
NEW APPROACHES TO SLOW DYNAMICS OF PROTEIN DOMAINS

S.O. YESYLEVSKYY, V.N. KHARKYANEN

UDC 577.3

©2009

Institute of Physics, Nat. Acad. of Sci. of Ukraine
(46, Nauky Prosp., Kyiv 03039, Ukraine)

Proteins are very familiar objects for biologists and biophysicists. However, from the point of view of the physicist the proteins are mysterious objects, which cannot be compared with any other object in nature. What really set the proteins apart from any other physical system are their motions. The time scale of protein motions covers 15–16 orders of magnitude and extends from 10^{-13} s to minutes or even hours. This is probably the broadest spectrum of motions observable in any other physical system of comparable size. The fastest motions are localized and mostly harmonic. In contrast, the slowest motions are collective (delocalized), strongly inharmonic and dissipative (diffusive). Functioning of the proteins in living cells require subtle balance between the motions with very different time scales. Large scale slow motions are critical for the functioning of numerous enzymes, transport proteins, molecular motors, ion channels and other proteins. In this work we provide systematic analysis of our recent developments in the field of slow protein dynamics based on the concepts of dynamic domains and fuzzy domains.

1. Introduction

Proteins are very familiar objects for biologists and biophysicists. However, from the point of view of the physicist, the proteins are mysterious objects which cannot be compared with any other object in the nature. What really set the proteins apart from any other physical system is their motions. The time scale of protein motions covers 15–16 orders of magnitude and extends from 10^{-13} s to minutes or even hours. This is probably the broadest spectrum of motions observable in any other physical system of comparable size. The fastest motions are localized and mostly harmonic. In contrast, the slowest motions are collective (delocalized), strongly inharmonic, and dissipative (diffusive). The functioning of the proteins in living cells requires a subtle balance between the motions with very different time scales. Large-scale slow motions are critical for the functioning of numerous enzymes, transport proteins [1–3], molecular motors [4], ion channels, and other proteins.

In this work, we carry on the systematic analysis of our recent developments in the field of slow protein

dynamics based on the concepts of dynamic domains [5, 6] and fuzzy domains.

2. Theory and Methods

2.1. The concept of protein domains

In many cases it is sufficient to determine the general character of slow motions of the protein without going into atomic details. The most fruitful concept in this respect is the concept of protein domains.

Domains can be loosely defined as quasi-independent parts of protein molecules serving as the structural blocks and functional units [7, 8]. Primarily, the term “domain” means a distinct structural block of the protein, but quite different criteria are presently used to identify this block [8–18]. Different definitions of domains and methods of their identification can be grouped around three key concepts: 1) Domain is a recognizable (often visually) substructure within a protein as a compact folded part of the molecule connected to other domains by very few structural elements such as a loop or a helix [13, 16–23]. 2) Domain is a part of the protein molecule that behaves in a quasi-independent manner with respect to the action of different factors inducing structural transitions in protein. 3) Domain is a relatively compact part of the protein that is characterized by its own pattern of intramolecular collective dynamics [5, 6]. Such domains will be referred as dynamic domains and will be discussed below.

The following terms are used hereon: The residue is the amino acid residue of the primary sequence of a protein. The cluster is the collection of residues which share a certain character of motion and behave more-or-less as a rigid body. The domain is the large cluster which often coincides with the protein domains identified by other techniques. The detailed discussion of these terms is given below.

2.2. The Gaussian Network Model

One of the most popular methods of the protein dynamics studies is the normal modes analysis (NMA) [17, 24]. However, all-atom NMA appears to be too slow, expensive, and excessively detailed for the majority of applications. Recently, the greatly simplified Gaussian Network Model (GNM) [25–29] became a popular method of choice in determining the character of large-scale motions in the folded proteins.

GNM can be viewed as an extremely simplified version of NMA, where realistic atom-atom interactions are substituted by the residue-level harmonic potentials [26]. GNM describes a protein as the network of identical springs which connect the C_α atoms of the residues located within the cut-off distance r_c . The equilibrium lengths of springs are assumed to be equal to the distances between C_α atoms in the x-ray structure. Normal modes of such a network of elastic interacting particles can be computed [26, 30].

Using the computed normal modes, the cross-correlations between the motions of any residue i and another one j , (c_{ij}), can be easily calculated in GNM [5, 25, 26]. Here, c_{ij} is a square matrix of size N , where N is the number of residues in the protein.

