Journal of Molecular Graphics and Modelling 49 (2014) 99–109

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Effect of ligand binding on the dynamics of trypsin. Comparison of different approaches

Elena Ermakova*, Rauf Kurbanov

Kazan Institute of Biochemistry and Biophysics RAS, P.O. Box 30, Kazan 420111, Russia

ARTICLE INFO

Article history:
Received 2 January 2014
Received in revised form 7 February 2014
Accepted 8 February 2014
Available online 17 February 2014

Keywords:
Trypsin
Signal transduction
Molecular dynamics
GNM
PCA

ABSTRACT

The intramolecular signal transduction induced by the binding of ligands to trypsin was investigated by molecular dynamics simulations. Ligand binding changes the residue-residue interaction energies and suppresses the mobility of loops that are in direct contact with the ligand. The reduced mobility of these loops results in the altered flexibility of the nearby loops and thereby transmits the information from ligand binding site to the remote sites. The analysis of the flexibility of all residues confirmed the coupling between loops L1 (185–188) and L2 (221–224) and the residues in the active center. The significance of S1 pocket residues for the signal transduction from the active center to the substrate-binding site was confirmed by the dynamical network and covariance matrix analyses. Gaussian network model and principal component analysis demonstrated that the active center residues had zero amplitude in the slowest fluctuations acting as hinges or anchors. Overall, our results provide a new insight into protein–ligand interactions and show how the allosteric signaling may occur.

1. Introduction

Ligand binding is an important step in many biological processes. It can affect the structure and dynamics of a protein [1–3]. Although the process starts via local interactions with residues in direct contact, it often has an effect on remote protein regions. The mechanism, by which the information is passed from one site to another within a folded protein, is not clear. In some proteins, binding of ligand induces structural changes. However, in many proteins, ligand binding changes protein dynamics without inducing significant structural changes. It has been suggested that ligand binding initiates perturbations that propagate to the remote sites of protein [4–6]. Signal transduction induced by the ligand binding may affect the functional properties of protein, e.g. cause an allosteric effect. Initially, the allosteric effect was associated with oligomeric proteins. However, the allosteric behavior has been also observed within a single protein domain [5]. Recently it was suggested that all dynamic proteins could demonstrate allosteric properties [6].

The mechanisms of information transfer range from enthalpically driven conformational changes to merely entropically driven motions or can be a combination of both effects [7]. Inter-residue communication, the change of intrinsic dynamics of protein, and long-range correlated motions can contribute to signal transduction in proteins depending on the properties of a particular system. Dynamics, flexibility and ligand-induced conformational changes in biological macromolecules have become an active area of research [8–10]. The role of long-range correlated motions and the mechanism of signal transduction were studied for ubiquitin [11], calmodulin [3], maltose/maltodextrin-binding protein [12], lipase [8], and other proteins [13]. It has been recognized that the conformational and global dynamics correlates to enzymatic function [13–15]. Several theoretical and experimental studies have shown that the catalysis and specificity are not controlled by a few residues, but rather are the property of entire protein framework [16], and the enzymatic activity depends on a subtle interplay between chemical kinetics and molecular motions [13].

Here, we use molecular dynamics (MD) simulations to investigate the intramolecular signal transduction induced by the ligand binding to trypsin. Trypsin belongs to the family of serine proteases, which are intensively investigated experimentally and theoretically [17–22] and they are well suited for the detailed analysis of protein dynamics and for testing the new methods of analysis. Signal transduction is a key phenomenon of the allosteric regulation of protein activity. The allosteric effect was found in trypsin-like serine proteases, including factor IXa and thrombin [23–25]. For example, the binding of inhibitor alters the conformation of the 60-loop near the active site of thrombin that leads to enhanced

* Corresponding author. Tel.: +7 843 2319037; fax: +7 843 2927347.
E-mail address: ermakova@mail.knc.ru (E. Ermakova).

http://dx.doi.org/10.1016/j.jmgm.2014.02.001
1093-3263/

reactivity with bovine pancreatic trypsin inhibitor (BPTI) [23]. Heparin modulates the 99-loop of factor IXa affecting its functional activity [24]. Allosteric inactivation of a trypsin-like serine protease points to new therapeutic strategies aimed at inhibiting or activating the enzyme [22,25].

Trypsin catalyzes the peptide or ester bonds hydrolysis of substrates containing Arg or Lys residues. Trypsin has 223 amino acid residues arranged in two six-stranded beta barrels. The residues H57, D102, and S195 form a catalytic triad, whereas the residues 189–195, 214–219, and 224–228 form a primary substrate-binding pocket called S1 binding pocket. The specificity of trypsin is determined by the amino acid residues at positions 189, 216, and 226. The residues 185–188 and 221–224 form two loops near the S1 pocket, called L1 and L2, respectively [16].

The dynamics of trypsin alone and in complex with ligands was studied before [19,22,26,27]. Experimental and theoretical studies have shown that the S1 substrate-binding pocket and the two loops outside the binding pocket (L1 and L2) significantly affect the enzyme activity and substrate specificity of trypsin [17,20]. Targeted molecular dynamics was used to demonstrate that the activation and deactivation of a-chymotrypsin begins with the movement of loop (215–225) pulling on loop (186–194) [20].

Here, we analyze the several components of signal transduction mechanism: how the binding of ligand changes the residue-residue interaction energy and the dynamics of protein and what role the long-range correlated motions play in the mechanism of signal transduction. These findings provide additional information for the development of new specific drugs.

MD simulation is a powerful instrument for studying the dynamics of different systems. However, the analysis of MD trajectories is a complicated and time-consuming procedure. We use several different computational approaches to analyze protein motions and compare their advantages and shortcomings. We use the Gaussian network model (GNM) [13,14,28] and anisotropic network model (ANM) [29,30] to determine the motions occurring at low-frequency modes. The dynamic cross-correlation matrix (DCCM) and principal component analysis (PCA) [31,32] based on the analysis of covariance matrix obtained from MD trajectory are used to investigate long-range correlated motions in protein. The analysis of residue-residue interaction energies and their correlation [33,34] provides new insights into the pathway of signal transduction.

2. Materials & methods

2.1. *MD simulations*

The coordinates of trypsin in complex with bovine pancreatic trypsin inhibitor (BPTI) were obtained from the Protein Data Bank (PDB ID: 3OTJ.pdb) [35]. The structure of free trypsin was obtained by removing the BPTI molecule from the complex. Comparison of the structures of unbound trypsin (PDB entry code 1S0Q.pdb) and trypsin in the complexes with BPTI (3OTJ.pdb) and with small inhibitor (1O2T.pdb) reveals that change of the protein backbone is very little. The root-mean square deviation (RMSD) of backbone atoms of ligand-bound trypsin with respect to unbound protein is about 0.3–0.4 Å. BPTI is a native inhibitor of trypsin and N^{α}-benzoyl-L-arginine ethyl ester (BAEE) is a specific trypsin substrate, which is widely used for experimental study of trypsin activity.

The structure of substrate N^{α}-benzoyl-L-arginine ethyl ester (Figure S1) was constructed and was optimized using PM3 method [36]. The structure of the complex of trypsin and BAEE was constructed by manual substitution of the BPTI molecule with BAEE so that N, Cα, C atoms of backbone and Cβ, Cγ, Cδ atoms of the side chain of K15 residue of BPTI were superimposed with corresponding atoms of Arg residue of BAEE. In that case, carbonyl oxygen atom of the substrate is fixed at the oxyanion hole, the guanidine group is located near to D189 and the ester C—O bond is in vicinity of S195.

MD simulations were performed using NAMD 2.6 [37] software with CHARMM27 force field parameters for the proteins [38], CHARMM36 all-atom carbohydrate force field [39] for BAEE and TIP3P for water molecules. Simulations were carried out under NPT ensemble conditions with periodic boundary conditions and Particle Mesh Ewald (PME) electrostatics using a 2 fs time step. The Langevin piston method was used to maintain a constant pressure of one atm with a piston period of 100 fs and a piston decay of 50 fs. The temperature was maintained at 300 K. Langevin dynamics maintained the temperature at 300 K with a damping coefficient of 1.0 ps^{-1}. The trypsin alone or in the complex with ligands was solvated in a box of $80 \times 94 \times 74$ $Å^3$. Following the initial minimization, heating, and equilibration stages for 2 ns, the production phase of each MD simulation ran for 20 ns. Coordinates were saved every 10 ps for further analysis. The overall stability of each simulation system was evaluated by monitoring the time evolution of the root-mean square deviation (RMSD) of backbone atoms with respect to the initial structure (Figure S2). Trajectories were analyzed using the VMD [40] and CHARMM [41] program packages. GNM, ANM and PCA were carried out with ProDy procedure [42] implemented in VMD program. Default values of force constants and distance cut-offs were used for the ANM (15 Å), for the GNM (10 Å) analyses. All molecular images were prepared using the VMD.

