The Change of Protein Intradomain Mobility on Ligand Binding: Is It a Commonly Observed Phenomenon?

Semen O. Yesylevskyy,* Valery N. Kharkyanen,* and Alexander P. Demchenko†‡
*Department of Physics of Biological Systems, Institute of Physics, National Academy of Science of Ukraine, Kiev, Ukraine; †A.V. Palladin Institute of Biochemistry, Kiev, Ukraine; and ‡Research Institute for Genetic Engineering and Biotechnology, TUBITAK, Gebze-Kocaeli, Turkey

ABSTRACT Analysis of changes in the dynamics of protein domains on ligand binding is important in several aspects: for the understanding of the hierarchical nature of protein folding and dynamics at equilibrium; for analysis of signal transduction mechanisms triggered by ligand binding, including allostery; for drug design; and for construction of biosensors reporting on the presence of target ligand in studied media. In this work we use the recently developed HCCP computational technique for the analysis of stabilities of dynamic domains in proteins, their intrinsic motions and of their changes on ligand binding. The work is based on comparative studies of 157 ligand binding proteins, for which several crystal structures (in ligand-free and ligand-bound forms) are available. We demonstrate that the domains of the proteins presented in the Protein DataBank are far more robust than it was thought before: in the majority of the studied proteins (152 out of 157), the ligand binding does not lead to significant change of domain stability. The exceptions from this rule are only four bacterial periplasmic transport proteins and calmodulin. Thus, as a rule, the pattern of correlated motions in dynamic domains, which determines their stability, is insensitive to ligand binding. This rule may be the general feature for a vast majority of proteins.

INTRODUCTION

Dynamic domains are usually defined as relatively independent units, which maintain their integrity during the large-scale motions of proteins. This means that the interactions inside the dynamic domains are on average stronger than the interdomain interactions and that the motions within the domains are highly correlated in comparison with the motions of the elements of protein structure belonging to different domains. Domains can integrate and disintegrate in response to change of environment conditions, and this can change the whole pattern of protein motions. Identification of dynamic domains can be provided by several proposed techniques based on computer-assisted analysis of protein three-dimensional structures (1–6). One of the most advanced is the recently developed hierarchical-clustering-of-the-correlation-patterns (HCCP) method (7,8). This technique was especially designed to eliminate the influence of small random variations of the protein structure on the identification of dynamic domains. Thus, any changes of dynamic domains detected by this method are likely to be caused by systematic changes of conformation of polypeptide chains forming these domains.

One of the most important functions of proteins is ligand binding. In many proteins the ligand binding occurs in clefts formed by two or more structural domains and often leads to dramatic changes in their positions and orientations (9,10). However, it is not known if these changes also lead to significant alteration of the protein dynamics on the levels of the whole protein and of individual dynamic domains. Internal flexibility and stability of individual domains may change upon the ligand binding, but it is still obscure how extended and how general are these changes. Particularly it is not known if the presence of the ligand can change the effective number of dynamic domains (by merging several domains to a single rigid unit or splitting the domain into several independent blocks). It is necessary to emphasize that these changes should cause the alterations of the whole pattern of protein dynamics. It is hard to describe them based on the concept of static structural domains.

Intramolecular protein dynamics has been the subject of high interest over the last decades. Conventional experimental techniques have provided information about this dynamics, but this information has been quite limited on time- and length scales. For instance, in the studies using fluorescent or spin labels, only the dynamics of the residues in the close proximity to the labeled residue can be observed directly. Other methods, such as neutron scattering, provide only a very coarse-grained picture of overall protein dynamics. However, recent advances in NMR techniques, particularly the heteronuclear NMR relaxation methods, have allowed us to obtain dynamic data on an atomic level of detail, for a broad range of timescales (11–17). In particular, these methods allowed us to study the dynamics of such ligand binding proteins as human protein kinase B (11), ribosomal protein L11 (14), retinoid X receptor (18), chicken liver bile acid-binding protein (15), and many others. The dynamics of well-defined protein domains was also studied in the case of two-domain protein Pin1 (12). However, this latest technique remains quite complex, and can only be applied to a limited number of proteins.

Computer simulations may offer the fast and easy way of studying the whole pattern of protein dynamics on an

Submitted April 26, 2006, and accepted for publication June 30, 2006.

Address reprint requests to S. O. Yesylevskyy, Tel.: 044-525-9851; E-mail: yesint3@yahoo.com.

0006-3495/06/10/3002/12 $2.00

doi: 10.1529/biophysj.106.087866

extended timescale and with an atomic level of detail (2–4, 19,20). The additional advantage of the computational approach is the ability to work with a large number of diverse proteins to extract the universal principles of protein dynamics. In this line, the dynamics of several well-known ligand-binding proteins with and without the ligands was investigated by various computational techniques (21–23). However, to our knowledge, there were no systematic computational studies that aimed at investigating the dynamics of the large number of ligand-binding proteins with the same approach and in a comparative manner. The problems here are in applied methodologies. It is clear that this task cannot be accomplished by molecular dynamics (MD) simulations, because this technique can hardly be automated to study large number of diverse proteins, and to study slow dynamics requires a prohibitive amount of time.

Luckily, numerous ligand-binding proteins are crystallized in several conformations, typically with and without the ligand, or with several different ligands. The low-amplitude dynamics of these alternative conformations can be determined by the normal mode analysis (3,4,20) or its coarse-grained modifications such as Gaussian network model (GNM), anisotropic network model (24–29), or the rotations-translations of blocks approach (5). The GNM approach proved itself to be especially useful. Particularly, GNM was successfully applied to the analysis of different conformations of HIV-1 protease (30). This study revealed the robustness of dynamics and its weak dependence on the studied conformation. It was shown that the ligand binding changes the direction of motions in HIV-1 reverse transcriptase (31). Several other ligand-binding proteins were analyzed in different works (32,33). In principle, the normal modes themselves reflect any changes of low amplitude dynamics on any hierarchical level (individual residues, intra- and interdomain motions) caused by the change of conformation. However, there are several complications that make direct comparison of the normal modes of a protein in different conformations impractical:

1. Individual normal modes can be quite sensitive to small variations of the protein structure, which are caused by random perturbations or by the influence of crystallization conditions rather than by extensive conformational changes (8).
2. It is not clear how to compare the sets of normal modes. The normal modes are usually arranged by frequency in ascending order. However, the frequencies of the modes can be quite different in different conformations. So the modes, which are topologically equivalent, possess different numbers, and this makes their automatic comparison impossible (32). A similar problem is also observed in essential dynamics simulations (21).

In this respect, the comparison of dynamic domains seems to be advantageous, since it eliminates these complications. During the identification of dynamic domains in the HCCP method, each normal mode contributes to the values of residue-residue correlations of motion according to its frequency and amplitude, but the exact index of the mode is not important (see (8,24–29) for details). The contributions of all normal modes are additive, eliminating the problem of mode-swapping. The boundaries of obtained dynamic domains are insensitive to small random variations of the protein structures, thus any significant changes of domain boundaries are likely to be systematic effects of extensive conformational changes (8). Finally, HCCP technique allows us to estimate the degree of domain independence and stability not only objectively but expressed in quantitative terms (7). If the conformational change occurs within the domain and does not influence the position of the domain boundary it can still change the degree of correlated motions inside the domain. This will be detected by HCCP analysis.

