International Journal of Biotechnology and Biochemistry
ISSN 0973-2691 Volume 6 Number 2 (2010) pp. 231–258
© Research India Publications
http://www.ripublication.com/ijbb.htm

Protease-Inhibitor Interactions – A structural Insight

Aparna Laskar[1]*, C.N. Mandal[1] and Aniruddha Chatterjee[2#]

[1]Indian Institute of Chemical Biology (CSIR Unit, Govt. Of India), Jadavpur, Kolkata-700032, West Bengal, India
Correspondence author E-mail: aparnalaskar@gmail.com
[2]M.Sc Biotechnology, School of Biotechnology, Chemical & Biomedical Engineering, VIT University, Vellore-632014, Tamilnadu, India
[#]Presently at Department of pathology, Dunedin School of Medicine, University of Otago, New Zealand

Abstract

Proteases and their specific inhibitors are ubiquitously distributed in the animal, plant and micro organism kingdoms, and play key regulatory roles in many biological processes, including the blood coagulation system, the complement cascade, apoptosis and the hormone processing pathways. In present work we have performed structural analysis of the interactions of proteases with their inhibitors in known structures of the proteases-inhibitors complexes. Trypsin and thrombin, the two best studied proteases are the focal point of the study. We have formulated a mechanistic structural perspective about the properties that define the specificity of the binding interface between these inhibitors and serine proteases. We have interpreted the nature of interaction of serine proteases with their inhibitors at atomic level and also analyzed energy perspectives of these interactions. Finally, we compared and found our modeled proteases of pathogenic organisms have almost no homology with the proteases of human genome.

Keywords: Trypsin; Proteases; Inhibitor; Interaction; Specificity; Energy

Introduction

Proteolytic enzymes (best termed peptidases) are essential for the survival of all kinds of organisms, and are encoded for by approx. 2% of all genes (Barrett *et al*., 2001). Despite their life-giving functions, enzymes that break down proteins are potentially very damaging in living systems, so their activities need to be kept strictly under control. Several distinct mechanisms exist for the control of excessive peptidase

activity, important amongst which are the interactions of the enzymes with proteins that inhibit them. It is likely that all of the proteins considered have the potential to attenuate the activities of peptidases both *in vitro* and *in vivo* by the formation of complexes with the enzymes. Valuable proposals have been made as to how one can assess the physiological relevance of an inhibitor (Turk *et al*., 2002). The scientific study of the peptidase inhibitors is nearly as old as that of the peptidases themselves. Hundreds of protein inhibitors of peptidases are now known and they are the subjects of thousands of research communications. The research is driven by the many potential applications of knowledge about the inhibitors in medicine, agriculture and biotechnology. At the most fundamental level, an understanding of the mode of interaction of protein inhibitors with enzymes may suggest novel approaches to the design of synthetic inhibitors for use as drugs. Many naturally occurring inhibitors, such as the anticoagulant hirudin, are being used as the basis of engineered proteins for injection in their own right (De Filippis *et al*., 2002). There are a number of inherited diseases that are attributable to abnormalities in peptidase inhibitors. These include forms of emphysema, epilepsy and hereditary angioneurotic oedema. Netherton syndrome (Lomas *et al*., 2001, Ritchie, B. C 2003, Lehesjoki.A.E 2003, Bitoun *et al*., 2002) such diseases may be susceptible to treatment with the inhibitors administered as drugs, with synthetic inhibitors that take over their function, or with the natural inhibitors made available by gene therapy. Excessive proteolytic activities may well contribute to a number of disease conditions and, again, gene therapy to introduce inhibitors is under consideration (Krol *et al*., 2003 and McKay *et al*., 2003). In agriculture, genetically modified crop plants expressing inhibitors of the digestive enzymes of their insect pests are already under study (Telang *et al*., 2003). This active field of research generates a rapid flow of information, but the storage and retrieval of all the new information that is being obtained about the peptidase inhibitors are handicapped by difficulties of nomenclature.

In what was perhaps the most significant review that has been written on the peptidase inhibitors, (Laskowski, M., Jr. and Kato, I.1980) deplored the confusion of nomenclature that existed in the field in 1980. They pointed out that inhibitors are commonly discovered by their activity against readily available enzymes, most commonly trypsin, chymotrypsin or subtilisin, and then are named after the source organism or tissue, as '*Streptomycin* subtilisin inhibitor' or 'pancreatic trypsin inhibitor'. Such names give no clue to the relationships of the inhibitors, and make it difficult to know whether information that is available about the mechanism of action of one inhibitor can correctly be applied to another. It was evident to (Laskowski, M., Jr. and Kato, I.1980) that peptidase inhibitors could best be classified in their homologous families, but the sequence information then available allowed only about a dozen families to be recognized. The names used for peptidase inhibitors have not improved since 1980, but there is now a wealth of sequence data for these proteins, and the time seems right to make a new attempt at a systematic classification of them.

Clans of Inhibitors

A clan of peptidase inhibitors contains all the modern-day inhibitors that have arisen from a single evolutionary origin of inhibitors. It is a group of inhibitors in one or more families that show evidence of their evolutionary relationship by their similar tertiary structures. Each clan of inhibitors has a two-letter identifier, in which the first letter is "I".

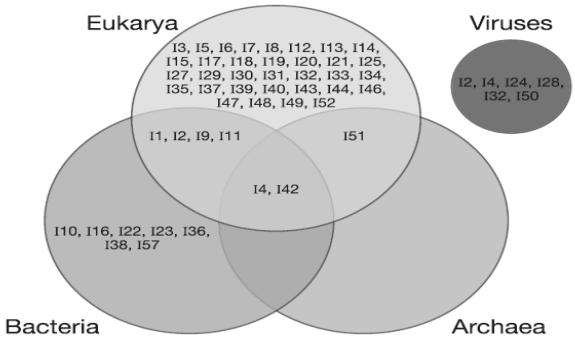


Figure 1: Distribution of inhibitor families throughout the super kingdoms of cellular organisms.

Specificity of serine protease-inhibitor interactions

Trypsin and chymotrypsin are both serine proteases. The two enzymes have high sequence identity and their tertiary structures are very similar. In the chymotrypsin index, His-57, Asp-102, and Ser-195 form the catalytic triad, residues 189–195, 214–220, and 225–228 form the primary substrate-binding pocket called S1 binding pocket (Ma *et al*., 2005). Residues 185–188 and 221–224 form two loops near the S1 pocket, called L1 and L2, respectively. Catalytic mechanisms of these two proteases are similar, but their substrate specificities are different. Trypsin favors basic residues like lysine and arginine; chymotrypsin favors aromatic residues like phenylalanine, tyrosine, and tryptophan. The S1 binding pocket in trypsin and chymotrypsin are almost identical in primary sequences and backbone tertiary structures. An important difference is that residue 189 is a negatively charged Asp in trypsin and a polar Ser in chymotrypsin. This residue lies at the bottom of the S1 binding pocket and determines different S1 pocket chemical properties. This difference was once used to explain the different specificity of trypsin and chymotrypsin. But the mechanism is not that simple other residues are also involved in the specificity of enzyme-inhibitor interactions (Ma *et al*., 2005).

Residue Asp-189 plays an important dual role in trypsin; it defines the primary specificity for Arg side chains. This role is shared by other proteases with trypsin-like specificity. Replacement of Asp-189 with Ala, Asn, Glu, and Ser drastically reduces the specificity toward substrates carrying Arg or Lys at P1, whereas it has little or no effect toward the hydrolysis of substrates carrying Phe at P1. These findings confirm the important role of Asp-189 in substrate recognition by trypsin like proteases

(Prasad *et al*., 2004). Specificity of thrombin-like serine proteases is usually categorized in terms of the P1-S1 interaction. The S1 site is a pocket adjacent to Ser-195, formed by residues 189-192, 214-216, and 224-228. Specificity pocket is usually determined by the residues at positions 189, 216, and 226 (Perona, J. J. and Craik, C. S. 1995, Czapinska, H. and Otlewski, J.1999). For example, the specificity of chymotrypsin correlates with the hydrophobicity of the P1 residue, with P1-Phe preferred over Ala (Knowles 1965). The combination of Ser-189, Gly-216, and Gly-226 create a deep hydrophobic pocket in thrombin that accounts for this specificity (Blow, D. M. 1997). Asp-189, Gly-216, and Gly-226 create a negatively charged S1 site that accounts for trypsin's specificity for substrates containing Arg or Lys at P1 (Huber *et al*., 1974). Elastase prefers substrates with small aliphatic residues at P1; the S1 site of elastase is smaller than the S1 sites of chymotrypsin and trypsin due to the presence of Val-216 and Thr-226 (Shotton, D. M. and Watson, H. C. 1970).

