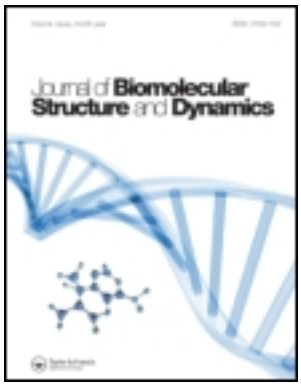

Journal of Biomolecular Structure and Dynamics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tbsd20

Structural Features of the Interfaces in Enzyme-Inhibitor Complexes

Alexei N. Nekrasov [a] & Alexei A. Zinchenko [a]
[a] Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry Russian Academy of Sciences, ul. Miklukho-Maklaya, 16/10, Moscow, 117997, Russia
Published online: 15 May 2012.

To cite this article: Alexei N. Nekrasov & Alexei A. Zinchenko (2010): Structural Features of the Interfaces in Enzyme-Inhibitor Complexes, Journal of Biomolecular Structure and Dynamics, 28:1, 85-96

To link to this article: http://dx.doi.org/10.1080/07391102.2010.10507345

Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 28, Issue Number 1, (2010)

Structural Features of the Interfaces in Enzyme-Inhibitor Complexes

Alexei N. Nekrasov[1,2]
Alexei A. Zinchenko[1,3]

[1]Shemyakin-Ovchinnikov
Institute of Bioorganic Chemistry
Russian Academy of Sciences
ul. Miklukho-Maklaya, 16/10,
Moscow, 117997 Russia.

http://www.jbsdonline.com

Abstract

Specific protein-protein interaction is essential for the function of life systems. A variety of computational methods are being extensively used now-a-days to investigate this interaction and to identify structural features of binding sites. In this paper, the informational structure analysis method was applied to the study of protein-protein interaction interfaces in enzyme-inhibitor complexes. The analysis of amino acid sequence by informational structure analysis method reveals three types of sites (ADD+, NORMAL and ADD-) which differ in the density of first rank elements in the informational structure. ADD+, NORMAL and ADD- sites also differ in their ability towards adaptive conformational reorganization which contributes to the formation of protein-protein interaction interfaces in enzyme-inhibitor complexes. The study of hydrolytic enzymes in complex with their protein inhibitors shows that at least one of the interaction interface sites is of ADD- type. ADD- sites possess an increased ability towards adaptive conformational changes thus enabling effective protein interaction.

Key words: Informational structure of proteins; Enzyme-inhibitor complex; Protein-protein interaction.

Introduction

The study of protein complexes is one of the major trends in modern molecular biology, and now-a-days it is developing rapidly. There are multiple types of molecular systems in cell: enzyme/inhibitor, signal protein/receptor, antigen/antibody complexes, *etc*. Such interactions of proteins form the core of all biological regulation. Computational methods that are becoming of wide use to study such interactions are molecular modeling, docking and molecular dynamics simulations (1-15). In this paper we provide an alternate method, the informational structure analysis method to the study the interfaces in enzyme-inhibitor complexes.

The identification of the contacting groups involved in the formation of interaction interfaces is a fundamental aspect of protein complex study (16-18, http://ppidb.cs.iastate.edu/). The formation of protein associate is the result of effective cooperative interaction of all functional groups at the interface. The ability of polypeptide chains to undergo spatial reorganization enables the effective interaction

Abbreviations: BPTI: Bovine pancreatic trypsin inhibitor; Subt/CI-2A: Subtilisin-chymotrypsin inhibitor – 2; PME 1: Pectin methylesterase 1; PMEI: Pectin methylesterase inhibitor; RNase A: Ribonuclease A; RI: Ribonuclease inhibitor; PPE: Porcine pancreatic elastase; α_1-PI: α_1-proteinase inhibitor; MMP-3: Matrix metalloproteinase-3; TIMP-1: Tissue inhibitor of metalloproteinases 1.

E-mail: [2]alexei_nekrasov@mail.ru
[3]alezina@mail.ru

of functional groups and, in our opinion, plays the key role in the formation of protein-protein complexes.

In the present work we have applied the ANIS method (19) to introduce the computable parameter allowing to evaluate the ability of polypeptide chain to undergo adaptive conformational reorganization.

Methods

The previously described method of protein primary structure analysis (ANIS method) allows to consider amino acid sequence as the system of hierarchically organized sites - **EL**ements of **I**nformational **S**tructure (ELIS). It is shown that high rank ELIS correspond to domains in case of proteins with distinct domain architecture (19), the subsequent study has indicated (20) that high rank ELIS are involved in the enzyme catalytic activity manifestation.

The ANIS method is based on a representation of protein sequence as a set of informational units (IU). Such approach was proposed according to the study of positional information entropy (21). In that paper we described two fundamental features of information encoding in protein sequences:

- The correlation between amino acid residues which can be observed on great distances in protein sequences
- High and constant level of correlation between residues on distances less then 6 positions.

The latter allows introducing the new approach to protein sequence representation which considers the primary structure as a set of short overlapping peptide fragments (IU). Such a method can be used to discover the hierarchical organization of protein sequences.

Informational Structure Analysis

Graphic representation of IS (IDIC-diagrams) can be composed using following algorithm:

- the amino acid sequence of a given protein is encoded as a set of informational units (IU);
- the population profile of the target protein structure by IU is determined;
- IDIC-sites are localized within the protein sequence;
- the graphic representation of IS is constructed.