2.2. *Binding energy calculation*

The binding energy was calculated as the difference between the energy of the complex and the energy of the individual components from a single trajectory of the complex according to:

$$< \Delta E_{\text{bind}} > = < \Delta E_{\text{VdW}} + < \Delta E_{\text{elec}} + \Delta E_{\text{polar}} >$$

where < > denotes an MD-averaged quantities; E_{VdW} stands for the van der Waals interaction energy; E_{elec} stands for the electrostatic interaction energy and G_{polar}, the polar contribution of solvation energy, is the energy required to move the system from a vacuum dielectric ($\epsilon_{\text{in}} = 1.0$) to an aqueous dielectric ($\epsilon_{\text{out}} = 78$). The sum of electrostatic and polar solvation energies (PB)

$$\Delta E_{\text{PB}} = \Delta E_{\text{elect}} + \Delta G_{\text{polar}}$$

was calculated using a Poisson–Boltzmann implicit continuum solvent model [43] and the CHARMM [41] program. In CHARMM calculations, the grid spacing was set to 0.45 Å, and the longest linear dimension of the grid was extended at least 20% beyond the protein. Despite the MM/PBSA method is sensitive to simulation protocols, such as charge models, force fields, the solute dielectric constant it is widespread used for the study of protein–ligand interaction. Applicability of the method is intensively discussed in the literature [44,45]. Entropy analysis was not included in this work.

2.3. *Dynamic cross-correlation matrix*

Motional correlations between/among residues were analyzed by examining the dynamic cross-correlation map (DCCM) of Cα atoms for trypsin in the ligand-bound and unbound states. Protein structures from the trajectory were superimposed onto the reference structure in order to remove an overall translational and rotational motion of protein.

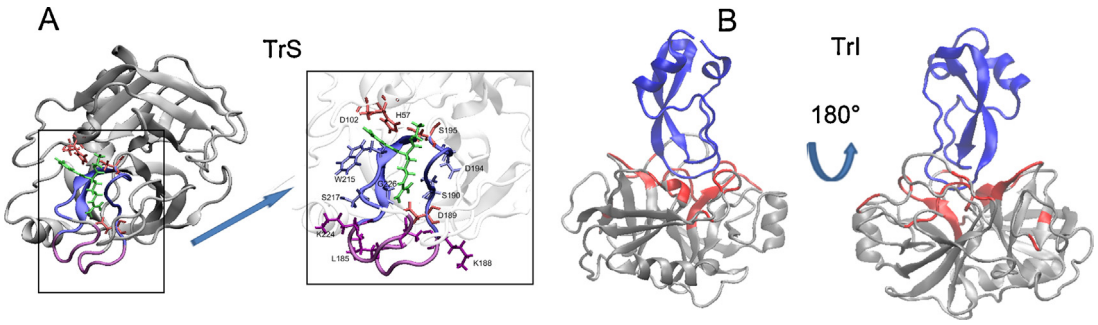

Fig. 1. A. The structure of the trypsin–substrate complex (TrS). The substrate is shown in green. The residues of the active center (H57, D102, and S195) and D189 are shown in red, the S1 pocket is shown in blue, and the loops L1 and L2 are shown in magenta. B. The structure of the complex trypsin–inhibitor. The inhibitor is shown in blue; the residues of trypsin that directly interact with the inhibitor are colored in red.

The cross-correlation matrix element C_{ij} of atoms i and j was calculated as:

$$C_{ij} = \frac{< \Delta\vec{r}_i(t)\Delta\vec{r}_j(t) >}{(< (\Delta\vec{r}_i(t))^2 >< (\Delta\vec{r}_j(t))^2 >)^{1/2}}$$

where < > denotes an MD-averaged quantity. The displacement from the average MD position of atom i is given by:

$$\Delta\vec{r}_i(t) = \vec{r}_i(t) - < \vec{r}_i(t) >$$

2.4. Interaction energy correlation

The interaction energies between amino acid residue pairs were calculated as a function of time using the 20 ns MD trajectory. Only residue-residue interactions with average interaction energies greater than 1 kcal/mol were analyzed [33,34]. The correlation between two sets of residue–residue interaction energies *i,j* and *k,l* is defined as:

$$CE_{ij|kl} = \frac{< \Delta E_{ij}(t)\Delta E_{kl}(t) >}{(< (\Delta E_{ij}(t))^2 >< (\Delta E_{kl}(t))^2 >)^{1/2}}$$

where

$$\Delta E_{nm}(t) = E_{nm}(t) - < E_{nm}(t) >$$

and $E_{nm}(t)$is the electrostatic energy between protein residues m and n.

We used the method of a projection of the interaction energy correlation matrix on the residue space (residue correlation matrix), which is based on calculating energy correlations between all interacting residue pairs in the MD ensemble of protein conformers [33,34]. We slightly modified this approach. To create this interaction correlation matrix, we first calculated electrostatic interaction energies between all residue pairs as a function of time during the MD trajectory and then determined equal-time correlations for the interaction energies. To visualize these pairwise correlations, we created a matrix, in which each element was equal to the average, absolute value of the correlation coefficient for any pair of residues. For example, if the interaction energy between residues 102 and 79 correlates with interaction energy between 117 and 16 residues (102,79|117,16) with coefficient of 0.8, and for pair (102,153|16,72) correlation coefficient equals 0.6, the projection on the residue space between residues 102 and 16 would be equal to 0.7. In the residue correlation matrix, the columns and rows represent a specific residue in the protein. Accordingly, we obtained a 223 × 223 normalized matrix.

2.5. Dynamical network analysis

The dynamical network analysis was carried out as described in ref. [46]. Two nodes were considered connected if any heavy atoms from the two residues were within 4.5 Å from each other during of 75% length of the trajectory. To define shortest pathways, the length of edges was calculated using the correlation coefficient between two residues. The distance between two nodes decreases as the correlation between them increases. The communication between different communities occurs through a smaller number of critical edges. Although Sethi et al. [46] have shown that the changes in the parameters lead to minor changes in the community distribution, the critical nodes can vary. The proposed algorithm [46] does not include the edges between nodes from adjacent residues so that the closest neighbors can belong to two different communities. For example, the residues 194, 196 can belong to one community, while the residues 195, 197 can be included in a different community. Therefore, we included these neighbor nodes into the consideration of critical nodes, because the interactions between them are strong and correlated. However, we included only those residues that "pretend" to be critical in all three studied systems.

3. Results and discussion

3.1. Ligand–trypsin complexes.

We used equilibrium molecular dynamics simulation to compare the conformational dynamics of trypsin (Tr) and of its complexes with substrate N^{α}-benzoyl-L-arginine ethyl ester molecule (TrS) and with bovine pancreatic trypsin inhibitor molecule (TrI). The structures of TrS and TrI complexes are shown in Fig. 1.

The TrI complex is stabilized by noncovalent protein – ligand interactions and by the hydrogen bonds between the residues Y39-I19, F41-R17, N97-R39, D189-K15, S190-K15, G193-K15, S195-K15,

Table 1
Averaged along trajectory some geometric parameters and standard deviations (SD) for free trypsin (Tr), trypsin–inhibitor (TrI) and trypsin–substrate (TrS) complexes (in Å).

Distance	Tr	SD	TrI	SD	TrS	SD
His57N – Ser195H	2.3	0.6	2.0	0.1	2.0	0.1
His57N – Ser195O	3.1	0.4	2.9	0.1	2.9	0.1
Asp189CA – Asp194CA	10.2	0.2	9.8	0.2	9.9	0.2
Asp189CA – Ser195CA	12.2	0.3	11.7	0.3	11.7	0.2
Asp189CA – His57CA	19.9	0.4	18.7	0.5	18.5	0.4
Ser190CA – Ser195CA	8.5	0.3	8.1	0.3	8.1	0.2
Glu192CA – Gly216CA	8.4	0.5	10.6	0.4	9.6	0.5
Gly193CA – Ser214CA	9.7	0.4	10.1	0.3	9.9	0.3
Trp215CA – Gly219CA	9.3	0.4	7.4	0.7	9.1	0.4

S214-K15, and G219-K15. In the TrS complex, the substrate forms hydrogen bonds with the H57, D189, S190, G193, S195, and G219 residues. In both complexes, the oxygen atom of ligands is fixed at the oxyanion hole with averaged distances equal to 2.3 Å, 1.9 Å, and 2.2 Å, 1.9 Å for distances O-G193NH, O-S195NH for complexes with inhibitor and substrate, respectively. The RMSD calculated for all backbone atoms of averaged structures for the free and ligand-bound trypsin is about 1 Å. As expected, the largest RMSD difference is observed for the loops. Additionally, the active center shows small but measurable changes. Formation of the complexes makes the hydrogen bond between H57 and Ser195 stronger (Table 1), the hydrogen bond slightly decreases due to ligand binding but the standard deviation decreases significantly (from 0.4 to 0.1 Å). This result agrees with the earlier observations of Katz et al. [47], who did not find significant conformational changes in trypsin upon inhibitor binding but detected small adjustments in 50 residues. The authors analyzed 7 complexes of trypsin-like proteases with inhibitors and 38 structures of free trypsin-like serine proteases and found that the distance between H57 and S195 was equal to 2.60 Å for complexes versus 3.06 Å, on average, for free serine proteases.