2.3. Identification and analysis of dynamic domains

2.3.1. The hierarchical clustering of correlation patterns (HCCP)

The HCCP method was designed as a technique which could allow a reliable identification of domains regardless of their spatial position and orientation in the complex proteins [5]. HCCP utilizes the correlation matrix c_{ij} obtained from GNM calculations or from other source. The matrix c_{ij} contains all information about the correlation of motions which can be extracted from the normal mode vibrations, but contains only pairwise correlations. Therefore, even the small changes of a protein structure can lead to the changes in the c_{ij} matrix. In order to eliminate this problem, we considered, instead of pairwise correlations, the *correlation patterns*. A single k -th column of the c_{ij} matrix contains the correlations of the given residue k with all other residues in the system. We will call such column-vector the *correlation pattern* of the residue k . The new matrix, the *correlation matrix of correlation*

patterns p_{ij} , can be defined as

$$p_{ij} = \frac{\frac{1}{N} \sum_{k=1}^N c_{ik} c_{jk} - \bar{c}_i \bar{c}_j}{\sigma_i \sigma_j},$$

where \bar{c}_i is the mean of the i -th column of the matrix c , and σ_i is the root mean square deviation of the i -th column of the matrix c . The p_{ij} matrix is of dimension $N \times N$, and its elements show to what extent the correlation patterns of elements i and j are similar in terms of linear correlation. The matrix p_{ij} provides a much more robust way of comparing the motions of residues than the conventional correlation matrix c_{ij} does.

On the next step, the residues with similar correlation patterns can be combined to larger clusters which share the same character of motion. Several such clusters can be further combined as having weaker motion similarities and so on. This idea is utilized in our hierarchical clustering procedure used to identify the domains. For this purpose, we developed a modified agglomerative clustering scheme with average linkage. In this scheme, the most similar clusters are merged (agglomerated) on each step to produce larger clusters. The pairwise similarity criterion is applied to all intercluster pairs and then averaged to calculate the similarity between the clusters. The details of the clustering algorithm are the following:

1. Each amino acid residue of the protein is assigned to be the simplest cluster of size 1.
2. Minimal v_{\min} and maximal v_{\max} elements of p_{ij} are found. The interval $(v_{\min}:v_{\max})$ is divided into M bins $v_{\max} > v_1 > v_2 > \dots > v_{M-1} > v_{\min}$ ($M = 1000$ in this study). The index of the current bin is set to $k = 1$.
3. The pair of residues, whose correlation $p_{ij} > v_k$, is found. If no such pairs exist, then the current bin index k is increased by 1, and step 3 is repeated.
4. Residues from the matching pair of residues are merged into a single cluster. The matrix p_{ij} is recalculated by the following rule:

$$p_{ij} = \frac{1}{m_i m_j} \sum_{k \in \{M_i\}} \sum_{l \in \{M_j\}} p_{kl},$$

where m_i and m_j are the numbers of elements in the clusters i and j ; M_i and M_j are the vectors of sizes m_i and m_j , respectively, which contain the indices of the residues in these clusters. In other words, the average correlation of all intercluster pairs is calculated. (In the

present study, we will use only p values as a measure of the correlation between two residues or clusters of residues. Thus, for the sake of simplicity, we will use the term “correlation of two clusters” instead of “correlation of the correlation patterns of two clusters” hereon.)

5. Step 3 is continued until all the residues are merged and the whole protein becomes the single cluster.

2.3.2. Domain stability criterion and determination of the most plausible number of domains

In the course of the HCCP clustering, the system goes through the stages with different numbers of clusters – from N to 1. In the course of clustering, the value of correlation gradually reduces from 1 to -1 in a series of small discrete intervals (bins). Each bin corresponds to a particular number of clusters in the system, which is stable on the current level. Let us assume that the state with M clusters appears on the bin number K_1 . Some of these M clusters can merge only if the correlation threshold becomes smaller than their cross-correlation. This happens on the bin number K_2 ($K_2 > K_1$). We call the length of the region ($K_1:K_2$) the *stability gap* defined as $g = K_2 - K_1$. It is possible to show that the stability gap indicates to what extent the corresponding effective number of clusters is resistant to environmental changes. We define the *most plausible number of domains* (N_{MPN}) as the number of clusters which is observed in the region of the largest stability gap.

