE III. Rigid Body Movement in Aspartic Proteinases*

Proteins compared	Screw axis†		Magnitude‡		Rigid body rms (Å)			rms ratio§ 2 after/ 2 before
	Direction	Displacement (Å)	θ (°)	t (Å)	1 after**	2 before††	2 after‡‡	
4APE–5APE	−.608, −.775, −.175	2.6, −2.9, 3.4	4.0	−0.29	0.43 (221:221)	1.47	0.46 (111:111)	.31
4APE–APP	−.339, −.046, −.940	1.6, −1.6, −0.5	9.9	0.44	0.80 (215:209)	2.02	1.17 (104:85)	.58
4APE–MON	−.344, −.191, −.919	2.2, −0.6, −0.6	17.6	−0.14	1.08 (209:196)	3.36	1.71 (105:82)	.51
5APE–PGEN	−.446, .207, −.871	0.9, −0.9, −0.7	17.6	−0.24	1.15 (207:184)	3.25	1.58 (105:78)	.49
2APP–2APR	.706, −.491, .511	−2.3, 2.3, 5.4	2.9	−0.06	0.99 (212:200)	1.67	1.41 (103:91)	.84
2APP–CHY	−.193, −.135, −.972	5.8, −2.8, −0.8	6.1	−0.33	1.11 (208:194)	1.80	1.28 (98:84)	.71
MON–PGEN	−.664, .638 .389	−2.7, 0.4, −5.3	6.5	−0.45	0.77 (211:193)	1.99	1.22 (111:105)	.61
MON–CHY	.255, .663, .704	−8.3, 5.1, −1.8	4.0	−0.12	0.82 (215:213)	1.33	0.91 (104:104)	.69
PGEN–CHY	.931, −.344, .121	0.2, 0.9, 0.9	5.7	−0.58	0.96 (211:191)	2.06	0.97 (104:103)	.47

*All screw axes are directly comparable: they were transformed applying the transformation from the superposition of the first rigid body of the first molecule in the comparison to the endothiapepsin's (4APE) first rigid body.

†Direction cosines and location of the axes are specified. The location is described by a displacement vector of the point on the axis closest to point P where point P was found by minimizing the sum of squared distances from it to all the axes. In the 4APE coordinate system point P is (24.0, −10.9, 7.2) and is only 6 Å away from C_α atoms of residues 212, 215, and 303.

‡A rotaton (θ) around, and translation (t) along the screw axis are specified.

§C_α atoms found equivalent in the local fit of carboxyl-terminal rigid bodies were used in both rms calculations for the second rigid body.

**The rms deviation for the first rigid bodies as obtained from the local superposition of the first rigid bodies only. The two numbers in parentheses (N:M) are the numbers of equivalent C_α atoms: N is the number of initially equivalent residues as inferred from the alignment obtained by the program COMPARER[40] and M is the final number of equivalences obtained from the refinement of the rigid body superposition by iterative least-squares fitting (see also the section Definition of rigid body movement in the Methods). Relatively small differences between the two numbers demonstrate that the larger parts of the enzymes do move as rigid bodies.

††The rms deviation for the second rigid bodies as obtained from the superposition of the first rigid bodies only.

‡‡The rms deviation for the second rigid bodies as obtained from the superposition of the second rigid bodies only.

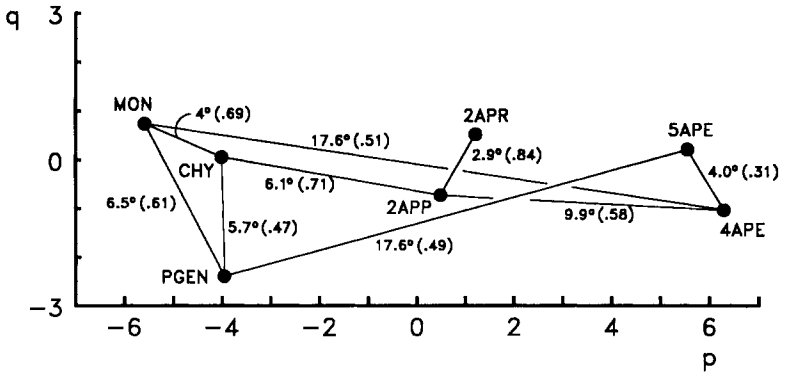

Fig. 5. Principal components analysis of the rigid body orientations in the aspartic proteinase family. The first two principal components, p and q, account for 91.4 and 4.9%, respectively, of the original variance among the seven orientations examined. For both principal components, loads of original rotation and translation coordinates are approximately the same; therefore, no simplification by neglecting some rotations or translations is possible. Relationships detailed in Table III are indicated by a line connecting the two proteins in the comparison. Angle of rotation and fractional improvement in the rms deviation for the second rigid body when rigid body movement is allowed are also shown—see Table III for further details.