The goal of the present work is to identify the dynamic domains in large number of ligand-binding proteins, which are crystallized in several different conformations, and to study the effects of the ligand binding on the properties of dynamic domains. For this purpose we selected 221 ligand binding proteins and scanned the PDB to find alternative crystal structures for each of them. One-hundred-and-fifty-seven proteins possessing ligand-bound and ligand-free crystal structures, or structures with several different ligands, were investigated. We show that the ligand binding has no significant effect on these properties in the majority of studied proteins. We have found only six proteins, whose domains possess a significantly different degree of intradomain correlation in the free and the ligand-bound forms. These proteins are known as classical examples of significant conformational changes associated with the ligand binding. Our data allows the conclusion that significant change of the properties of dynamic domains caused by ligand binding is most probably an exception observed in only few proteins, while the general rule is that the dynamic domains are almost insensitive to the ligand binding.

METHODS

The Gaussian network model

The Gaussian network model (GNM) (24,25,27–29) is a popular method of choice in determining the character of large-scale motions in the folded proteins. The detailed description of GNM can be found elsewhere (24,25). Here we present only its basic description.

GNM can be viewed as an extremely simplified version of NMA, where realistic potentials of the atom-atom interactions are substituted by the residue-level harmonic potentials (25). GNM describes the protein as a network of identical harmonic springs, which connect the C_α atoms of the residues located in close spatial proximity (within cutoff distance r_c) regardless of their positions in the sequence. Equilibrium lengths of these springs are assumed to be equal to the distances between C_α atoms in the x-ray structure and deviations from these distances are considered to be purely harmonic. Normal modes of this network of interacting particles can be computed easily. It was shown that GNM describes harmonic motions of the folded proteins surprisingly well, and produces results that are often indistinguishable from those of full-scale NMA (25,26).

Using the computed normal modes, the cross-correlations between the motions of any residue i with the other j (c_{ij}), can be easily calculated in GNM. This procedure is described in detail in original GNM articles (24,25) and in our previous work (8). Here, c_{ij} is a square matrix of size N, where N is the number of residues in the protein. This matrix is used for domain identification in our HCCP method.

The theory of HCCP

The HCCP method was designed as a technique for reliable identification of domains regardless of their spatial position and orientation in the complex proteins. It was shown that HCCP produces essentially identical domain assignments for different conformations of the same multidomain protein, providing that the domains maintain their integrity, but change relative position and orientation (8). HCCP allows obtaining quantitative description of correlations of motions inside the domains and cross-correlation of motions of different domains. This feature makes the HCCP method especially attractive for the study of domain stability and interdependence. Only a brief description of HCCP is given here. The details are described in our previous articles (7,8).

HCCP utilizes the correlation matrix c_{ij}, obtained from GNM calculations or from another source (full-scale NMA, molecular dynamics, essential dynamics analysis, etc.). To extract more information from the correlation matrices and to increase the robustness of the results, the HCCP uses not the initial pair correlations, but the correlation patterns—the essence of which is in the following. A single k^{th} column (or row) of c_{ij} matrix contains the correlations of the given residue k with all other residues in the system. We will call such a column-vector the correlation pattern of the residue k. The new matrix, the correlation matrix of correlation patterns p_{ij}, can be defined as

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

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 does the conventional correlation matrix c_{ij}. (The details can be found in (7,8).)

In the next step, the residues with similar correlation patterns can be combined into larger clusters that 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 for identifying the domains. For this purpose, we developed the modified agglomerative clustering scheme with the average linkage. In this scheme the most similar clusters are merged (agglomerated) on each step to produce larger clusters. A pairwise similarity criterion is applied to all intercluster pairs and then averaged to calculate the similarity between the clusters. The scheme of the HCCP algorithm is given in Fig. 1. (Further details can be found in (7,8).)

Domain stability criterion

HCCP provides a unique opportunity to estimate the degree of domain stability and interdependence. In our previous work we showed that the mean intradomain correlation p_{dom} can be considered as a quantitative measure of domain stability (7). This quantity can be calculated as

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

where N_k is the number of residues in the k^{th} domain, D_k is the vector that contains the indexes of the residues from the k^{th} domain, and N_D is the number of domains. In this work, $N_D = 2$ (see below).

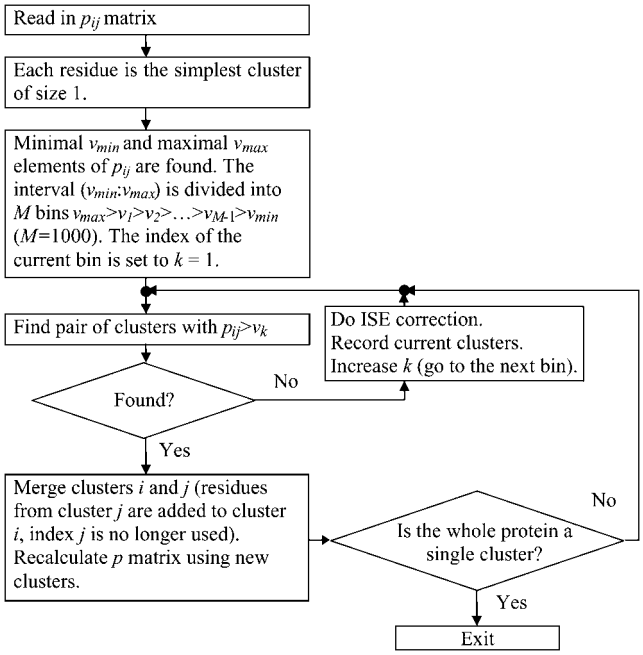

FIGURE 1 The basic scheme of the HCCP algorithm.

If p_{dom} is large, then the intradomain interactions are strong, while the interdomain bonds are weak. As a result, each domain moves independently in the diffusive manner. In contrast, if p_{dom} is small, then the interdomain bonds are strong and comparable to intradomain bonds. As a result, the domains move in an interdependent manner.

The number of domains

In the course of HCCP clustering the system goes through the stages with different number of clusters—from N to 1. On which stage can the clusters be identified as domains? In our previous work, we developed and applied an automatic procedure that determines the most plausible number of domains in the system (7). It was shown that the vast majority of the proteins in the PDB databank, including those used in this study, have one or two dynamic domains. Single-domain proteins have zero intradomain correlation, because, in this case, the domain is the entire protein, which is assumed not to move as a whole (see (7) for details). This means that the comparison of single- and multidomain proteins becomes meaningless. To avoid this complication we forced our program to calculate all correlations at the level of two domains for all proteins. The single-domain proteins were artificially split into two parts, which are less stable than the single native domain. Such treatment allows our describing both single and double-domain proteins by the same parameter—an intradomain correlation.

Selection of proteins

There are several publicly available databases of the ligand-binding proteins, such as PLD (34), Binding MOAD (35), or AffinDB (36). However, none of them contain information about the existence of multiple crystal structures for particular protein. To find such proteins, the following scheme was used:

1. The list of PDB identifiers of the proteins, which bind natural ligands, was obtained from the PLD database (34). This list contains 221 proteins (see Supplementary Material).
2. The database from the standalone PISCES protein culling package (37) was used to extract information about the chains currently stored in

PDB. This information includes the annotated list of all chains in PDB, the database of the pairwise BLAST sequence alignment scores of all nonredundant chains and the list of redundant chains, which correspond to each nonredundant one.