The active site of thrombin is occluded by the B and C insertion loops, which impede docking of macromolecular substrates and inhibitors to the active site pocket. These insertion loops are unique to thrombin. In the crystal structure of thrombin, Tyr60a, Pro60b, Pro60c, and Trp60d of the B-insertion loop form a lid over the S2 specificity pocket of thrombin. There are also several variant residues in the extending binding pockets of thrombin and factor Xa that can influence specificity of these enzymes in their reactions with antithrombin. Residue 192, at the base of the active site pocket, is another variant residue that is known to influence the S3 specificity of thrombin, factor Xa, and other coagulation proteases. The critical role of Glu-192 in restricting the reactivity of thrombin with the serpin, α_1-antitrypsin, and the Kunitz inhibitors, bovine pancreatic trypsin inhibitor and tissue factor pathway inhibitor is well studied (Rezaie, A. R.1998).

A virus (HAV) 3C peptidase was the first structure identified for a viral 3C enzyme that exhibited the three-dimensional fold of the chymotrypsin family of serine peptidases but had a cysteine sulfur atom instead of the serine oxygen as the nucleophile. The structure of HAV 3C was unusual in that the Asp residue expected as the third member of the catalytic triad did not interact with the general base His (James, M. N 2006**).** The 3Cpro is distinguished from most other proteases by the fact that it has a cysteine nucleophile but with a chymotrypsin-like serine protease folding. This unique protein structure together with its essential role in viral replication made the 3Cpro an excellent target for antiviral intervention. (Wanga, Q. M. and Chen, S. H. 2007)

Computational methods are needed to exploit the structural information to understand specific molecular recognition events and to elucidate the function of the target macromolecule. This information should ultimately lead to the design of small molecule ligands for the target, which will block/activate its normal function and thereby act as improved drugs.

Materials and Methods

In this work we have focused on the studies of the interactions of proteases with their inhibitors in known structures of the proteases-inhibitors complexes. For this purpose

we focused our attention on two well-studied proteases, trypsin and thrombin, with large number of known structures of the complexes. The protein data bank files were downloaded as listed in Table.I for thrombin and in Table. II for trypsin.

Table I: The list of experimental structures of the complexes of thrombin with the inhibitors, which were used for superposition and analysis of enzyme-inhibitor interactions.

PDB CODE	Source	Inhibitor
1A2C	*Homosapiens*	Aeruginosin298-A
1A4W	*Homosapiens*	2EP:2-ethylpiperidinekth2-ketothiazole
1A5G	*Homosapiens*	BIC:4-amino-4-benzyl-5-oxo-1,6-diazabicyclo[4.3.0] nonane-7-carbaldehyde
1AY6	*Homosapiens*	HHO:[1-(hydroxymethyleneamino)-8-hydroxy-octane]
1BA8	*Homosapiens*	Hirugen, CVS1578
1CA8	*Homosapiens*	3GA:[3-piperidyl-n-guandino-l-alaninal]
1FPC	*Homosapiens*	EPI:[4-ethylpiperidine]
1H8D	*Homosapiens*	HirudinI, Lepirudin
1H8I	*Homosapiens*	HirudinI, LepirudinTys
1LHC	*Homosapiens*	DP7[AC(D) phe-pro-boro-arg-OH]
1LHD	*Homosapiens*	DI2[AC-(D) phe-pro-borolys-OH]
1LHG	*Homosapiens*	DI5[AC-(D) phe-pro-borohomoornithine-OH]
1PPB	*Homosapiens*	Chloro Methyl Ketone
1TMT	*Homosapiens*	CGP50,856 (synthetic)
1TOM	*Homosapiens*	MIN: Methyl-Phe-Pro-Amino-Cyclohexylglycine
1UMA	*Homosapiens*	IN2: [N,N-Dimethylcarbamoyl-Alpha Azalysine]
1UVS	*Homosapiens*	BM12.1700

Table II: The list of experimental structures of the complexes of trypsin with the inhibitors, which were used for superposition and analysis of enzyme-inhibitor interactions.

PDB CODE	Source	Inhibitor
1AQ7	*Bos taurus*	Aeruginosin 98-B
1AZ8	*Bos taurus*	BIS-Phenylamidine Inhibitor
1EJM	*Bos taurus*	BPI Aprotinin
1JRS	*Bos taurus*	Leupeptin
1MTW	*Bos taurus*	DX9:(+)-2-[4-[((s)-1-acetimidoyl-3-pyrrodinyl) oxy]-3-7-amidino-2-napthyl) propionic acid
1QL8	*Bos taurus*	ZEN: [4-(6-Chloro-Naphthalene-2-sulfonyl)-
1TPP	*Bos taurus*	APPA: (p-amidino-phenyl-pyruvate
1TPA	*Bos taurus*	BPTI: Bovine Pancreatic Trypsin Inhibitor
1TYN	*Bos taurus*	CTA:[cyclotheonamide A]
1XUF	*Bos taurus*	BAZ:BIS(5-amidino-benzimidazolyl)methane zinc
3PTB	*Bos taurus*	Benzamidine

Structural superposition of protease along with the inhibitors

All the structures (Table.I) of a selected set of thrombin, containing the inhibitor bound at their active sites, were superposed with a reference structure a structure of thrombin (1A2C) using our in-house software, MODELYN, with respect to all the Cα atoms of the proteases common to both the structures. This process led the superposition of the inhibitor bound to these enzymes allowing us to analyze the common areas of the inhibitors participating in the interaction. Similarly, another set of structures (Table.II) for another well-known protease, trypsin and superposition was done on its x-ray structure (1AQ7).

Identification of the atoms involved in the interactions

Distances of all the atoms of the inhibitor within the interacting distance of 4 Å were calculated using MODELYN. This helped us to identify the atoms involved in protease-inhibitors interactions. The nature of chemical forces involved in binding was analyzed using these data.

Analysis of hydrogen-bonding pattern and calculation of interactions energies

Hydrogen bonds between the protease and inhibitors were calculated using InsightII software to identify the atoms involve in hydrogen bonding. Complexes with the modeled structures were predicted by repeated energy minimization and molecular dynamics.The DOCKING module of InsightII was used to calculate the free energy of interactions between the proteases and their inhibitors both in water and protein water environment.

Calculation of free energies of inhibitor binding, Δ*G (*bind)

The free energy of binding, ΔG (bind), was calculated according to linear interaction energy (LIE) method developed by Aqvist and Samuelsson (1994) and subsequently used for characterization protein ligand interactions (Luzhkov). The equation is given below:

$$\Delta G\,(\text{bind}) = \alpha\Delta\ V_{1\text{-}s}^{\ vdW} + \beta\Delta\ V_{1\text{-}s}^{\ el}$$

where $V_{1\text{-}s}^{\ vdW}$ and $V_{1\text{-}s}^{\ el}$ denote the Lenard-Jones and electrostatic interactions between the ligand and its surroundings. The symbol, **Δ**, denotes the difference between the energies of the ligand in the protein-water and water environments. The values of **α** and **β** were taken as 0.16 and 0.5 respectively as used for inhibitor binding to the Plasmepsin IV, a protease from *Plasmodium falciparum* (Luzhkov, *et al*., 2006). The Lenard-Jones and electrostatic interactions were calculated using the DOCKING module of InsightII.

Results and Discussion

Interaction of Thrombin with its Inhibitors

Common inhibitor binding pockets

In search of the common inhibitor binding pockets of thrombin, sets of experimental structures with bound inhibitors (Table.I) were selected. All the enzyme structures containing the bound inhibitor was superposed with respect to all the corresponding Cα atoms common to both the enzymes of complexes. This led to indirect superposition of all the inhibitors sitting in the specificity pockets of the enzyme. The mutually superposed structures of thrombin showed a deviations (RMSD) ranging from 0.25 to 0.53Å. These deviations are partly due to differences in the 3-D structure determinations in different experimental conditions at different laboratories and partly due to influence of inhibitor binding on the enzyme structures. Hence, such deviations are in a reasonable limit of around 0.5Å.