2.3.3. Analysis of the properties of dynamic domains

The following steps were performed after the hierarchical clustering procedure:

1. The N_{MPN} and g_{MPN} were determined as described above.

2. The mean intradomain correlation was computed as

$$p_{\text{dom}} = \frac{1}{N_{\text{MPN}}} \sum_{k=1}^{N_{\text{MPN}}} \frac{1}{(N_k^2 - N_k)/2} \sum_{i,j \in \{D_k\}, i > j} p_{ij},$$

where N_k is the number of residues in the k -th domain; D_k is a vector which contains the indices of residues from the k -th domain.

3. The interdomain correlation was computed as

$$p_{\text{int}} = \frac{1}{(N_{\text{MPN}}^2 - N_{\text{MPN}})/2} \times$$

$$\times \sum_{k=1}^{N_{\text{MPN}}-1} \sum_{l=k+1}^{N_{\text{MPN}}} \frac{1}{N_k N_l} \sum_{i \in \{D_k\}, j \in \{D_l\}} p_{ij}.$$

4. Mean intradomain energy per residue was calculated as

$$E_{\text{dom}} = \frac{1}{N_{\text{MPN}}} \sum_{k=1}^{N_{\text{MPN}}} \frac{1}{N_k} \sum_{i,j \in \{D_k\}, i > j} E_{\text{DFIRE}}(s_i, s_j, r_{ij}),$$

where N_k is the number of residues in the k -th domain; E_{DFIRE} is the DFIRE-SCM energy between residues of the types s_i and s_j situated at the distance r_{ij} between their force centroids.

5. Mean interdomain energy per residue was calculated as

$$E_{\text{int}} = \frac{1}{N} \sum_{k=1}^{N_{\text{MPN}}-1} \sum_{l=1}^{N_{\text{MPN}}} \sum_{i \in \{D_k\}, j \in \{D_l\}} E_{\text{DFIRE}}(s_i, s_j, r_{ij}),$$

where N is the total number of residues in the protein.

All these energies were computed per residue to eliminate the effect of the variable size of the protein.

2.3.4 Modeling motion of the clusters in HCCP

Here, we present a method of simulating the motions of dynamic domains called *the method of hierarchical rotations*. HCCP subdivides the protein into the hierarchy of clusters defined on different levels of correlations of infinitesimal motions around a local energy minimum. The clusters of a particular hierarchical level are assumed to move as a whole relatively independently of one another. Thus, the internal degrees of freedom of the clusters can be considered frozen. For the sake of simplicity, each residue is represented by the C_α atom and the center of geometry of the side chain. We assume further that those clusters, which are physically connected by linkers, can move relative to one another if the number of linkers is one or two. If there is only one linker, it can serve as a pivot for almost free rotation characterized by two coordinates. If there are two linkers, the motion is the rotation about the axis connecting linkers or the rotation around a virtual pivot located between them [31].

Our procedure transforms the protein into a flexible system of rigid bodies of different sizes connected by pivots and axes. The number of moving clusters N_{cl} depends on the number of hierarchical levels N_L taken into account.

We model the evolution of this mechanical system by means of standard Metropolis Monte-Carlo simulations

performed using a custom software. For each protein, we performed several simulations with different numbers of hierarchical levels N_L . Each run starts from the crystallographic open state of the corresponding protein.

2.4. The fuzzy domains

In all conventional domain identification techniques (including HCCP technique [5,6]), each residue can belong only to one particular domain. This presumes that the boundaries between the domains are sharp – one domain ends abruptly on residue i and another domain starts on the next residue $i+1$. However, it is quite obvious that, in real proteins, some transition region can exist, where the residues are “shared” by two (or more) domains and both domains influence their motion. In other words, the boundary between the domains becomes “fuzzy” – each residue can belong to several domains at the same time.

Let us formalize this intuitive picture. Let us assume that the protein consists of N residues and can be subdivided into M domains which will be called *fuzzy domains* hereafter. Let us define the matrix of weights σ_{ij} which shows the extent to which residue i belongs to domain j . Each residue should belong to at least one domain, which leads to the obvious normalization condition

$$\sum_{j=1}^M \sigma_{ij} = 1. \quad (1)$$

In order to determine the values of all elements of σ_{ij} , we approximate the correlation matrix p_{ij} by the expression

$$p'_{ij}(\sigma) = \sum_{k,l=1}^M \sigma_{ik}\sigma_{jl} (2\delta_{kl} - 1), \quad (2)$$

$$\text{where } \delta_{kl} = \begin{cases} 1, & k = l, \\ 0, & k \neq l. \end{cases}$$

If residues i and j belong to the same domain ($k = l$), then their motions are *correlated* to the extent which is determined by their corresponding weights. If residues i and j belong to different domains ($k \neq l$), then their motions are *anticorrelated* to the same extent. The weights σ_{ij} should be chosen in such a way to give the best approximation of the real correlation matrix p . In order to do this, one should minimize the functional

$$F(\sigma) = \sum_{i \geq j}^N (p_{ij} - p'_{ij}(\sigma))^2 \quad (3)$$

in respect to the weights σ_{ij} . For the purpose of analysis, we also define the weight of each domain as $w_j = \frac{1}{N} \sum_{i=1}^N \sigma_{ij}$.

