Volume 271, number 1,2, 9–13 FEBS 08889 October 1990

Hypothesis

Could domain movements be involved in the mechanism of trypsin-like serine proteases?

M.J. Dufton

Department of Pure & Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow G1 1XL, UK

Received 11 July 1990

It is hypothesised that the characteristic twin domain structure of serine proteases permits important allosteric responses in the molecule when peptide and protein substrates bind. Such movement would be ideal for stressing the scissile bond in the substrate, thereby making the task of hydrolysis substantially easier. The control of the domain movement can be closely associated with substrate binding, via the N- and C-terminal regions of the enzyme. The hypothesis also suggests that certain inhibitory peptides exert their effect by binding without inducing the domain movement.

Domain movement; Serine protease

1. INTRODUCTION

Many enzymes possess distinct domains in their superstructure and in several instances, relative movement of the domains is thought to be important for the entrance and exit of substrates and products [1], or the stressing of bonds that are required to be broken [2]. The trypsin-like serine proteases, a large family of peptide and protein cleaving enzymes, are superficially promising candidates for one to suspect such a contribution to their mechanism. To begin with they have two very marked domains in their structures. In a chain of approximately 240 residues, these two domains comprise approximately residues 1–120 and 121–240 (Fig. 1). Secondly, perhaps the largest single obstacle to be overcome by the enzyme in its catalytic mechanism is the resonance energy stabilisation of the peptide bond (approx. 20 kcal/mol). The severity of conditions for the non-enzymic hydrolysis of peptide bonds is ample testimony to this. These enzymes are thought to surmount this obstacle by contriving to distort the peptide bond to be broken out of planarity [3], thereby producing an ester-like character and consequent increased susceptibility to hydrolysis. In view of the precedents seen in other enzyme types, a reasonable possibility might be that the necessary peptide bond distortion is achieved mechanically in the serine proteases via a relative movement of the two large domains.

Correspondence address: M.J. Dufton, Department of Pure and Applied Chemistry, University of Strathclyde, 295, Cathedral Street, Glasgow G1 1XL, UK

This latter suggestion runs contrary to accepted wisdom, largely because there is such a wealth of structural detail on these enzymes as complexed to a variety of peptide and protein inhibitors. These complexes are widely regarded as accurate models for the interaction with true substrates [4] and it is clear from them that the domains have not moved upon complexation as compared to the free enzyme. Consequently, the possibility of a large scale allosteric element to the mechanism is not entertained.

However, as will be described below, there is reason to believe that the differences between a long-lived serine protease/inhibitor complex and a transient serine protease/substrate complex could be far more marked than the subtleties currently envisaged.

2. IS THE SITUATION FAVOURABLE FOR DOMAIN MOVEMENT TO BE USEFUL?

In complexes such as those of bovine pancreatic trypsin inhibitor with kallikrein [5], and turkey ovomucoid inhibitor with chymotrypsin [6], the inhibitor protein is bound across the two domains with the would-be 'scissile' bond forming the point of cross-over from one domain to the other. As shown in Table I, virtually all of the inhibitor molecule bound N-terminal to the 'scissile' bond interacts with the 121–240 domain while most interactions C-terminal to this bond occur with the 1–120 domain. In the usual parlance adopted for these enzymes [7], the S_1, S_2, S_3, etc. subsite components are largely located within the 121–240 domain while the S_1', S_2', S_3' etc. subsite components are largely

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

located within the 1–120 domain. Thus, in terms of the essential requirements of a domain movement-induced bond stressing mechanism, the basic situation is favourable – the substrate/inhibitor is bound across both domains in such a way that if movement did occur, the 'scissile' bond would be the natural focus of any stress that resulted. Nonetheless, the fact remains that the cross domain situation observed in the inhibitor complexes has not altered domain position. Thereby, the domain movement hypothesis can only remain tenable if it is considered that the true Sn or Sn′ subsites for substrates have not been correctly inferred. In other words, have the inhibitors that bridge the domains bound 'wrongly' to one of them as part of a modus operandi designed to prevent the stressing mechanism?