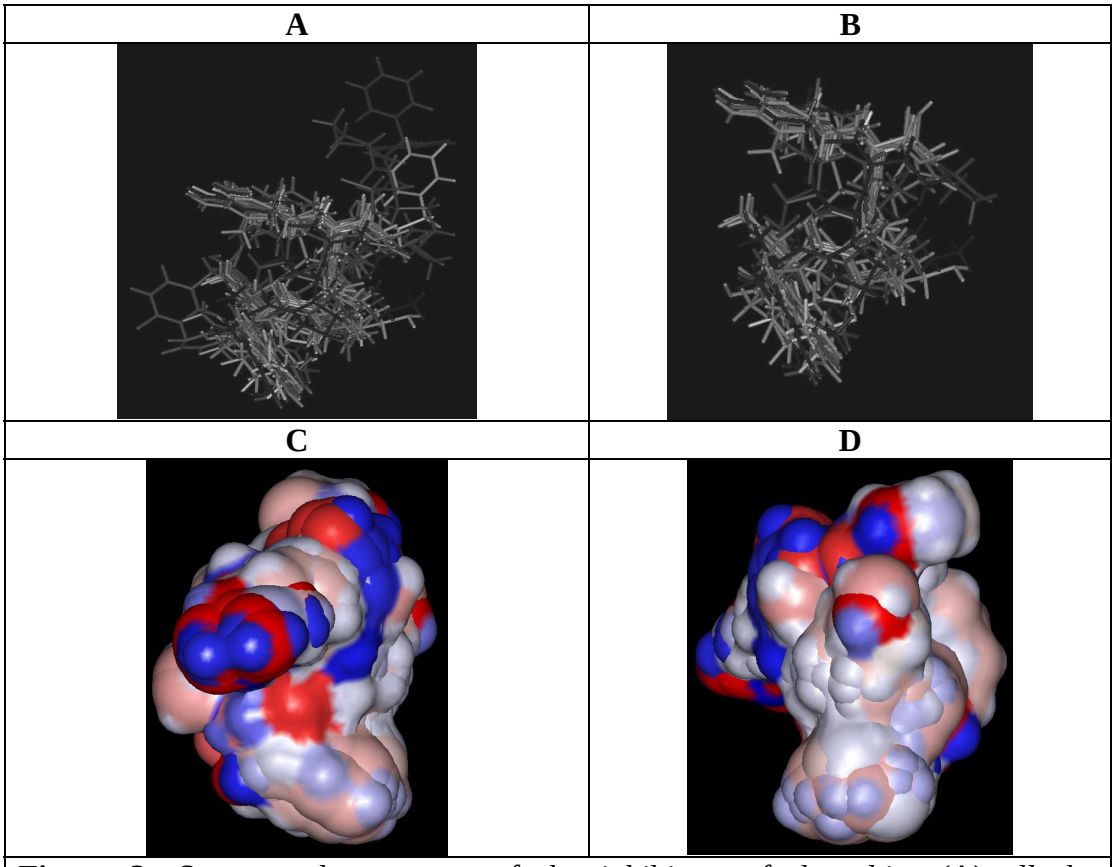


Figure 2: Superposed structures of the inhibitors of thrombin: (A) all the superposed inhibitors in their complete structures in stick representations; (B) shows the structures of only those inhibitors which are very compactly fits into the active site in stick representations; (C & D) respectively show the front and back view of the composite surfaces of the superposed inhibitors in space-filling representations, which is colored by electrostatic potentials on the surface, Blue colour represents positive, red negative and white neutral electrostatic potentials. The back view was obtained by rotating the structures by 180^{o} along the y-axis.

Fig.2A shows the multiple structural alignments of all the inhibitors involve in enzyme binding. It may be noted that all the inhibitors ensembles in a common space giving the replica of the binding site. Only three enzymes had some groups extended out side the common core of the rest of the inhibitors.

When these inhibitors were taken out of the ensemble the superposed inhibitors occupied a very compact space (Fig.2B), resembling the complementary space of the binding pocket. The composed surfaces of these inhibitors and their electrostatic potentials are shown in Fig.2C & D. The complimentary surfaces show patches of varied electrostatic potentials, some patches of strong positive (blue), negative (red) and neutral segments (white or mixed). This reflects the composite chemical nature of the binding pocket on the protease, thrombin. It is well known that the protease-inhibitor interactions are mediated through a number of sites and sub-sites on the enzyme (Hedstrom, L. 2002). Therefore, we identified all the atoms of the enzyme as well as on the inhibitors, which are in close proximity of the experimentally determined structures of the selected set complexes.

Table.III presents the atoms involved in the interaction of the inhibitors with thrombin in reference to the specificity pockets S1 and S2. The interaction between two atoms, which are in close proximity are said to be polar when both the partners are polar in nature which have either hydrogen bonding, dipole-dipole or charge-charge interactions. The S1 site of thrombin contains Asp-189 as the key specificity determining group; the other two residues in S1 pocket, Gly-216 and Gly-226 (14, 15) are essential for maintaining the conformation around the pocket and providing space to accommodate the complementary groups. Examination of atoms interacting at the S1 site as shown in Table. III reveals the polar nature of the interactions. These findings would help in the process of rational inhibitor design to incorporate groups with proper polarity.

Table III: Interactions in S1 and S2 specificity sites of thrombin in experimental structures of complexes with inhibitors.

PDB Code	**Atoms in S1 (AA: D-189, G-216, G-226) specificity site; polar interactions**		**Atoms in S2 (Y-60A, P-60B, P-60C, Y-60D) specificity site; non-polar interactions**	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
1A4W	189:CG	350:NE2	60A:OE	373:CM1
	189:OD1	350:NE2	60D:CE3	377:C5
	189:OD21	350:NE2	60D:CE3	377:C4
	216:N	350:NE1	60D:CE3	377:N3
	216:CA	350:CZ	60D:CZ2	373:C2
	216:C	350:N	60D:CE2	375:C3
	216:O	350:CG		
	226:CA	350:NE2		
1AG6	189:CG	372:CZ	60A:CE2	370:C3

PDB Code	Atoms in S1 (AA: D-189, G-216, G-226) specificity site; polar interactions		Atoms in S2 (Y-60A, P-60B, P-60C, Y-60D) specificity site; non-polar interactions	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
	189:OD1	372:CZ	60A:CZ	370:C3
	189:OD2	372:CZ	60A:OE	370:C3
	216:N	372:NE	60D:CZ2	370:C3
	216:CA	372:NE	60D:CZ2	370:N1
	216:C	370:N3	60D:CE2	370:N1
	216:O	370:C6		
	226:CA	372:NE1		
1AY6	189:CG	7:CZ	60A:CE1	6:CG
	189:CG	7:NE1	60A:CE2	6:CG
	189:OD1	7:NE1	60A:CZ	6:CG
	189:OD2	7:NE1	60A:OE	6:CG
	216:N	7:NE	60D:CZ2	9:C4
	216:CA	7:NE2	60D:CZ2	9:C6
	216:O	5:CB	60D:CZ3	9:C1
	226:CA	7:NE1	60D:CE2	9:C5
1CA8	189:CG	4:NE1	60A:CE2	2:CG
	189:OI1	4:NE2	60A:CZ	2:CG
	189:OI2	4:CZ	60A:OE	2:CG
	216:N	4:CI	60I:CZ2	2:CE
	216:CA	4:C3	60I:CZ2	2:CG
	216:C	4:C3	60I:CE2	2:CE
	216:O	4:C3		
	226:CA	4:NE2		
1FPC	189:CG	371:NE2	60A:CB	372:C2'
	189:OD1	371:NE2	60A:CG	372:C1'
	189:OD2	371:NE2	60A:CD	372:C2'
	216:N	371:O	60A:CE2	372:C1'
	216:CA	371:NE2	60A:OE	370:C3
	216:C	371:O	60D:CZ2	370:C2
	216:O	371:CB	60D:CZ2	370:C3
			60D:CE2	372:C5
1H8D	216:N	1:C55	60A:CE1	1:C8A
	216:CA	1:C55	60A:CE2	1:C8A
	216:C	1:C55	60A:CZ	1:C8B
	216:O	1:O15	60A:OE	1:C8A