The number of variables in the optimization problem defined by (2) is NM , which is typically of the order of thousands. It is well known that the majority of multidimensional optimization problems with a large number of variables are usually ill-posed [32, 33]. In order to solve this problem, we used the Maximum Entropy Method (MEM) which is one of the most powerful and robust optimization techniques available to date [32, 34]. We follow the simplified formulation of MEM of Steinbach et al. [33]. In order to obtain the unique solution of the optimization problem (3), the auxiliary functional Q is composed as

$$Q = S(\sigma) - LF(\sigma), \quad (4)$$

where $S(\sigma) = \sum_{i=1}^N \sum_{j=1}^M \left[\sigma_{ij} - \sigma_{ij}^0 - \sigma_{ij} \ln \left(\frac{\sigma_{ij}}{\sigma_{ij}^0} \right) \right]$ is the Shannon–Jaynes entropy, L is the unknown Lagrange multiplier, σ^0 is the initial guess for σ . It is possible to prove [32, 33] that the maximization of functional (4) produces a unique solution of the initial optimization problem (3). The resulting set of values σ_{ij} possesses the minimal information content among all sets which are able to approximate the matrix p with the given accuracy [32, 33].

3. Results and Discussion

3.1. Test of the HCCP technique

In order to test the HCCP technique, we studied the relation between the intradomain correlation and the stability gap, by using the collection of 2022 non-homologous proteins from all major protein families [6]. Figure 1 shows the stability gap as a function of the intradomain correlation for the proteins from the test collection. The proteins with one domain, which are artificially split into two parts, form a tight group with the stability gaps below 300 and intradomain correlations scattered around 0.25–0.3. In contrast, the proteins with two and three domains are scattered along a well-defined line and exhibit the very strong correlation between p_{dom} and g_{MPN} . The three-domain proteins have systematically lower stability gaps in comparison with those of the double-domain proteins with the same p_{dom} . This shows that “artificial” domains in the single-domain proteins are different from “natural” domains –

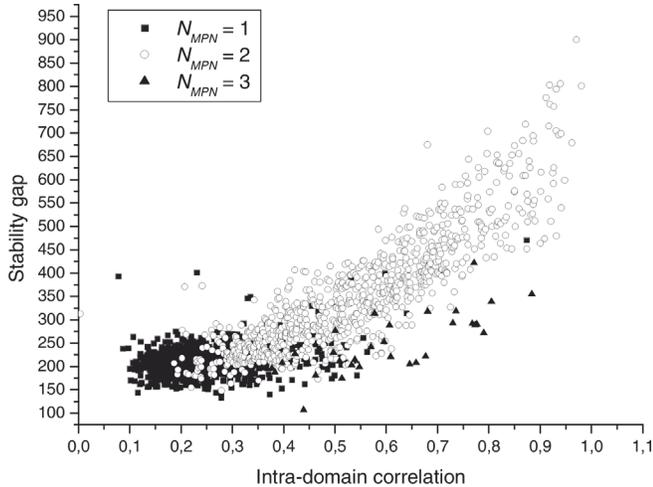


Fig. 1. Stability gap g_{MPN} as a function of the mean intradomain correlation p_{dom} for the proteins from the test collection

they are less compact and less stable. This allows concluding that our procedure of finding the most plausible number of domains allows distinguishing between real and spurious domains.

3.2. Effects of ligand binding on domain dynamics

The concept of dynamic domains allows determining the proteins which possess dynamic domains sensitive to ligand binding events [36]. We used 157 ligand-binding proteins which are crystallized in several forms, including ligand-free and ligand-bound forms. The procedure of selection of these proteins is described in details in our previous paper [36]. In order to study changes of the stability of domains, we computed the mean intradomain correlations p_{dom} for all available structures. The maximal difference in p_{dom} between the alternative structures of the same protein Δp_{dom} was computed. We also computed the root mean square differences (RMSDs) of C_{α} atoms between the crystal structures which possess maximal Δp_{dom} values.