3. The list of PDB identifiers from the PLD was converted into the list of corresponding nonredundant chains. The latter was filtered to eliminate the chains, which share >99% sequence homology according to the BLAST database. The final list contained 178 chains, which were used as ''seeds'' for the groups of chains, which represent the same protein.
4. For each seed all sequences, which share >99% homology with it, were extracted from the BLAST database and added to this group.
5. For each chain in the group all redundant chains were added to the same group. This step is necessary because the redundancy is traditionally determined using the amino-acid sequence similarity, but not the three-dimensional structure. As a result, very different conformations of the same protein are often treated as redundant. For example, a closed form of the dipeptide-binding protein (PDB identifier 1WDN) is erroneously considered redundant and substituted by the open form of the same protein (1GGG).
6. The groups were filtered to eliminate multiple identical chains from the same PDB entry.
7. The median number of residues (the number found for majority of the chains) was determined for each group. Chains with lengths deviating from the median for >5% were eliminated.
8. Finally, 157 groups with more than one chain were maintained (see Supplementary Material).
9. All structures from each group were subjected to HCCP domain identification.

Calculation details

Steps 1–8 of the protein selection scheme above were performed using our own Perl scripts. Step 9 was accomplished using the HCCP program, described in our previous works (7,8) (available at http://www.geocities.com/yesint3/hccp.html). The cutoff of 7 Å was adopted for GNM Kirchhoff matrices (other values were also used to verify the sensitivity of results to the cutoff radius; see The Role of the GNM Cutoff Radius for details). The force constant in GNM is assumed to be 1 (this value only scales all of the eigenvectors and does not influence the normalized correlation matrices). All eigenvectors with nonzero eigenvalues were used for computing the correlation matrices. The intercalating-segments elimination procedure was used in HCCP as described in (7). This procedure is needed to avoid the ambiguous assignment of the residues in the proximity of the interdomain interface. Root mean-square difference (RMSD) calculations were performed in VMD (38) using our own TCL script, which reads and interprets the HCCP domain assignment data. All molecular graphics were rendered by VMD.

RESULTS AND DISCUSSION

Identification of the proteins possessing dynamic domains sensitive to the ligand binding

Only 157 out of 219 studied ligand-binding proteins appear to be represented by several distinct crystal structures. The number of alternative structures is very different for various proteins and ranges from 2 (glutamine binding protein, enolase, etc.) to 227 (trypsin), which is the direct consequence of redundancy existent in the Protein DataBank. The total number of structures subjected to HCCP calculations is 3336 (see Supplementary Material for full list). The list of proteins is provided in Supplementary Material. The data for each protein were analyzed separately. The 157 selected proteins belong to very different structural classes and folds. They bind a wide variety of ligands—from ions, drugs, and small metabolites (NADH, nucleosides, sugars, etc.) to peptides and other proteins. Therefore, these proteins form a representative subset of ligand-binding proteins stored in PDB.

To study the changes of stability and interdependence of the domains, we computed the mean intradomain correlations p_{dom} for all available structures. This parameter is different for different crystal structures of the same protein. The maximal difference in p_{dom} between the alternative structures of the same protein Δp_{dom} was computed. If Δp_{dom} is small, then all crystal structures are very similar and the differences are likely to be produced by unaccountable factors, such as variations in contacts and packing in the crystal lattice, preparation procedures for crystallization, etc. In contrast, if Δp_{dom} is large, then there are at least two crystal structures, which possess the domains of very different stability.

We also computed root mean-square deviations (RMSDs) of C_α atoms between the crystal structures, which possess maximal Δp_{dom} values. It is necessary to emphasize that the information expressed in the values of Δp_{dom} and $RMSD_{\mathrm{max}}$ is complementary, and cannot be considered as interchangeable. The stability and independence of domains, and as a consequence, Δp_{dom}, can change only if the underlying GNM contact matrix changes significantly. In contrast, RMSD values of the structures with essentially the same contact matrices can vary dramatically. Such limitation of the RMSD values is well known. This stimulated the development of alternative ways of measuring the similarity of protein structures (39). Nevertheless, the RMSD remains the most common and intuitive measure of conformational changes and thus was used in this study.

Fig. 2 shows sorted Δp_{dom} values of all studied proteins. All studied proteins could be visually (and thus somewhat arbitrarily) classified into three groups, where the values of Δp_{dom} are insignificant, significant, or anomalously large. The majority of proteins have Δp_{dom} values below the magic-threshold of 0.2. The distribution of the Δp_{dom} values for these proteins is almost exponential (Fig. 2 *b*). Such distribution can be observed if Δp_{dom} values deviate from zero in a purely random manner; thus deviations can be considered as insignificant. In striking difference are the results for four studied proteins, for which anomalously large Δp_{dom} values were found (Table 1). There are also seven proteins, which possess the Δp_{dom} values distributed almost evenly between 0.2 and 0.3. These proteins are referenced as having significant Δp_{dom}, since uniform distribution of such large values is unlikely to be caused by random factors.

We inspected the crystal structures of the proteins from both anomalous and significant groups to find the reason for large differences in p_{dom} between different conformations. Ligand binding, which leads to the change of strength of intra- and interdomain interactions, is only one of the possible reasons. Other possibilities include different packing of the proteins in the crystal unit cell, swapped domains and

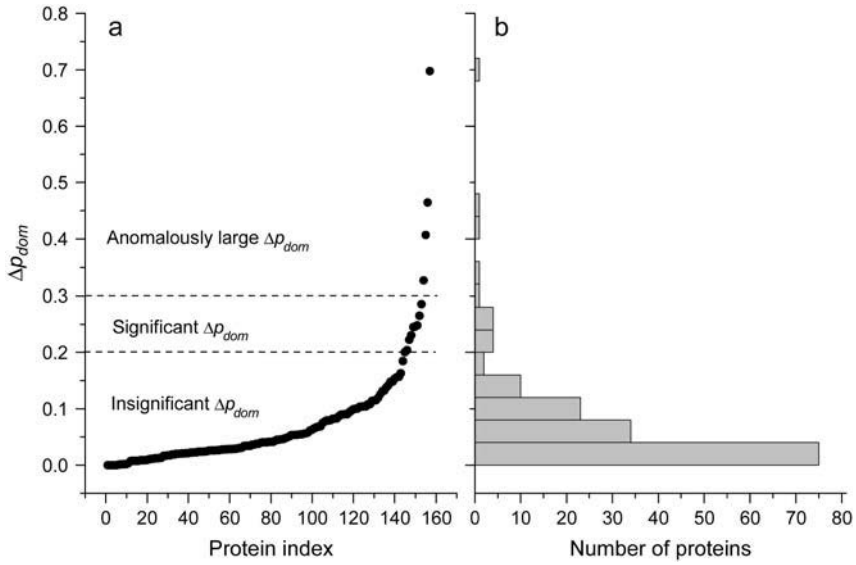

FIGURE 2 The maximal difference in p_{dom} between the alternative structures of the same protein Δp_{dom} obtained for 157 studied proteins. (*a*) The diagram in which the studied proteins are sorted by Δp_{dom} in ascending order. (*b*) The distribution of these proteins according to their Δp_{dom} values.

existence of disordered loops, which cannot be treated correctly by the GNM and HCCP. The results of inspection are summarized in Table 1.

Our analysis shows that the ligand binding is the reason of large Δp_{dom} values for only five proteins. Four of them—the lysine-, arginine-, ornithine-binding protein (LAOBP) from *Salmonella typhimurium* (40), phosphate-binding protein (PhBP) from *E. coli* (41), dipeptide binding protein (DpBP) from *E. coli* (42), and glutamin binding protein (GlnBP) from *E. coli* (43) belong to the same superfamily of bacterial periplasmic binding proteins (c.94.1 according to SCOP classification). They possess two distinct domains, which undergo significant hinge-bending motions to capture the ligand in the binding cleft between the domains. Obtained results demonstrate that the dynamics of their domains is very different in the ligand-bound and the ligand-free forms. Human calmodulin (44) is probably the most prominent example of the protein, which undergoes dramatic conformational change upon the binding of calcium ions. These proteins are identified as double domain proteins in which the ligand binding leads to significant relative motion of domains. In the case of dipeptide-binding protein, this motion even leads to the change in the number of dynamic domains. The open state of this protein has two domains, while the closed state has only one.