The S1 pocket of trypsin also changes upon ligand binding. In TrI and TrS complexes, the distances between the Cα atoms of the residues located on the opposite pocket walls of the S1 pocket (e.g. distance between 191 and 216 (191/216), 192/215, 193/214) increase, while the distances between Cα atoms of the residues located on the side walls of the pocket (e.g. distances 189/195, 189/57, 215/219) become smaller. Table 1 lists some of those distances. Guvench et al. [27] have also shown that the trypsin-binding pocket residues remain relatively rigid regardless of whether there is no ligand, a high-affinity ligand (benzamidine), or a low-affinity ligand (tranylcypromine).

3.2. Ligand binding-induced changes in residue-residue interactions.

Small structural changes can be difficult to evaluate but they can be very important for the protein activity. Even small structural changes can induce significant modifications in the residue interaction energies. Vice versa, the changes of residue-residue interaction energies reflect the conformational rearrangements of the protein structure. Thus, we calculated the van der Waals (VdW), the electrostatic interaction energies, and the polar contribution of solvation energy to the ligand-protein interactions. Table 2 lists the trypsin residues giving a large contribution to ligand binding. The Poisson–Boltzmann (PB) characterizes the sum of electrostatic energy and the polar contribution of solvation energy. The substrate primarily interacts with the residues of active center, i.e. H57, S195, with residue I16, and with the S1 pocket residues 189–194, 213–219. Our observations agree with the results of Chi et al. [48]. Authors showed that the residues H57, L99, S195, T215, and C220 play a major role in the binding of oxytetracycline to trypsin. Additionally, inhibitor interacts with the loops (residues 58, 60, 96–99, 151, 193) and β -strands (residues 39–42, 216–217), further stabilizing the complex. The inhibitor forms more stable Michaelis–Menten complex than the substrate. Van der Waals interactions give the main contribution to protein–ligand interaction energy for TrS and TrI complexes. The Poisson–Boltzmann energy gives a smaller contribution. It appears to show that the electrostatic and solvation energies nearly cancel each other. D189 has the largest contribution to the electrostatic interaction energy for both complexes, TrI and TrS.

The ligand binding leads to significant changes of inter-residue interaction energies in the ligand-binding site of trypsin (e.g. 189, 190, 195, and 219). Interestingly, it also leads to the changes of interaction energies of equal and even greater magnitude between the residues outside the binding site, e.g. 71/77, 60/37, 102/94, 186/222. The changes of residue-residue interaction energy upon ligand binding reflect the changes in trypsin structure.

Fig. 2A demonstrates the difference of the electrostatic interaction energies of the trypsin–inhibitor complex and free trypsin. The largest differences are indicated by boxes. Fig. 2B exhibits the amino acid residues, whose interaction energies change due to ligand binding. The residues whose interaction energy alters greater than 2 kcal/mol are colored yellow. The residues in direct contact with ligand are colored red. The residues distant from the binding site that show significant energy changes are colored orange. Similar changes are induced by substrate binding (Figure S4), but the substrate has smaller effect on the energetic landscape of protein. However, the changes of residue-residue interaction energy are observed in both complexes. First, the active center residues change their interaction energy and interaction energy with the residues located in S1 pocket. Second, residues in direct contact with ligands

Table 2
Energetic contributions and standard deviations (SD) to ligand binding for individual amino acids of trypsin in the complexes with ligands.

Residue number	TrI (kcal/mol)				TrS (kcal/mol)			
	VdW	SD	PB	SD	VdW	SD	PB	SD
Ile16			−3.1	0.3			1.3	0.3
Tyr39	−7.2	1.1	−0.8	0.4				
Phe41	−4.6	0.8	−0.5	0.4	−0.6	0.3		
His57	−8.1	1.0	−1.0	0.9	−2.9	1.1	0.6	0.8
Cys58	−1.8	0.4			−0.7	0.3		
Lys60	−2.2	0.6	2.8	0.7				
Ser96	−2.0	1.3						
Leu99	−3.2	0.9						
Asp102	−0.5	0.1	−1.5	0.7			−0.5	0.4
Tyr151	−4.8	1.3						
Asp189	2.1	1.7	−6.7	2.7	3.2	2.2	−9.3	1.6
Ser190			−2.7	1.3			1.4	0.8
Cys191	−2.7	0.5			−3.0	0.5	−0.5	0.3
Gln192	−10.3	1.2	−0.6	0.9	−3.9	1.0		
Gly193	−2.4	0.9	−0.6	0.5			−0.9	0.5
Asp194	−1.6	0.3	−3.4	0.4	−1.1	0.2	−1.4	0.3
Ser195	−2.4	0.7			−1.7	0.7		
Ser214	−2.0	0.8	0.9	0.8	−1.0	0.6	0.6	0.4
Trp215	−5.3	1.0	−1.1	1.0	−3.6	1.0		
Gly216	−2.4	0.6			−2.7	0.6	0.6	0.4
Ser217	−1.9	0.4			−0.9	0.4		
Gly219	−2.4	1.1	−0.8	1.3			−1.6	1.2
Cys220	−1.3	0.3			−1.5	0.3		

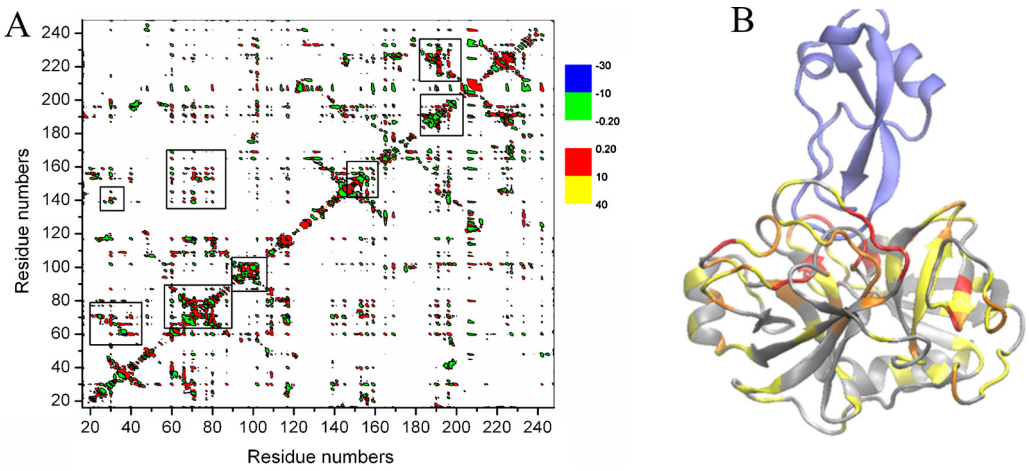

Fig. 2. A. The difference of the average interaction energies between the residues of trypsin caused by the inhibitor binding (complex–free state). Zones with largest difference are indicated by boxes. B. The residues of trypsin whose interaction energies changed due to ligand binding. All residues whose interaction energy changes are greater than 2 kcal/mol are colored in yellow. The red color shows the residues in direct contact with ligand, and the orange color shows the residues that are far from the binding site, but their energy changes are greater than 5 kcal/mol. Numeration of residues in trypsin begins from 16.

also change the interaction energy with the rest of protein, especially with the residues in close proximity. Detailed analysis of the ligand binding-induced changes in the residue-residue interactions one can see in the Supplementary.

The most interesting change is that of the interaction energy between the loops 185–188 and 221–225. The binding of substrate slightly enhances the interaction energy between L1 and L2 loops, whereas the binding of BPTI decreases it significantly. In the trypsin–inhibitor complex, the interaction energy of D189 with charged residues of the loop 219–225 decreases. The effect of the surface loops L1 and L2 on the enzymatic activity was shown before in experiments of conversion of trypsin to chymotrypsin [16,17]. Surprisingly, a strong increase of the van der Waals and electrostatic interaction energies between Q30 and S139 is observed in TrI complex although these residues are located on the β -strands of different barrels. Role of S139 in trypsin is not well investigated, but mutation of neighboring residue I138 is an important substitution for converting the trypsin to elastase [49].