3. SUBSTRATE INTERACTIONS IN THE 1–120 DOMAIN

From the extensive detail discovered about the hydrolysis of model substrates and the character of peptide inhibitors, the only feasible suggestion is that inhibitors like bovine trypsin inhibitor and turkey ovomucoid inhibitor are giving a false impression of both the nature and extent of Sn′ interactions (i.e., interactions that occur with the 1–120 domain). As also shown in Table I, the degree of interaction in the S_1, S_2, S_3 etc. subsites is much more extensive and exclusive with respect to the twin domain structure than is the S_1', S_2', S_3' etc interaction, especially with regard to the S_1 subsite. The importance of Sn′ interactions is often overlooked because they are clearly not essential for the hydrolysis of ester and primary amide substrates. However, when dealing strictly with peptides and proteins, there is a body of work that shows that Sn′ interactions are essential if a peptide bond proper is to be successfully dealt with [8,9]. A peptide/protein substrate is required to have at least two residues beyond the scissile bond (i.e., S_1' and S_2' occupancy) and rates of hydrolysis increase with successive site occupancy (i.e., S_3' and beyond). Significantly, aminopeptidase and carboxypeptidase actions require completely different mechanisms – the serine proteases in question are strictly endopeptidases.

There is thus little doubt that successful cleavage of a peptide bond depends on the substrate binding properly to both the 121–240 domain (Sn sites) and the 1–120 domain (Sn′ sites). If the Sn′ occupancy extrapolated from the inhibitor complexes is misleading in extent and location, what is the alternative?

Table I

Inhibitor-enzyme intermolecular contacts

	Complex			
	BPTI/kallikrein		Ovomucoid/ chymotrypsin	
Inhibitor residues	P_6–P_1	P_1'–P_3'	P_6–P_1	P_1'–P_3'
Enzyme domain 1 (residues 1–120)	4%	24%	6%	22%
Enzyme domain 2 (residues 121–240)	58%	14%	57%	15%

Intermolecular contacts expressed as percentage of total number of contacts. P_6–P_1 are the six residues N-terminal to the 'scissile' bond, while P_1'–P_3' are the three residues C-terminal to this bond. Data from Chen and Bode [5] (kallikrein), and Fujinaga et al. [6] (chymotrypsin)

4. THE UNUSUAL SYMMETRY OF SERINE PROTEASES

One of the most remarkable and enigmatic features of the twin domain structure of the serine proteases is the presence of an approximate two-fold symmetry with respect to the domain architecture. As a consequence, the chain fold of one domain can be superimposed directly on to that of the other after a rotation through 180°. What makes this symmetry so unusual is that if the superimposition is carried out, the respective polypeptide chains of the two domains are found to run in opposite directions. In other words, the chain fold is palindromic in 3 dimensions either side of the crossover from one domain to the other (Fig. 1). This curiosity was noticed with the advent of the crystal structure of elastase [10] and was explored by later authors [11], but has not been fully explained. As one of these works showed [11], the strength of the approximate two-fold symmetry is best appreciated from α-carbon interatomic distance matrices, which contain a very obvious 'mirror plane' in the matrix pattern. It was also shown that the folding palindromy was not matched by

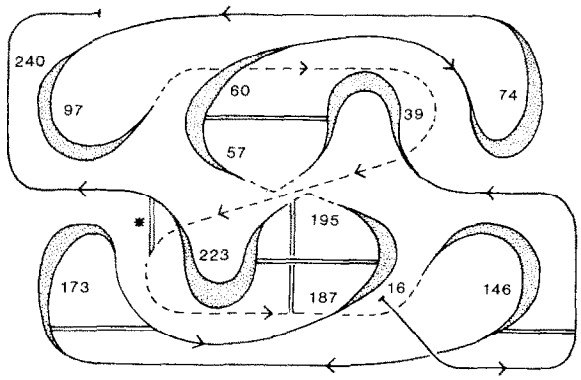

Fig. 1. A schematic representation of the chain fold in the trypsin-like serine proteases, drawn to emphasise the two-fold domain symmetry present. A view down on to the active site is presented, the broken lines denoting portions of the chain furthest from the viewer while the shaded portions are those closest. The double lines are disulphide bonds (that marked with an asterisk is absent in some variants). Note that this orientation is essentially unchanged in Fig. 2.

