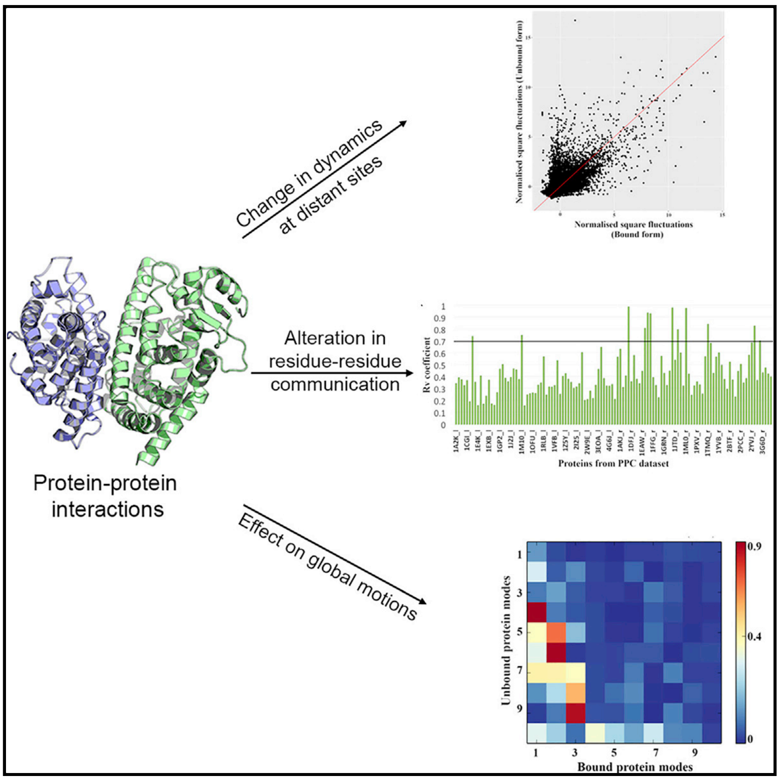
Structure

Article

Transient association between proteins elicits alteration of dynamics at sites far away from interfaces

Graphical Abstract



Authors

Himani Tandon,
Alexandre G. de Brevern,
Narayanaswamy Srinivasan

Correspondence

ns@iisc.ac.in

In Brief

Tandon et al. perform systematic analyses on 3D structures of proteins available in free form and bound to another protein. They observe that protein-protein association leads to change in dynamics at sites away from interfaces, alters inter-residue communication, and affects the global motions accessible to the protein in its unbound form.

Highlights

- Dynamics changes are common in proteins at distant sites upon binding to a protein
- Alteration in dynamics need not be accompanied by conformational changes
- Protein binding alters inter-residue communication within tertiary structure
- Global motions in the unbound protein are not always maintained in its bound form

Tandon et al., 2021, Structure *29*, 371–384
April 1, 2021 © 2020 Elsevier Ltd.
https://doi.org/10.1016/j.str.2020.11.015

Structure

Article

Transient association between proteins elicits alteration of dynamics at sites far away from interfaces

Himani Tandon,[1] Alexandre G. de Brevern,[2,3,4,5] and Narayanaswamy Srinivasan[1,6,*]
[1]Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India
[2]INSERM, U 1134, DSIMB, 75739 Paris, France
[3]Univ Paris, UMR_S 1134, 75739 Paris, France
[4]Institut National de la Transfusion Sanguine (INTS), 75739 Paris, France
[5]Laboratoire d'Excellence GR-Ex, 75739 Paris, France
[6]Lead Contact
*Correspondence: ns@iisc.ac.in
https://doi.org/10.1016/j.str.2020.11.015

SUMMARY

Proteins are known to undergo structural changes upon binding to partner proteins. However, the prevalence, extent, location, and function of change in protein dynamics due to transient protein-protein interactions is not well documented. Here, we have analyzed a dataset of 58 protein-protein complexes of known three-dimensional structure and structures of their corresponding unbound forms to evaluate dynamics changes induced by binding. Fifty-five percent of cases showed significant dynamics change away from the interfaces. This change is not always accompanied by an observed structural change. Binding of protein partner is found to alter inter-residue communication within the tertiary structure in about 90% of cases. Also, residue motions accessible to proteins in unbound form were not always maintained in the bound form. Further analyses revealed functional roles for the distant site where dynamics change was observed. Overall, the results presented here strongly suggest that alteration of protein dynamics due to binding of a partner protein commonly occurs.