the fungal group, endothiapepsin forms are separated from rhizopuspepsin and penicillopepsin, which are also significantly closer to mammalian enzymes than endothiapepsin forms. This could contribute to the failure of endothiapepsin and success of penicillopepsin and rhizopuspepsin as the search models for the molecular replacement technique in the X-ray analysis of pepsin.[9,10] In the whole family, penicillopepsin and rhizopuspepsin are the only enzymes with identical rigid body orientation, as judged by the small rotation angle and only small improvement in the rms distance when the movement is allowed. All other enzyme pairs have significantly different rigid body orientations. The magnitude of the rigid body rotation in the family ranges from 2.9° for penicillopepsin–rhizopuspepsin to 17.6° for endothiapepsin–pepsin comparisons. The latter is equivalent to 11 Å positional differences for C_α atoms distant from the axis. In all cases, translation is negligible compared to rotation. To put these numbers into perspective, a survey of the families of multidomain proteins showed that usually the differences in relative orientation of domains involve less than 1.0° rotation and improve the rms distance

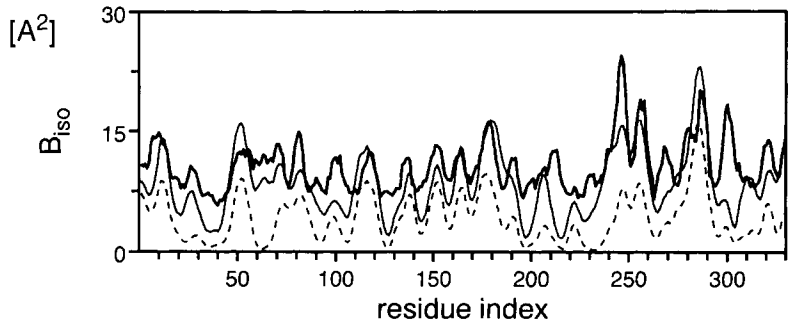

Fig. 6. TLS analysis of the rigid body motion. Three different types of main chain isotropic temperature factors smoothed by the running window of five residues are shown. Thick line, isotropic temperature factors refined directly.[45] Thin line, isotropic temperature factors calculated from the *T*, *L*, and *S* tensors refined by RESTRAIN [Eqs. (3)]. Dashed line, isotropic temperature factors calculated from the single screw motion fitted to the refined *T*, *L*, and *S* tensors [Eqs. (2)–(3)]. The average isotropic temperature factors for the three models are 11.2, 8.9, and 4.4 Å^2, respectively. The Pearson cross-correlation coefficients between the first two, the first and the last, and the last two curves are 0.69, 0.62, and 0.86, respectively.

by less than 20% (A. Šali, unpublished results; compare with Table III).

The complex between an oligopeptide inhibitor and rhizopuspepsin[15] is the only aspartic proteinase–inhibitor complex besides endothiapepsin complexes that was deposited to Brookhaven Protein Databank. Rigid body movement in rhizopuspepsin (Table III), if any, is comparable in magnitude to the endothiapepsin complexes with the native-like orientation of rigid bodies (Fig. 4). The only large conformational difference involves the "flap" region moving relative to the rest of the enzyme. The rigidity in rhizopuspepsin may be provided by the crystal packing: the rhizopuspepsin complex was prepared by soaking the inhibitor in the formed enzyme crystal whereas all endothiapepsin complexes were co-crystallized.

Although the precise geometry and magnitude of the rigid body movement in pairwise comparisons of aspartic proteinases vary from case to case (Table III), the main features of the movement are the same. First, the same division in the two rigid bodies applies. Second, screw axes always pass roughly through the interface between the rigid bodies in the vicinity of the two main chain segments that connect the two parts (Table III). These two parts also happen to be near the tips of the two loops carrying the essential aspartates 32 and 215. Third, the movement is associated with geometrical changes that are largest around the S_3 pocket.

TLS Analysis

The *T*, *L*, and *S* tensors, refined by RESTRAIN,[43,45] for both rigid bodies in endothiapepsin–CP-69,799 complex were obtained as described in the Methods. Figure 6 shows that the isotropic temperature factors calculated from these tensors [Eqs. (3)] have very similar values to those obtained directly from an isotropic temperature factor refinement.[45] This indicates that a significant fraction of the protein dynamics and static disorder can be described by a rigid body motion.

The single screw motion calculated from Eqs. (2) accounts for the isotropic temperature factors of the enzyme [Eqs. (3)] almost as well as the original refined *T*, *L*, and *S* tensors (Fig. 6). This indicates that the rigid body motion can be at least kinematically approximated by a single screw oscillation.

The orientations and positions of the screw axes obtained from the TLS analysis for the two rigid bodies are very similar to these of the axis used to describe the static change: for the first rigid body, the angle between the two screw axes is only 16° and the distance 1.2 Å (the standard deviation of rotation around the screw is 3.5°, translation is negligible); the numbers for the second rigid body are 45° and 0.9 Å (the standard deviation for rotation is 2.8°). These results are consistent with the two rigid bodies of the molecule oscillating in a manner that leads most directly to the other equilibrium orientation (Figs. 3 and 4) when enough energy is accumulated to climb the activation energy barrier.

The single screw-axis model derived from the refined *T*, *L*, and *S* tensors describes atomic displacements in the enzyme with only six parameters for each of the two rigid bodies. From Figure 6 it can be seen that this very simple model accounts for 39% of the mean-square displacement observed with the full isotropic least-squares model which requires 2389 thermal parameters.