Encoding Protein Sequences as IU Sets

Let $\Re$ be the set of all known amino acid residues forming primary structure of native protein sequences, *i.e.*, $\Re = \{A,C,D,E,F,G,H,I,K,L,M,N,P,Q,R,S,T,V,W,Y\}$. In traditional form the protein sequence is written as $B:\{1,...,N\} \rightarrow \Re$, where N is the number of residues in the sequence.

In our previous work (21) it was shown that in native proteins neighboring residues have high level of correlation on R_{HLCD} distance. We encoded a protein sequence by a set of Informational Units (IU) $U_i = (B_{i-\delta},...,B_{i+\delta})$ each representing a group of ε neighboring residues, where $\varepsilon = 2\delta + 1$, $\delta = 1,...N$. All the protein sequences included in the PIR database rel. 8.0 (22) were encoded in this manner and have formed the set $\Im = \{U_\varepsilon\}$. All elements in the set $\Im$ are characterized by the occurrence frequencies P_{Uk} of informational units of type k (U_k). The occurrence frequencies of IUs were used for the analysis of informational structure of studied protein sequences.

The next step in the study of protein information structure is the composition of the profile $F = \{F(j)\}$ of the protein sequence populated by IU:

$$F(j) = \left\{ \begin{array}{cc} \sum\limits_{i=j\pm\varepsilon/2} P_{U_i} & if \;\; U_j \in \Im \\ 0 & if \;\; U_j \notin \Im \end{array} \right\} \quad [1]$$

where $j = 1 + \delta,...N - \delta$, $\varepsilon = 2\delta + 1$ and $\Im$ is the set of all possible IU from all native protein sequences. Information Units belong to $\Im$ if 80% of IU forming residues are equal to the elements of $\Im$.

Localization of IDIC-Sites

IU encoding allows to compute a value of IU correlation in protein sites having different size. Sites with local high coordination level were called IDIC-sites (19). To identify the location of IDIC-sites we introduced a function $f(j')$ for each $j = 1,2...N$

$$f(j') = De^{-\frac{(j'-j)^2}{\rho^2}} \quad [2]$$

which is determined for all values $j = 1,2...N$ and which complies with the condition

$$F(j) - f(j') \leq 0 \quad [3]$$

The latter condition is the limitation of $f(j')$ function D parameter value. For each $j = 1,...N$ the $D = D(j)$ parameter is chosen to be the maximal of D satisfying the condition (3). During the computation the ρ values were matching interval $\delta < \rho < N - \delta$, where N is the length of target sequence. At the same time,

$$\omega(j,\rho) = \sum_{j'} D(j) e^{-\frac{(j'-j)^2}{\rho^2}} \quad [4]$$

can be computed for each $f(j')$ function.

Graphic Representation of a Protein IS (the IDIC-Diagram)

The IS of a protein can be graphically represented as a surface Γ_ω which is the function $\omega = \omega(j, \rho)$ determined for all possible ρ and j values. A simpler and more convenient representation of IS (IDIC-diagram) can be constructed by considering only sites with a locally increased degree of coordination between residues [4], thus operating with a restricted set of Γ_ω surface values, which are local maximums complying with the condition:

$$\omega(j-1,\rho) \leq \omega(j,\rho) \geq \omega(j+1,\rho) \quad [5]$$

As the result of such substitution the structurally complicated Γ_ω surface is reduced to a limited set of points $K = \{\omega(j,\rho) : \omega(j-1,\rho) \leq \omega(j,\rho) \geq \omega(j+1,\rho)\}$ $1+\delta \leq j \leq N-\delta$, $\delta \leq \rho \leq N-\delta$. By connecting the closest points which correspond to positions of IDIC-sites of different length ρ, hierarchical graphs - ELIS are produced. The ELIS at the given node point is characterized by the rank which is equal to the number of node points along the most distant path from lowest-level element to target node point. The ELIS located at the lowest level of hierarchy and corresponding to the shortest IDIC-sites are assigned a rank equal to 1.

Results

Upon encoding the protein primary structure as a set of IU we discovered the hierarchy of sites with increased degree of information coordination (IDIC-sites) between amino acid residues. We designated the ensemble of hierarchically organized IDIC-sites as informational structure (IS) and the individual parts of IS as ELIS. ELIS is characterized by its location in the primary structure and its rank – the position inside IS hierarchy. Each IDIC-site of various length corresponds to its own ELIS. The shortest IDIC-sites correspond to first rank ELIS. The length of IDIC-sites conforming to high rank ELIS, ranges from several tens up to hundreds of residues, depending on the total size of protein sequence. In the previous study (20) we discussed the functional role of high rank ELIS in the enzyme structure. In the present work we consider the structure and functional role of first rank ELIS.

We used a small protein pheromone ER-1 (23) as an illustration of the role of first rank ELIS in the formation of protein secondary structure. The pheromone IS is presented on Figure 1A, the protein spatial configuration with marked central residues of IDIC-sites, forming first rank ELIS, is on Figure 1B. The spatial structure of ER-1 consists of three packed helixes of different length. Figure 1B shows that the longer is helix the higher is quantity of first rank ELIS it incorporates.