PDB Code	Atoms in S1 (AA: D-189, G-216, G-226) specificity site; polar interactions		Atoms in S2 (Y-60A, P-60B, P-60C, Y-60D) specificity site; non-polar interactions	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
			60D:CZ2	1:C8B
			60D:CZ3	1:C3
			60D:CE2	1:C1
			60D:CE2	1:C2
			60D:CE2	1:C3
1LHG	216:N	400:C3	60A:CE2	400:C8
	216:CA	400:O3	60A:CZ	400:C8
	216:C	400:N3	60A:OE	400:C15
	216:O	400:O3	60D:CZ2	400:C8
			60D:CE2	400:C8
1LHC	189:CG	400:N3	60A:CE1	400:C10
	189:OD1	400:N4	60A:CE2	400:C10
	189:OD2	400:C6	60A:CZ	400:C10
	216:N	400:C5	60A:OE	400:C17
	216:N	400:N2	60D:CZ2	400:C10
	216:CA	400:N3	60D:CE2	400:C10
	216:C	400:N6		
	216:O	400:O4		
	226:CA	400:N4		
1PPB	189:CG	3:CZ	60A:CE2	2:CG
	189:OD1	3:NE2	60A:CZ	2:CG
	189:OD2	3:CZ	60A:OE	1:CE1
	216:N	1:O	60D:CZ2	2:CD
	216:CA	3:NE2	60D:CE2	2:CD
	216:C	1:N		
	216:O	1:CB		
	226:CA	3:NE1		
1TOM	189:CG	1:N1	60A:CE1	1:C11
	189:OD1	1:N1	60A:CE2	1:C11
	189:OD2	1:N1	60A:CZ	1:C11
	216:N	1:C2	60A:OE	1:C11
	216:CA	1:C2	60D:CZ2	1:C12
	216:C	1:N13	60D:CE2	1:O8
	216:O	1:C2		
	216:O	1:C16		

PDB Code	Atoms in S1 (AA: D-189, G-216, G-226) specificity site; polar interactions		Atoms in S2 (Y-60A, P-60B, P-60C, Y-60D) specificity site; non-polar interactions	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
1TMT	189:CG	3:NE1	60A:CE2	2:CG
	189:OD1	3:NE2	60A:CZ	2:CG
	189:OD2	3:CZ	60A:OE	2:CG
	216:N	3:NE		
	216:CA	3:NE2		
	216:C	1:N		
	216:O	1:CB		
	226:CA	3:NE1		
1UMA	189:CG	600:N1	60A:CB	500:N1
	189:OD1	600:N1	60A:CG	500:N1
	189:OD2	600:N1	60A:CD1	500:N1
	216:O	600:C2	60A:CD2	500:N1
			60A:CE1	500:C3
			60A:CE2	500:N1
			60A:CZ	500:C2
			60A:OE	500:C5
			60D:CZ3	600:C11
			60D:CE2	600:C11
1UVS	189:CG	11:N31	60A:CD2	11:C39
	189:OD1	11:E31	60A:CE1	11:C38
	189:OD2	11:C30	60A:CE2	11:C38
	216:N	11:O42	60A:CZ	11:C39
	216:CA	11:C30	60A:OE	11:C38
	216:C	11:N20	60A:EE	11:C5
	216:O	11:E20	60D:CZ2	11:C40
	216:E	11:E27	60D:CZ3	11:C40
			60D:CE2	11:C40

On the other hand, the S2 specificity pocket of thrombin provides the non-polar type of interaction. A non-polar interaction results when at least one of the interacting partners is non-polar. These non-polar interactions are mainly van der Waals type of close atomic contacts or stacking interaction with ring systems containing π-electron clouds. In thrombin the S2 specificity site, the main residues are Tyr-60A, Pro-60B, Pro-60C, Tyr-60D; tyrosine though has a polar hydroxyl group, are very good in staking interactions. Two proline residues serve the duel purpose of providing a

highly rigid pocket and supplying polar van der Waals contacts. Atomic level identification of the interacting partners in S2 site as listed in Table.III establishes this non-polar nature of the site for the selected set of thrombin-inhibitor complexes.

Hydrogen-bonding patterns of Thrombin-Inhibitor Interactions

Hydrogen-bonding interactions play major roles in the specificity determination of binding pockets in biological systems, hence, we analyzed the hydrogen-bonding patterns in the binding of thrombin with its inhibitors and shown in Table.IV. It may be noted that all the complexes exhibit a network of hydrogen bonding which include many important specificity pocket residues (shown in bold letters). In rational design of protease inhibitors it would be a good idea to maintain these hydrogen bonds and the aim will also be to incorporate more such interaction wherever possible. Judicial inclusion of newer groups in the designed inhibitors should be placed in such positions so as to maximize these interacting groups.

Table IV: Analysis of the hydrogen-bonding pattern of the complexes of thrombin with its selected inhibitors with known experimental structures in reference to the specificity pockets. Atom name convention is the same as in protein data bank (PDB). Atoms of thrombin in the specificity pockets, which are involved in hydrogen bonding with the inhibitor, are shown in bold letters.

PDB ID (Enzyme-Inhibitor)	HYDROGEN BONDING Enzyme/Inhibitor: Residue: Atom	
	DONOR	ACCEPTOR
1A5G (A5GE- A5GI)	**A5GI:R372:HH11**	**A5GE:H189:OD1**
	A5GI:R372:HH11	**A5GE:H189:OD2**
	A5GI:R372:HH21	A5GE:H189:OD2
	A5GI:R370H:H2	A5GE:H192:OE2
	A5GE:H193:HN	A5GI:R372:O
	A5GE:H195:HN	A5GI:R372:O
	2A5GE:H195:HG	A5GI:R372:O
	A5GI:R370H:HN31	**A5GE:H216:O**
	A5GI:R372:HE	A5GE:H219:O
	A5GI:R372:HH22	A5GE:H219:O
1AY6 (AY6E-AY6I)	**AY6I:J7:HH11**	**AY6E:H189:OD1**
	AY6I:J7:HH21	**AY6E:H189:OD2**
	AY6I:J7:HH12	AY6E:H190:O
	AY6E:H195:HG	AY6I:J7:O
	AY6E:H216:HN	**AY6I:J5H:O**
	AY6I:J7:HE	AY6E:H219:O
	AY6I:J7:HH22	AY6E:H219:O
1CA8 (CA8E-CA8I)	CA8E:B57:HE2	CA8I:D4H:O
	CA8I:D4H:HH2	**CA8E:B189:OD1**
	CA8I:D4H:HN	CA8E:B189:OD2

	CA8I:D4H:HH11 CA8E:B216: **CA8I:D1H:H1S** **CA8I:D2H:HN** CA8I:D4H:HH12	CA8E:B214:O CA8I:D2H:O **CA8E:B216:O** **CA8E:B216:O** CA8E:B219:O
1FPC (FPCE-FPCI)	**FPCI:R371:HH21** **FPCI:R371:HH22** FPCE:H216:HN **FPCI:R371:HN** FPCI:R371:HE	**FPCE:H189:OD1** **FPCE:H189:OD2** FPCI:R371:O **FPCE:H216:O** FPCE:H219:O
1H8D (H8DE-H8DI)	H8DE:H57:HE2 **H8DI:K1H:H59** H8DE:H219:HN	H8DI:K1H:O1A **H8DE:H216:O** H8DI:K1H:O15
1LHC (LHCE-LHCI)	LHCE:H57:HE2 **LHCI:400H:HN41** LHCI:400H:HN31 LHCE:H193:HN LHCE:H195:N LHCE:H195:HN LHCI:400H:HN1 **LHCE:H216:HN** LHCI:400H:HN6 LHCI:400H:HN32	LHCI:400H:O1 **LHCE:H189:OD1** LHCE:H189:OD2 LHCI:400H:O2 LHCI:400H:HO2 LHCI:400H:O2 LHCE:H214 **LHCI:400H:O4** LHCE:H216:O LHCE:H219:O
1PPB (PPBE-PPBI)	**PPBI:I3:HH11** **PPBI:I3:HH21** PPBE:H193:HN PPBE:H195:HN PPBI:I3:HN **PPBE:H216:HN** **PPBI:I1:HN2** PPBI:I3:HH22	**PPBE:H189:OD1** **PPBE:H189:OD2** PPBI:I3:O PPBI:I3:O PPBE:H214:O **PPBI:I1:O** **PPBE:H216:O** PPBE:H219:O
1THM (THME-THMI)	TMTE:H57:HE2 **TMTI:I3:HH11** **TMTI:I3:HH21** TMTE:H193:HN TMTE:H195:HN TMTI:I3:HN TMTE:H216:HN **TMTI:I1:HN2** TMTE:H219:O	TMTI:I3:O **TMTE:H189:OD1** **TMTE:H189:OD2** TMTI:I3:O TMTI:I3:O TMTE:H214:O TMTI:I1:O **TMTE:H216:O** TMTI:I3:HH22
1TOM	**TOMI:1H:H7** **TOME:H216:HN**	**TOME:H214:O** **TOMI:1H:O13**

Empirical Energies of Thrombin-Inhibitor Interactions

The values of energies of interaction provide an important index of binding affinity between the protease and inhibitors. The DOCKING module of InsightII was used to calculate the empirical energy showing the contributions from the van der Waals and electrostatic components of the interactions between thrombin and its inhibitors and presented in Table.V. It is to be noted that all the interaction energies are negative indicating stable complex formation; van der Waals contributions ranged from -76 to -25 Kcals/mole and the electrostatic contributions ranged from -213 to -20 Kcals/mole. It is to be noted that these interaction parameters are the primary guides in the process of inhibitor design. The inhibitors shown in Table.V are those of known inhibitors in the experimental structures of the complexes. The idea is to enhance the binding affinity by decreasing the free energy of interaction. These experimental structures may serve as the starting point in the process of drug design to make better inhibitors followed by improvements in other properties of drugs using the advanced computer-aided techniques. It should be admitted that these energy parameters, though provide initial guidance, there are many more steps in the overall process of drug design. Methods are being constantly improved for the prediction of binding affinity in terms of ΔG and K_d values which needs calculations involving aqueous environment (Luzhkov, *et al*., 2006). In this study we have used these techniques for a few complexes and compared with wet-lab experimental results as presented later in this chapter.