It is generally the case that the isotropic temperature factor in proteins correlates very strongly with residue solvent accessibility or distance from the molecular gravity centre. The reason for this correlation is thermal flexibility of exposed smaller parts of the structure such as loops and not motion of a larger rigid body as a whole. TLS analysis cannot distinguish on its own between the two phenomena. As expected, the cross-correlation coefficient between the directly refined isotropic temperature fac-

tors and the smoothed fractional residue contact area in endothiapepsin–CP-69,799 complex is just as high (0.67) as the coefficient between the former and the isotropic temperature factors calculated from the refined T, L, and S tensors (0.69). However, two alternative geometrical models for rigid body movement, rotation of the two rigid bodies around their gravity centers and rotation around the axes through their gravity centers that are also perpendicular to the static screw axis (Table III), result in cross-correlation coefficients of only 0.41 and at most 0.31, respectively. These are significantly weaker correlations than the one provided by the TLS model and they indicate that rigid body motion does contribute to the overall mobility in endothiapepsin in addition to uncorrelated motions of smaller parts of the structure.

Structural Background for the Rigid Body Movement

The two rigid bodies of aspartic proteinases comprise separate parts of the structure and the boundary between them is evident, both from the trace of C_α backbone (Fig. 1) and from the two-dimensional plot of hydrophobic contacts (Fig. 2). The interface between the rigid bodies in endothiapepsin consists of 12 potentially charged residues (9 negatively, 3 positively), 15 polar and 53 nonpolar residues (Fig. 7). Altogether, these 80 residues represent a significant fraction of the total of 330 residues. The rigid bodies are held together by two mainchain segments at residues 190 and 303, by van der Waals contacts, 1 side chain–side chain hydrogen bond, 3 main chain–main chain hydrogen bonds, 10 side chain–main chain hydrogen bonds, and a number of hydrogen bonds mediated by 13 water molecules. There are no ion pairs at the cutoff distance of 5 Å. Differences between the CP-69,799 complex and the native enzyme are limited to absence of three of the 13 bridging water molecules, to addition of four new water molecules and to minor changes in some van der Waals contact distances.

The most intimate contact is made between the tips of the two active site loops containing catalytic aspartates 32 and 215. The two loops are engaged in a side chain–main chain hydrogen bonding network called a "fireman's" grip.[25] Additionally, there is a contact between the helix 303–309 from the first rigid body and the 3_{10} helix 269–274 from the second rigid body. While the disposition of the active site loops remains essentially unchanged during the movement as a result of the screw axis passing through them, a change in the relative orientation of the two helices can be rationalized by the helix interface shear mechanism. The shear mechanism was proposed to be a general way of accomplishing the conformational variability in α-helical proteins by allowing differences of up to 1.5 Å in a relative helix orientation brought about by small adjustments between the side chains in contact.[50] This description fits precisely the changes in the relative orientation of the two helices as well as the whole rigid bodies (Figs. 6 and 8). Additionally, the unfavorable atomic contacts resulting from the rigid body movement are dissipated further by very small movement of the 303–309 helix relative to the first rigid body (Fig. 2).

Consequences of the Rigid Body Movement for the Ligand Position and Conformation

Comparison of the positions and conformations of the inhibitors bound to different enzyme forms, each with a different orientation of the rigid bodies, provides an opportunity to study how the rigid body movement affects the relative position and conformation of ligands in the active site cleft.

To increase the accuracy of the comparison, the complexes were first superposed on their C-domains and the averages of positions for mainchain atoms within the two groups of inhibitors were calculated (Fig. 9). For comparison, these averages were used instead of the individual structures. The averaging was justified a posteriori by the deviations of the mainchain atoms from the average being in the order of the experimental error in coordinates or smaller ($\approx$ 0.25 Å).

Relatively large rms deviation of 0.73 Å for the two averages means that the positions and/or conformations of the two ligand averages relative to the C-domain are significantly different. The movement of the inhibitor mainchain relative to the C-domain can be described as a screw motion involving a −0.34 Å translation along and a 4.7° rotation around the screw axis. When the movement relative to the larger rigid body is considered the 0.13 Å translation and 3.8° rotation occur. A relatively small rms deviation of 0.33 Å after the superposition of the inhibitor averages shows that the conformations of the two ligand types are not very different. However, that there are real differences is supported by the fact that this rms deviation is still larger than the one expected by chance when the rms error in the atomic positions of individual inhibitors is 0.25 Å: then the expected rms deviation in the positions of the averages of five inhibitors is 0.11 Å ($0.25/\sqrt{5}$) and the expected rms deviation between two such averages is only 0.16 Å ($\sqrt{2} \times 0.11$). The movement of the inhibitor as a whole and the internal distortion result in relative positional changes of up to 1 Å for the main chain atoms of the P_3 residue.