This example together with molecular modeling of IDIC-sites sequences allows to conclude that IDIC-sites forming first rank ELIS possess determined helical-like conformation. This allows postulating that the system of first rank ELIS forms the protein secondary structure elements by enabling the conformational states of polypeptide chain in the reverse turns of β-sheets and also in helices. The presence of sequence fragments with determined helical-like conformation in the protein primary structures correlates with the previously described periods in oscillating term of positional information entropy (21).

One should expect first rank ELIS to be uniformly distributed along the protein sequence. However, our study of first rank ELIS distribution function in the PIR database sequences (24) has shown that it is a function with several maximums (Figure 2). Such function can be approximated by the superposition of three Gaussian functions. These functions reflect the distances between first rank ELIS in sequences. It is important to note that distance equal by 5 positions in polypeptide

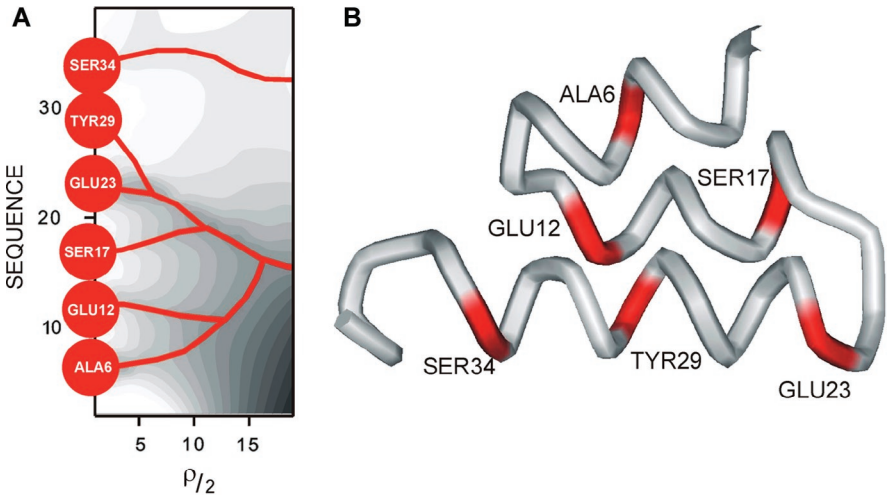

Figure 1: The information (**A**) and spatial (**B**) structures of ER-1 pheromone (2ERB.PDB). On the informational structure plot the central residues of IDIC-sites corresponding to first rank ELIS are shown with red round marks, the residues designations are given. The same residues are marked with red color on the spatial structure chart.

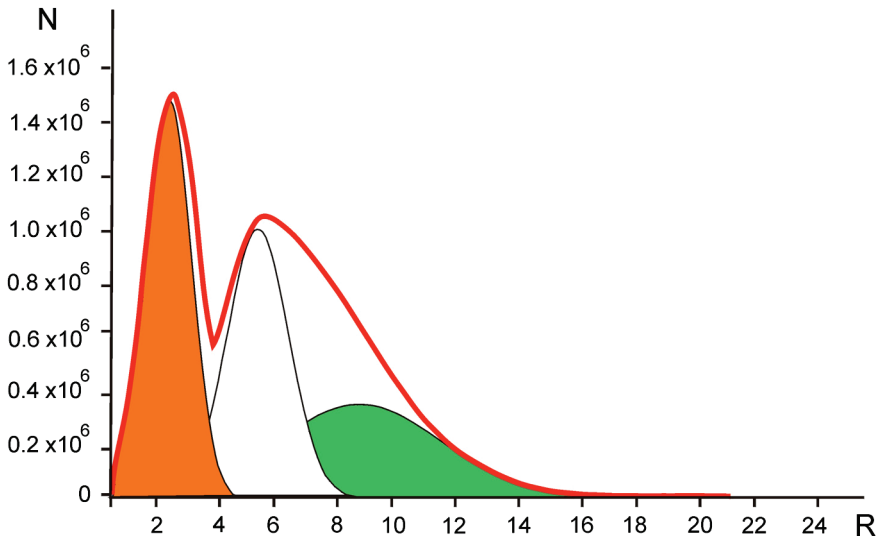

Figure 2: The first rank ELIS distance distribution graph built for the PIR database representatives (the distribution density of first rank ELIS). The distances (R) between central residues of IDIC-sites corresponding to first rank ELIS are on X-axis, while the occurrence rates of such distances (N) in the database are on Y-axis. The first rank ELIS distribution density curve is represented by the thick line on a chart. The Gauss functions approximating first rank ELIS distribution density are shown as dotted lines. The central Gauss function (white color) corresponds to first rank ELIS densities which are typical for NORMAL sites. The left Gauss function (orange color) corresponds to first rank ELIS densities which are characteristic for ADD+ sites. The right Gauss function (green color) corresponds to first rank ELIS densities typical for ADD- sites. The intervals conforming to different type of sites are designated with braces.

chain (HLCD value) match the range on which the highest level of amino acid coordination is observed (21).