The human P21/H-RAS-1 protein (the product of the well-known HRAS1 proto-oncogene) (45) is identified as a single-domain protein. It can bind to several larger proteins (SOS-1, protein kinase BYR2, etc.). However, it is not clear if it is correct to analyze the dynamics of this small protein without its tightly bound ligand, which is comparable in size to, or even larger than, the RAS protein itself. This question can be clarified by analyzing the intradomain and the whole-structure RMSDs, as shown below.

The role of the GNM cutoff radius

The GNM cutoff radius r_{c} is an adjustable parameter in our approach. In the original works on GNM, the value of r_{c} was assumed to be 7.0 Å, because this is essentially the radius of the first coordination shell of an average residue (25). In later

TABLE 1 Proteins with large Δp_{dom} values for $r_{\mathrm{c}} = 7$ Å

Protein	Δp_{dom}	Reason for large Δp_{dom}
The group of anomalously large Δp_{dom} values		
Calmodulin	0.697	Effects of ligand binding.
LAO binding protein	0.464	Effects of ligand binding.
Pancreatic ribonuclease	0.407	Existence of domain swapped forms in crystal structures.
Phosphate-binding protein	0.326	Effects of ligand binding.
The group of significant Δp_{dom} values		
Dipeptide-binding protein	0.285	Effects of ligand binding.
Aspartate carbamoyltransferase regulatory chain	0.264	Part of the multichain complex. HCCP analysis of the single chain is meaningless.
Cyclin-dependent protein kinase 2	0.247	Existence of long unstructured loops with unresolved residues.
T4 lysozyme	0.245	Additional segment in one of the structures.
Glutamine-binding protein	0.230	Effects of ligand binding.
P1/Mahoney poliovirus capsid subunit	0.222	Several chains are arranged into capsid monomer. HCCP analysis of this single chain is meaningless.
P21/H-RAS-1	0.203	Unclear since the ligand is protein, which is larger than RAS protein itself.

works, the value of 10.0 Å was implemented because it was expected to provide a somewhat-better description of the cohesiveness of the protein globule (33). Furthermore, the value of 13 Å is typically used for anisotropic network model (24), which is closely related to our approach. The choice of cutoff in the present study is not trivial, since the domain stability depends strongly on r_c. This can be understood from very general considerations. One can consider the residue, which is close to the interdomain interface. According to GNM it is connected to some residues of another domain, which are inside the cutoff radius. With the increase of cutoff the number of such residues increases, which means that the interdomain connections become more numerous. As a result, the motion of domains becomes more interdependent and the interdomain correlations increase, while the intradomain correlations decrease accordingly. To study this behavior in detail, we performed the HCCP calculations on several representative proteins (LAOBP, calmodulin, isocitrate dehydrogenase, RAS protein, malate dehydrogenase, HIV protease, beta amylase, xylose isomerase, trypsin, pepsin) varying the cutoff radius from 5.5 to 13 Å. Chosen proteins represent all three groups of anomalously large, significant, and insignificant Δp_{dom} found in the present study for $r_c = 7.0$ Å. Fig. 3 shows the results of this analysis. It is clearly seen that in the case of proteins from anomalous group (such as LAOBP), Δp_{dom} decreases rapidly with the increase of cutoff. In contrast, Δp_{dom} changes rather chaotically, but remains very small for the proteins from the insignificant Δp_{dom} group (such as malate dehydrogenase) for all cutoff values. One can expect that the gap in Δp_{dom} between the anomalous group and the rest of the proteins would be well defined for commonly implemented cutoff values of 7 and 10 Å; however, it would nearly vanish for the 13 Å cutoff, as evident from Fig. 3. In that case, small chaotic variations in Δp_{dom} are comparable to the gap itself, which can introduce artifacts that are extremely hard to control. That is why we conclude that our analysis requires the cutoff values to be <13 Å.

To verify whether the anomalous and significant groups according to Δp_{dom} criteria change with the change of cutoff, we repeated our calculations for $r_c = 10$ Å. The distribution of Δp_{dom} values in this case are very similar to those shown in Fig. 3, but all the Δp_{dom} values (with few exceptions) become smaller in accord with the results discussed above (data not shown). The magic-threshold of Δp_{dom} values, which separates significant and insignificant groups, is also lower, and it was chosen to be 0.1. The proteins, which have large Δp_{dom} values due to the ligand binding, are essentially the same as in the case of $r_c = 7$ Å (Table 2), but their assignment to anomalous and significant groups is different. From the comparison of Tables 1 and 2 it is possible to conclude that the change of GNM cutoff radius from 7 to 10 Å does not change the set of proteins, which have large Δp_{dom} values due to the ligand binding. The smaller cutoff values produce more pronounced gaps in p_{dom} between different conformations (Fig. 3), which makes the analysis easier and more reliable. That is why we limit the subsequent analysis to the case of $r_c = 7$ Å.

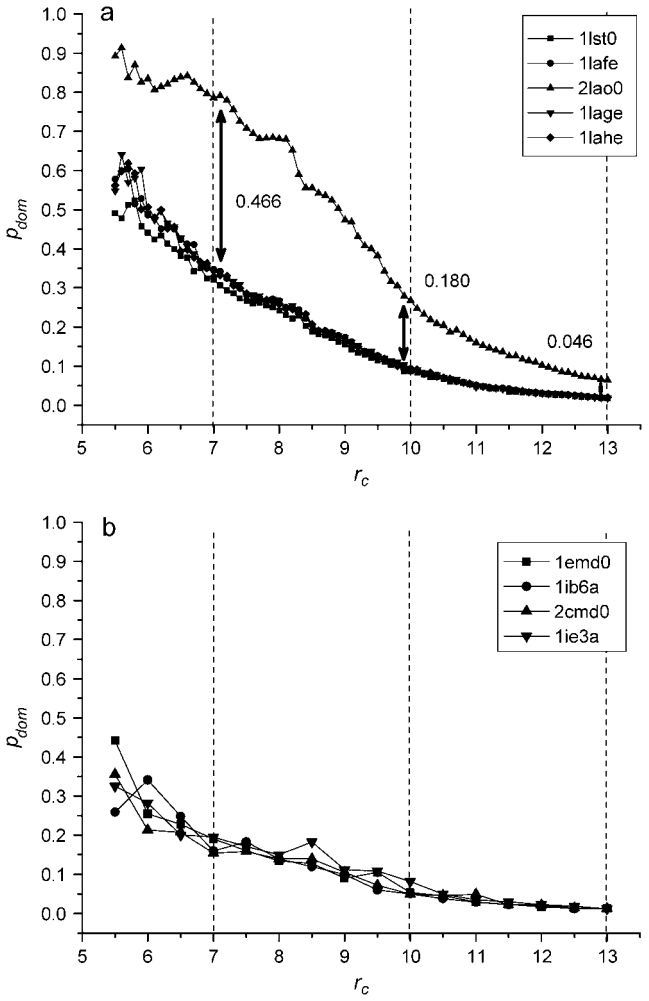

FIGURE 3 Dependence of the p_{dom} on GNM cutoff radius r_c for two representative proteins: LAOBP (*a*) and malate dehydrogenase (*b*). In panel *a*, the gaps between the open and closed forms are shown by arrows and the corresponding Δp_{dom} values are indicated. PDB identifiers of different conformations are indicated. Last character indicates the chain (''0'' means that only one chain is present).