All studied proteins can be classified into three groups, where the values of Δp_{dom} are insignificant, significant, or anomalously large. We inspected the crystal structures of the proteins from the “anomalous” and “significant” groups. Our analysis shows that the ligand binding is the reason of large Δp_{dom} values for only 5 proteins. Figure 2 shows the relationship between the whole-structure $RMSD_{dom}$ and the intradomain $RMSD_{max}$. There are few proteins which are clearly out of the regular trend. These proteins are calmodulin,

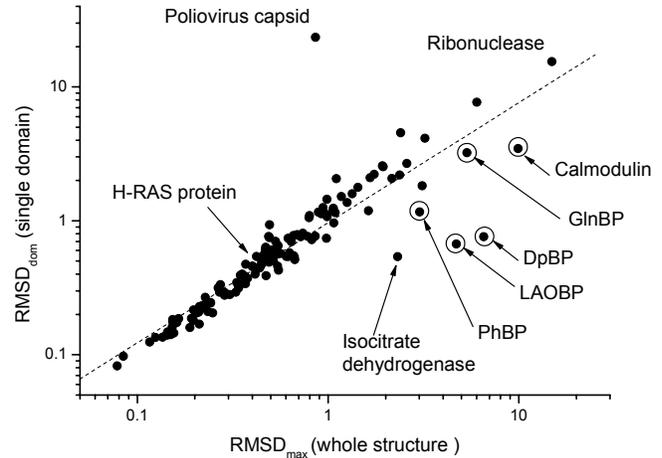


Fig. 2. Comparison of RMSD ($RMSD_{max}$) of C_{α} atoms between alternative structures of the same protein with the largest Δp_{dom} and the largest pairwise C_{α} RMSD between individual domains of alternative structures ($RMSD_{dom}$) of these proteins. The data are for 157 studied proteins. The proteins, for which the ligand binding is the reason for large Δp_{dom} values, are indicated by open circles. Linear fit is shown by the dashed line

DpBP, LAOBP, GlnBP, and isocitrate dehydrogenase. All these proteins, except isocitrate dehydrogenase, are identified above as having significant Δp_{dom} occurring upon the ligand binding. It was shown that these proteins can be classified as a promising target for the design of protein-based biosensors [36].

3.3. Simulations of the domain dynamics

We simulated the domain closure in several hinge-bending proteins using the method of hierarchical rotations [37]. We used six hinge-bending proteins which are crystallized in both open and closed states. The proteins LAOBP, GlnBP, DpBP, and EPSP can be classified as “easy targets”. The compact structure which emerges in our simulations is much closer to the crystallographic closed state than to the initial open state. The calmodulin and PGK protein can be classified as “hard targets”. In the case of calmodulin, such a behavior is expectable, because this protein is known to undergo a dramatic conformational change. Only 3% of simulations reach the closed state. In the case of PGK, no simulations reached the closed state. The reason for such a behavior is not clear in details, however several possible explanations were suggested [37].