We have termed "NORMAL" the sites which were characterized by such distances between central residues, meaning that they possess a normal density of first rank ELIS (5-7 sequence positions). The positions of centers of two other Gaussian functions deviate from "NORMAL" value. There are sites in the protein sequences with the distance between central residues less than 5 positions, and we term them "ADD+ sites", *i.e.*, the sites with anomalously high density of first rank ELIS. In the opposite, there are sites with the distance between central residues exceeding 7 positions, and we designate them as "ADD- sites", *i.e.*, sites with anomalously low density of first rank ELIS (Figure 2). Such classification based on the first rank ELIS density allows to represent a protein structure as an ensemble of NORMAL, ADD+ and ADD- sites. Figure 3 shows the mutual arrangement of first rank ELIS at the ADD+, NORMAL and ADD- sites. Above we discussed the role of first rank ELIS in the development of secondary structure elements by providing the determined helical conformation of short regions of polypeptide chains. Such short fragments correspond to IDIC-sites of first rank ELIS. The first rank ELIS fragments overlap at the ADD+ sites (Figure 3), thus enabling the determined conformation of polypeptide chain in ADD+ site. In the opposite, ADD- sites are conformationally labile. This allows ADD- sites to exist as a series of different conformational states with similar values of potential energy. Such regions possess topological variability and significant ability towards adaptive conformational reorganizations. The conformational and adaptive properties of NORMAL type sites lay between characteristics of ADD- and ADD+ sites.

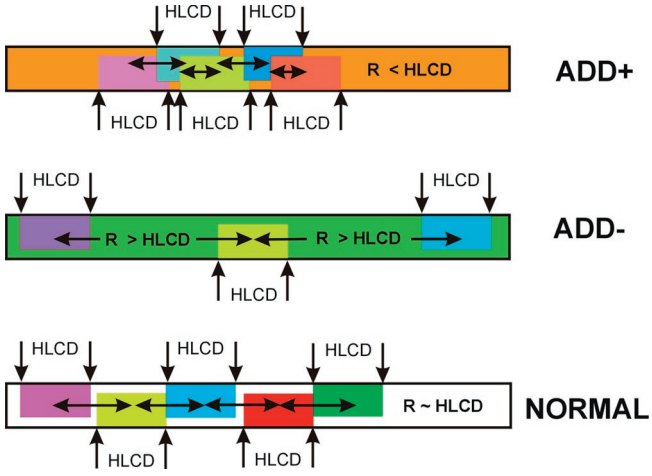

Figure 3: The localization of IDIC-sites corresponding to first rank ELIS in the ADD+, NORMAL and ADD- sites. The IDIC-sites are marked by different colors. "R" is the distance between central residues of IDIC-sites, which is designated by horizontal arrows.

According to the proposed classification of protein sites by the first rank ELIS density (ADD-, ADD+, NORMAL) there are six types

of possible contacts between residues in protein-protein interaction interfaces, depending on the type of site each residue is affiliated with:

Type I ('+-'). ADD- –	ADD+ or (ADD+ – ADD-)
Type II ('N-').	ADD- – NORMAL or (NORMAL – ADD-)
Type III ('--').	ADD- – ADD-
Type IV ('NN').	NORMAL – NORMAL
Type V ('N+').	NORMAL – ADD+ or (ADD+ - NORMAL)
Type VI ('++').	ADD+ - ADD+

The object of present research is the study of properties of polypeptide chain regions composing protein-protein interfaces and the role of adaptive conformational reorganization in the formation of contact surfaces. We used the previously proposed protein informational structure approach (19) and the classification of protein sites based on distribution density of first rank ELIS, which is described above. We studied the complexes of hydrolytic enzymes with their inhibitors of different source and molecular weight (Figure 3):

- Trypsin with bovine pancreatic trypsin inhibitor (Trps/BPTI), 3TGI.PDB;
- Subtilisin with Subtilisin-Chymotrypsin inhibitor - 2 (Subt/CI-2A), 1A10.PDB;
- Pectin methylesterase 1with Pectin methylesterase inhibitor (PME 1/PMEI), 1XG2.PDB;
- Ribonuclease A with Ribonuclease inhibitor (RNase A/RI), 1DFJ.PDB;
- Porcine Pancreatic Elastase with α_1-Proteinase inhibitor (PPE/α_1-PI), 2D26.PDB;
- matrix metalloproteinase-3 with tissue inhibitor of metalloproteinases 1 (MMP-3/TIMP-1), 1OO9.PDB.

The initial information on protein sequences and spatial structures of complexes were obtained from PDB database (http://www.rcsb.org/pdb/home/home.do). The spatial structures of studied enzyme-inhibitor complexes are shown on Figure 4. The IS of protein complexes was computed, the locations of ADD-, ADD+ and