INTRODUCTION

Transient interactions among proteins govern vital cellular processes and maintain functional integrity of the cell (Acuner Ozbabacan et al., 2011; Levy and Pereira-Leal, 2008; Schreiber and Keating, 2011). These interactions may activate signaling process, recruit components of bigger complexes, and inhibit or trigger molecular function (Tsai et al., 2009). Generally, a conformational change is induced by a partner protein, and an allosteric communication between protein interface and non-interface regions seems responsible for these functions (Swapna et al., 2012a; Tsai and Nussinov, 2014). Allostery is defined as alteration of protein function due to binding of an effector molecule at a site away from its functional site (Nussinov and Tsai, 2013). The effector can either be a small molecule, another protein or DNA/RNA, a mutation, or a post-translational modification (Tsai et al., 2009). This alteration can either be a change in local/global protein conformation or a change in fine dynamic equilibrium between different alternative conformations. Two conceptual models for allostery, namely induced fit and conformational selection, were proposed as early as 1959 and 1965, respectively (Koshland, 1959; Monod et al., 1965). The conformational selection term coined by Monod-Wyman-Changeux in 1965 relates to a two-state model. Furthermore, Frauenfelder et al. (1991) discussed the existence of static broad ensembles of states. Later, a dynamic landscape, which is the basis of the modern view of conformational selection, was proposed (Kumar et al., 2000).

Traditionally, allostery was associated with a change in protein structure upon effector binding. Earlier studies have analyzed the extent of structural changes in a protein because of binding of another protein (Betts and Sternberg, 1999; Grant et al., 2010; Martin et al., 2008a, 2008b; Swapna et al., 2012a). Furthermore, Swapna et al. (2012a) showed that the structural changes occurring in a protein due to binding are not just limited to protein-protein interfaces but are also widespread in regions distant from the interfaces. The idea of allostery without a significant structural change was proposed in 1984 (Cooper and Dryden, 1984), but it is only recently that it has been fully appreciated. In their seminal paper, Cooper and Dryden (1984) laid the theoretical groundwork for the possibility of dynamic allostery, i.e., allosteric changes happening without a conformational change. It was argued that dynamic allostery operates through altered entropy mediated by changes in frequency and amplitude of thermal or vibrational fluctuations. This changing view of allostery has been appreciated in the last decade with popularization of nuclear magnetic resonance methods to study protein motions. Proteins such as catabolite activator protein have been shown to exhibit dynamic allostery, arising from changes in intrinsic dynamics of the structure upon cyclic AMP binding (Louet et al., 2015; Popovych et al., 2006). Similar observations have been made for a few small molecule-protein and peptide-protein interactions (Kern and Zuiderweg, 2003; Mercier et al., 2001; Olejniczak et al., 1997; Wang et al., 2001; Zidek et al., 1999). For PPIs, an increase in backbone flexibility upon partner binding has been reported for certain complexes (Arumugam et al., 2003; Fayos et al., 2003). In the past, short molecular dynamics simulations on 17 protein-protein complexes (PPCs) in their bound and unbound form suggested that the flexibility associated with protein structures changes upon binding, with a redistribution of dynamics within the complex (Grünberg et al., 2006). This important study countered the idea of increased rigidity of proteins upon complex formation, but not much is known about the functional relevance of these observations.