DISCUSSION

In the previous section, it was shown that a relatively large change in the shape of the active site cleft in endothiapepsin occurs as a result of the rigid body movement of the C-domain relative to the rest of the molecule. It was also demonstrated that the

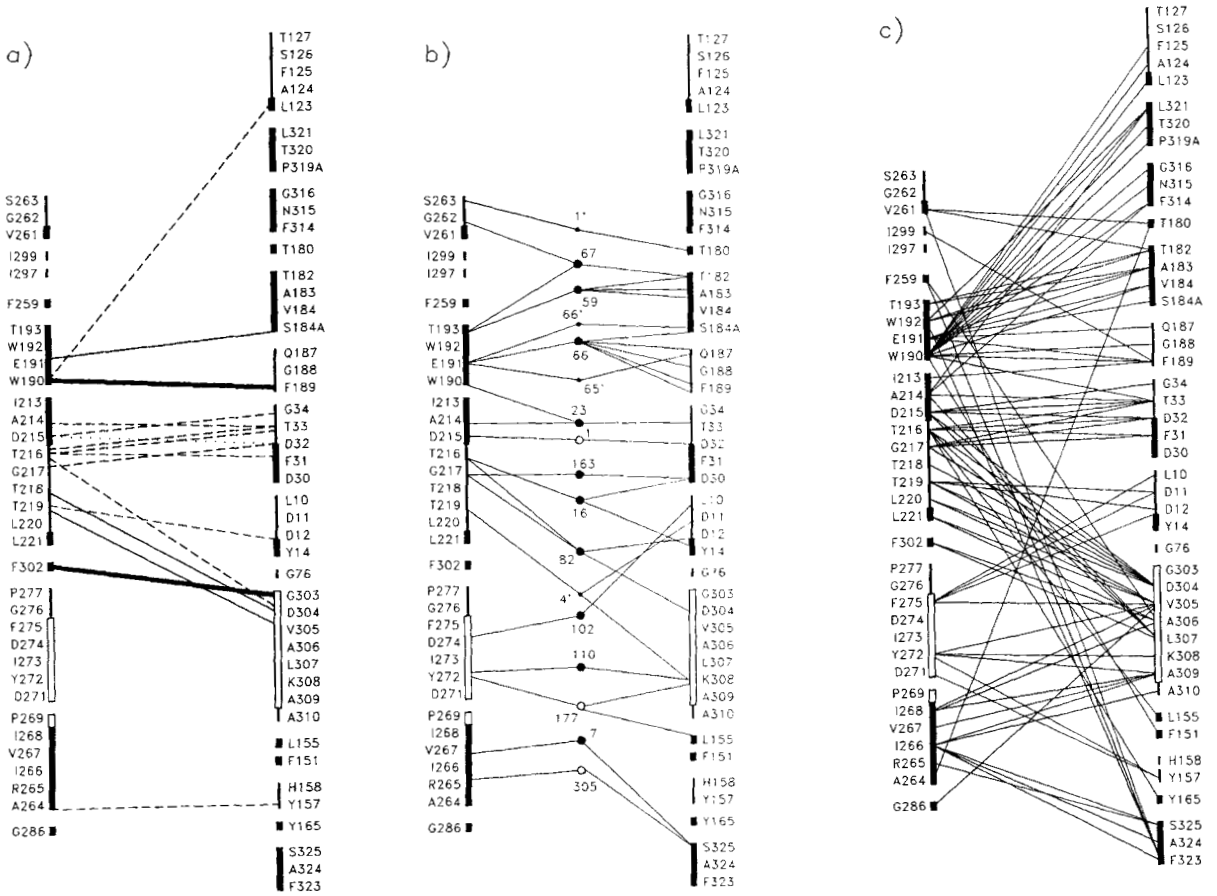

Fig. 7. Schematics of the rigid body interface in endothiapepsin. Residues forming the interface between the two rigid bodies are shown together with their secondary structure assignments: empty rectangle for helix, filled bold line for β-strand, and a line for unordered structure. For clarity, different types of interaction are separated into three diagrams. **(a)** Two main chain connections are shown by bold lines, main chain–main chain hydrogen bonds by thin continuous lines, one side chain–side chain hydrogen bond by a dotted line, and main chain–side chain hydrogen bonds by dashed lines. **(b)** Hydrogen bonds between the two rigid bodies mediated by water molecules. Large filled circles are water molecules present in the native and inhibitor bound endothiapepsin whereas large empty and small filled circles correspond to water molecules that occur, respectively, in native and inhibitor bound endothiapepsin only. **(c)** van der Waals contacts are shown for native endothiapepsin only.

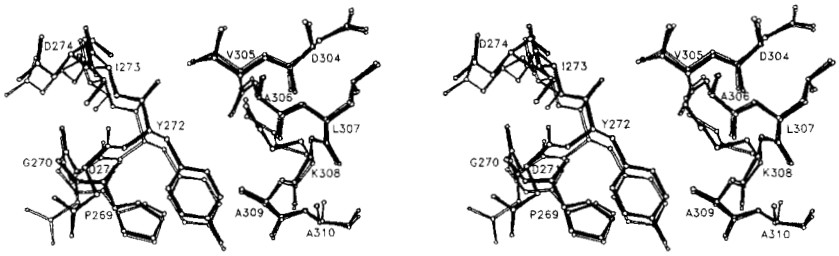

Fig. 8. The helix interface shear mechanism of the two helical segments at rigid body interface. The two pairs are superposed on the helix 303–309, which is from the first rigid body. 3_{10} helix 269–274 is from the second rigid body. Filled line, native endothiapepsin; open line, endothiapepsin–CP-69,799 complex.

same type of a conformational change accounts for about 50% of the variability in three-dimensional structures among different enzymes in the aspartic proteinase family. Further, the TLS analysis of thermal motion in the endothiapepsin–CP-69,799 crystal showed that atomic mobility is consistent with the division into two rigid bodies oscillating around the screw axes that are similar to the screw axis transforming one endothiapepsin form into the other. And finally, this rigid body movement may be