Table V: Calculated values of empirical energies of protein-ligand interactions of the experimental structures of complexes of thrombin with its inhibitors with the contributions from van der Waals and electrical components.

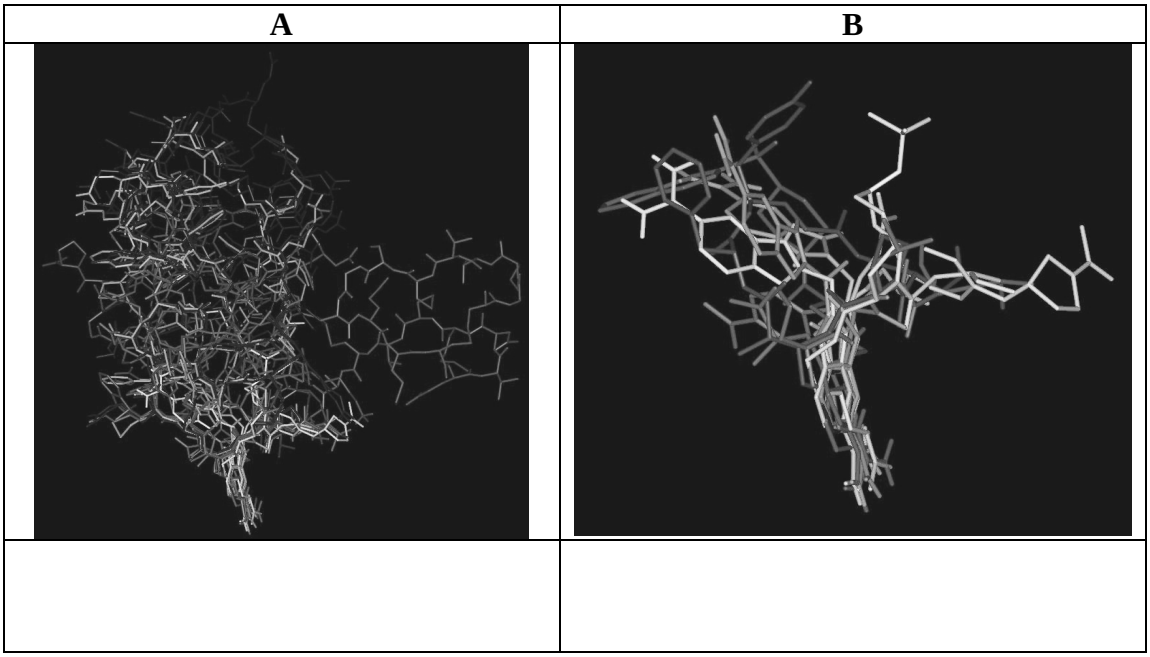
PDB ID	Empirical energies of protein-ligand interactions		
	van der Waals	Electrical	Total
1A4W	– 41.20	-118.12	–159.32
1A5G	– 76.14	–90.66	–166.81
1AY6	– 53.28	–153.57	–206.86
1CA8	– 57.74	–28.09	–85.83
1FPC	– 48.49	–125.16	–173.65
1H8D	– 83.76	–28.12	–110.88
1LHC	– 49.96	–57.71	–107.68
1LHD	– 40.96	–52.71	–93.67
1LHG	– 35.23	–50.71	-85.94
1PPB	– 36.66	–212.76	–249.43
1TOM	– 50.57	–20.13	–70.71
1TMT	– 40.76	–189.28	–230.05
1UMA	– 30.23	–54.71	–84.94
1UVS	– 25.57	–43.04	–68.61

Interaction of Trypsin with its Inhibitors

Common Inhibitor Binding Pocket of Trypsin

Common inhibitor binding pockets of trypsin were also analyzed, as in case of thrombin; sets of experimental structures with bound inhibitors (Table.II) were selected. All the enzyme structures containing the bound inhibitor was superposed with respect to all the corresponding Cα atoms in a common set of enzyme-inhibitor complex. This process led to indirect superposition of all the inhibitors sitting in the specificity pockets of trypsin. The mutually superposed structures of trypsin showed a deviations (RMSD) ranging from 0.25 to 0.56Å, values being similar to thrombin complexes. Hence, these deviations are partly due to differences in the 3-D structure determinations in different experimental conditions at different laboratories and partly due to influence of inhibitor binding on the enzyme structures. Thus, such deviations are in a reasonable limit of 0.5Å.

Fig.3A shows the multiple structural alignments of all the inhibitors involve in binding to trypsin. In case of the trypsin complexes, many inhibitors are protein inhibitors with extended structures, but at contact point with the enzyme all the inhibitors converged in a common space giving the replica of the binding site. When only the small the inhibitors, which also occupied the common binding pocket were considered, the ensemble of the superposed inhibitors occupied a very compact space (Fig.3B), resembling the complementary space of the binding pocket. The superposed molecules of these inhibitors are shown in space-filling models coloured by atoms (Fig.2 C&D, front and back views). We also identified all the atoms of the enzyme as well as on the inhibitors, which are in close proximity of the experimentally determined structures of the selected set complexes of trypsin with its inhibitors.



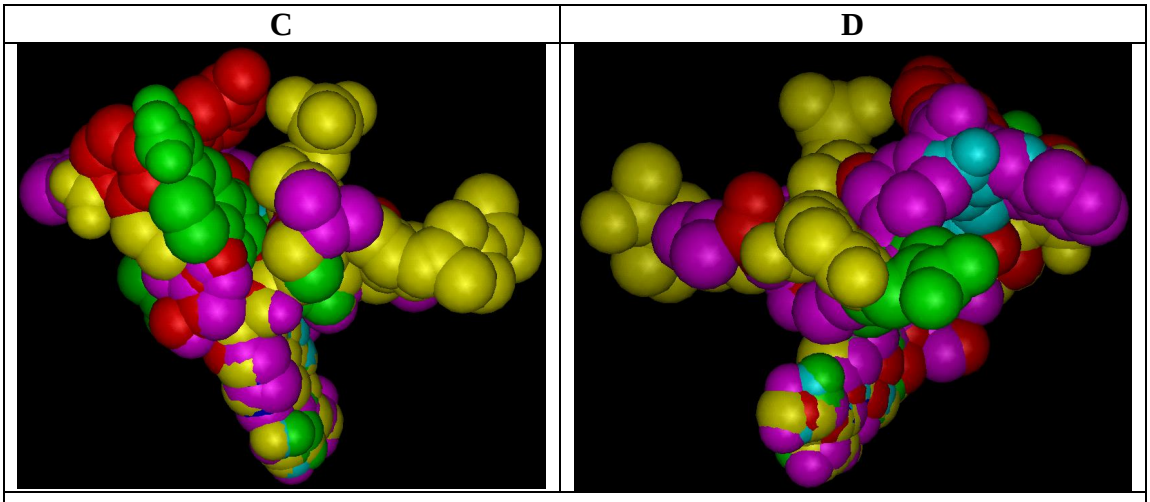


Figure 3: Superposed structures of the inhibitors of trypsin: (A) all the superposed inhibitors in their complete structures in stick representations; (B) shows the structures of only those inhibitors which are very compactly fits into the active site in stick representations; (C & D) respectively show the front and back view of the small inhibitors. The back view was obtained by rotating the structures by 180^{o} along the y-axis.

Table. VI presents the atoms involved in the interaction of the inhibitors with trypsin in reference to the specificity pockets S1 and S2. As in case of thrombin, the interactions between atoms in close proximity are thought to be polar when both the partners are polar in nature and have either hydrogen bonding, dipole-dipole or charge-charge interactions. The S1 site of trypsin contains Asp-189 as the key specificity determining group; the other two residues in S1 pocket, Gly-216 and Gly-226 are essential for maintaining the conformation around the pocket and providing space to accommodate the complementary groups, being the identical amino acids as thrombin. Examination of atoms interacting at the S1 site as shown in Table.VI reveals the polar nature of the interactions.