Developments in the field of dynamic allostery have led to a renewed interest in understanding the effect on intrinsic dynamics when two proteins interact. The aforementioned reports in support of dynamic allostery are based on the analysis of selected individual proteins, and it is not clear how far these are prevalent in PPCs. The present study is a systematic attempt to explore the prevalence, extent, location, and functional relevance of the dynamics change in proteins due to transient protein-protein interactions (PPIs). A detailed structural and dynamics analysis was performed on a non-redundant dataset of 120 and 58 complexes, respectively. The datasets consist of proteins in their bound and unbound states. We first demonstrate that it is common to observe alteration of dynamics at a distant site upon binding of a partner protein. Second, we show that alteration in dynamics need not be accompanied by conformational change at the distant site. Third, we demonstrate that communication between the site of perturbation and allosteric site happens by alteration in inter-residue interactions within the structure. Fourth, we demonstrate that, in many cases, global motions accessible to a protein in its unbound form are not always maintained in the bound form. We also show that even if they are maintained, in most of them the modes are reordered. We further highlight with examples how an alteration in dynamics is related to function. Taken together, the results presented here strongly suggest that the alteration of dynamics, upon interaction of two proteins, occurs more frequently than previously thought.

RESULTS

Binding of proteins influences the conformations of associated proteins

In an earlier study, Swapna et al. (2012a) showed that proteins bound to other proteins undergo larger structural changes compared with proteins in the unbound form. With the availability of far more structures of proteins in bound and unbound forms in the Protein Databank (PDB), it is worthwhile exploring the extent of structural changes induced by PPIs. Therefore, a comprehensive dataset of 120 protein complexes was prepared for this study as described in STAR methods. The root-mean-square deviation (RMSD) and global distance test—total score (GDT-TS) scores for interacting partners in the PPC dataset showed that, for 91/120 complexes (75.8%), at least one of the interacting partners in the bound form show significant differences in structure (Figure 1A). RMSD values greater than the standard deviation from the mean of RMSDs for proteins in control dataset 1 were considered significant. For 52 of these 91 complexes (57.1%), one binding partner showed significant structural change and other partner showed no change in its conformation (Figure S1). To account for the effect of crystal packing on the structure, we compared RMSD distribution for 120 complexes with that of control dataset 1 (for details, see STAR methods) (Figure 1B). The two distributions (mean values 0.45 Å and 1.47 Å for control dataset 1 and PPC dataset, respectively) were found to be significantly different (two-sample Kilmogorov-Smirnov [KS] test, $p < 2.2 \times 10^{-16}$), suggesting that the observed differences in global protein conformation is mainly due to binding of another protein and not due to crystallization artifacts.

To identify the local regions of structural difference between complexed and free forms, we classified residues into interface, near-interface, and far-from-interface residues as described in STAR methods. RMSD was calculated for these stretches of residues separately. RMSD distribution for interface (mean value 1.3 Å), near-interface (mean value 0.9 Å) and far-from-interface or non-interface (mean value 1.2 Å) residues was found to be significantly different from the values in control dataset 1 (two-sample KS test, $p < 2.2 \times 10^{-16}$) (Figure 1B). Careful analysis of the plots suggested that not only the regions at the interface show significant deviations upon binding, but regions away from the interfaces also deviate significantly between the bound and unbound forms. RMSD values greater than the standard deviation from the mean of RMSDs for proteins in control dataset 1 were considered significant. At least one of the interaction partners for 53 complexes is reported to show deviation in the region away from interfaces (Table S1). For 37 complexes, both partners showed deviations in the residues away from interfaces (Table S2). Examples of cases that show structural changes away from the interface are shown in Figure 1C.

Residue dynamics is altered upon protein-protein complex formation

To understand the extent of change in residue dynamics between the bound and unbound form of a protein, we analyzed normalized square fluctuations obtained from normal mode analysis (NMA) for 58 complexes. These were obtained after filtering cases with missing residues in the structure in either bound or unbound form (see STAR methods). Normal modes pertaining to 80% of the variance were considered for calculation of squared fluctuations. These fluctuations are equivalent to thermal motions or vibrational motions of the residues around a mean position and define the flexibility of a protein. The distributions of normalized square fluctuations for the proteins in bound and unbound forms were found to be significantly different (two-sample KS test, $p < 2.2 \times 10^{-16}$) (Figure 2A). A higher variance in the fluctuation distribution of the bound form was observed, suggesting that many residues in the bound form show a change in flexibility (Figure 2A). To impose confidence on the results, we performed two control studies. First, the difference between normalized square fluctuations of the