The alteration of the residue-residue interaction energies is observed throughout the protein (Fig. 2B). The residues in direct contact with ligands also change their interaction energy with the rest of protein, especially with the adjacent residues. The changes

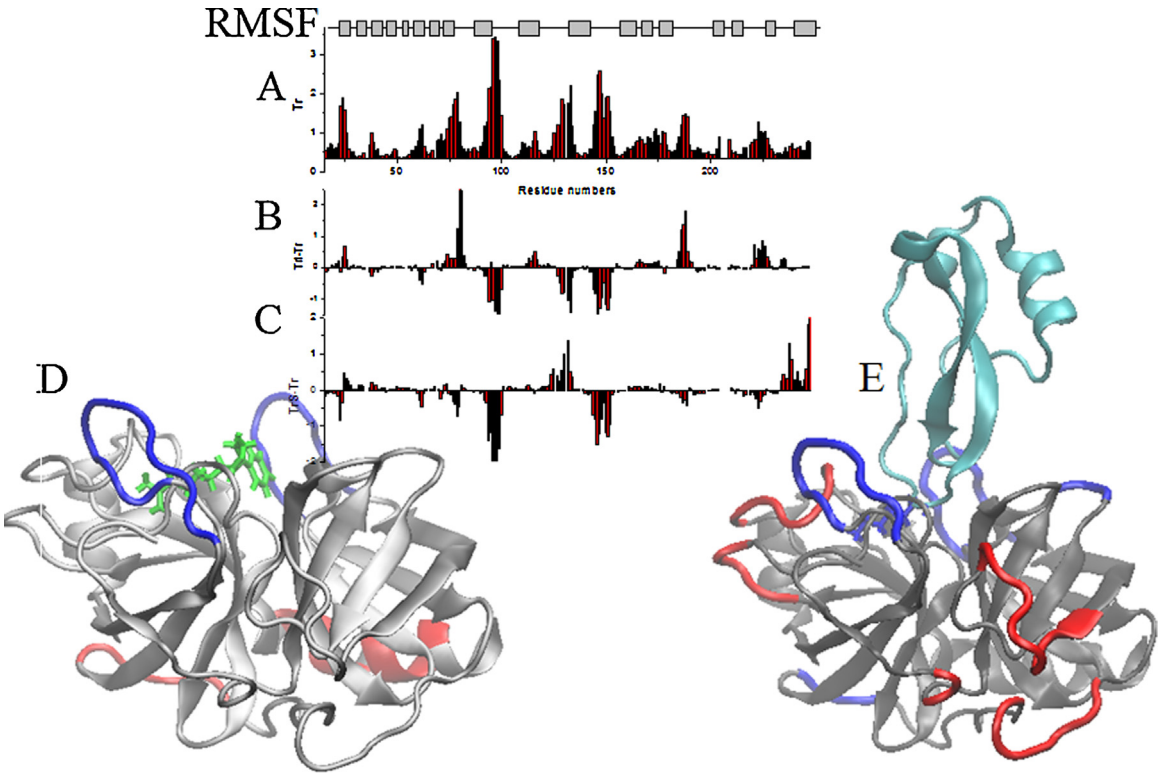

Fig. 3. A. RMSF values for Cα atoms are shown for the ligand-free trypsin. B and C. The difference of RMSF values for the trypsin–ligand complex and for free trypsin. The difference of the RMSF values for the trypsin–inhibitor complex and free trypsin (B). The difference of the RMSF values for the trypsin–substrate complex and free trypsin (C). D and E. The changes of RMSF values are highlighted for the trypsin–substrate complex (D) and for the trypsin–inhibitor complex (E). Blue and red colors indicate the decreasing and increasing RMSF values, respectively, as the result of ligand binding.

of electrostatic contribution are not very strong, confirming that structure does not demonstrate significant conformational changes. Nevertheless, these changes propagate from the ligand-binding site to the remote regions and may give a contribution into intramolecular signal transduction mechanism. The analysis of the protein energetic landscape shows that the residues located in the loops and β -strands are involved in this process.

3.3. Ligand binding-induced changes of backbone dynamics.

Fig. 3A shows the root-mean square fluctuations (RMSF) of Tr about the average atomic positions of the backbone. The average positions of trypsin backbone atoms were calculated from the MD snapshots after the least-square alignment of the trypsin backbone atoms. As expected, the loops represent the most mobile regions of the structure. Free trypsin demonstrates the large flexibility of the loops 94–99 and 146–151. The ligand binding reduces the flexibility of these loops, but increases it in other regions of protein. The binding of inhibitor enhances the flexibility of the loops 23–26, 79–81, 185–189, 221–227 (Fig. 3B). The binding of substrate increases the flexibility of the loop 124–132 (Fig. 3C). Fig. 3D and E demonstrate the regions of protein whose flexibility decreases (colored in blue) or increases (colored in red) as a result of ligand binding. Ligand binding decreases the flexibility of some regions of protein, whereas increases in the other regions (entropy-entropy compensation). The averaged RMSF values of backbone slightly decrease upon ligand binding.

The RMSF describes the fluctuations of amino acids averaged over MD trajectory. There are several methods to extract functionally relevant collective motions from the molecule structure. The information obtained by these methods partially overlaps and partially is complementary. Gaussian network and the anisotropic network models allow one to decompose the global dynamics into a set of modes, to identify different types of motions allowed by the structure and to separate the high- and low-frequency vibrations. In contrast to RMSF, these methods reflect the coupling/interdependence of residue motions and characterize correlated motions. These methods are based on the elastic network model, which describes the protein as a network of nodes. In both methods, the interaction between the nodes is modeled by a harmonic potential if the distance between them is less than the critical distance (*R* cutoff). A detailed explanation of both methods is given in refs. [10,13,14,28]. In this study, the Cα atoms of protein were chosen as nodes that describe the motion of the corresponding residues. The diagonalization of the matrix of the interaction potentials gives eigenvalues and eigenvectors, which are the orthogonal and normal vectors characterizing the contribution of each residue to the corresponding mode. The analysis of eigenvectors allows the estimation of the coupling between the different parts of the protein.

We performed the GNM and ANM analyses of free trypsin and its complexes with ligands. Fig. 4 displays the results of the GNM analysis of the equilibrium dynamics of trypsin (see also the Supplementary). Panel 4A compares B-factor predicted by the GNM (red curve) for TrI with the X-ray crystallographic B-factor (black line) taken from 3OTJ.pdb [35] file. The correspondence between the experimental and the GNM-predicted magnitude of fluctuation is remarkable (correlation coefficient is 0.77), indicating that the GNM model correctly predicts/reproduces the dynamics of trypsin.

The GNM gives $N-1$ eigenvalues and eigenvectors, where N is the number of nodes. The eigenvector of the first slowest mode of free trypsin is shown in Fig. 4B. The first mode characterizes the global motion of protein, i.e. the relative motion of the two barrels. Almost all residues participate in this motion; all loops demonstrate high mobility. It is very important to examine how the active site residues move in this mode. The residues 57, 102, 195 are indicated by blue arrows in Fig. 4B. Yang et al. [13] conducted the GNM analysis of several group of proteins and assumed that the catalytic sites should occupy the positions near the global hinges in the global modes. Serine proteases were not included in the set of the investigated proteins [13], but our analysis of the slowest mode of Tr demonstrates that catalytic residues 57, 102, and 195 have a restricted mobility in agreement with the results of Yang et al. [13]. Furthermore, according to work of Bahar et al. [28], the residues that are responsible for ligand distinction and binding demonstrate high mobility in slow modes in unliganded state. It is well known that the residue 189 plays a key role in serine proteases. D189 is located in the ligand-binding site of trypsin and demonstrates a moderate flexibility to accommodate the ligand. Fig. 4B also demonstrates that the residues 216–217 and 96–100 can play the role of recognition residues (red arrows). It is known that the residues at the positions 216 and 189 define the specificity of serine proteases, while the residue L99 is important for the S2 pocket [16]. G216 and S217 may play a role in the recognition of the ligand, whereas the less exposed residue D189 aids to fix the position of substrate in the active site. Thus, the predictions based on GNM analysis are in very good agreement with known experimental data and this method can be used for the analysis of other members of serine protease family.