sequence palindromy. The presence of this symmetry element in the molecular fold creates a major connundrum, for it is not possible with conventional wisdom to envisage a 'deliberate' evolutionary path that could have arrived at such a situation. Consequently, this symmetry is considered by some to be fortuitous, or accidental [12], and therefore of no real relevance to any catalytic mechanism that has arisen in the molecule.

However, whether or not the origins of this symmetry can be satisfactorily accounted for, its presence in the molecule means that it is worthy of serious consideration. When a representative α-carbon interatomic distance map of a serine protease is considered, the approximate two-fold domain symmetry in the molecule means that for any point or locality in one domain, an analogous point or locality can be identified in the opposite domain by means of the 'mirror plane' present in the matrix pattern. In view of the lack of sequence palindromy in the two domains, the analogy is in terms of the surrounding main chain architecture, not chemical character. Thus for any substrate/inhibitor binding site found in one domain, there is a similarly shaped 'twin' in the opposite domain which differs in residue composition and polypeptide chain direction. This only applies fully to the secondary binding sites because the catalytic site itself (i.e., the triad of histidine 57, aspartate 102 and serine 195, as they are generally numbered) is constituted at the domain interface from residue contributions from both sides. The Sn and Sn′ subsites, in contrast, tend to be wholly in one domain or the other.

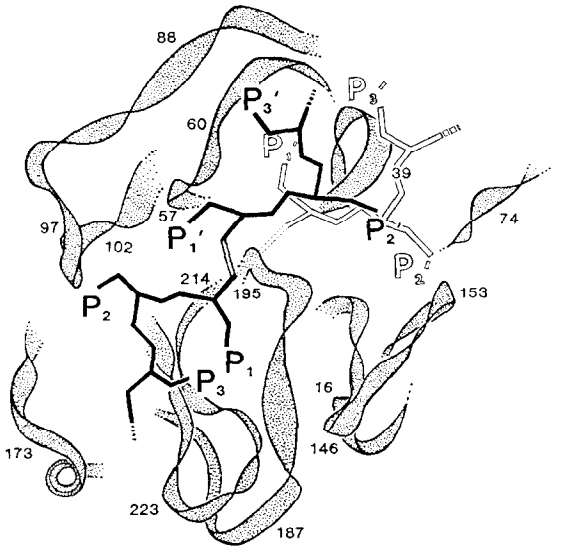

Fig. 2. A simplified representation of inhibitor binding, showing the enzyme in a computer-generated ribbon form (only those segments nearest the viewer are shown) and the inhibitor as a carbon-nitrogen backbone without side chains (solid lines). It is based on the chymotrypsin/ovomucoid complex elucidated by Fujinaga et al. [6] (see Fig. 16 in this reference). P_1–P_3 are the side chains on the N-terminal side of the 'scissile' bond (double lined) and P_1'–P_3' are on the C-terminal side. The alternative positioning of residues P_1'–P_3' (white line) shows where binding would have to occur to complement the domain symmetry. Note the stretched 'scissile' bond (dotted lines).

5. THE SYMMETRY AND SUBSTRATE/-INHIBITOR BINDING

If the Sn and Sn′ binding sites and the 'catalytic triad' are appraised with respect to the approximate two-fold chain symmetry, it can be noted that: (a) The two most directly involved catalytic residues, histidine 57 and serine 195, occupy almost equivalent positions in their respective domain folds. In other words, if the 1–120 and 121–240 domains are superimposed according to the approximate two-fold symmetry, histidine 57 is approximately superimposed on serine 195; and (b) by way of contrast, the Sn and Sn′ subsites as currently identified are not superposed by virtue of such a domain superimposition. Thus, inhibitors that bind across the domains do not do so in a way that complements the two-fold symmetry – the Sn′ interactions are less extensive and appear misplaced with respect to the Sn interactions.