Relationship between the conformational changes and the intradomain dynamics

It is interesting to compare the changes of dynamics described by Δp_{dom} with the differences of conformation between alternative structures of the same protein. Fig. 4 shows the $RMSD_{\max}$ values as a function of Δp_{dom} for all studied proteins. There is substantial degree of correlation between these quantities (the linear fit is presented by the *dotted line*; linear correlation is 0.703). However, the scattering of the data points become larger for large RMSD values. This agrees with the well-known fact that RMSD fails to discriminate very distant conformations correctly (46,47). It is particularly interesting to compare the Δp_{dom} and $RMSD_{\max}$

TABLE 2 The proteins with large Δp_{dom} values for $r_c = 10$ Å

Protein	Δp_{dom}	Reason for large Δp_{dom}
The group of anomalously large Δp_{dom} values		
Calmodulin	0.890	Effects of ligand binding.
P1/Mahoney poliovirus capsid subunit	0.553	Several chains are arranged into capsid monomer. HCCP analysis of this single chain is meaningless.
Isocitrate dehydrogenase	0.251	Existence of unstructured segments in one of the conformations.
Influenza virus hemagglutinin	0.211	Several chains are arranged into functional protein. HCCP analysis of this single chain is meaningless.
The group of significant Δp_{dom} values		
LAO binding protein	0.181	Effects of ligand binding.
Glutamine-binding protein	0.170	Effects of ligand binding.
D-xylose isomerase	0.161	Difference in chain packing in the crystal cell and unstructured segments.
T4 lysozyme	0.155	Additional segment in one of the structures.
Dipeptide-binding protein	0.153	Effects of ligand binding.
Mannose-binding protein	0.137	Difference in chain packing in the crystal cell.
Aspartate carbamoyltransferase regulatory chain	0.124	Part of the multichain complex. HCCP analysis of the single chain is meaningless.
Cyclin-dependent protein kinase 2	0.113	Existence of long unstructured loops with unresolved residues.
Phosphate-binding protein	0.104	Effects of ligand binding.

values of the proteins, which have large Δp_{dom} values due to the ligand binding (marked by *circles* in Fig. 4). It is clearly seen from Fig. 4 that there is some correlation between Δp_{dom} and $RMSD_{max}$ for these proteins, but the points are scattered significantly.

Large Δp_{dom} values indicate that the correlations of motions *inside* the dynamic domains differ significantly in two conformations (ligand-bound and ligand-free or conformations with different ligands) of the same protein. In contrast, large $RMSD_{max}$ values provide no information about the conformational changes inside the domain. $RMSD_{max}$ can be large if the domains are rigid, and only change their relative position or orientation. However, $RMSD_{max}$ can also be large, if substantial rearrangements occur inside the domains but the positions of the domains do not change.

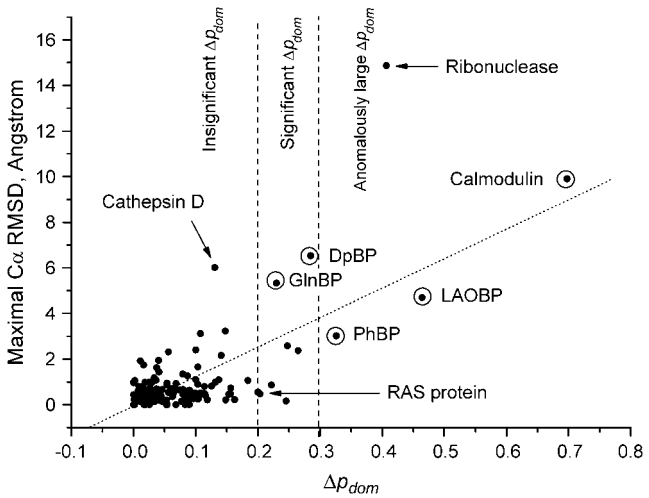

FIGURE 4 The comparison of RMSD ($RMSD_{max}$) of C_α atoms between alternative structures of the same protein with the largest Δp_{dom}. The data are for 157 studied proteins. Those proteins for which ligand-binding is the reason for large Δp_{dom} values are indicated by open circles. The reason for large Δp_{dom} for other proteins is the existence of long unstructured loops, swapped domains, missed segments, etc. Linear fit is shown by the dashed line.

There are other proteins with quite small Δp_{dom} but with RMSDs comparable to those of the periplasmic binding proteins. The most pronounced case is cathepsin D (marked by the *arrow* in Fig. 4). Cathepsin D is characterized by the presence of extended unfolded loops, which interact closely with other protein chains in the crystal cell and thus cannot be analyzed reliably by GNM and HCCP.

The stability of dynamic domain can change upon the ligand binding due to two factors. The first factor is an indirect effect caused by the changes of interdomain interactions produced by the ligand. This has to change the pattern of motions inside the domains, but will not change the equilibrium conformations of the domains themselves. The second factor is a conformational change within the domain itself caused by the interaction with the ligand. The observed change of dynamics in this case is a direct consequence of this conformational change. The parameters Δp_{dom} and $RMSD_{max}$ do not discriminate between these two factors. However, for practical applications it is important to know the origin of detected changes in dynamics in more detail. For this purpose, we identified the pairs of structures used for Δp_{dom} calculations, and considered each of their dynamic domains found by HCCP separately. We provided the structural alignment of these two conformations and computed the RMSD between the C_α atoms using the first domain only. The same procedure was applied for the second domain. The larger of these two RMSD values—$RMSD_{dom}$ was used as a measure of conformational changes occurring inside the domains.

Fig. 5 shows the relationship between the whole-structure $RMSD_{dom}$ and the intradomain $RMSD_{max}$. We observe a pronounced correlation between $RMSD_{dom}$ and $RMSD_{max}$ values (linear correlation 0.89) when all the obtained data are on the log scale (shown by the *dashed line*). This correlation means that the conformational changes in the majority

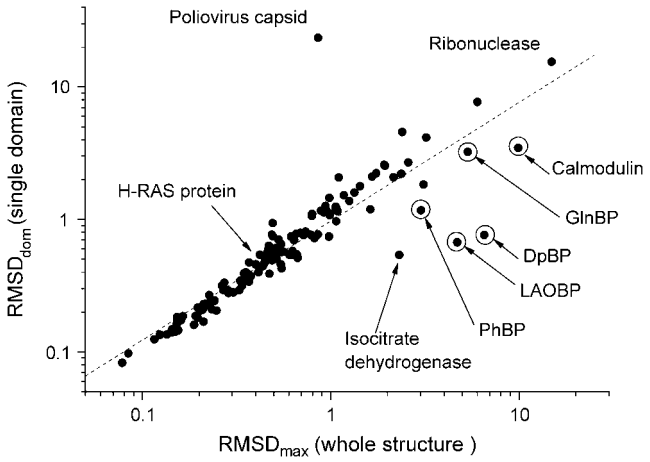

FIGURE 5 The comparison of RMSD ($RMSD_{max}$) of C_α atoms between alternative structures of the same protein with the largest Δp_{dom} and the largest pairwise C_α RMSD between individual domains of alternative structures ($RMSD_{dom}$) of these proteins. The data are for 157 studied proteins. Those proteins for which ligand-binding is the reason for large Δp_{dom} values are indicated by open circles. Linear fit is shown by the dashed line. The proteins discussed in the text are marked by arrows.

of studied proteins involve both intra- and interdomain rearrangements, which contribute to the whole-structure RMSD (regardless of whether or not the absolute value of RMSD is large). The contributions of the intra- and interdomain conformational changes to the conformational change of the whole globule are comparable.