The S2 specificity pocket of trypsin also provides the non-polar type of interactions although the amino acids are different from those of thrombin S2 specificity pocket. In trypsin the main residues are Ser-39, His-40, Phe-41 and Tyr-151 (Ma *et al*., 2005); the benzene rings of phenylalanine and tyrosine provides the p-electron cloud for non-polar interactions. The interacting partners in S2 site as listed in Table.VI, which establishes this non-polar nature of the site for the selected set of trypsin-inhibitor complexes. However, there are less non-polar interactions in this specificity pocket of trypsin compared to that of thrombin.

Table.VI: Interactions in S1 and S2 specificity sites of trypsin in experimental structure of complexes with inhibitors.

PDB Code	**Atoms in S1 (D-189, G-216, G-226) specificity site; polar interactions**		**Atoms in S2 (S-39, H-40, F-41 and Y-151) specificity site; non-polar interactions**	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
1ANI	189:CG	8:NZ	40:O	10:CD2
	189:OD1	8:NZ	41:CD1	10:O
	189:OD2	8:NZ	41:CD2	10:O
	216:N	6:O	41:CE2	10:O
	216:O	5:C	151:CE2	10:CD2
	226:CA	8:NZ	151:OE	10:CD1
1AQ7	189:CG	1:C71		
	189:OD1	1:N73		
	189:OD2	1:C71		
	216:N	1:N68		
	216:CA	1:N68		
	226:N	1:N72		
	226:CA	1:C71		
	226:CA	1:N72		
1AZ8	189:CG	1:C20		
	189:OD1	1:N4		
	189:OD2	1:C20		
	216:N	1:O3		
	216:CA	1:N3		
	216:O	1:O3		
	226:CA	1:N4		
1EJM	189:CG	515:CZ	151:OE	534:CG1
	189:OD1	515:NE1		
	189:OD2	515:NE2		
	216:N	513:C		
	216:CE	515:NE2		
	216:O	513:CE		
	226:CE	515:NE1		
1JRS	189:CI	B:HH12		
	189:CG	B:CZ		
	189:OD1	B:CZ		
	189:OD2	B:CZ		

PDB Code	Atoms in S1 (D-189, G-216, G-226) specificity site; polar interactions		Atoms in S2 (S-39, H-40, F-41 and Y-151) specificity site; non-polar interactions	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
	189:OD2	B:HH22		
	216:N	B:C		
	216:CE	B:HH21		
	216:C	B:O		
	216:C	B:NH2		
	216:C	B:HE		
	216:O	B:HH21		
	216:H	B:HE		
	226:N	I:HH12		
1MTW	189:CG	999:C2		
	189:OD1	999:N1		
	189:OD2	999:N1		
	216:N	999:C4		
	216:CE	999:N1		
	216:C	999:C9		
	216:O	999:C18		
	226:CE	999:N2		
1QL8	189:OD1	999:C2	151:CZ	999:C30
	216:N	999:C3	151:OE	999:C30
	216:CE	999:C3		
	216:C	999:C3		
	216:O	999:C8		
1TAW	189:CG	15:CZ	39:CE2	17:O
	189:CG	15:NE1	39:CZ	19:CI
	189:OD1	15:NE2	39:OE	19:CI
	189:OD2	15:CZ	40:O	17:CG
	216:N	13:O	41:CE	17:O
	216:CE	15:NE2	41:C	17:N
	216:O	13:CE	41:O	16:CE
	226:CE	15:NE1	41:CI	17:O
			151:CD1	17:SD
			151:CE1	17:CE
			151:CE2	17:CE
			151:CZ	17:CE
			151:OE	17:CE
1TPA	189:CG	15:NZ	39:CD2	19:CD1

PDB Code	Atoms in S1 (D-189, G-216, G-226) specificity site; polar interactions		Atoms in S2 (S-39, H-40, F-41 and Y-151) specificity site; non-polar interactions	
	Enzyme	**Inhibitor**	**Enzyme**	**Inhibitor**
	189:OD1	15:NZ	39:CE2	19:CD1
			Continued...	
	189:OD2	15:NZ	39:CZ	19:N
	216:N	13:O	39:OE	19:O
	216:O	13:C	40:C	17:NE1
	226:CE	15:NZ	40:O	17:NE1
			41:CE	17:O
			41:C	17:N
			41:O	16:CE
			151:CE	17:CG
			151:CZ	17:CG
			151:OE	17:CG
1TYN	189:CG	246:C46	39:OE	246:C37
	189:OD1	246:N48	41:O	246:C34
	189:OD2	246:N48	41:C	246:C39
	216:N	246:N45		
	216:CE	246:N48		
	216:O	246:C52		
	226:CE	246:N47		
1XUF	189:CG	246:N1	41:CE	246:E2'
	189:OD1	246:C7	41:C	246:E2'
	189:OD2	246:EN2	41:O	246:C2'
	216:N	246:C6	41:CB	246:N2'
	216:CE	246:N1	41:CB	246:E2'
	216:C	246:EN1		
	216:E	246:C6		
	226:N	246:EN2		

Hydrogen-bonding patterns of Trypsin-Inhibitor Interactions

We also analyzed the hydrogen-bonding patterns in the binding of trypsin with its inhibitors and shown in Table.VII. Like those of thrombin, all the complexes exhibited a network of hydrogen bonding which include many important specificity pocket residues (shown in bold letters).

Table VII: Analysis of the hydrogen-bonding pattern of the complexes of trypsin with its selected inhibitors with known experimental structures in reference to the specificity pockets. Atom name convention is the same as in protein data bank (PDB). Atoms of trypsin in the specificity pockets, which are involved in hydrogen-bonding with the inhibitor, are shown in bold letters.

PDB ID (Enzyme-Inhibitor)	**HYDROGEN BONDING Enzyme/Inhibitor: Residue: Atom**	
	DONOR	**ACCEPTOR**
1ANI (ANIE-ANII)	ANII:I10:HN **ANII:I8:HZ3** ANIE:E192:HE21 ANIE:E193:HN ANIE:E195:HN **ANII:I8:HN** ANIE:E216:HN **ANII:I6:HN** ANIE:E219:HN	ANIE:E41:O **ANIE:E190:O** ANII:I7:O ANII:I8:O ANII:I8:O **ANIE:E214:O** ANII:I6:O **ANIE:E216:O** ANII:I4:O
1AQ7 (AQ7E- AQ7I)	AQ7B:1H:H721 AQ7B:1H:H731 AQ7A:192:HE21 **AQ7B:1H:H57** AQ7B:1H:HO31 **AQ7A:216:HN** **AQ7B:1H:H18** AQ7A:219:HN AQ7B:1H:H68 AQ7B:1H:H732	AQ7A:189:OD1 AQ7A:189:OD2 AQ7B:1H:O56 **AQ7A:214:O** AQ7A:216:N **AQ7B:1H:O31** **AQ7A:216:O** AQ7B:1H:O16 AQ7A:219:O AQ7A:219:O
1AZ8 (AZ8E-AZ8I)	AZ8I:1H:H35 AZ8I:1H:H32 AZ8I:1H:H34 AZ8E:149:HG1 AZ8I:1H:H20 AZ8I:1H:H18 **AZ8I:1H:H21** **AZ8E:216:HN** AZ8I:1H:H19	AZ8E:146:O AZ8E:148:O AZ8E:148:O AZ8I:1H:N2 AZ8E:189:OD1 AZ8E:189:OD2 **AZ8E:190:OG** **AZ8I:1H:O3** AZ8E:219:O
1EJM (EJME- EJMI)	**EJME:B519:HN** **EJME:B517:HE** **EJME:B517:HN** EJMEI:A57:ND1 EJMEI:B539:HE EJMEI:B539:HH22 EJMEI:B515:HH11 EJME:B515:HH21	**EJMI:A39:OH** **EJMI:A40:O** **EJMI:A41:O** EJMI:B514:SG EJMI:A97:O EJMI:A97:O EJMI:A189:OD1 EJMI:A189:OD1