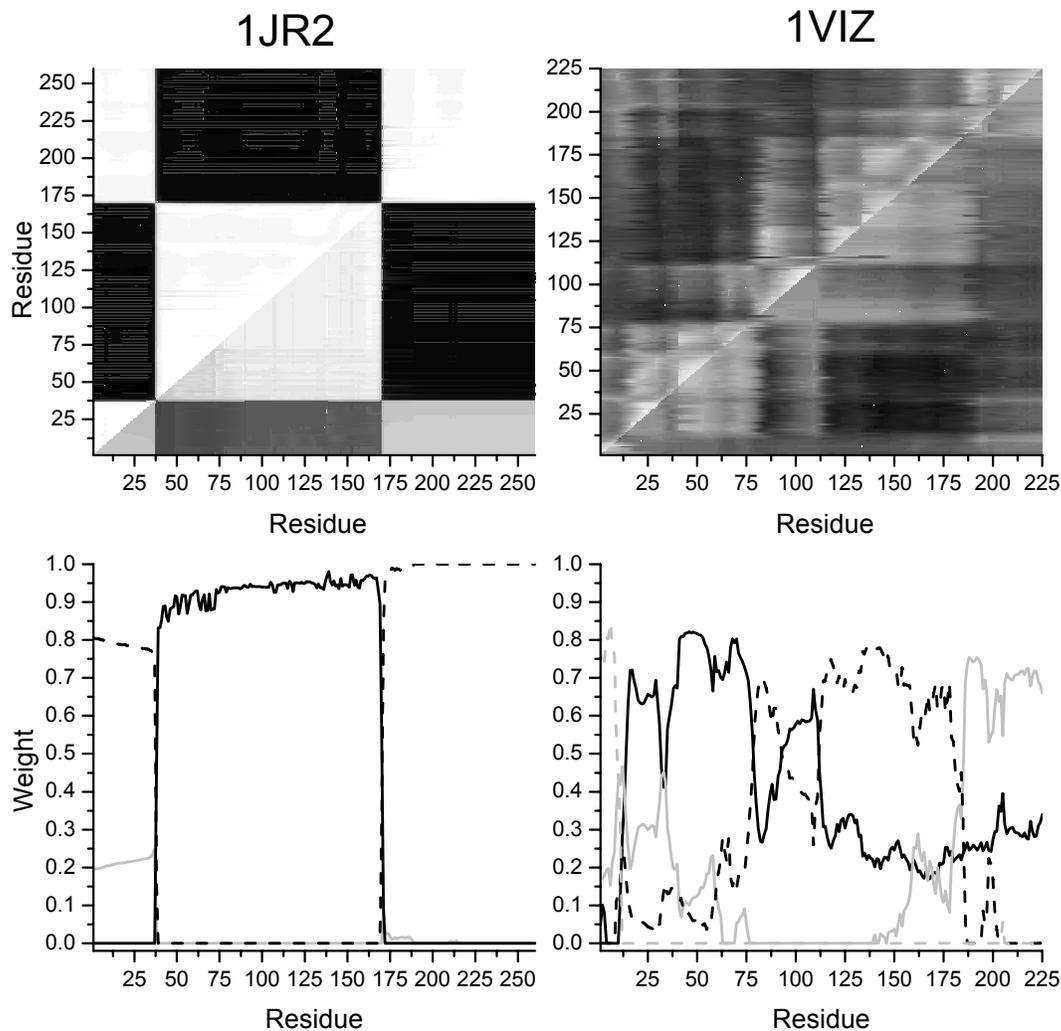


Fig. 3. Comparison of the correlation matrices p shown in the upper triangles of the matrices and approximated matrices p' shown in the lower triangles of the matrices (upper panels) and the weights of the fuzzy domains (lower panels) for two proteins with extreme patterns of their dynamics. The PDB codes of the corresponding proteins are shown. Different line styles correspond to different fuzzy domains in lower panels

3.4. Fuzzy domains

We identified the fuzzy domains in more than 1500 proteins and analyzed the general performance of the technique. Figure 3 shows that the general patterns of the correlation matrices are reproduced correctly by our method. The upper triangles of the matrices shown in Fig. 3 correspond to the initial correlation matrices p , while the low triangles to the fitted matrices p' . Two extreme cases of the correlation matrix patterns are shown. The lower panel of the Fig. 3 shows the weights of significant domains which correspond to the matrices shown in the upper panel. It is clearly seen that the

domains of 1JR2 protein are very well defined. There is almost no overlap between them, except for residues 1-35, where two domains coexist. The domain boundaries match the boundaries of the visually detectable blocks on the correlation matrix perfectly. In contrast, the domains of 1VIZ protein overlap with each other significantly for all residues. Particular residues can belong to up to three fuzzy domains simultaneously.

The visualization of fuzzy domains is quite challenging, because the domains “penetrate” into one another. However, it is possible to show each domain separately. Figures 4 and 5 illustrate this idea for proteins 1VIZ and 1LST. For each protein, four signifi-



Fig. 4. Visualization of four significant fuzzy domains of protein 1VIZ. Black color corresponds to the value of domain weight of 1, white to 0

cant fuzzy domains are shown in individual panels. Each residue is colored according to the weight of the given domain. Black corresponds to 1, white to 0. The ring-like structure of protein 1VIZ is subdivided into three fuzzy domains which gradually substitute one another. The fourth domain occupies a protruding loop. The structure of 1LST shows the transitions between the domains which are more independent and constitute the parts of well-defined dynamic domains divided by the dashed line in Fig. 5. These examples show that the concept of fuzzy domains allows one to model and to visualize the pattern of internal flexibility of the protein domains in a very clear and visually appealing way.

4. Conclusions

In this work, we have summarized our results in the area of the slow dynamics of protein domains. We have described two concepts for the physically consistent identification of the protein domains and their motions – the concepts of dynamic and fuzzy domains.