The GNM analysis of the complexes shows the change of intrinsic dynamics of protein in the presence of ligand. The GNM ranges obtained results with eigenvalues, but the difference between eigenvalues can be small and the order of the eigenvectors can change in different systems. We compare the eigenvectors of different systems using the correlation coefficient. Fig. 4C depicts the dependence of the absolute value of the correlation coefficients

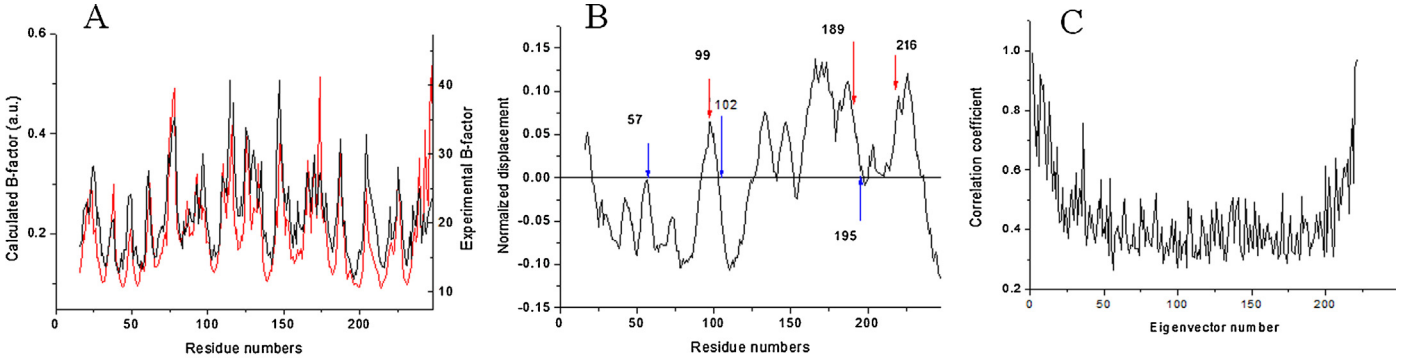

Fig. 4. GNM analysis of trypsin and its complexes. A. B-factor predicted by the GNM (red curve) and X-ray crystallographic B-factor (black line) for TrI. B. Eigenvector of the first slowest mode of free trypsin. Residues 57, 102, and 195 are indicated by blue arrows and the residues 99, 189, and 216 are shown by red arrows. C. Dependence of the absolute values of the correlation coefficients between the eigenvectors of free trypsin and the complex trypsin–substrate on the eigenvector number.

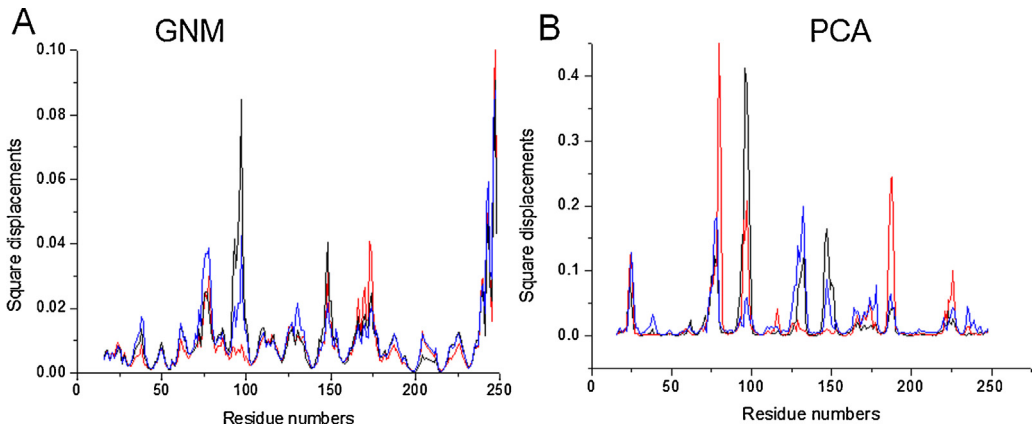

Fig. 5. A. The sum of the five slowest modes predicted by the GNM analysis for trypsin (black), the complex TrI (red), and the complex TrS (blue). B. The sum of the five first modes with the largest amplitudes predicted by PCA. Arrows point to the active center residues (57, 102, 189, and 195).

between the eigenvectors of free trypsin and trypsin–substrate complex on the eigenvector number. A strong correlation is observed for slow modes and for the fastest modes. While the vibrational low frequency mode corresponds to collective motions, the higher frequency modes correspond to local deformations and describe the local motion of separate residues or the groups of residues. Fast motions were shown to be associated with the stability of protein [28]. In the case of trypsin, these motions characterize the symmetric and anti-symmetric fluctuations of the β -strands residues. Here we analyze only slow modes.

Fig. 5A and Figure S6 demonstrate the low-frequency vibrations in free trypsin (black line) and in the complexes with substrate (blue line) and with inhibitor (red line) calculated as a sum over the first five modes by GNM and ANM, respectively. Figure S7 depicts the three slowest modes of ligand-free and ligand-bound trypsin obtained by ANM. Both methods show that the loops 76–78, 92–99, 146–151 and the C-terminal residues of free trypsin exhibit the larger amplitude of motion than other loops. The sum of five first modes of the complexes differs from that of the free trypsin only by the amplitudes observed for the loops 92–99, 146–151. In free trypsin the loops 92–99, 146–151 move almost independently. Both ligands decrease the contribution of these loops to the slowest modes.

In general, GNM and ANM methods based on the analysis of the structure of the ligand-free and ligand-bound trypsin predicts that dynamics of the protein should be changed due to ligand binding. The ligand binding to free trypsin suppresses the motion of the loops 146–149 and 92–99 moving independently, gives rise to the redistribution of energy, and thereby the amplitude of the motion of other loops increases and, most importantly, the motion of all loops becomes more correlated. Therefore, the ligand-binding initiates the perturbations of inter-residue interaction energy and the dynamics which propagate to the remote sites of the protein.

Recently [50] it was shown that anisotropic network model predictions concur with global motions exhibited by proteins in micro- to milliseconds simulations and ANM modes favor transitions between most probable substates of protein. Considering the simplicity of the GNM and ANM methods and the speed of calculations, we believe that these methods are valuable for the preliminary analysis of protein dynamics, especially of large multidomain proteins or proteins that undergo significant structural changes upon the complex formation.

3.4. *Correlated fluctuations between/among residues.*

We investigated motional coupling between the residues by using the covariance analysis of coordinate displacements averaged over the simulation trajectory. This analysis provides an insight into the correlated fluctuations and the long-range dynamic coupling of residues. In order to analyze the fluctuations of Cα atoms about their average positions during the course of simulation, we conventionally classified them as the “high frequency” (HF) and “low-frequency” (LF) fluctuations. LF fluctuations were defined as fluctuations with frequencies less than 25 GHz and were extracted from the signal using the Fourier transformation technique, and HF fluctuations were defined as the difference between the total signal and LF. That separation allows one to improve the signal/noise ratio.

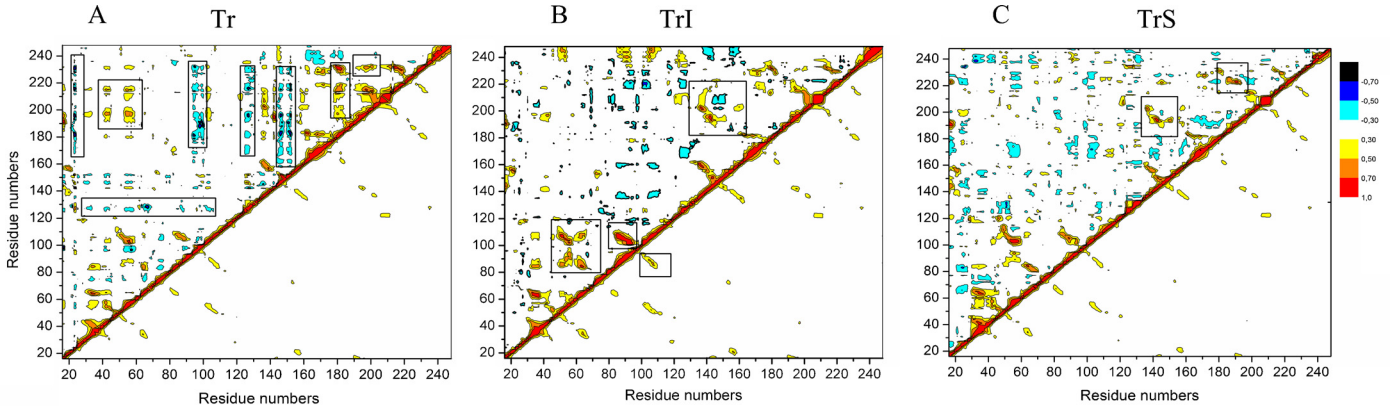

Fig. 6. Dynamics cross-correlation matrices for free trypsin (A), the complex with inhibitor (B), and the complex with substrate (C). The lower triangles indicate the DCCM for high frequency fluctuations and the upper triangles show the DCCM for low frequency fluctuations.