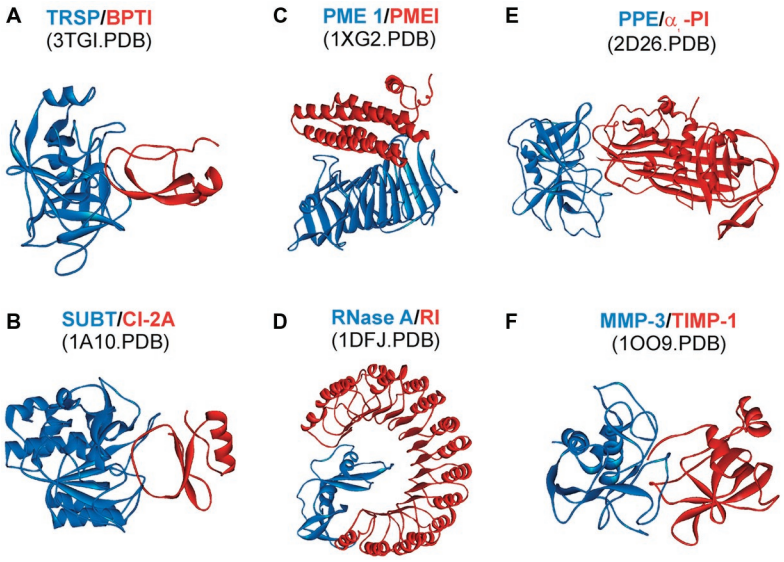

Figure 4: The spatial structures of enzyme-inhibitor complexes. (**A**) trypsin with bovine pancreatic trypsin inhibitor (Trps/BPTI), 3TGI.PDB; (**B**) Subtilisin Carlsberg with Subtilisin-chymotrypsin inhibitor - 2 (Subt/CI-2A), 1A10.PDB; (**C**) pectin methylesterase 1 with pectin methylesterase inhibitor (PME 1/PMEI), 1XG2.PDB; (**D**) ribonuclease A with ribonuclease inhibitor (RNase A/RI), 1DFJ.PDB; (**E**) porcine pancreatic elastase with α_1-proteinase inhibitor (PPE/α_1-PI), 2D26.PDB; (**F**) matrix metalloproteinase-3 with tissue inhibitor of metalloproteinases 1 (MMP-3/TIMP-1), 1OO9.PDB. The spatial structures of enzymes are in blue color. The spatial structures of inhibitors are in red color.

NORMAL sites were determined (Figure 5). To determine the localization of residues situated on boundaries of sites with different first rank ELIS density, we have introduced the following rule: the central residue of IDIC-site situated in a border area between ADD-, NORMAL or ADD+ sites is attributed to the one with maximal density of first rank ELIS. According to such rule the central residues of border IDIC-sites are assigned to adjacent NORMAL or ADD+ sites.

We chose the interatomic distance of 3.7 angstrom or less as a criterion of contact between a pair of residues in the protein-protein interaction interface. The analysis of each contact considered the type of the site (ADD-, NORMAL, ADD+) to which the each residue belonged. Thus, we composed a table of interresidual contacts which were obtained from X-ray data for the hydrolase-inhibitor complexes (Supplementary Material).

Figure 5A depicts the sequences of trypsin and bovine pancreatic trypsin inhibitor (BPTI) with marked locations of ADD-, ADD+ and NORMAL sites. The

Figure 5: The amino acid sequences of enzymes and their inhibitors forming protein-protein complex. The central residues of IDIC-sites conforming to first rank ELIS are in bold. ADD+ and ADD- sites are noted with orange and green background, accordingly. NORMAL sites are not marked with color background. (**A**) trypsin and bovine pancreatic trypsin inhibitor (Trps/BPTI); (**B**) subtilisin Carlsberg and subtilisin-chymotrypsin inhibitor - 2 (Subt/CI-2A); (**C**) pectin methylesterase 1 and pectin methylesterase inhibitor (PME 1/PMEI); (**D**) ribonuclease A and ribonuclease inhibitor (RNase A/RI); (**E**) porcine pancreatic elastase and α_1-proteinase inhibitor (PPE/α_1-PI); (**F**) matrix metalloproteinase-3 and tissue inhibitor of metalloproteinases 1 (MMP-3/TIMP-1).

X-ray data (25) have shown that the Trps/BPTI interaction interface is formed by 27 pairs of residues: 18 residues in structure of trypsin and 9 residues in structure of BPTI. 24 contacts are of type I (88.9%), there are also unitary contacts of types III, V and VI (3.7% each). The contact between amino acid residues usually involves several atoms. Type I contacts are characterized 59 interatomic distances (93.6%) which are less then 3.7 angstroms. There are 2 cases (3.2%) of such close mutual atomic dislocation in type IV contacts, and one case in each of type III and type VI contacts (1.6% each). The data obtained is summarized in Table I.

Figure 5B depicts the sequences of subtilisin and subtilisin-chymotrypsin inhibitor - 2 (Subt/CI-2A) with marked locations of ADD-, ADD+ and NORMAL sites. The X-ray data (http://www.rcsb.org/pdb/files/1a10.pdb) have indicated that the Subt/CI-2A interaction interface is formed by 30 pairs of residues: 23 residues in structure of subtilisin and 11 residues in structure of CI-2A. 11 interresidual contacts are of type I (36,7%), 9 are of type II (30%), 3 are of type IV (10%), 6 are of type V (20%) and 1 is of type VI (3,3%). The analysis of interatomic contacts have indicated that there are 30 (39%), 26 (33.7%), 9 (11.7%), 10 (13.0%) and 2 (2.6%) cases of interatomic distances being less then 3.7 angstrom among the atoms involved in type I, type II, type IV, type V and type VI contacts, respectively. The data obtained is summarized in Table I.