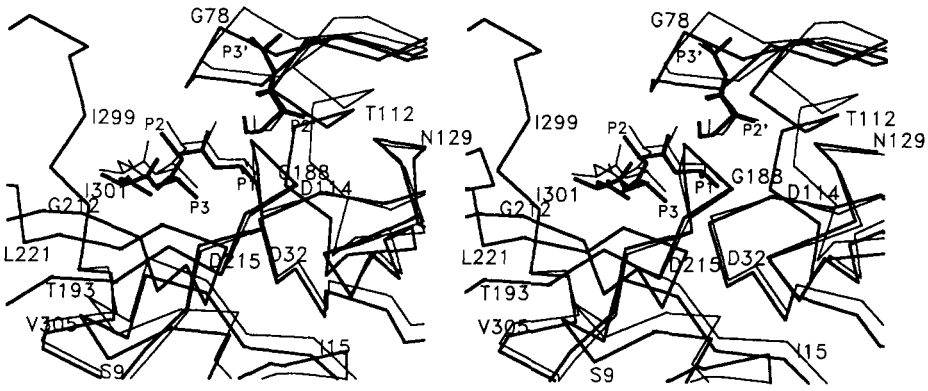

Fig. 9. Plot of the endothiapepsin active site cleft with the average atomic positions for the two inhibitor groups. The native endothiapepsin is shown in thick line. All other structures are superposed with their second rigid bodies on the second rigid body of the native endothiapepsin: the first rigid body of endothiapepsin in the CP-69,799 complex (thin line), average of positions of main chain atoms N, CA, CB, C, and O for the inhibitors bound to endothiapepsin with the native-like orientation of the rigid bodies (thick line), and the main chain average for the other group of inhibitors (thin line). Atoms of the scissile bond surrogate are missing because of the variable stereochemistry at this position.

associated with changes in the conformation and relative position of up to 1 Å of the inhibitor in the active site cleft. These differences may be even larger in a substrate that does undergo a change in the stereochemistry of the scissile peptide bond.

In endothiapepsin and pepsin, the conformational change is associated with the presence of the oligopeptide in the active site cleft, inhibitor in endothiapepsin and propart in pepsin, although other factors such as crystalline packing forces may also be responsible. Unfortunately, it is not yet possible to choose between the two alternatives for the driving force of the conformational transition: ligand binding or crystal environment. The crystallization of the BW625 complex with endothiapepsin in the nonisomorphous unit cell, yet with the native like orientation of rigid bodies, indicates that it is the occupation of the S_3 and/or S_2 pockets that induces the conformational change (Table I, Fig. 4) since the absence of the P_3 and P_2 residues is the only outstanding feature of the BW625 inhibitor. The facts that small conformational changes similar to the rigid body movement occur in all endothiapepsin complexes (Fig. 4) and that the monoclinic and hexagonal pepsin forms have the same conformation (A. Šali, unpublished) also point to the ligand binding as a likely factor contributing to the conformational change. On the other hand, the crystallization of the H261 complex in the isomorphous and nonisomorphous unit cell with native and nonnative like orientation of the rigid bodies, respectively, indicates that the conformational change is either a cause, a consequence, or an accompanying feature of a certain crystal environment.

Even if crystal packing influences the relative orientation of the two rigid bodies, the fact remains that the fold of endothiapepsin and probably of the other aspartic proteinases can exist in a delicate equilibrium between two forms differing considerably in the shape of the active site cleft. It is also a fact that the two forms are able to interchange as a result of a stimulus from the environment, possibly as a result of filling in the S_3 pocket. In this context, it is reasonable to propose a working hypothesis that this structural flexibility plays a part in the function of the aspartic proteinases. It could serve to increase both the turnover and specificity in the hydrolysis of peptide bonds.

It has been suggested before that some of the steps in the catalysis by aspartic proteinases involve a conformational change in the enzyme.[24,35,51] This was based mainly on the $\approx$ 1000-fold increase in k_{cat} and virtually no change in K_m when the P_3 residue is added to the substrate. However, the same effect could also be due to nonproductive binding of short substrates[15] or distortion of the substrate itself.[52]

The enzyme distortion hypothesis was that the P_3 residue, by occupying the S_3 pocket, uses most of its binding energy to trigger the conformational change in the enzyme, a change which decreases the activation energy of the rate-limiting step and thereby increases k_{cat}. There are several pieces of evidence which support this view. The Arrhenius plots of the hydrolysis of long substrates have a sharp break which has been shown to correlate with a break in the temperature dependence of the intensity of a circular dichroism band at 242 nm in a pepstatin complex.[53] The Arrhenius plot is linear for short substrates, as is the temperature dependence of the 242 nm circular dichroism band of the free enzyme.[53] The break in the Arrhenius plot is thus most likely due to a conformational change in the enzyme. The 242 nm band is an unusual band which is common to all three fungal aspartic proteinases as well as pepsin and originates from a structural feature that is very sensitive to raised temperature,

pH, and urea concentration (T. Hofmann, unpublished results). It is therefore associated with parts of the molecule that are fairly flexible. The intensity of this band also changes upon binding of pepstatin.