However, there are a few proteins clearly out of this regularity. They have large overall RMSD, but anomalously small intradomain RMSD. In other words, the major contribution to the conformational change in these proteins is the change of domain position and orientation but not the intradomain conformational rearrangements. These proteins are calmodulin, DpBP, LAOBP, PhBP, and isocitrate dehydrogenase. All these proteins, except isocitrate dehydrogenase, are identified above as having significant Δp_{dom} occurring upon ligand binding. The protein GlnBP cannot be reliably classified to be out of trend in Fig. 5, although it also has large Δp_{dom} due to the ligand binding. Most probably this is explained by significant intradomain conformational changes in GlnBP, which exceed those found in other periplasmic ligand binding proteins.

Another protein that deviates dramatically from the general trend is the poliovirus capsid protein, already mentioned in Table 1. The HCCP domain assignment for this protein is unreliable due to the fact that several tightly bound chains are present in the crystal structure. Thus its out-of-trend position is probably an artifact. The RAS protein, which has quite large Δp_{dom} of unclear origin (Table 1), does not deviate from the main trend depicted in Fig. 5. This allows us to conclude that the quite-large Δp_{dom} value for this protein is also an artifact. The probable reason for this is that RAS protein is commonly found in complexes with larger proteins, so to derive correct conclusions, the dynamics of the whole complexes should be studied.

The data of Fig. 5 allow us to conclude that the proteins with significant Δp_{dom} possess the domains, which perform large relative motions upon the ligand binding. This domain rearrangement produces the major part of the RMSD between two distinct conformations. The conformational changes inside the domains can range from nearly nonexistent (LAOBP, DpBP) to a profound reorganization of the secondary structure (calmodulin). It should be stressed that these changes do not contribute much to the overall RMSD because they are masked by much larger RMSD changes caused by the changes of relative positions of domains.

Such significant conformational changes in the proteins with unusually high Δp_{dom} are well seen in x-ray crystal structures. Their structures are visualized in Fig. 6. In the group of periplasmic proteins the relative motion of domains is clearly seen. In calmodulin, in addition to the motion of domains, substantial internal rearrangements inside the domains are observed. In other studied proteins with insignificant Δp_{dom} values the alternative structures can hardly be distinguished visually, so the comparison of their structures is not shown.

Features of the proteins possessing dynamic domains sensitive to the ligand binding

It is remarkable that although we operated with a large and rather representative selection of two-domain ligand-binding proteins, the number of proteins with internal dynamics of domains that are sensitive to the ligand binding is extremely small—only 5 out of 157. Analysis of their ligand-bound and ligand-free crystallographic structures showed that the stabilities of domains of remaining the 151 proteins are essentially the same, regardless of the presence or absence of the ligand. This allows our deriving an unexpected conclusion that the character of domain motion is a strongly conservative feature of protein structure that can change only in exceptional cases. In these cases the mechanism of ligand binding involves tight mechanical closure of the binding pocket situated between the domains. This process results in significant change in the stability of domains, which is detected with our approach. In the open protein conformation, the domains are rather from each other and they move almost independently in a diffusive manner. In the closed form, the domain interfaces become very close to each other and the interdomain interactions become quite strong. This results in substantial decrease of intradomain correlations and increase of interdomain ones. Such picture is observed in the behavior of hinge-bending periplasmic proteins—LAOBP, PhBP, DpBP, and GlnBP. However, regarding another member of the same family, maltodextrin-binding protein, the stability of its domains is almost insensitive to the ligand binding.

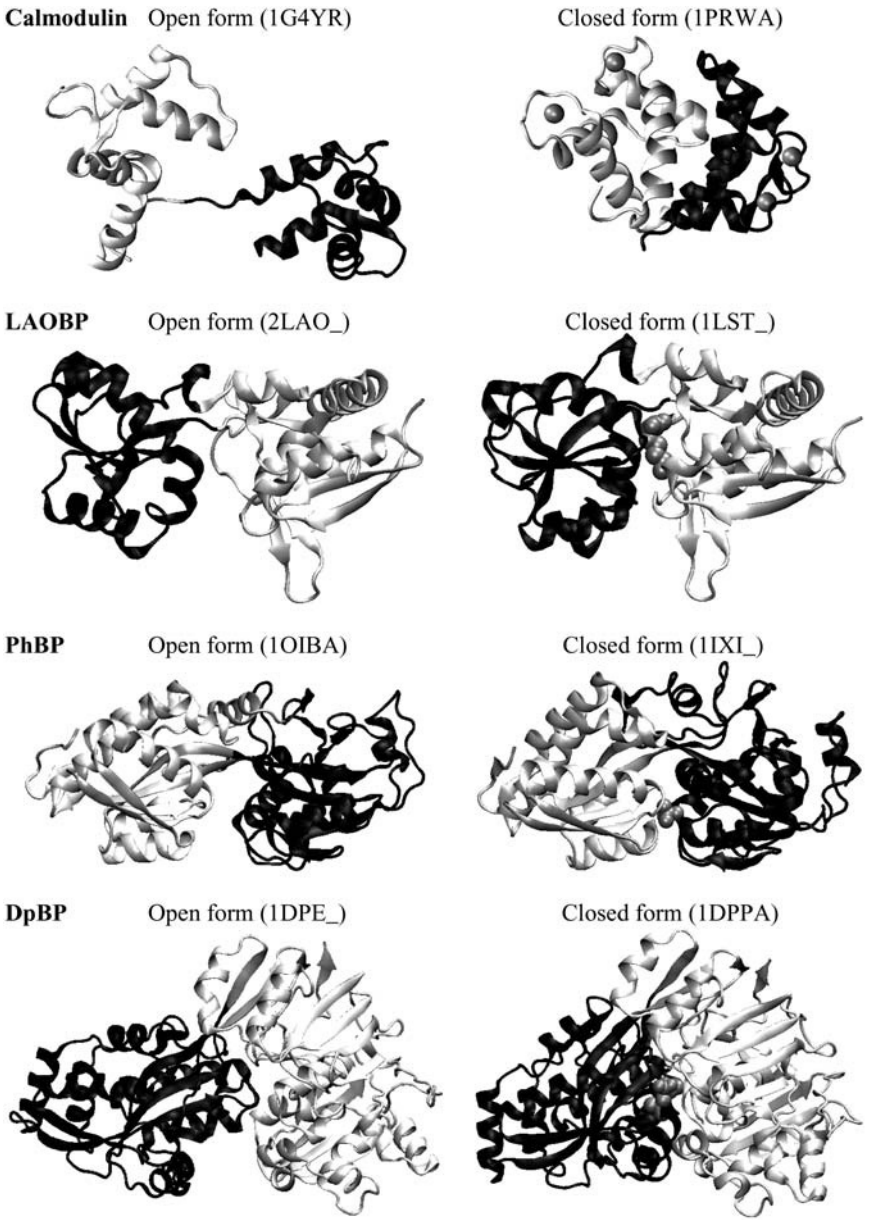

FIGURE 6 The proteins with significant Δp_{dom} caused by the ligand binding. Two structures, which correspond to the largest value of Δp_{dom}, are shown for each protein. The dynamic domains, computed by HCCP, are marked by black and white coloring. The ligands are shown in the space-fill representation. The structure of the GlnBP protein is visually very similar to LAOBP and thus not shown. The PDB codes for each conformation are given in the parentheses. The last letter of the code indicates the chain used. The underscore sign (_) means that the PDB entry contains a single chain.