The results are presented in the two dynamics cross-correlation matrices (DCCM), expressed as the mean-squared correlation from −1.0 to 1.0. In Fig. 6, the lower triangles depict the DCCM for the HF fluctuations and the upper triangles show the DCCM for LF fluctuations. The correlations can be positive or negative as relative motions can occur in the same or in opposite directions, respectively. The covariance maps for Tr, TrI and TrS are shown in Fig. 6A, B, and C, respectively. In all complexes, several short lines of high correlation across and perpendicular to the diagonal are observed. These lines correspond to the β -sheets in the protein structure. Additionally, there are two small nodes corresponding to the alpha helixes.

At HF zone, the correlated motion of β -strands and the correlation between the hydrogen-bonded loops and the loops connected by ion pairs 16/189, 72/153, 142/193, 123/234, and 186/222 are observed. This is consistent with the finding of GNM and ANM analyses that the high frequency fluctuations describe the correlated motions of β -strands. In the high frequency zone, the correlation coefficients are not large and are equal to 0.3–0.4.

The LF fluctuations of the residues in β -strands and the hydrogen-bonded loops are larger by the absolute values than HF fluctuations. Additionally, the correlation between β -strands of different barrels becomes apparent. The electrostatic interactions and cysteine bridges lead to the correlated motion of the loops. We consider the fluctuations with the correlation coefficient greater than 0.5. The most interesting fluctuation is the fluctuation of the active center residues, the residues of S1 pocket and of the L1 and L2 loops. Figure S8 shows the active center residues and the residues correlated with them in free trypsin. A strong correlation is observed between the active center residues H57, D102, and S195. Furthermore, long-range correlations between the active center residues and the residues 16, 40–45, 180–182, 210, and the S1 pocket residues become evident. Not only the loop residues, but also the residues of β -strands are involved in the correlated motion of active center (Figure S8). In free trypsin, the long blue zones in Fig. 6A characterize the correlated motion of the loops 22–24, 96–101, 125–129, 146–152 relative to the protein. The correlation coefficient between D189 and the active center residues is less than 0.5. However, there is a strong correlation between the active center residues and the S1 pocket residues, as well as between the S1 pocket residues and D189, and the loop L1. The residues of S1 pocket can transfer the information from D189 to the active center residues and vice versa. From a catalytic point of view, presumably, the correlated motion of residues of the active center and S1 pocket is an efficient way to optimize the signal transduction and facilitate the catalytic process.

The comparison of Fig. 6A, B, and C indicates the changes that occur in protein as a result of ligand binding. The HF fluctuations undergo small changes. At LF, the loops 22–24, 91–99, and 146–152 and the residues 177–199 demonstrate the largest alterations. The binding of ligand decreases the 91–99 and 146–152 loop motions, and the correlation between these loops and other residues disappears. In trypsin–ligand complexes, the negative correlation between the residues of the active center, H57 and D102, and the loop 96–99 is replaced by the positive correlation with the residues 88–96 resulting from the increase of the interaction energy between D102 and Y94. The increasing correlation between the adjacent β -strands 103–108 and 84–91 is observed in the complexes. Following ligand binding, the motional correlation of D189 with the residues of S1 pocket becomes stronger. The inhibitor slightly decreases, while the substrate slightly increases the correlation between H57 and D102. Interestingly, there is a strong correlation between the loop L2, and loop L1, and D189 in the complex with substrate. This correlation does not exist in the complex with the inhibitor or in free trypsin. It was shown that the cooperative motions of these two loops and the substrate-binding sites contribute to the activity and substrate specificity of trypsin [18,20].

Similarly to the DCCM, the principal component analysis is based on the analysis of the covariance matrix obtained from MD trajectory. PCA analyzes the motion of Cα atoms, but their motion results from all interactions, including the long-range electrostatic interactions and the interactions of side chains. The diagonalization of the covariance matrix constructed using the Cα atom displacements produces a set of eigenvectors and eigenvalues, which indicate the direction and the amplitude of motions, respectively. Each eigenvalue divided by the sum of all eigenvalues represents the relative contribution of a given mode to the total conformational variance observed during the simulation. In free trypsin, the ten first modes characterize 66% of all protein motions.

It was shown that only the motions characterized by several largest eigenvalues are important for protein functionality [31,32]. The first mode corresponds to the motion of the largest amplitude. PCA reveals that the functionally important loops are involved in the correlated motion due to the electrostatic interaction between residues. In addition to the loops 93–100, 143–148, and 150–153, the loops 126–133, 22–25, 184–188 are also involved in the first and second modes in Tr. The third mode shows an almost independent motion of the loop 94–98. The fourth and the fifth mode demonstrate an increasing contribution of the loops 74–80 and 218–225 in the correlated motion. The sum of the five first modes (Fig. 5B) reveals the leading role of the loop 93–100 in the fluctuation of free trypsin. The C-terminal alpha helix does not participate in the motion of large amplitude.

The inhibitor binding changes the profile of the first mode dramatically. The loops 184–188 and 78–82 have the largest contribution to the first and the second mode and participate in the correlated motion together with the loops 94–99, 218–225, 22–25. The sum of five modes displays the leading role of the loops 78–82 and 184–188 with a notable contribution from the loops 94–99, 218–224, and 22–25. The binding of inhibitor decreases the fluctuation of the loop 145–150 significantly and that of the loop 94–99 to a lesser degree, whereas it increases the fluctuation of the loops 78–82, 184–188, and 218–225.

The substrate binding exhibits the increasing mobility of the loops 124–133, 73–78, and 22–27 in the first mode. In comparison to the ANM result, the PCA reveals that the loops L1 and L2 give a noticeable contribution to the large-scale motion of trypsin. The ANM also demonstrated the motion of these loops, but their contribution was very small. The sum of the five first modes shows that the binding of substrate strongly affects the motion of the loop 94–99, but not the loop 145–150.

Overall, the ligand binding suppresses the characteristic motions in free trypsin and enhances the motion that includes the residues 186, 188, 222, and 225 and the loop 78–82. Taking into account the electrostatic interactions changes the contribution of main loops into the functionally important motions as compared to the analysis done with the GNM and ANM. Similarly to the GNM and ANM, the PCA reveals an increasing mobility of the loop 78–82 and the decreasing mobility of the loops 94–99 and 145–150 due to ligand binding. It is interesting to note that the active center residues have zero amplitude in the first modes (Fig. 5B) and can act as hinges or anchors in these motions.

3.5. Ligand binding-induced changes to the side chain dynamics.

The analysis of the backbone motion only can underestimate global dynamic events and most dynamics may be hidden in the side chains [2,3]. The analysis of the side chain dynamics can provide the new information on the correlated motion of residues and the couplings between distant amino acids. To analyze the dynamics of side chains, we calculated the correlations between

the residue-residue interaction energies using the method of Kong and Karplus [33,34]. This method is based on the calculation of energy correlations between all interacting residue pairs, and thus it reflects the coupling between residues due to the long-range electrostatic interaction of side chains. Additional advantage of this method is that the relative motion of molecules in complex does not affect the intramolecular pair interaction energy. Figure S9 illustrates the projection of the interaction energy correlation matrix on the residue space (residue correlation matrix) [33,34] for free trypsin and for the complexes of trypsin with inhibitor or substrate (for details, see Materials and methods). We calculated the electrostatic interaction energies between all residue pairs as a function of time over the 20 ns MD trajectories and used the results to determine the equal-time correlations of the interaction energies of any two pairs of residues throughout the molecule. Two residues are included in the analysis only if the average interaction energy between them is greater than 1 kcal/mol by magnitude. Covalently bonded residues and pairs, showing the absolute value of the correlation coefficient smaller than 0.5, were not included in the analysis. The residue correlation matrix shows the coupling between any two residues in the system. It visualizes the correlation between distant residues independent on the origin of this correlation.

Figure S9 reveals that electrostatic interaction between the side chains provides the correlation between the residues on the remote loops of protein (e.g. 60–66 and 153–156, 70–80 and 94–99) that was not detected by DCCM analysis. The correlation can be induced by the direct electrostatic interaction or can be passed over via the correlated motion with other residues, e.g. the amino acids of S1 pocket correlate with D102 and D189 and can transfer the information between them. Additionally to the information obtained by the analysis of DCCM, the projection of correlation matrix of the interaction energies on the residue space of free trypsin demonstrates that the signal transduction can be realized through alternate pathways. For example, the charged residues of the loop 70–80 have a correlated motion with D102, at the same time, with the residues of the loops L1 and L2, and with some residues in the S1 pocket. This correlation strongly increases in the complexes. Moreover, the ligand binding enhances the correlation of the functionally important amino acids. In the complexes, a strong correlation of the side chains of D102 and D189 is observed, which was not detected for backbone atoms. The comparison of the residue correlation matrices for Tr, TrI, and TrS shows that the ligand binding increases the correlation between functionally important residues D189 and D102 despite of more than 10 Å separation between them. The residues 70, 80, and 117, whose fluctuation is increased in the complexes, demonstrate an increasing coupling with the functionally active residues 102, 194, and 189. The substrate has a stronger effect on the correlation between D102 and D189 than the inhibitor despite the small size of the molecule.