	EJME:B515:HH11 **EJME:B515:HH12** EJME:A192:HE21 EJME:A193:HN EJME:A195:HN EJME:A195:HG **EJME:B515:HN** **EJME:A216:HN** EJME:B515:HH22	EJMI:A189:OD2 **EJMI:A190:OG** EJMI:B514:O EJMI:B515:O EJMI:B515:O EJMI:B515:O **EJMI:A214:O** **EJMI:B513:O** EJMI:A219:O
1JRS (JRSE- JRSI)	JRSE:A57:HE2 JRSI:B3:HH11 JRSI:B3:HH22 **JRSI:B3:HH12** JRSE:A192:HE21 **JRSI:B3:HN** **JRSE:A216:HN** JRSI:B3:HH21	JRSI:B3:O JRSE:A189:OD1 JRSE:A189:OD2 **JRSE:A190:OG** JRSI:B2:O **JRSE:A214:O** **JRSI:B1:O** JRSE:A219:O
1MTW (MTWE-MTWI)	MTWI:999H:N28 MTWI:999H:N28 MTWI:999H:H322 MTWI:999H:H321 MTWE:51:HN MTWI:999H:O23 MTWI:999H:O23	MTWE:48:O MTWE:48:OD1 MTWE:49:N MTWE:49:O MTWI:999H:N28 MTWE:242:O MTWE:245:O
1QL8 (QL8E-QL8I)	QL8I:A600H:H4 QL8E:A57:HE2 QL8E:A57:HE2 **QL8I:A999H:H1** QL8E:A193:HN	QL8E:A57:NE2 QL8I:A600H:O1 QL8I:A600H:O4 **QL8E:A190:OG** QL8I:A600H:O2
1TYN (TYNE- TYNI)	TYNE:57:HE2 TYNI:246H:H471 TYNI:246H:H481 **TYNI:246H:H472** TYNE:192:HE21 TYNE:193:HN TYNE:195:HN **TYNI:246H:H12** **TYNI:216:HN** **TYNI:246H:H51** TYNI:246H:H482	TYNI:246H:O40 TYNE:189:OD1 TYNE:189:OD2 **TYNE:190:OG** TYNI:246H:O49 TYNI:246H:O41 TYNI:246H:O41 **TYNE:214:O** **TYNI:246H:O50** **TYNE:216:O** TYNE:219:O
1XUK (XUKE- XUKI)	XUKE:246H:H11 **XUKE:190:HG** XUKE:246H:H12	XUKE:189:OD2 **XUKI:246H:N2** XUKI:219:O

Many protease inhibitors have been developed using proteases as target for rational design (Turk *et al*., 2005). It would be useful if these hydrogen bonds are maintained and more are incorporated wherever possible. Judicial inclusion of newer groups in the designed inhibitors should be placed in such positions so as to maximize these hydrogen-bonding interactions.

Empirical Energies of Trypsin-Inhibitor Interactions

The index of binding affinity between the protease and inhibitors is very useful in predicting the stronger inhibitors. The calculated empirical energies interim of contributions from the van der Waals and electrostatic components of the interactions between trypsin and its inhibitors presented in Table.VIII. In case of two selected trypsin-inhibitor complexes showed positive values of empirical energies indicating very unstable complex formation. This may be due some imperfection in the structure of these complexes or due to unfavourable charge distribution as the electrical components were positive while van der Waals component had very high negative values (shown in bold in Table.VIII); however, it was not further analyzed. The interaction energies of other complexes were negative indicating stable complex formation with van der Waals contributions ranging from -87 to -28 Kcals/mole of the same order as those of thrombin-inhibitor complexes and the electrostatic contributions ranged from -322 to -14 Kcals/mole. These experimental structures also may serve as the starting point in the process of drug design to make better inhibitors followed by improvements in other drug-like properties.

Table VIII: Calculated values of empirical energies of protein-ligand interactions of the experimental structures of complexes of trypsin with its inhibitors with the van der Waals and electrical contributions.

PDB ID	Empirical energies of protein-ligand interactions		
	van der Waals	Electrical	Total
1ANI	– 83.91	- 67.51	–151.42
1AQ7	– 53.86	– 42.43	–96.29
1AZ8	– 38.40	– 18.87	– 57.27
1EJM	**– 95.59**	**125.63**	**30.04**
1JRS	– 35.46	– 82.59	–118.06
1MTW	– 43.03	– 16.45	–59.48
1QL8	– 46.24	– 18.20	–64.45
1TAW	– 87.56	– 322.81	– 410.37
1TPA	**– 88.11**	**101.95**	**13.83**
1TYN	– 57.21	– 31.47	– 88.68
1XUF	– 28.11	– 14.34	– 42.45

Empirical binding energies of modeled protease-inhibitor complexes

We predicted the structures of the protease-inhibitor complexes using our predicted models and calculated their interaction energies using the DOCKING module of InsightII. Table.IX presents the values of the energies of modeled complexes along with the calculated values of energies of the experimental structures of complexes (PDB Codes: 1GBI, 1GBD and 1GBM), which were used to predict the structure of these complexes with the same inhibitor bound to them. Structures of three complexes with three different inhibitors from PDB structures were predicted by homology-based method for the protease from *P. furiosus* of the SA clan. It may be noted that the values of the empirical interaction energies are comparable to those of the experimental structures (Table.IX). Two other threading based models of the proteases from *P. falciparum* and *N. crassa* of the SA clan were used to predict the structures of the complexes using the x-ray structure 1GBM with the inhibitor present in the PDB file. In theses cases also the values of calculated energies are comparable to each other. Thus, it may be stated that the modeled structures predicted by us are suitable for design of inhibitors.

Table IX: Calculated values of empirical energies of protein-ligand interactions of the modeled structures of complexes in comparison to the reference experimental structures with the van der Waals and electrical contributions.

PDB Code	**Structure of Complex**	**Empirical energies of protein-ligand interactions**		
		van der Waals	**Electrical**	**Total**
1GBI	x-ray	-34.35	-77.12	-111.47
1GBD	x-ray	-32.38	-90.71	-123.29
1GBM	x-ray	-41.34	-90.48	-131.82
1GBI	Model (*P. furiosus*)	-28.78	-123.80	-152.59
1GBD	Model (*P. furiosus*)	-27.91	-43.96	-71.87
1GBM	Model (*P. furiosus*)	-22.47	-24.42	-46.90
1GBM	Model (*P.falciparum*)	-13.09	-30.63	-43.72
1GBM	Model (*N. crassa*)	-41.24	-100.93	-142.18

Free energies (ΔG) of the protease-inhibitor interactions

Now methods are evolving for the calculation of binding affinities in terms of ΔG and K_d values which can be compared with experimental data obtained from wet laboratory experiments using physico-chemical techniques (Mitchell *et al*., 1996a, Mitchell *et al*., 1996b). In this study we have used one such technique, linear interaction energy (LIE) method, for calculation of ΔG values of a few complexes and compared them with wet-lab experimental results as well as with results calculated using other method (Table.X). Biomolecular Ligand Energy Evaluation Protocol (BLEEP) is a knowledge-based method that derives potentials of mean force (PMF)

by converting the distribution of atom distances between 2.5Å and 8.0Å in a protein-ligand system into pair potential like functions.

Table X: Binding free energies (ΔG) calculated by linear interaction energy (LIE) method and comparison with the values available in website http://www.mitchell.ch.cam.ac.uk/pld/background_energy_bleep.html calculated by BLEEP using potentials of mean force (PMF) method and determined experimental techniques.

PDB Code	Inhibitors	Values of ΔG in (in Kcals/mol)				
		Calculated by us using LIE			BLEEP**	Experimental*
		vdW	Elec.	Total		
1ARC	TLCK	-4.22	-6.38	-3.86		
1BCR	antipain	-5.00	-4.67	-1.53		
1BCS	chymostatin	-7.87	-12.89	-7.70		
1BRA	Benzamidine	-7.70	-8.08	-5.27	-2.49	-2.49
1D3D	Hirugen	-2.14	-7.68	-4.18	-12.38	-12.38
1DWC	argatroban	-2.80	-3.40	-2.15	-10.09	-10.09
1EQ9	PMSF	-6.39	-4.22	-4.11		
1PPH	NAPAP	-9.61	-7.12	-5.06	-8.48	-8.48
1TNG	2-amino-methyl-cyclohexane	-6.42	-3.70	-2.87	-4.00	-4.00
1TNH	2,4-fluorobenzy-lamine	-5.96	-6.62	-2.09	-4.59	-4.59
1TNK	2,3-phenylpropyl-amine	-8.94	-2.16	-2.51	-2.03	-2.03
1KZD	tyrostatin	-5.25	-14.65	-8.16		
1VGC	L-para-chloro-1-acetamido boronic acid	-5.13	-7.24	-4.42		

Search of protease sequences in the human genome

Current drug discovery efforts are based on rationally identifying chemical compounds that will bind to therapeutic target molecules such as proteins. These efforts are based on the 3D structure of the target molecule and use a variety of computer-based molecular modeling techniques to exploit the 3D structural information. Of the approximately 400 known human proteases, approximately 14% are under investigation as drug targets. The initial annotation of the approximately 30,000 human proteome set includes approximately 500 proteases. Bioinformatics analysis can now be performed on complete human protease families. New sequences will require evaluation of their function in normal physiology and human disease. Genomic sequence information will have a central role in the validation of protease drug targets (Southan, C. 2001).