The following earlier observations are also consistent with a role of enzyme conformational changes in the function of aspartic proteinases. First, a multistep inhibitor binding occurs when the P_3 residue is present while the kinetics are simpler for inhibitors without a P_3 residue.[36] Second, pre-steady-state fluorimetric measurements indicate conformational changes.[51] And third, small nonsubstrate peptides can act as activators of the hydrolysis of poor short substrates, probably by simulating the binding in S_3 pocket, an event that is accompanied by the changes in the CD spectrum which are probably due to conformational changes.[54]

We propose that at least some of these solution experiments may be rationalized with the help of the crystallographically observed rigid body movement described in this paper.

Apart from the role that the rigid body movement in aspartic proteinases may play in the hydrolysis, there are two additional potential applications. The first one is in the design of potent inhibitors: an inhibitor that did not induce a conformational change in the enzyme would presumably bind better, since it would not spent part of its binding energy on distorting the enzyme. The second application is in modeling of the structures of as yet unknown aspartic proteinases and their zymogens: when a rigid body movement is taken into account considerably better models are obtained. These models can be used, for example, in a molecular replacement technique of X-ray crystallography.

NOTE

Two other groups independently published papers that also describe conformational changes in aspartic proteinases (Abad-Zapatero, C., Rydel, T.J., Erickson, J. Revised 2.3 Å structure of porcine pepsin: Evidence for a flexible subdomain. Proteins 8: 62–81, 1990; Sielecki, A.R., Federov, A.A., Boodhoo, A., Andreeva, N.S., James, M.N.G. Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 Å resolution . J. Mol. Biol. 214:143–170, 1990).

ACKNOWLEDGMENTS

We are indebted to Dr. Matthew Newman and Drs. Jim Remington and Jean Hartsuck for providing us with the structure of chymosin and pepsinogen, respectively, prior to their publication. A.Š. was funded by ORS Awards Scheme, Research Council of Slovenia, Merck Sharp & Dohme and J. Stefan Institute, Ljubljana. B.V. was funded by the SERC and J.B.C. by the MRC.

REFERENCES

1. Blundell, T.L., Jenkins, J., Pearl, L.H., Sewell, T. The high resolution structure of endothiapepsin. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. Berlin: Walter de Gruyter, 1985:151–161.
2. Blundell, T.L., Jenkins, J.A., Sewell, B.T., Pearl, L.H., Cooper, J.B., Wood, S.P. Veerapandian, B. X-ray analyses of aspartic proteinases. The three-dimensional structure at 2.1 Å resolution of endothiapepsin. J. Mol. Biol. 211:919–941, 1990.
3. Hsu, I.N., Delbaere, L.T.J., James, M.N.G., Hofmann, T. Penicillopepsin from *Penicillum janthinellum* crystal structure at 2.8 Å and sequence homology with porcine pepsin. nature (London) 266:140–145, 1977.
4. James, M.N.G., Sielecki, A. Structure and refinement of penicillopepsin at 1.8 Å resolution. J. Mol. Biol. 163:299–361, 1983.
5. Bott, R., Subramanian, E., Davies, D. Three-dimensional structure of the complex of the *Rhizopus chinensis* carboxyl proteinase and pepstatin at 2.5 Å resolution. Biochem. 21:6956–6962, 1982.
6. Suguna, K., Bott, R.R., Padlan, E.A., Subramanian, E., Sheriff, S., Cohen, G.E., Davies, D.R. Structure and refinement at 1.8 Å resolution of the aspartic proteinase from *Rhizopus chinensis*. J. Mol. Biol. 196:877–900, 1987.
7. Andreeva, N.S., Zdanov, A.S., Gutschina, A.E., Federov, A.A. Structure of ethanol-inhibited porcine pepsin at 2 Å resolution and binding of the methyl ester of phenylalanyldiiodotyrosine to the enzyme. J. Biol. Chem. 259:11353–11365, 1984.
8. Cooper, J.B. Unpublished results.
9. Cooper, J.B., Khan, G., Taylor, G., Tickle, I.J., Blundell, T.L. X-ray analyses of aspartic proteinases. The 3D-structure of the hexagonal crystal form of porcine pepsin at 2.3 Å resolution. J. Mol. Biol. 214:199–222, 1990.
10. Newman, M., Khan, G., Frazao, C., Tickle, I., Blundell, T., Safro, M., Andreeva, N. Personal communication.
11. Gilliland, G.L., Winborne, E.L., Nachman, J., Wlodawer, A. The three-dimensional structure of recombinant bovine chymosin at 2.3 Å resolution. Proteins 8:82–101, 1990.
12. James, M.N.G., Sielecki, A. Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 Å resolution. Nature (London) 319:33–38, 1986.
13. Remington, J., Hartsuck, J. Personal communication.
14. Sielecki, A.R., Hayakawa, K., Fujinaga, M., Murphy, M.E.P., Fraser, M., Muir, A.K., Carilli, C.T., Lewicki, J.A., Baxter, J.D., James, M.N.G. Structure of recombinant human renin, a target for cardiovascular-active drugs, at 2.5 Å resolution. Science 243:1346–1350, 1989.
15. Suguna, K., Padlan, E.A., Smith, C.W., Carlson, W.D., Davies, D. Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis*: Implications for a mechanism of action. Proc. Natl. Acad. Sci. U.S.A. 84:7009–7013, 1987.
16. Tang, J., James, M.N.G., Hsu, I.N., Jenkins, J.A., Blundell, T.L. Structural evidence for gene duplication in the evolution of the acid proteases. Nature (London) 217:618–621, 1978.
17. Antonov, V.K., Ginodman, L.M., Kapitannikov, Y.V., Barshevskaya, T.N., Gurova, A.G., Rumsh, L.D. Mechanism of pepsin catalysis: General base catalysis by the active-site carboxylate ion. FEBS Lett. 88:87–90, 1978.
18. Antonov, V.K., Ginodman, L.M., Rumsh, L.D., Kapitannikov, Y.V., Barshevskaya, T.N., Yavashev, L.B., Gurova, A.G., Volkova, L.I. Studies on the mechanism of action of proteolytic enzymes using heavy oxygen exchange. Eur. J. Biochem. 117:195–200, 1981.
19. Antonov, V.K., Chemical approaches to the mechanism of the aspartic proteinases. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. Berlin: Walter de Gruyter, 1985:263–220.
20. Silver, M.S., James, S.L.T. Enzyme catalyzed condensation reactions which initiate rapid peptic cleavage of substrates. 2. Proof of mechanism for three examples. Biochem. 20:3183–3189, 1981.
21. Blum, M., Cunningham, A. Bendiner, M., Hofmann, T. Penicillopepsin, the aspartic proteinase from *Penicillum janthinellus*: Substrate binding and intermediates in transpeptidation reactions. Biochem. Soc. Trans. 13:1044–1046, 1985.