Correlated motion of side chains can be an important component of the signal transduction and can affect the catalytic process. The mechanism, by which such correlations operate, is likely to involve the long-range electrostatic interactions between atoms of the side chains of protein.

3.6. Dynamical network analysis

The community network analysis [51–53] of free trypsin and its complexes was carried out using the VMD and CARMA programs [54]. The community network analysis defines the protein as the system of nodes connected in physical network. The simplest network consists of the sets of nodes and links/edges that connect the pairs of nodes. Each residue forms one node. The two nodes are considered connected if any of the heavy atoms from the two residues are within 4.5 Å of each other during the 75% of the trajectory. The Girvan–Newman algorithm [55] joins nodes together into communities. By definition, each community contains nodes that are highly intra-connected, while the different communities are poorly interconnected through a few critical pairs of nodes.

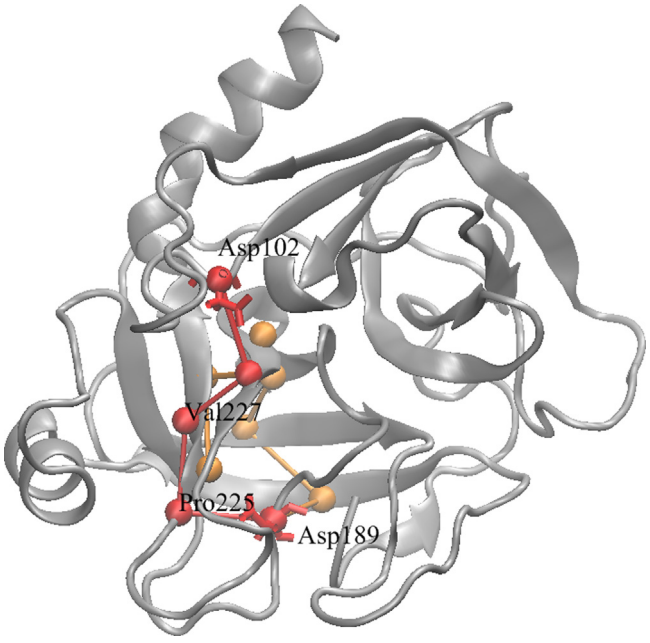

Fig. 7. Short pathways between the residues D102 and D189 of trypsin. The shortest path is colored in red.

The Girvan–Newman algorithm defines six communities for free trypsin and for the substrate–trypsin complex and eight communities for the TrI complex. There are three most interesting community. The first community of all three systems includes the active center residues despite they are located on the two different barrels. This community also contains the well-correlated residues of the neighbor β -strands. Both ligands add the loop 90–98 into the first community demonstrating the increase of communication between this loop and the active center residues. The second community incorporates I16, the residues 189–194, and the residues of S1 pocket and the loops 73–77, 139–157. Both ligands partially exclude the loop 73–77 from the second community confirming a more independent motion of this loop. The third community joins the residues of the loops L1, L2 and the residues of alpha helix. The inhibitor disrupts the communication between L1 and L2 and sets them into the two different communities.

The communication between different communities occurs through a smaller number of critical edges. Blocking of the communication between the critical nodes can interrupt the information transfer in the network. The analysis of the communication between the communities of three independent trajectories (see Materials and methods) reveals that the residues 30, 138, 141, 189, 194–198, 214, 215, 227, 228, and 237 have a large contribution to the communication in trypsin network and can be labeled as the critical nodes. The importance of some residues (i.e., the residue 189 and the residues of S1 pocket) is known and expected, whereas the role of other residues (i.e., 30, 138, 141, and 237) is not so evident. However, our results are in a good agreement with the observations of del Sol et al. [52] demonstrated that the substrate-binding specificity is regulated by a set of distributed residues in the structure of trypsin, acting in a cooperative manner by interchanging information. The authors noticed that in addition to well-known residues 189, 194, 227–228, the residues 212, 213, 138, 30, 46, and 141 also play key roles in long-range interactions.

Based on the community network analysis, short paths that connect any two residues by a series of noncovalent interactions can be calculated. Several short pathways between the residues D102 and D189 are shown in Fig. 7. For example, the shortest path can go through the residues 214, 227, 225 or 214, 227, 138, and 16. The existence of several pathways suggests that the signal transduction cannot be interrupted by a mutation. The involvement of the residues of S1 pocket in the signal transduction path between the active center residues and the substrate-binding site explains their importance for the catalysis. In the complexes, the shortest pathway can be decreased if the atoms of ligand participate in signal transduction, e.g. through the 102–214- ligand- 189 path. The participation of ligand atoms in signal transduction pathways and the increasing correlation between the active center and the substrate-binding site upon ligand binding allows us to assume that the ligands organize the space around them and prepare the protein for the reaction.

4. Conclusion

The structure and dynamics of trypsin alone and in complex with ligands were investigated using molecular dynamics simulations. Ligand–trypsin interactions have small effect on the structure of trypsin, but change the dynamics of several loops. The reduced mobility of the loops 90–99, 142–149 (these residues directly interact with the ligands) and increased mobility of the loop 74–78 (no direct interaction) due to ligand binding was predicted by GNM and ANM methods and was further confirmed by MD simulations. The reduced mobility of the loops due to ligand binding alters the flexibility (mobility) of the nearby loops, and thereby transmits the information to the surface loops.

We confirmed the coupling between the loops L1 (185–188) and L2 (221–224) and the residues in the active site and indicated an importance of other loops. Dynamical network analysis shows that in addition to well-known residues of S1 pocket (189, 194–198, 214–215, 227–228), the residues 30, 138, 141 also play a key role in communication between distant residues. The importance of S1 pocket in signal transduction between the active center and the substrate-binding site was further confirmed by the covariance matrix analysis. The residues of S1 pocket demonstrate strong correlation with the residues in the active center and the residue D189, and thus the information can be transferred from D189 to the active center and back.

The increased correlation between the active center and the substrate-binding site residues upon the binding of ligand suggests that the ligand organizes the space around it, prepares the protein for the reaction, and facilitates the catalytic process. The substrate enhances the interaction energy and the correlation between the loops L1 and L2, while the inhibitor attenuates them. We believe that mechanism of inhibition based on the regulation of flexibility of functionally important residues can be considered.

The comparison of several methods used in this study for the analysis of MD trajectory illustrates that these methods agree and complement each other. Given the simplicity of GNM and ANM methods and the speed of calculations, these methods are useful for the preliminary analysis of protein dynamics. In contrast, the analysis of the correlation between residue-residue interaction energies is a complicated and time-consuming method. However, at the same time, it is a unique method analyzing the conformational motion of side chains. In addition, the advantage of this method is that the relative motion of molecules in complex does not affect the intramolecular pair interaction energy. The comparative analysis of covariance matrices of backbone atom displacement done using the DCCM and PCA provides the insight into the long-range dynamic couplings of residues and the structure-dynamics relationship.

Despite small conformational changes caused by ligand binding, we believe that the combined, enthalpy–entropy mechanism of signal transduction takes place in trypsin. The ligand binding not only changes the residue-residue interaction energy, but also suppresses the mobility of loops leading to the propagation of perturbations in the protein. Additionally, the increased long-range correlated motions of functionally important residues contribute to the catalytic activity of protein, supporting the idea that the catalysis is defined by the entire framework of protein.

Acknowledgments

The support of Russian Academy of Sciences for Basic Research under Russian Foundation for Basic Research Grant N 12-04-01286-а is acknowledged. The authors thank the Minnesota Supercomputing Institute (University of Minnesota) and Joint Supercomputer Center of Russian Academy of Sciences for providing computer resources.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2014.02.001.