Fig. 2 is a representation of an enzyme inhibitor complex which shows the actual binding mode and where the P′n residues would have to bind to complement the two-fold domain symmetry. If the Pn residues are regarded as optimally bound, it is clear that symmetrical binding could only be achieved if either the domains were rotated with respect to one another and moved closer, or the 'scissile' bond were stretched and distorted. It is this that leads to the suspicion that by binding wrongly/inadequately to the 1–120 domain, inhibitors prevent a mechanism in which the domains adjust to bind a true substrate and contrive a situation in which the scissile bond becomes the focus of constraint against subsequent relaxation. A stressed peptide bond restraining the motion of two large domains is clearly a situation favourable for the loss of resonance energy and assisted bond cleavage.

6. HOW MIGHT DOMAIN MOVEMENT BE CONTROLLED?

The viability of such a hypothesis also depends on whether a suitable allosteric control can be envisaged. An inspection of the chain fold on purely mechanical principles shows that the major restraints on relative domain motion are likely to be the main chain termini, each of which protrudes from its parent domain to be bedded on the outer face of the other (Fig. 1). The possibility that these termini effectively constitute the 'locks' against domain movement has been strongly intimated in theoretical analyses by Segawa and Richards [13], albeit in another context. In the native enzyme, they detected no chain flexibility near residue 120, where the main chain crosses from one domain to the

other (and which lies directly under the scissile bond location). However, when the N and C termini sections were removed for their analysis, marked flexibility did appear near residue 120. Thereby, a prerequisite of domain movement is a change in disposition of the N and C termini. A simple bringing together and/or slewing of the domains (as induced by a substrate) might achieve this, but there is more suggestive evidence from studies on the activation mechanisms of these enzymes. The crystal structures of trypsinogen, trypsin and bovine pancreatic trypsin inhibitor complexed to trypsinogen reveal an allosteric interplay between the disposition of the N-terminus of the enzyme and the conformational status of the S_1 binding subsite [14,15]. In the trypsinogen-trypsin transition, removal of some N-terminal residues allows the new N-terminus to occupy a special binding site in the 121–240 domain, the consequence of the binding being that the nearby S_1 substrate binding pocket is induced to adopt its receptive form [14]. The two-fold symmetry in these enzymes also means that the C-terminal α-helix resides in a similar position relative to the 'symmetry-twin' of the S_1 pocket. When bovine pancreatic trypsin inhibitor binds to trypsinogen, conformational changes in the enzyme ensue which are very similar to those that occur when N-terminal activation occurs [15]. Aspartate 194 is seen to rotate and the so-called 'activation domain' becomes rigid, forming both the S_1 specificity pocket and the binding cleft for the enzyme's N-terminus. Hence, not only can the N-terminal of the enzyme change the character of the S_1 pocket, but also binding in the S_1 pocket can change the character of the N-terminal binding cleft. This potential two-way communication between the S_1 pocket and the N-terminus raises a very interesting question, namely: Could full binding of the substrate in the S_1 and S_1' pockets influence the positions of the enzyme's nearby N and C termini and thereby help to precipitate domain movements which ultimately contribute to cleavage? It is implicit in the overall mechanism proposed here that the substrate is responsible for bringing about domain movement and consequently its own breakdown. An 'induced fit' aspect to the mechanism is therefore required, so any suggestion of an allosteric interplay between substrate side chain binding and the induction of domain movement is highly relevant to the issue.