Figure 5C depicts the sequences of pectin methylesterase (PME) and its inhibitor (PMEI) with marked locations of ADD-, ADD+ and NORMAL sites. The X-ray data (26) have indicated that the PME/PMEI interaction interface is formed by 31 pairs of residues: 17 residues in structure of PME and 22 residues in structure of PMEI. 13 contacts are of type I (41.9%), 8 are of type II (25.8%), 5 are of type III (16.1%), 2 are of type IV (6.5%) and 3 are of type V (9.7%). The analysis of interatomic contacts have indicated that there are 24 (33.8%), 22 (31.0%), 15 (21.1%), 4 (5.6%) and 6 (8.5%) cases of interatomic distances being less then 3.7 angstrom among the atoms involved in type I, type II, type III, type IV and type V contacts, respectively. The information on PME/PMEI complex interresidual and interatomic contacts is summarized in Table I.

Figure 5D depicts the sequences of RNase A (RNaseA) and its inhibitor (RI) with marked locations of ADD-, ADD+ and NORMAL sites. The X-ray data (27) have indicated that the RNaseA/RI interaction interface is formed by 30 pairs of residues: 21 residues in structure of RNase A and 21 residues in structure of RI. 8 interresidual contacts are of type I (26.7%), 7 are of type II (23.3%), 13 are of type III (43.3%), 2 are of type V (6.7%). The analysis of interatomic contacts have indicated that there are 16 (31.4%), 12 (23.5%), 21 (41.2%) and 2 (3.9%) cases of interatomic distances being less then 3.7 angstrom among the atoms involved in type I, type II, type III and type V contacts, respectively. The information on RNase A/RI complex interresidual and interatomic contacts is summarized in Table I.

Figure 5E depicts the sequences of porcine pancreatic elastase (PPE) and α_1-proteinase inhibitor (α_1-PI) with marked locations of ADD-, ADD+ and NORMAL sites. The X-ray data (28) have indicated that the PPE/α_1-PI interaction interface is formed by 5 pairs of residues: 5 residues in structure of PPE and 5 residues in structure of α_1-PI. 1 interresidual contact is of type I (20.0%), 1 is of type II (20.0%), 2 are of type III (40.0%), 1 is of type V (20.0%). The analysis of interatomic contacts have indicated that there are 1 (20.0%), 1 (20.0%), 2 (40.0%) and 1 (20.0%) cases of interatomic distances being less then 3.7 angstrom among the atoms involved in type I, type II, type III and type V contacts, respectively. The information on PPE/α_1-PI complex interresidual and interatomic contacts is summarized in Table I.

Table 1
The types of intermolecular and interatomic contacts in the interfaces of enzyme-inhibitor complexes.

Contact type	Enzyme-inhibitor complex																		Average data for 6 complexes, %		
	Trps/BPTI			Subt/CI-2A			PME/PMEI			RNase A/RI			PPE/α_1-PI			MMP3/TIMP-1					
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
I (+–)	88.9 (24)	93.6 (59)		36.7 (11)	39.0 (30)		41.9 (13)	33.8 (24)		26.7 (8)	31.4 (16)		20.0 (1)	14.3 (2)		5.1 (3)	9.6 (49)		36.5	36.9	
II (–N)	0 (0)	0 (0)	95.2	30 (9)	33.7 (26)	72.7	25.8 (8)	31.0 (22)	85.9	23.3 (7)	23.5 (12)	96.1	20.0 (1)	14.3 (2)	92.9	54.2 (32)	46.9 (238)	62.8	25.5	24.9	84.2
III (--)	3.7 (1)	1.6 (1)		0 (0)	0 (0)		16.1 (5)	21.1 (15)		43.3 (13)	41.2 (21)		40.0 (2)	64.3 (9)		6.8 (4)	6.3 (32)		18.3	22.4	
IV (NN)	0 (0)	0 (0)		10 (3)	11.7 (9)		6.5 (2)	5.6 (4)		0 (0)	0 (0)		0 (0)	0 (0)		27.1 (16)	24.6 (125)	37.2	7.3	7.0	
V (+N)	3.7 (1)	3.2 (2)	4.8	20 (6)	13.0 (10)	27.3	9.7 (3)	8.5 (6)	14.1	6.7 (2)	3.9 (2)	3.9	20.0 (1)	7.1 (1)	7.1	6.8 (4)	12.6 (64)		11.2	8.1	15.8
VI (++)	3.7 (1)	1.6 (1)		3.3 (1)	2.6 (2)		0 (0)	0 ()		0 (0)	0 (0)		0 (0)	0 (0)		0 (0)	0 (0)		1.2	0.7	

1 - The number of interacting pairs of residues in enzyme-inhibitor interface (% of total amount of contacting residue pairs). The number of contacting residue pairs is in parenthesis.
2 - The number of interatomic contacts with the distances less then 3.7Å in enzyme-inhibitor interface (% of total amount of contacts). The number of interatomic contacts is in parenthesis.
3 - The total number of interatomic contacts taking place during adaptive (contact types I-III) and non-adaptive (contact types IV-VI) interactions of polypeptide chains in enzyme-inhibitor interface (% of whole amount of interatomic contacts).