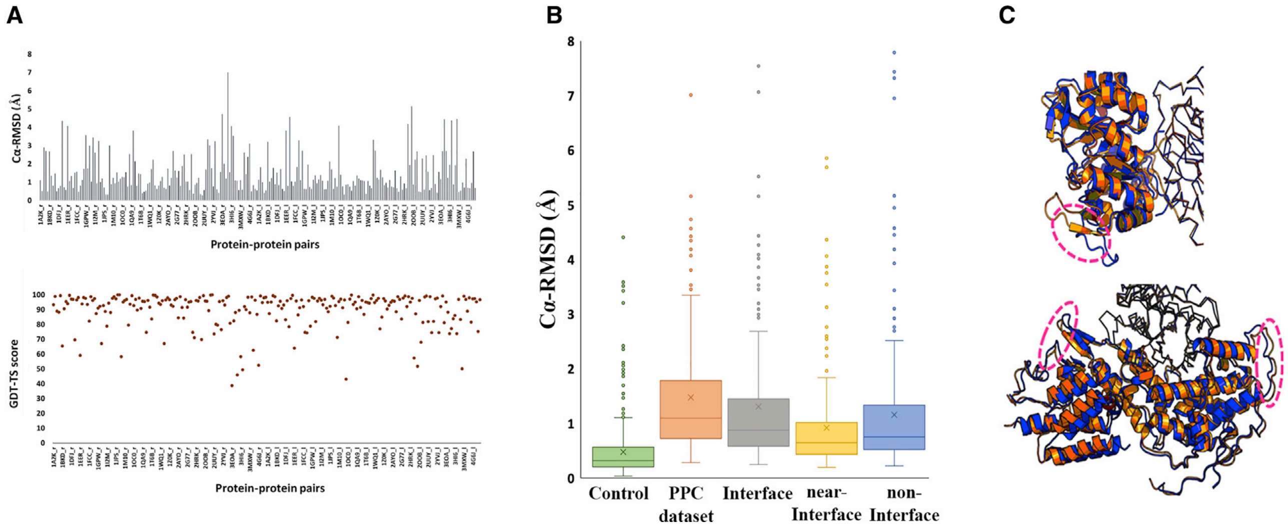


Figure 1. Structural analyses of the PPCs

(A) (Top panel) Cα-RMSD for protein pairs is plotted as bar plot. The higher the RMSD, greater is the structural difference between bound and unbound form. (Bottom panel) GDT-TS score is plotted. Here, the higher the score, higher is the similarity between two structures.

(B) Box plots showing distribution of Cα-RMSD (Å) for control dataset 1, PPC dataset, interface, near-interface, and non-interface regions. Residues are classified as either of the types based on distance between atoms of the two proteins. Distributions are significantly different from each other (two-sample KS test, $p < 2.2 \times 10^{-16}$).

(C) Two examples with significant structural changes away from the PPIs are shown. Proteins undergoing structural changes away from the interface are rendered as cartoon and partner protein as ribbon. Bound form is shown in orange and unbound form in blue. Top panel shows the HISF protein in its bound (PDB: 1gpw) and unbound (PDB: 1thf) form. Bottom panel shows SOS (Son of Sevenless) protein in its bound (PDB: 1bkd) and unbound (PDB: 2iio) form.

See also Figure S1; Tables S1 and S2.

bound and unbound form of proteins from PPC dataset was compared with the difference between normalized square fluctuations of the pairs from control dataset 1 (Figure 2B). The differences between two distributions were found to be statistically significant (two-sample, KS test, $p < 2.2 \times 10^{-16}$). Second, the normalized square fluctuations of the unbound form were compared with those of the fictitious unbound dataset, i.e., control dataset 2 (for details see STAR methods) (Figure 2C). The fluctuation profiles of both the datasets were not significantly different from each other (two-sample KS test, $p = 0.10$). These results suggest that differences in fluctuations between bound and unbound forms (Figure 2A) are indeed due to binding of a partner protein and not due to crystal packing effects. It was further observed that residues, in general, showed higher fluctuations in the bound form (Figure 2D). To identify the percentage of residues showing significant change, we calculated the difference between residue fluctuations. This difference was considered significant only if it was more than twice the standard deviation from the mean of fluctuation difference of control dataset 1. We found that ~10% of the residues show significantly higher fluctuations in the bound form than in the unbound form. On the other hand, ~11% of the residues showed significantly higher fluctuations in unbound form when compared with the bound form. To ascertain that these differences are insensitive to distance cutoff of 15 Å used for NMA calculations, we also calculated normalized fluctuations of the bound and unbound proteins at 12 Å and 10 Å cutoffs. We observed that differences in distributions of normalized square fluctuations were significant irrespective of the cutoffs used (Figures S2A and S2B). For both cutoffs (i.e., 12 Å and 10 Å), ~7% of the residues showed significantly higher fluctuations in the bound form than in the unbound form and ~6% of the residues showed significantly higher fluctuations in the unbound form than in the bound form.