7. SUMMARY OF HYPOTHESIS

The general mechanistic hypothesis that emerges from the above discussion is one in which domain movement is paramount. It is suggested that in its uncomplexed form, the enzyme has its domains 'locked' together by virtue of the enzyme's N and C termini interacting with their binding sites on the outer faces of the two domains. Upon proper substrate binding, however, there are two factors which could lead to a disengagement of these locks. Firstly, S_1 and S_1' occupancy by the substrate could allosterically alter the nature of the enzyme's N and C terminal binding sites and reduce their affinity for the termini themselves. Secondly, the substrate's desire to bind symmetrically across the two domains could bring the domains together. Because of the chain fold, this could transmit a direct mechanical impetus to the N and C termini of the enzyme to disengage. Both factors could augment each other, the net consequence being that the substrate induces its own fit and its scissile bond becomes one of only two peptidic links restraining the relative motion of the two large domains. However, whereas the enzyme's own peptidic domain connection describes a spring-like zig-zag (Fig. 1), the substrate's scissile bond would be held diagonally taught between the domains, flanked by histidine 57 and serine 195. Thereby, any subsequent desire of the two domains to move apart focusses stress on the scissile bond, and were the stress to distort the peptide bond geometry enough to significantly lessen the double-bond character of the carbon-nitrogen bond, successful attack by the 'catalytic' residues would be greatly facilitated. As soon as the scissile bond breaks, the domains are free to move apart. By a reversal of the previous pathway, a mechanically forced re-engagement of the enzyme's N and C termini in their normal binding sites could adversely affect the chemical hospitability of the S_1 and S_1' binding pockets to the extent that the products might be encouraged to leave. With the dissociation of the products, the enzyme is restored to its original status with the domains locked. In contrast, inhibitor peptides/proteins may fail to induce the stressing mechanism because there is insufficient and/or incorrect binding to one of the domains to 'unlock' the domain movement. Consequently, the enzyme remains more or less in its resting state and the 'stress factor' contribution to the mechanism is lost. The efficiency of inhibitor hydrolysis is thus drastically reduced.

8. CONCLUSION

The arguments presented herein suggest that a significant part of the serine protease mechanism lies concealed behind an experimental 'blindspot' and, unfortunately, it may be some time before the extremely transient events envisaged can be properly tested in the laboratory. Proof one way or the other will require protein engineering experiments on the enzyme itself or an exhaustive computer simulation of all induced-fit possibilities between substrate and enzyme. However, by accepting that the potential for domain movement exists in the fold itself, and by following the symmetry aspects to their logical conclusion, there seems to be good prima facie case for an investigation along these lines.

Acknowledgement: I thank Dr P. Bladon, University of Strathclyde for helpful discussion.

REFERENCES

[1] Karplus, M. and McCammon, J.A. (1983) Annu. Rev. Biochem. 53, 263–300.
[2] Sali, A., Veerapandian, B., Cooper, J.B., Foundling, S.I., Hoover, D.J. and Blundell, T.L. (1989) EMBO J. 8, 2179–2188.
[3] Delbaere, L.T.J. and Brayer, G.D. (1985) J. Mol. Biol. 183, 89–103.
[4] Polgar, L. (1987) in: Hydrolytic Enzymes (Neuberger, A. and Brocklehurst, K. eds) pp. 159–200, Elsevier, Amsterdam, New York.
[5] Chen, Z. and Bode, W. (1983) J. Mol. Biol. 164, 283–311.
[6] Fujinaga, M., Sielecki, A.R., Read, R.J., Ardelt, W., Laskowski, M. and James, M.N.G. (1987) J. Mol. Biol. 195, 397–418.
[7] Schecter, I. and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157–162.
[8] Morihara, K. and Oka, T. (1973) Arch. Biochem. Biophys. 156, 764–771.
[9] Fersht, A.R., Blow, D.M. and Fastrez, J. (1973) Biochemistry 12, 2035–2041.
[10] Shotton, D.M. and Watson, H.C. (1970) Nature 235, 811–816.
[11] Nishikawa, K. and Ooi, T. (1974) J. Theor. Biol. 43, 351–374.
[12] McLachlan, A.D. (1979) J. Mol. Biol. 128, 49–79.
[13] Segawa, S. and Richards, F.M. (1988) Biopolymers 27, 23–40.
[14] Fehlhammer, H., Bode, W. and Huber, R. (1977) J. Mol. Biol. 111, 415–438.
[15] Bode, W., Schwager, P. and Huber, R. (1978) J. Mol. Biol. 118, 99–112.