Figure 5F depicts the sequences of matrix metalloprotease-3 (MMP-3) and tissue inhibitor of metalloproteinases 1 (TIMP-1) with marked locations of ADD-, ADD+ and NORMAL sites. The X-ray data (29) have indicated that the MMP-3/TIMP-1 interaction interface is formed by 59 pairs of residues: 21 residues in structure of MMP-3 and 22 residues in structure of TIMP-1. 3 interresidual contact are of type I (5.1%), 32 are of type II (54.2%), 4 are of type III (6.8%), 16 are of type IV (27.1%), and 4 are of type V (6.8%). The analysis of interatomic contacts have indicated that there are 49 (9.6%), 238 (46.9%), 32 (6.3%), 125(24.6%) and 64 (12.6%) cases of interatomic distances being less then 3.7 angstrom among the atoms involved in type I, type II, type III, type IV and type V contacts, respectively. The information on MMP-3/TIMP-1 complex interresidual and interatomic contacts is summarized in Table I.

The analysis of data from Table I indicates that interactions between residues of ADD+ sites ('++' contacts) are minor in case of six examined complexes and averages 0.7% of all interatomic contacts. The impact of interatomic contacts of the residues from ADD+ and NORMAL sites ('N+' contacts) does not exceed 13% and averages 8.1% for all studied complexes. The impact of interatomic contacts of the residues located in the sites with normal distribution of first rank ELIS ('NN' contacts) does not exceed 12% in 5 cases of 6, and the average impact of 'NN'-type interactions is about 7%. In the case of MMP-3/TIMP-1 complex the impact of 'NN' contacts reaches 24.6%.

A completely different picture is observed in the case of three remaining types of possible contacts which involve an ADD- site as imperative component. As it was mentioned, the ADD-- sites possess the significant ability towards adaptive conformational reorganization. The impact of interatomic contacts of residues located in ADD- sites ('--' contacts) averages 22.4%. The actual impact of such contact type varies widely in different complexes. The percentage of '--' contacts does not exceed 6.3% in case of 3 complexes we've studied, and for the rest of 3 cases it ranges from 21.1% (PME/PMEI) up to 64.3% (PPE/α_1-PI).

The interatomic contacts involving residues from ADD- and NORMAL sites ('-N' contacts) average 24.9% of total number of interactions. There are no such contacts in Trps/BPTI complex, as for other complexes, the impact of '-N'-type interactions ranges from 14.3% (PPE/α_1-PI) to 46.9% (MMP3/TIMP1). Finally, the interatomic contacts between residues belonging to ADD- and ADD+ sites ('+-' contacts) are major for the studied complexes and their average impact is about a third of all interactions (36.9%). The percentage of such contacts varies from 9.6% (MMP3/TIMPI) to 93.6% (Trps/BPTI). The average impact of contacts involving ADD- site residue as a member of interacting pair amounts to 84.2%. This value ranges from 62.8% (MMP3/TIMP-1) to 95.2% (Trps/BPTI). More then 90% of such interactions are observed in the case of RNAse/RI (96.1%) and PPE/α_1-PI (92.9%) complexes. However, the possible variants of contacts ('--', 'N-', '+-') are not represented equally in the studied proteins. As for Trps/BPTI complex, the '+-' contacts are virtually predominant, while there are all three possible variants of interaction in the structure of RNAse/RI and PPE/α_1-PI complexes, yet the '--' contacts are major. It is important to note that Trps/BTPI and PPE/α_1-PI complexes represent the different mechanisms of enzyme inhibition: α_1-PI forms a covalent bound with a residue of PPE active site, whereas BPTI possesses a peptide bound with distorted geometry, which mimics the substrate (30). As for other three enzyme-inhibitor complexes, the percentage of contacts involving ADD- sites is slightly lower and ranges from 62.8% (MMP3/TIMP-1) to 85.9% (PME/PMEI). In PME/PMEI complex all three possible contact types are uniformly present, while in MMP3/TIMP-1 the 'N-' type contacts are dominant (46.9%). As it was mentioned above, the impact of NN-type contacts to the formation of MMP3/TIMP-1 complex is also significant. This can provide an explanation of the fact that this complex is characterized by lowest binding constant among the others complexes (29).

It is significant that '+-' contacts are the most common among the three dominant interaction types ('+-', 'N-', '--'). In our opinion this allows to propose a set of criteria to which members of the protein complex should fit to provide an effective and rapid interaction:

- one polypeptide chain must possess the determined topology
- the partner molecule has to be inclined to adaptive conformational reorganization.

The presence of such contacting pairs of sites in the interaction interfaces provides evidence for the regular nature of such interaction developed during molecular evolution.

Thus, our results have indicated that the formation of interaction interfaces in enzyme-inhibitor complexes complies with a set of regularities which can be described by ANIS-method. It is shown that the protein interaction interfaces are mostly formed by contacting pairs of sites which include at least one site (ADD-site) which is characterized by the decreased density of first rank ELIS. The presence of ADD- sites reflects the necessity of adaptive conformational reorganization for the effective interaction of polypeptide chains. As informational structure is computed according to protein sequence and because the density of first rank ELIS reflects the ability of polypeptide chains to undergo conformational reorganization, we can state that the first rank ELIS density is an effective measure of interaction interface formation capacity.

Conclusion

The results obtained provide strong evidence of crucial role of adaptive conformational mobility of at least one of the polypeptide chains involved in the formation of protein complex. The ANIS method involves the computable parameter derived from protein sequence to detect sites with different abilities towards adaptive conformational reorganizations and can be an effective tool for protein interactions study, not only in complexes, but also inside independent polypeptide globule.

Supplementary Material

Supplementary material dealing with inter-residue contacts is available at no charge from the authors directly; the supplementary data can also be purchased from Adenine Press for US $50.0.