We have described the HCCP technique which allows identifying and analyzing the dynamic domains. We show how this technique identifies the most plausible number of domains and their boundaries with high accuracy. The mean intradomain correlation p_{dom} can be used as a *quantitative criterion* of domain independence and stability.

Using our HCCP technique, we analyzed the stabilities of dynamic domains of 157 ligand-binding proteins, for which several crystal structures are available. We

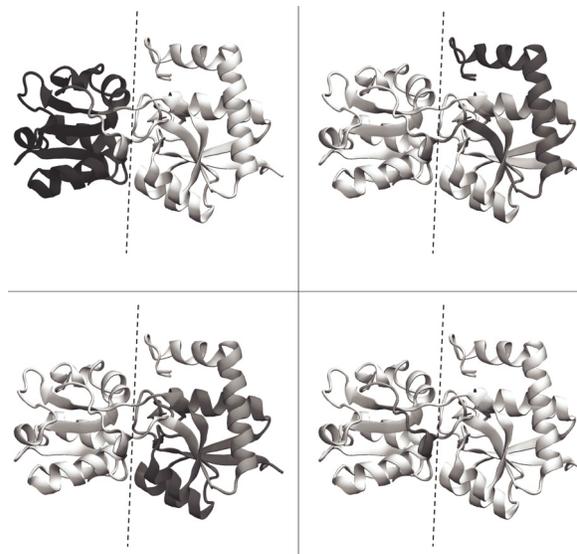


Fig. 5. Visualization of four significant fuzzy domains of protein 1LST. Black colour corresponds to the value of domain weight of 1, white to 0

demonstrate that five proteins (LAOBP, phosphate-binding protein, dipeptide-binding protein, glutamine-binding protein, and calmodulin) are likely to be a good target for the biosensor design, which opens a good biotechnological perspective for our method.

We developed the computational technique, which allows simulating the slow large-scale motions of proteins which lead to large-amplitude conformational changes. We performed the blind search for the closed states of 6 hinge-bending proteins starting from the crystallographic open states. Four proteins reach the closed state with almost 100% success rate. The obtained results show that our technique can be useful in the coarse-grained modeling of large-scale slow motions in proteins, in the search of alternative protein conformations, and in other areas of proteins science.

Finally, we have proposed the innovative method of domain identification based on the concept of fuzzy domains. Our method allows describing the fuzzy extended boundaries between the domains and gradual changes of the motion pattern from one domain to another. In our knowledge, our method is the first method of domain identification which allows describing fuzzy domain boundaries. The concept of the fuzzy domains can be used for the visualization purposes and in the coarse-grained simulations of the domain dynamics in order to account for the internal protein flexibility.