22. Hofmann, T., Fink, A.L. Cryoenzymology of penicillopepsin. Biochem. 23:5247–5256, 1984.
23. Hofmann, T., Hodges, R.S., James, M.N.G. Effect of pH on the activities of penicillopepsin and rhizopuspepsin and a proposal for the productive substrate binding mode in penicillopepsin. Biochemistry 23:635–643, 1984.
24. Fruton, J.S. The mechanism of the catalytic action of pepsin and related acid proteinases. Adv. Enzymol. 46:1–36, 1976.
25. Pearl, L.H., Blundell, T.L. The active site of aspartic proteinases. FEBS Lett. 174:96–101, 1984.
26. James, M.N.G., Sielecki, A. Stereochemical analysis of peptide bond hydrolysis catalyzed by the aspartic proteinase penicillopepsin. Biochemistry 24:3701–3713, 1985.
27. James, M.N.G., Sielecki, A.R., Hofmann, T. X-ray diffraction studies on penicillopepsin and its complexes: The hydrolytic mechanism. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. Berlin: Walter de Gruyter, 1985:163–177.
28. Pearl, L.H. The catalytic mechanism of aspartic proteinases. FEBS Lett. 214:8–12, 1987.
29. Polgar, L. The mechanism of action of aspartic proteinases involves 'push-pull' catalysis. FEBS Lett. 1:1–4, 1987.
30. Takahashi, M., Wang, T.T., Hofmann, T. Acyl intermediates in pepsin and penicillopepsin catalyzed reactions. Biochem. Biophys. Res. Commun. 57:39–46, 1974.
31. Newmark, A.K., Knowles, J.R. Acyl- and amino-transfer routes in pepsin-catalyzed reactions. J. Am. Chem. Soc. 97:3557–3559, 1975.
32. James, M.N.G., Sielecki, A.R., Salituro, F., Rich, D.H., Hoffmann, T. Conformational flexibility in the active sites of aspartyl proteinases revealed by a pepstatin binding to penicillopepsin. Proc. Natl. Acad. Sci. U.S.A. 79:6137–6142, 1982.
33. Hallett, A., Jones, D.M., Atrash, B., Szelke, M., Leckie, B., Beattle, S., Dunn, B.M., Valler, M.J., Rolph, C.E., Kay, J., Foundling, S.I., Wood, S., Pearl, L.H., Watson, F.E., Blundell, T.L. Inhibition of aspartic proteinases by transition state substrate analogues. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. Berlin: Walter de Gruyter, 1985:467–478.
34. Blundell, T.L., Cooper, J., Foundling, S.I. On the rational design of renin inhibitors: X-ray studies of aspartic proteinases complexed with transition-state analogues. Biochemistry 26:5585–5590, 1987.
35. Hofmann, T., Allen, B., Bendiner, M., Blum, M., Cuningham, A. Effect of secondary substrate binding in penicillopepsin: Contributions of subsites S_3 and S_2 to k_{cat}. Biochemistry 27:1140–1146, 1988.
36. Rich, D.H., Bernatowicz, M.S. Synthesis of analogues of the carboxyl protease inhibitor pepstatin. Effect of structure in subsite P_3 on inhibition of pepsin. J. Med. Chem. 25:791–795, 1982.
37. Šali, A., Veerapandian, B., Cooper, J.B., Foundling, S.I., Hoover, D.J., Blundell, T.L. High-resolution X-ray diffraction study of the complex between endothiapepsin and an oligopeptide inhibitor: Analysis of the inhibitor binding and description of the rigid body shift in the enzyme. EMBO J. 8:2179–2188, 1989.
38. Cooper, J.B., Foundling, S.I., Blundell, T.L., Boger, J., Jupp, R.A., Kay, J. X-ray studies of aspartic proteinase-statine inhibitor complexes. Biochemistry 28:8596–8602, 1989.
39. Richards, F.M., Kundrot, C.E. Identification of structural motifs from protein coordinate data: Secondary structure and first level supersecondary structure. Proteins 3:71–84, 1988.
40. Šali, A., Blundell, T.L. The definition of topological equivalence in homologous and analogous structures: A procedure involving comparison of local properties and structural relationships through simulated annealing and dynamic programming. J. Mol. Biol. 212:403–428, 1990.
41. Muirhead, H., Cox, J., Mazzarella, L., Perutz, M.F. Structure and function of haemoglobin: III. A three-dimensional Fourier synthesis of human deoxyhaemoglobin at 5.5 Å resolution. J. Mol. Biol. 28:117–156, 1967.