Long-range communication between the interface and other regions in protein-protein complexes

The fluctuation profiles of the bound and unbound proteins for the interface, near-interface, and far-from-interface residues were analyzed separately (Figure 3). The distributions were found to be significantly different for the three regions (two-sample KS test, $p < 2.2 \times 10^{-16}$). Intuitively, interface residues showed higher fluctuations in the unbound form with ~28% of the residues showing significantly higher fluctuations (Figure 3A). The remaining 72% showed comparable fluctuations. The interface residues were further divided into "core" and "rim" (Figure 3A). It must be noted that while only ~2% of "core" interface residues showed significantly higher fluctuations in the unbound form, ~30% of "rim" interface residues showed higher fluctuations in the unbound form. This underlines the fact that many core interface residues remain rigid in their unbound forms.

The non-interface regions (near-interface and far-from-interface), show both increase and decrease in fluctuations in bound forms. For the near-interface region, ~5% of residues showed significantly higher fluctuations in the bound form and ~10% residues showed higher fluctuations in the unbound form (Figure 3B). Interestingly, for regions far from PPIs, at least 11% of residues showed higher fluctuations in the bound form than in the unbound form, whereas ~5% of residues showed higher fluctuations in the unbound form (Figure 3C). This result counters the general idea that flexibility of interacting partners should