References

[1] D. Long, R. Bruschweiler, Atomistic kinetic model for population shift and allostery in biomolecules, J. Am. Chem. Soc. 133 (2011) 18999–19005.
[2] D. Kern, E.R.P. Zuiderweg, The role of dynamics in allosteric regulation, Curr. Opin. Struct. Biol. 13 (2003) 748–757.
[3] K.H. DuBay, J.P. Bothma, P.L. Geissler, Long-range intra-protein communication can be transmitted by correlated side-chain fluctuations alone, PLoS Comput. Biol. 7 (2011) e1002168.
[4] B. Alakent, S. Baskan, P. Doruker, Effect of ligand binding on the intraminimum dynamics of proteins, J. Comput. Chem. 32 (2011) 483–496.
[5] N.M. Goodey, S.J. Benkovic, Allosteric regulation and catalysis emerge via a common route, Nat. Chem. Biol. 4 (2008) 474–482.
[6] K. Gunasekaran, B. Ma, R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? Proteins: Struct. Funct. Genet. 57 (2004) 433–443.
[7] I. Rivalta, M.M. Sultan, N.S. Lee, G.A. Manley, J.P. Loria, V.S. Batista, Allosteric pathways in imidazole glycerol phosphate synthase, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) E1428–E1436.
[8] G.H. Peters, R.P. Bywater, Essential motions in a fungal lipase with bound substrate, covalently attached inhibitor and product, J. Mol. Recognit. 15 (2002) 393–404.
[9] D.-W. Li, D. Meng, R. Bruschweiler, Short-range coherence of internal protein dynamics revealed by high-precision in silico study, J. Am. Chem. Soc. 131 (2009) 14610–14611.
[10] Z.N. Gerek, S.B. Ozkan, Change in allosteric network affects binding affinities of PDZ domains: analysis through perturbation response scanning, PLoS Comput. Biol. 7 (2011) e1002154.
[11] R.B. Fenwick, S. Esteban-Martín, B. Richter, D. Lee, K.F.A. Walter, D. Milovanovic, S. Becker, N.A. Lakomek, C. Griesinger, X. Salvatella, Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition, J. Am. Chem. Soc. 133 (2011) 10336–10339.
[12] H.X. Kondo, N. Okimoto, G. Morimoto, M. Taiji, Free-energy landscapes of protein domain movements upon ligand binding, J. Phys. Chem. B 115 (2011) 7629–7636.
[13] L.-W. Yang, I. Bahar, Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes, Structure 13 (2005) 893–904.
[14] E. Marcos, R. Crehuet, I. Bahar, On the conservation of the slow conformational dynamics within the amino acid kinase family: NAGK the paradigm, PLoS Comput. Biol. 6 (2010) e1000738.
[15] J. Roy, C.A. Laughton, Long-timescale molecular-dynamics simulations of the major urinary protein provide atomistic interpretations of the unusual thermodynamics of ligand binding, Biophys. J. 99 (2010) 218–226.
[16] L. Hedstrom, Serine protease mechanism and specificity, Chem. Rev. 102 (2002) 4501–4523.
[17] L. Hedstrom, L. Szilagyi, W.J. Rutler, Converting trypsin to chymotrypsin: the role of surface loops, Science 255 (1992) 1249–1253.
[18] B. Wroblowsk, J.F. Díaz, J. Schlitter, Y. Engelborghs, Modelling pathways of a-chymotrypsin activation and deactivation, Protein Eng. 10 (1997) 1163–1174.
[19] E.L. Wu, K. Han, J.Z.H. Zhang, Selectivity of neutral/weakly basic P1 group inhibitors of thrombin and trypsin by a molecular dynamics study, Chem. Eur. J. 14 (2008) 8704–8714.

[20] W. Ma, C. Tang, L. Lai, Specificity of trypsin and chymotrypsin: loop-motion-controlled dynamic correlation as a determinant, Biophys. J. 89 (2005) 1183–1193.
[21] M. Perakyla, P.A. Kollman, Why does trypsin cleave BPTI so slowly? J. Am. Chem. Soc. 122 (2000) 3436–3444.
[22] R. Goncalves, N. Mateus, I. Pianet, M. Laguerre, V. de Freitas, Mechanisms of tannin-induced trypsin inhibition: a molecular approach, Langmuir 27 (2011) 13122–13129.
[23] A.R. Rezaie, X. He, C.T. Esmon, Thrombomodulin increases the rate of thrombin inhibition by BPTI, Biochemistry 37 (1998) 693–699.
[24] P.F. Neuenschwander, S.R. Williamson, A. Nalian, K.J. Baker-Deadmond, Heparin modulates the 99-loop of factor IXa: effects on reactivity of protein with Kunitz-type inhibitor domains, J. Biol. Chem. 281 (2006) 23066–23074.
[25] H.L.N. de Amorim, P.A. Netz, J.A. Guimarães, Thrombin allosteric modulation revisited: a molecular dynamics study, J. Mol. Model. 16 (2010) 725–735.
[26] B.O. Brandsal, J. Aqvist, A.O. Smalas, Computational analysis of binding of P1 variants to trypsin, Protein Sci. 10 (2001) 1584–1595.
[27] O. Guvench, D.J. Price, C.L. Brooks III, Receptor rigidity and ligand mobility in trypsin–ligand complexes, Proteins: Struct. Funct. Bioinform. 58 (2005) 407–417.
[28] I. Bahar, A.R. Atilgan, M.C. Demirel, B. Erman, Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability, Phys. Rev. Lett. 80 (1998) 2733–2736.
[29] A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, Biophys. J. 80 (2001) 505–515.
[30] E. Eyal, L.W. Yang, I. Bahar, Anisotropic network model: systematic evaluation and a new web interface, Bioinformatics 22 (2006) 2619–2627.
[31] A. Amadei, A.B.M. Linssen, H.J.C. Berendsen, Essential dynamics of proteins, Proteins 17 (1993) 412.
[32] L. Skjaerven, A. Martinez, N. Reuter, Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit, Proteins 79 (2011) 232–243.
[33] Y. Kong, M. Karplus, The signaling pathway of rhodopsin, Structure 15 (2007) 611–623.
[34] Y. Kong, M. Karplus, Signaling pathways of PDZ2 domain: A molecular dynamics interaction correlation analysis, Proteins: Struct. Funct. Genet. 74 (2009) 145–154.
[35] F.C. Bernstein, T.F. Koetzle, G.J.B. Williams, E.F. Meyer, M.D. Brice, J.R. Rogers, O. Kennard, T. Shimanouchi, M. Tatsumi, The protein data bank: a computer-based archival file for macromolecular structures, J. Mol. Biol. 112 (1977) 535–542.
[36] J.J.P. Stewart, Optimization of parameters for semiempirical methods I. Method, J. Comput. Chem. 10 (1989) 209–220.
[37] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (2005) 1781–1802.
[38] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.
[39] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D. Mackerell Jr., CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, J. Comput. Chem. 33 (2012) 2451–2468.
[40] W. Humphrey, A. Dalke, K. Schulten, VMD – Visual Molecular Dynamics, J. Molec. Graphics 14.1 (1996) 33–38.
[41] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217.
[42] A. Bakan, L.M. Meireles, I. Bahar, ProDy: protein dynamics inferred from theory and experiments, Bioinformatics 27 (2011) 1575–1577.
[43] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D.A. Case, T.E. 3rd Cheatham, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[44] T. Hou, J. Wang, Y. Li, W. Wang, Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations, J. Chem. Inf. Model. 51 (2011) 69–82.
[45] D.P. Oehme, R.T.C. Brownlee, D.J.D. Wilson, Effect of atomic charge, solvation, entropy, and ligand protonation state on MM-PB(GB)SA binding energies of HIV protease, J. Comput. Chem. 33 (2012) 2566–2580.
[46] A. Sethi, J. Eargle, A.A. Blacka, Z. Luthey-Schulten, Dynamical networks in tRNA: protein complexes, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 6620–6625.
[47] B.A. Katz, J. Finer-Moore, R. Mortezaei, D.H. Rich, R.M. Stroud, Episelection: Novel Ki – nanomolar inhibitors of serine proteases selected by binding or chemistry on an enzyme surface, Biochemistry 34 (1995) 8264–8280.
[48] Zh. Chi, R. Liu, H. Zhang, Noncovalent interaction of oxytetracycline with the enzyme trypsin, Biomacromolecules 11 (2010) 2454–2459.
[49] S.H. Hung, L. Hedstrom, Converting trypsin to elastase: substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity, Protein Eng. 11 (1998) 669–673.
[50] M. Gur, E. Zomot, I. Bahar, Global motions exhibited by proteins in micro- to milliseconds simulations concur with anisotropic network model predictions, J. Chem. Phys. 139 (2013) 121912.
[51] Z.N. Gerek, S.B. Ozkan, Change in allosteric network affects binding affinities of PDZ domains: analysis through perturbation response scanning, PLoS Comput. Biol. 7 (2011) e1002154.
[52] del Sol, H. Fujihashi, D. Amoros, R. Nussinov, Residues crucial for maintaining short paths in network communication mediate signaling in proteins, Mol. Syst. Biol. 2 (2006) 0019.
[53] A. Ghosh, S. Vishveshwara, A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 15711–15716.
[54] N.M. Glykos, Software news and updates. CARMA: a molecular dynamics analysis program, J. Comp. Chem. 27 (2006) 1765–1768.
[55] M.M. Girvan, Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 7821–7826.