Acknowledgement

This work was supported by grant NSh-5207.2010.4 from the President of the Russian Federation for Support of Leading Scientific Schools, Federal Target Programme "Scientific and scientific-educational staff of innovative Russia for 2009-2013" and Russian Academy of Sciences Program "Molecular and Cell Biology".

References

1. G. H. Braun, D. M. M. Jorge, H. P. Ramos, R. M. Alves, V. B. da Silva, S. Giuliatti, S. V. Sampaio, C. A. Taft, and C. H. T. P. Silva. *J Biomol Struct Dyn 25,* 347-355 (2008).
2. D. Josa, E. F. F. da Cunha, T. C. Ramalho, T. C. S. Souza, and M. S. Caetano. *J Biomol Struct Dyn 25*, 373-376 (2008).
3. E. F. F. Da Cunha, T. C. Ramalho, and R. C. Reynolds. *J Biomol Struct Dyn 25,* 377-385 (2008).
4. K. V. Ramesh, M. Purohit, K. Mekhala, M. Krishnan, K. Wagle, and S. Deshmukh. *J Biomol Struct Dyn 25,* 481-493 (2008).
5. J. Dasgupta, and J. K. Dattagupta. *J Biomol Struct Dyn 25,* 495-503 (2008).
6. J. Yoon, J. Park, S. Jang, K. Lee, and S. Shin. *J Biomol Struct Dyn 25,* 505-515 (2008).
7. V. Rajakrishnan, V. R. Manoj, and G. Subba Rao. *J Biomol Struct Dyn 25,* 535-542 (2008).

8. B. P. Mukhopadhyay, B. Ghosh, H. R. Bairagya, T. K. Nandi, B. Chakrabarti, and A. K. Bera. *J Biomol Struct Dyn 25,* 543-551 (2008).
9. C. Arcangeli, C. Cantale, P. Galeffi, G. Gianese, R. Paparcone, and V. Rosato. *J Biomol Struct Dyn 26,* 35-47 (2008).
10. A. M. Andrianov. *J Biomol Struct Dyn 26,* 49-56 (2008).
11. S. Subramaniam, A. Mohmmed, and D. Gupta. *J Biomol Struct Dyn 26,* 473-479 (2009).
12. J.-H Zhao, H.-L Liu, Y.-F Liu, H.-Y Lin, H.-W. Fang, Y. Ho, and W. -B Tsai. *J Biomol Struct Dyn 26,* 481-490 (2009).
13. H. R. Bairagya, B. P. Mukhopadhyay, and K. Sekar. *J Biomol Struct Dyn 26,* 497-507 (2009).
14. L. Zhong and J. Xie. *J Biomol Struct Dyn 26, 525*-533 (2009).
15. M. Parthiban, M. B. Rajasekaran, S. Ramakumar, and P. Shanmughavel. *J Biomol Struct Dyn 26,* 535-547 (2009).
16. D. Douguet, H. C. Chen, A. Tovchigrechko, and I. A.Vakser. *Bioinformatics 22*, 2612-2618 (2006).
17. D. Reichmann, O. Rahat, M. Cohen, H. Neuvirth, and G. Schreiber. *Curr Opin Struct Biol 17*, 67-76 (2007).
18. H. X. Zhou and S. Qin. *Bioinformatics 23*, 2203-2209 (2007).
19. A. N. Nekrasov. *J Biomol Struct Dyn 21*, 615-623 (2004).
20. A. N. Nekrasov, and A. A. Zinchenko. *J Biomol Struct Dyn 25*, 553-561 (2008).
21. A. N. Nekrasov. *J Biomol Struct Dyn 20*, 87-92 (2002).
22. W. C. Barker, L. T. Hunt, D. G. George, L. S. Yeh, H. R. Chen, M. C. Blomquist, E. I. Seibel-Ross, A. Elzanowski, J. K. Bair, and D. A. Ferrick, *et al. Protein Seq Data Anal 1*, 129-175 (1987).
23. M. S. Weiss and D. Eisenberg. *Acta Crystallogr Sect D 52*, 469-480 (1996).
24. C. H. Wu, L. S. Yeh, H. Huang, L. Arminski, J. Castro-Alvear, Y. Chen, Z. Hu, P. Kourtesis, R. S. Ledley, B. E. Suzek, C. R. Vinayaka, J. Zhang, and W. C. Barker. *Nucleic Acids Res 1*, 345-347 (2003).
25. A. Pasternak, D. Ringe, and L. Hedstrom. *Protein Sci 8*, 253-258 (1999).
26. A. Di Matteo, A. Giovane, A. Raiola, L. Camardella, D. Bonivento, G. De Lorenzo, F. Cervone, D. Bellincampi, and D. Tsernoglou. *Plant Cell 17*, 849-858 (2005).
27. B. Kobe and J. Deisenhofer. *Nature 374*, 183-186 (1995).
28. A. Dementiev, J. Dobó, and P. G. Gettins. *J Biol Chem 281*, 3452-3457 (2006).
29. S. Arumugam and S. R. Van Doren. *Biochemistry 42*, 7950-7958 (2003).
30. P. Gettins. *Chem Rev 102,* 4751-4804 (2002).

Date Received: December 4, 2009

Communicated by the Editor Ramaswamy H. Sarma