Different periplasmic proteins are characterized by not only different stabilities of domains but also by the extent of internal conformational changes within the domains. They range from nearly nonexistent in LAOBP to quite significant in GlnBP and maltodextrin-binding protein. This shows that the sensitivity of dynamic domains to the ligand binding is a very individual feature related to specific structure and cannot be simply deduced from the structural or functional class of particular protein. However, it is possible to suggest that the domains of the periplasmic binding proteins are most likely to possess such sensitivity in comparison to other proteins.

A unique case described in this study is calmodulin. In the Ca^{2+}-free form this protein is compact and its two globular domains are bound tightly. Binding of Ca^{2+} ions leads to the straightening of extended α-helix, which drives two domains far apart and makes their motions much less interdependent. Internal structures of the domains are also changed significantly upon the ligand binding. However, the relative motion of domains makes more significant contribution into the overall RMSD between the Ca^{2+}-free and Ca^{2+}-bound forms.

Limitations

It is important to note that our approach has two main limitations:

1. Only single-chain protein structures can be analyzed with this approach. This excludes from the analysis those proteins that are functional as multimers or as parts of supramolecular structures. Since in these cases the interchain

contacts may contribute substantially to domain dynamics, they may not be treated correctly. The typical example is poliovirus capsid protein, for which we failed to obtain consistent data. Unfortunately outside the limit of our approach appeared allosteric enzymes, in which the ligand binding at one site provides conformational change extended to catalytic site. They usually possess several subunits that do not allow us to apply our single-chain approach.

2. Only the proteins but not the ligands are treated directly. We do not analyze the direct influence of the ligand on the dynamic properties of the protein, since the ligands are not included into the GNM elastic network. This means that instead of analyzing the dynamics in protein-ligand complex we analyze the dynamics of the protein itself. But this is done in that conformation, which is acquired in the protein-ligand complex.

It is also necessary to note that the set of ligand binding proteins available in PDB is already biased toward the relatively rigid proteins with high stability of domains, because these proteins are usually quite easy to crystallize. In contrast, very ''soft'' proteins with rather weak intradomain bonds usually impose serious problems upon crystallization attempts. The same is true for the large group of intrinsically disordered proteins (48,49), or the proteins that undergo partial unfolding or folding during the functioning (50). It is possible to reduce this bias by considering NMR structures and theoretical models; however, the performance of GNM for such structures should be tested first.

These limitations probably do not influence the main result of this study. The proteins analyzed in the present work still represent a variety of structures with a broad range of structural and functional diversity. Their ligands vary in size, structure, and properties from small ions and drugs to peptides and proteins, so we believe that the obtained regularities are rather general. Probably, our analysis reflects the universal property of the ligand-binding proteins—independence of domain stability on the ligand binding, except in a very few specific cases.

Biotechnological perspective

Many attempts have been made to apply the ligand-binding proteins as molecular sensors for detection of their ligands (51). The major problem here is to indicate the act of ligand binding by generating the measurable response signal. Usually this is done by incorporating a fluorescent reporter group into a protein structure (52). If one puts such a group into a ligand-binding site, it may interfere with ligand binding. Therefore it is preferable to locate a reporter group outside the ligand-binding site but allow for some mechanism of signal transduction to this group, which could be a conformational change. It is interesting to note that for calmodulin and several periplasmic ligand-binding proteins only, was this principle realized in practice, and it led to the design of direct molecular biosensors for detection of correspondent ligands with a fluorescent group at a remote site (53,54). The design of these sensors is based on the concept that the ligand binding changes the internal dynamics of the protein and induces conformational change, thereby allowing transduction of the signal to reporter groups without its direct contact with the ligand (53). This mechanism is in perfect agreement with the results of the present study. The small number of these successful cases of sensor design and the apparent absence of correlation with the practical importance of detection of their targets is probably an indication of difficulties in extending of this principle of sensing on other ligand-binding proteins. According to our results, it is very difficult to design the sensors with fluorescent group at remote site based on the majority of 157 studied proteins due to stability of their intra-domain dynamics, which is not changed upon the ligand binding.

Thus, analysis of distinguishing features of calmodulin and periplasmic ligand-binding proteins allows formulating the necessary conditions, which should be satisfied to design the molecular biosensor in which the responsive groups are not in direct contact with the ligand and located in one of the dynamic domains. The intradomain dynamics of such a biosensor should be sensitive to the ligand binding events to transfer the signal from the binding site to the reporter group. This can only be achieved if 1), the protein undergoes significant conformational change upon the ligand binding; 2), the intradomain conformational changes are small in comparison with the relative motion of domains; and 3), the strength of the interdomain contacts is significant in the ligand-bound conformation and very small in the ligand-free conformation. The technique developed in this work can be applied to the candidate proteins to see if these conditions are fulfilled for them. This will allow minimizing the possible design errors caused by the choice of the protein, which is not suitable for the design of sensors, where the reporter group is not in contact with the ligand.

CONCLUSIONS

In this work, using the HCCP technique, we analyzed the stabilities of dynamic domains of 157 ligand-binding proteins, for which several crystal structures are available. We demonstrate that, in the majority of the studied proteins (152 out of 157), the ligand binding does not lead to any significant change of domain stability. The exceptions from this rule are four periplasmic transport proteins (LAOBP, phosphate-binding protein, dipeptide binding protein, and glutamine binding protein) and calmodulin. Our results allow us to conclude that, in the vast majority of proteins, the pattern of correlated motions in dynamic domains, which determines their stability, is insensitive to ligand binding.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