42. Schomaker, V., Trueblood, K.N. On the rigid-body motion of molecules in crystals. Acta Crystallogr. B24:63–76, 1968.
43. Howlin, B., Moss, D.S., Harris, G.W. Segmented anisotropic refinement of bovine ribonuclease A by the application of the rigid-body TLS model. Acta Crystallogr. A45:851–861, 1989.
44. Ibers, J.A., Hamilton, W.C. "International Tables for X-Ray Crystallography," Vol. IV. Birmingham: Kynoch Press, 1974:316.
45. Haneef, I., Moss, D.S., Stanford, M.J., Borkakoti, N. Restrained structure-factor least squares refinement of protein structures using a vector processing computer. Acta Crystallogr. A41:426–433, 1985.
46. Šali, A. Ph.D. Thesis. University of London, 1991.
47. Davis, J.C. In: "Statistics and Data Analysis in Geology." New York: Wiley, 1973:473–532.
48. Kabsch, W., Sander, C. Dictionary of protein secondary structure. Pattern recognition of hydrogen-bonded and geometrical features. Biopolymer 22:2577–2637, 1983.
49. Richmond, T.J., Richards, F.M. Packing of α-helices: Geometrical constraints and contact areas. J. Mol. Biol. 119: 537–555, 1978.
50. Chothia, C., Lesk, A.M. Helix movements in proteins. TIBS 10:116–120, 1985.
51. Fruton, J.S. Fluorescence studies on the active sites of proteinases. Mol. Cell. Biol. 32:105–114, 1980.
52. Pearl, L.H. The extended binding cleft of aspartic proteinases and its role in peptide hydrolysis. In: "Aspartic Proteinases and Their Inhibitors." Kostka, V., ed. Berlin: Walter de Gruyter, 1985:189–195.
53. Allen, B., Blum, M., Cunningham, A., Tu, G.-G., Hofmann, T. A ligand-induced, temperature dependent conformational change in penicillopepsin: Evidence fron non-linear Arrhenius plots and from circular dichroism studies. J. Biol. Chem. 265:5060–5065, 1990.
54. Wang, T.T., Dorrington, K.J., Hofmann, T. Activation of the action of penicillopepsin on leucyl-tyrosyl-amide by a non-substrate peptide and evidence for a conformational change associated with a secondary binding site. Biophys. Res. Commun. 57:865–869, 1974.
55. Cooper, J.B., Foundling, S.I., Blundell, T.L., Arrowsmith, R.J., Harris, C.J., Champness, J.N. In: "Topics in Medicinal Chemistry," Leeming, P.R., ed. Royal Society Special Publication 65, 1988:308–313.
56. Foundling, S.I., Cooper, J.B., Watson, J., Cleasby, F.E., Pearl, L.H., Sibanda, B.L., Hemmings, A., Wood, S.P., Blundell, T.L., Valler, T.L., Norey, Kay, J., Boger, J., Dunn, B.M., Leckie, B.J., Jones, D.M., Atrash, B., Hallett, A., Szelke, M. High resolution X-ray analyses of renin inhibitor–aspartic proteinase complexes. Nature (London) 327:349–352, 1987.
57. Cooper, J.B., Foundling, S.I., Watson, F.E., Sibanda, B.L., Blundell, T.L. Inhibitors of aspartic proteinases and their relevance to the design of antihypertensive agents. Biochem. Soc. Trans. 15:751–754, 1987.
58. Cooper, J.B., Foundling, S.I., Hemmings, A., Blundell, T.L., Jones, D.M., Hallett, A., Szelke, M. The structure of a synthetic pepsin inhibitor complexed with endothiapepsin. Eur. J. Biochem. 169:215–221, 1987.
59. Bailey, D., Cooper, J.B., Hemmings, A. al Karadaghi, S., Quayle, W., Veerapandian, B. Unpublished results.
60. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanovichi, T., Tasumi, M. The protein data bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
61. Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F., Weng, J. Protein data bank. In: "Crystallographic Databases—Information, Content, Software Systems, Scientific Applications. Allen, F.H., Bergerhoff, G., Sievers, R., eds. Cambridge: Data Commission of the International Union of Crystallography, 1987:107–132.
62. Davies, D.R. The structure and function of the aspartic proteinases. Annu. Rev. Biophys. Biophys. Chem. 19:189–215, 1990.