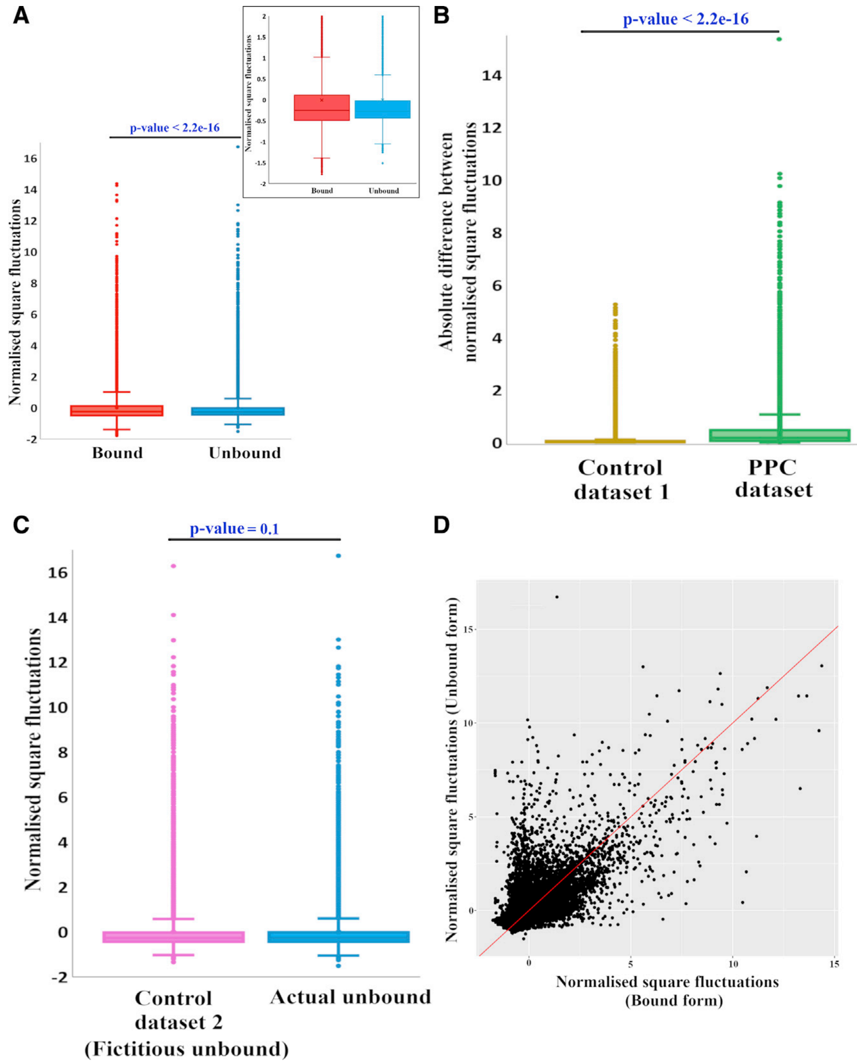


Figure 2. Analyses of dynamics of PPCs

(A) Box plots showing distribution of Cα fluctuations for all residues in bound and unbound forms. The two distributions are significantly different from each other (two-sample KS test, $p < 2.2 \times 10^{-16}$), showing variation in flexibility in bound and free proteins. Box plots with fluctuations only between the range of −2 and 2 are shown in the inset for clarity.

(B) Comparison of absolute difference between normalized fluctuations of control dataset 1 and PPC dataset. The distributions are significantly different (two-sample KS test, $p < 2.2 \times 10^{-16}$), suggesting no bias due to crystal packing.

(C) Comparison of normalized fluctuations of control dataset 2 and actual unbound proteins from PPC dataset. The distributions are not different from each other (two-sample KS test, $p = 0.1$), suggesting no effect of crystal packing.

(D) Scatterplot of square fluctuations for all residues in PPC dataset. The x axis represents normalized square fluctuations for bound proteins and the y axis represents normalized square fluctuations for unbound proteins. Solid red line is the unity slope line. Vibrational entropy was estimated using the fluctuations to calculate binding affinity presented in Data S1.

See also Figures S2A, S2B, and S8; Data S1; Table S7.

decrease when in complex. Our findings suggest that a loss of conformational entropy at the interface is likely compensated by an increase in flexibility at other regions of the complex. Such rearrangement is indicative of communication between the interface and regions away from it. Furthermore, the observed changes in fluctuations were found to be independent of the size of interacting proteins (Figure S2C) and interface area (Figure S2D and Table S3).

Changes in dynamics are not always accompanied by observed structural changes

To identify whether a change in dynamics is associated with a change in the local structure, we compared RMSD and root-mean-square difference of fluctuations ($RMSD^f$) values for all proteins (Figure 4A). A Pearson correlation coefficient of 0.1 suggests that binding of proteins can cause change in residue dynamics without undergoing a significant conformational change, and vice versa. To understand the trend for residues away from the interfaces, we divided the dataset ($58 \times 2 = 116$ proteins) into four categories: (1) proteins that only show significant structural changes away from the interfaces, (2) proteins that only show significant change in dynamics, away from the interfaces, (3) proteins that show significant changes in structure and dynamics, away from the interfaces, and (4) proteins that show no significant changes either in structure or dynamics, away from the interfaces (Figure 4B and Table S4). Fifteen of 116 cases (12.9%) belong to the first category, 17 of 116 (14.7%) to the second, 35 of 116 (30.2%) to the third, and 49 of 116 (42.2%) to the fourth category. It is interesting to note that although a structural change along with flexibility change was observed for bound forms in many cases, instances of significant change in residue dynamics away from the interfaces, without a structural change, were also observed.