1. B.H. Oh, J. Pandit, C.H. Kang *et al.*, *J. Biol. Chem.* **268**, 11348 (1993).
2. C.D. Hsiao, Y.J. Sun, J. Rose *et al.*, *J. Mol. Biol.* **262**, 225 (1996).
3. A.V. Nickitenko, S. Trakhanov, and F.A. Quioco, *Biochemistry* **34**, 16585 (1995).
4. R. Craig and J.L. Woodhead, *COSB* **16**, 204 (2006).
5. S.O. Yesylevskyy, V.N. Kharkyanen, and A.P. Demchenko, *Biophys. Chem.* **119**, 84 (2006).
6. S.O. Yesylevskyy, V.N. Kharkyanen, and A.P. Demchenko, *Biophys. J.* **in press** (2008).
7. J. Janin and S.J. Wodak, *Adv. Protein Chem.* **61**, 1 (2002).
8. J. Janin and S.J. Wodak, *Prog. Biophys. Mol. Biol.* **42**, 21 (1983).
9. L. Falquet, M. Pagni, P. Bucher *et al.*, *Nucleic Acids Res.* **30**, 235 (2002).
10. B. Nagar, W.G. Bornmann, P. Pellicena *et al.*, *Cancer Res.* **62**, 4236 (2002).
11. L. Schmitt and R. Tampe, *Curr. Opin. Struct. Biol.* **12**, 754 (2002).
12. K.F. Fischer and S. Marqusee, *J. Mol. Biol.* **302**, 701 (2000).
13. C. Anselmi, G. Bocchini, A. Scipioni *et al.*, *Biopolymers* **58**, 218 (2001).
14. W.L. Nichols, G.D. Rose, L.F. Ten Eyck *et al.*, *Proteins* **23**, 38 (1995).
15. W. Wriggers and K. Schulten, *Proteins* **29**, 1 (1997).
16. S. Hayward and H. J. Berendsen, *Proteins* **30**, 144 (1998).
17. K. Hinsen, *Proteins* **33**, 417 (1998).
18. K. Hinsen, A. Thomas, and M.J. Field, *Proteins* **34**, 369 (1999).
19. N. Alexandrov and I. Shindyalov, *Bioinformatics* **19**, 429 (2003).
20. L. Holm and C. Sander, *Proteins* **19**, 256 (1994).
21. S. Hayward and N. Go, *Annu. Rev. Phys. Chem.* **46**, 223 (1995).
22. A.D. Schuyler and G.S. Chirikjian, *J. Mol. Graph. Model* **24**, 46 (2005).
23. S. Kundu, D.C. Sorensen, and G.N. Phillips, jr., *Proteins* **57**, 725 (2004).
24. M. Levitt, C. Sander, and P.S. Stern, *J. Mol. Biol.* **181**, 423 (1985).
25. A.R. Atilgan, S.R. Durell, R.L. Jernigan *et al.*, *Biophys. J.* **80**, 505 (2001).
26. I. Bahar, A.R. Atilgan, and B. Erman, *Fold. Des.* **2**, 173 (1997).
27. Y. Yildirim and P. Doruker, *J. Biomol. Struct. Dyn.* **22**, 267 (2004).
28. O. Keskin, *J. Biomol. Struct. Dyn.* **20**, 333 (2002).
29. P. Doruker, A.R. Atilgan, and I. Bahar, *Proteins* **40**, 512 (2000).
30. I. Bahar and A. Rader, *Curr. Opin. Struct. Biol.* **15**, 586 (2005).
31. S.O. Yesylevskyy, V.N. Kharkyanen, and A.P. Demchenko, *Proteins: Structure, Function, and Bioinformatics* **71**, 831 (2008).
32. J. Skilling, *Classic Maximum Entropy* (Kluwer Academic, Norwell, MA, 1989).
33. P.J. Steinbach, K. Chu, H. Frauenfelder *et al.*, *Biophys. J.* **61**, 235 (1992).
34. P.J. Steinbach, A. Ansari, J. Berendzen *et al.*, *Biochemistry* **30**, 3988 (1991).
35. G. Wang and R.L. Dunbrack, jr., *Bioinformatics* **19**, 1589 (2003).
36. S.O. Yesylevskyy, V.N. Kharkyanen, and A.P. Demchenko, *Biophys. J.* **91**, 3002 (2006).
37. S.O. Yesylevskyy, V.N. Kharkyanen, and A.P. Demchenko, *Proteins: Structure, Function, and Bioinformatics* **9999**, NA (2007).
38. H. Park, J.L. Hilsenbeck, H.J. Kim *et al.*, *Mol. Microbiol.* **51**, 963–971 (2004).
39. W. Wriggers, E. Mehler, F. Pitici *et al.*, *Biophys. J.* **74**, 1622 (1998).
40. B.E. Bernstein, P.A. Michels, and W.G. Hol, *Nature* **385**, 275–278 (1997).
41. M.A. Mathews, H.L. Schubert, F.G. Whitby *et al.*, *EMBO J.* **20**, 5832 (2001).

НОВІ ПІДХОДИ ДО ПОВІЛЬНОЇ ДИНАМІКИ ДІЛЯНОК БІЛКА

С.О. Єсильєвський, В.М. Харкянен

Резюме

Білки – звичні об’єкти для біології, однак з точки зору фізики вони є дуже незвичайними об’єктами, які не мають аналогів у природі. Білки вирізняються з поміж усіх інших фізичних об’єктів завдяки унікальним властивостям їх рухів. Характерні часи рухів білків розподілені на інтервалі в 15-16 порядків від 10^{-13} с до хвилин і навіть годин. Це, можливо, найбільш широкий спектр характерних часів з усіх фізичних об’єктів зіставного розміру. Швидкі рухи білків є локальними та гармонічними. Навпаки, найбільш повільні рухи білків є колективними, ангармонічними та дисипативними (дифузійними). Функціонування білків в живих клітинах потребує точного балансу між рухами з різними характерними часами. Крупномасштабні повільні рухи є критичними для функціонування ферментів, транспортних білків, молекулярних моторів, іонних каналів та інших білків. В даній роботі наведено систематичний аналіз останніх розробок у галузі повільної динаміки білків, які ґрунтуються на концепції динамічних доменів та нечітких доменів.