REFERENCES

1. Wriggers, W., and K. Schulten. 1997. Protein domain movements: detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins.* 29:1–14.
2. Hayward, S., and H. J. Berendsen. 1998. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins.* 30:144–154.
3. Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins.* 33:417–429.
4. Hinsen, K., A. Thomas, and M. J. Field. 1999. Analysis of domain motions in large proteins. *Proteins.* 34:369–382.
5. Tama, F., F. X. Gadea, O. Marques, and Y. H. Sanejouand. 2000. Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins.* 41:1–7.
6. Kundu, S., D. C. Sorensen, and G. N. Phillips, Jr. 2004. Automatic domain decomposition of proteins by a Gaussian network model. *Proteins.* 57:725–733.
7. Yesylevskyy, S. O., V. N. Kharkyanen, and A. P. Demchenko. 2006. Dynamic protein domains: identification, interdependence and stability. *Biophys. J.* In press.
8. Yesylevskyy, S. O., V. N. Kharkyanen, and A. P. Demchenko. 2006. Hierarchical clustering of the correlation patterns: new method of domain identification in proteins. *Biophys. Chem.* 119:84–93.
9. Janin, J., and S. J. Wodak. 2002. Protein modules and protein-protein interaction. Introduction. *Adv. Protein Chem.* 61:1–8.
10. Janin, J., and S. J. Wodak. 1983. Structural domains in proteins and their role in the dynamics of protein function. *Prog. Biophys. Mol. Biol.* 42:21–78.
11. Auguin, D., P. Barthe, M. T. Auge-Senegas, M. H. Stern, M. Noguchi, and C. Roumestand. 2004. Solution structure and backbone dynamics of the pleckstrin homology domain of the human protein kinase B (PKB/Akt). Interaction with inositol phosphates. *J. Biomol. NMR.* 28: 137–155.
12. Bernado, P., M. X. Fernandes, D. M. Jacobs, K. Fiebig, J. Garcia de la Torre, and M. Pons. 2004. Interpretation of NMR relaxation properties of Pin1, a two-domain protein, based on Brownian dynamic simulations. *J. Biomol. NMR.* 29:21–35.
13. Clore, G. M., and C. D. Schwieters. 2004. Amplitudes of protein backbone dynamics and correlated motions in a small α/β protein: correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry.* 43:10678–10691.
14. Ilin, S., A. Hoskins, O. Ohlenschlager, H. R. Jonker, H. Schwalbe, and J. Wohnert. 2005. Domain reorientation and induced fit upon RNA binding: solution structure and dynamics of ribosomal protein L11 from *Thermotoga maritima*. *ChemBioChem.* 6:1611–1618.
15. Ragona, L., M. Catalano, M. Luppi, D. Cicero, T. Eliseo, J. Foote, F. Fogolari, L. Zetta, and H. Molinari. 2006. NMR dynamic studies suggest that allosteric activation regulates ligand binding in chicken liver bile acid-binding protein. *J. Biol. Chem.* 281:9697–9709.
16. Sakurai, K., and Y. Goto. 2006. Dynamics and mechanism of the Tanford transition of bovine β-lactoglobulin studied using heteronuclear NMR spectroscopy. *J. Mol. Biol.* 356:483–496.
17. Thormann, T., V. Soroka, S. Nielbo, V. Berezin, E. Bock, and F. M. Poulsen. 2004. Backbone dynamics of the first, second, and third immunoglobulin modules of the neural cell adhesion molecule (NCAM). *Biochemistry.* 43:10364–10369.
18. Lu, J., D. P. Cistola, and E. Li. 2006. Analysis of ligand binding and protein dynamics of human retinoid X receptor α-ligand-binding domain by nuclear magnetic resonance. *Biochemistry.* 45: 1629–1639.
19. Karplus, M. 2002. Molecular dynamics simulations of biomolecules. *Acc. Chem. Res.* 35:321–322.
20. Levitt, M., C. Sander, and P. S. Stern. 1985. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* 181:423–447.
21. Pang, A., Y. Arinaminpathy, M. S. P. Sansom, and P. Biggin. 2003. Interdomain dynamics and ligand binding: molecular dynamics simulations of glutamine binding protein. *FEBS Lett.* 550:168–174.
22. Wriggers, W., E. Mehler, F. Pitici, H. Weinstein, and K. Schulten. 1998. Structure and dynamics of calmodulin in solution. *Biophys. J.* 74:1622–1639.
23. Ababou, A., and J. R. Desjarlais. 2001. Solvation energetics and conformational change in EF-hand proteins. *Protein Sci.* 10:301–312.
24. Atilgan, A. R., S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.
25. Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Des.* 2:173–181.
26. Bahar, I., and A. Rader. 2005. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* 15:586–592.
27. Doruker, P., A. R. Atilgan, and I. Bahar. 2000. Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to α-amylase inhibitor. *Proteins.* 40:512–524.
28. Keskin, O. 2002. Comparison of full-atomic and coarse-grained models to examine the molecular fluctuations of c-AMP dependent protein kinase. *J. Biomol. Struct. Dyn.* 20:333–345.
29. Yildirim, Y., and P. Doruker. 2004. Collective motions of RNA polymerases. Analysis of core enzyme, elongation complex and holoenzyme. *J. Biomol. Struct. Dyn.* 22:267–280.
30. Kurt, N., W. R. Scott, C. A. Schiffer, and T. Haliloglu. 2003. Cooperative fluctuations of unliganded and substrate-bound HIV-1 protease: a structure-based analysis on a variety of conformations from crystallography and molecular dynamics simulations. *Proteins.* 51:409–422.
31. Temiz, N. A., and I. Bahar. 2002. Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. *Proteins.* 49:61–70.
32. Tama, F., and Y. H. Sanejouand. 2001. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* 14:1–6.
33. Keskin, O., R. L. Jernigan, and I. Bahar. 2000. Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* 78:2093–2106.
34. Puvanendrampillai, D., and J. B. O. Mitchell. 2003. Protein Ligand Database (PLD): additional understanding of the nature and specificity of protein-ligand complexes. *Bioinformatics.* 19:1856–1857.
35. Hu, L., M. L. Benson, R. D. Smith, M. G. Lerner, and H. A. Carlson. 2005. Binding MOAD (Mother of All Databases). *Prot. Struct. Func. Bioinformatics.* 60:333–340.
36. Beckstead, J. A., B. L. Block, J. K. Bielicki, C. M. Kay, M. N. Oda, and R. O. Ryan. 2005. Combined N- and C-terminal truncation of human apolipoprotein A-I yields a folded, functional central domain. *Biochemistry.* 44:4591–4599.
37. Wang, G., and R. L. Dunbrack, Jr. 2003. PISCES: a protein sequence culling server. *Bioinformatics.* 19:1589–1591.
38. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
39. Gilis, D. 2004. Protein decoy sets for evaluating energy functions. *J. Biomol. Struct. Dyn.* 21:725–736.
40. Oh, B. H., J. Pandit, C. H. Kang, K. Nikaido, S. Gokcen, G. F. L. Ames, and S. H. Kim. 1993. Three-dimensional structures of the periplasmic lysine-, arginine-, ornithine-binding protein with and without a ligand. *J. Biol. Chem.* 268:11348–11355.
41. Wang, Z., A. Choudhary, P. S. Ledvina, and F. A. Quiocho. 1994. Fine-tuning the specificity of the periplasmic phosphate transport receptor. *J. Biol. Chem.* 269:25091–25094.
42. Nickitenko, A. V., S. Trakhanov, and F. A. Quiocho. 1995. A resolution structure of DPPA, a periplasmic dipeptide transport/chemosensory receptor. *Biochemistry.* 34:16585–16595.
43. Hsiao, C. D., Y. J. Sun, J. Rose, and B. C. Wang. 1996. The crystal structure of glutamine-binding protein from *Escherichia coli*. *J. Mol. Biol.* 262:225–242.

44. Chattopadhyaya, R., W. E. Meador, A. R. Means, and F. A. Quiocho. 1992. Calmodulin structure refined at 1.7 Ångstroms resolution. *J. Mol. Biol.* 228:1177–1192.

45. Spoerner, M., C. Herrmann, I. R. Vetter, H. R. Kalbitzer, and A. Wittinghofer. 2001. Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc. Natl. Acad. Sci. USA.* 98:4944–4949.

46. Cohen, F. E., and M. J. Sternberg. 1980. On the prediction of protein structure: the significance of the root-mean-square deviation. *J. Mol. Biol.* 138:321–333.

47. Koehl, P. 2001. Protein structure similarities. *Curr. Opin. Struct. Biol.* 11:348–353.

48. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.

49. Tompa, P. 2005. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* 579:3346–3354.

50. Demchenko, A. P. 2001. Recognition between flexible protein molecules: induced and assisted folding. *J. Mol. Recogn.* 14:42–61.

51. de Lorimier, R. M., J. J. Smith, M. A. Dwyer, L. L. Looger, K. M. Sali, C. D. Paavola, S. S. Rizk, S. Sadigov, D. W. Conrad, L. Loew, and H. W. Hellinga. 2002. Construction of a fluorescent biosensor family. *Protein Sci.* 11:2655–2675.

52. Demchenko, A. P. 2005. Optimization of fluorescence response in the design of molecular biosensors. *Anal. Biochem.* 343:1–22.

53. Mizoue, L. S., and W. J. Chazin. 2002. Engineering and design of ligand-induced conformational change in proteins. *Curr. Opin. Struct. Biol.* 12:459–463.

54. Dattelbaum, J. D., L. L. Looger, D. E. Benson, K. M. Sali, R. B. Thompson, and H. W. Hellinga. 2005. Analysis of allosteric signal transduction mechanisms in an engineered fluorescent maltose biosensor. *Protein Sci.* 14:284–291.