The sequences of proteases from different clans, which were used to predict structures were submitted to Genome BLAST and compared with human genome for examining the suitability for use in target based drug design. Proteases from different pathogenic organism like protozoa, fungi and archea were selected for predicting their 3-D structures. Close sequence homology of these sequences with any human protease sequence may make these proteases unsuitable for using as a drug target.

As a positive control of the human genome search for identification of homologous sequences we used some human protease sequences from different clans, which were used in this study. Although the human protease genes are interrupted by introns, the genomic BLAST search, over all 6 reading frames, could identify homologous sequences in the genome. The identities of the portions of the sequence alignments varied from 34% to 100% in a number of human chromosomes like 1, 3, 4, 8, 12, 13, 16, 17, 18 and 19 with high bit score (229) and significant expect values ($3x10^{-58}$). Thus, this type search would be able to identify any significant homology with the pathogenic protease target sequences.

We used all the protease sequences of the clans of the lower organisms, which are potentially pathogenic, modeled by us in the genomic BLAST search in the human genome. It was found that in all the cases no significant match was found with any bigger segment (>10AA). The best expect value of $6x10^{-9}$ with 38% identity over a stretch of 100 AA was found with *Dictyostelium discoideum* protease of the SE clan. The next best expect value (6x10-6) with 40% AA identity over a stretch of 50AA with *Pyrococcus abyssi* protease of the same clan. These matches were due to some homologous segment from certain domain of the proteases, which may exhibit some structural similarity with the human protease. No other protease that we modeled showed any significant match (Expect value > 0.02). Thus, the designed inhibitors of proteases are very less likely to cause any interference with the human protease.

Conclusion

The aim of the work reported here is to develop a structural perspective about the properties that define the specificity of the binding interface between these inhibitors and trypsin-like serine proteases. Most of the serine protease inhibitors from different sources have been found to have considerable medical and industrial importance and they are being extensively studied to obtain an insight into mechanisms for understanding the specificity of inhibition of enzyme catalysis.

In this work, we have analyzed the nature interaction of serine proteases with their inhibitors in atomic details in the experimental structures. Superposition of the structures of the enzymes with respect to each other brought all the inhibitors in the binding sites of the well studied serine proteases, thrombin and trypsin, giving the complementary shapes of the active sites. The surfaces of the indirectly superposed inhibitors were used to assess the nature of the specificity pockets of these enzymes. Individual complexes were used to identify the atoms of the inhibitors which were in close proximity of the enzymes. Thus, it was established that the nature of S1 specificity site was mostly polar and that of S2 specificity site was mostly non-polar of both the serine proteases, thrombin and trypsin.

The calculation of the van der Waals and electrical components the empirical interaction energies showed that in most of the complexes were stabilized by higher electrostatic interactions. All the enzyme-inhibitor complexes were stabilized by an intricate network of hydrogen binding involving the specificity site residues. I calculated the free energy of binding using the linear interaction energy (LIE) method and compared with the reported values calculated using the BLEEP software and determined experimentally, which were in good agreement. Some of our modeled structures of proteases were used to predict the structures enzyme-inhibitor to demonstrate that these structures can be used for rational design of specific inhibitors. Finally it was established that our modeled proteases of pathogenic organisms bear very little homology with the proteases of human genome.

Acknowledgements

This work was carried out and supported by The Indian Institute of Chemical Biology, Kolkata-700032, West Bengal, India.

References

[1] Barrett, A. J., Rawlings, N. D. & O'Brien, E. A. (2001). The MEROPS database as a protease information system. J Struct Biol 134, 95-102

[2] Bitoun, E., Chavanas, S., Irvine, A. D., et al. (2002). Netherton syndrome: disease expression and spectrum of SPINK5 mutations in 21 families. J Invest Dermatol 118, 352-61

[3] Blow, D. M. (1997). The tortuous story of Asp. His. Ser: structural analysis of alpha-chymotrypsin. Trends Biochem Sci 22, 405-8

[4] Czapinska, H. & Otlewski, J. (1999). Structural and energetic determinants of the S1-site specificity in serine proteases. Eur J Biochem 260, 571-95

[5] De Filippis, V., Colombo, G., Russo, I., et al. (2002). Probing the hirudin-thrombin interaction by incorporation of noncoded amino acids and molecular dynamics simulation. Biochemistry 41, 13556-69

[6] Hedstrom, L. (2002). Serine protease mechanism and specificity. Chem Rev 102, 4501-24

[7] Huber, R., Kukla, D., Bode, W., et al. (1974). Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor. II. Crystallographic refinement at 1.9 A resolution. J Mol Biol 89, 73-10

[8] James, M. N. (2006). The peptidases from fungi and viruses. Biol Chem 387, 1023-9

[9] Knowles, J. R. (1965). Enzyme specificity: alpha-chymotrypsin. J Theor Biol 9, 213-28

[10] Krol, J., Kopitz, C., Kirschenhofer, A., et al. (2003). Inhibition of intraperitoneal tumor growth of human ovarian cancer cells by bi- and trifunctional inhibitors of tumor-associated proteolytic systems. Biol Chem 384, 1097-102

[11] Laskowski, M., Jr. & Kato, I. (1980). Protein inhibitors of proteinases. Annu Rev Biochem 49, 593-626
[12] Lehesjoki, A. E. (2003). Molecular background of progressive myoclonus epilepsy. Embo J 22, 3473-8
[13] Lomas, D. A., Lourbakos, A., Cumming, S. A., et al. (2002). Hypersensitive mousetraps, alpha1-antitrypsin deficiency and dementia. Biochem Soc Trans 30, 89-92
[14] Luzhkov, V. B., Almlof, M., Nervall, M., et al. (2006). Computational study of the binding affinity and selectivity of the bacterial ammonium transporter AmtB. Biochemistry 45, 10807-14.
[15] McKay, T. R., Bell, S., Tenev, T., et al. (2003). Procaspase 3 expression in ovarian carcinoma cells increases survivin transcription which can be countered with a dominant-negative mutant, survivin T34A; a combination gene therapy strategy. Oncogene 22, 3539-47
[16] Ma, W., Tang, C. & Lai, L. (2005). Specificity of trypsin and chymotrypsin: loop-motion-controlled dynamic correlation as a determinant. Biophys J 89, 1183-93
[17] Mitchell, J.B.O., et al. (1999a). BLEEP - Potential Of mean force describing protein - Ligand Interactions: I. Generating Potential. Journal of Computational Chemistry 20, 1165-1176
[18] Mitchell, J.B.O., et al. (1999b). BLEEP - Potential Of mean force describing protein - Ligand Interactions: II. Calculation of Binding Energies and Comparison with Experimental Data. Journal of Computational Chemistry 20, 1177-1185
[19] Perona, J. J. & Craik, C. S. (1995). Structural basis of substrate specificity in the serine proteases. Protein Sci 4, 337-60
[20] Prasad, S., Cantwell, A. M., Bush, L. A., et al. (2004). Residue Asp-189 controls both substrate binding and the monovalent cation specificity of thrombin. J Biol Chem 279, 10103-8
[21] Rezaie, A. R. (1998). Reactivities of the S2 and S3 subsite residues of thrombin with the native and heparin-induced conformers of antithrombin. Protein Sci 7, 349-57
[22] Ritchie, B. C. (2003). Protease inhibitors in the treatment of hereditary angioedema. Transfus Apher Sci 29, 259-67
[23] Shotton, D. M. & Watson, H. C. (1970). Three-dimensional structure of tosyl-elastase. Nature 225, 811-6
[24] Southan, C. (2001). A genomic perspective on human proteases as drug targets. Drug Discov Today 6, 681
[25] Telang, M., Srinivasan, A., Patankar, A., et al. (2003). Bitter gourd proteinase inhibitors: potential growth inhibitors of Helicoverpa armigera and Spodoptera litura. Phytochemistry 63, 643-52
[26] Turk, B., Turk, D. & Salvesen, G. S. (2002). Regulating cysteine protease activity: essential role of protease inhibitors as guardians and regulators. Curr Pharm Des 8, 1623-37

[27] Wanga, Q. M. & Chen, S. H. (2007). Human rhinovirus 3C protease as a potential target for the development of antiviral agents. Curr Protein Pept Sci 8, 19-27