Correlation between residue motions is altered upon protein binding

It is known that the information within a protein can be relayed through correlated fluctuations (DuBay et al., 2011; Goodey and Benkovic, 2008; Kern and Zuiderweg, 2003; Zhang et al., 2014). Furthermore, residues coupled in motion are helpful in constructing a pathway between the allosteric site and the functional site (Gerek and Ozkan, 2011). Therefore, we addressed the point of alteration in residue-residue communication within a protein upon binding of a partner protein. Cross-correlation matrices were plotted to understand the extent to which the residue-residue coupling is affected. The R_v coefficient was calculated between matrices of bound and unbound forms (Figure 5A). Coupling between the

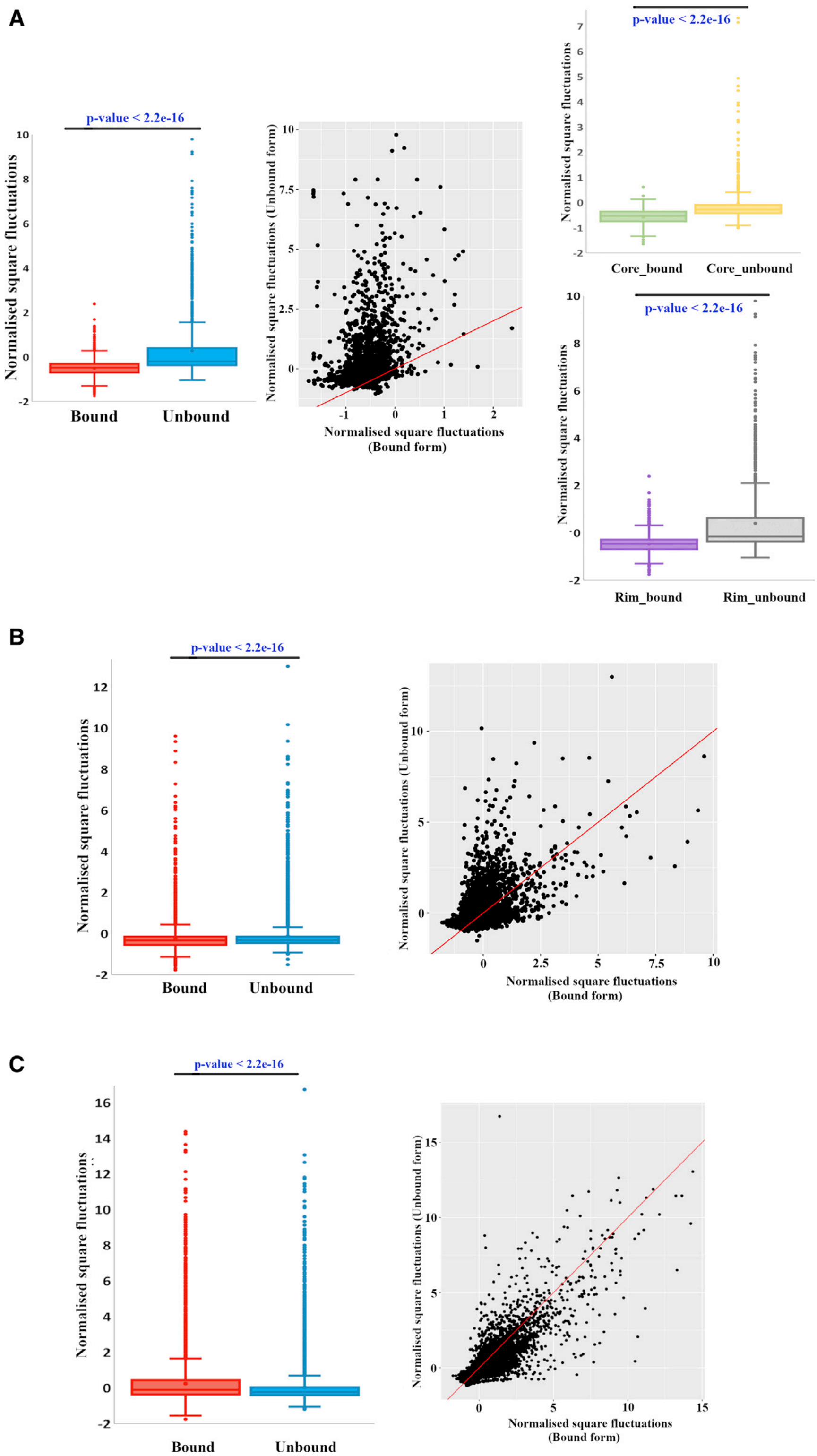


Figure 3. Comparison of fluctuations for different types of residues in proteins

(A) Box plot showing distribution of Cα fluctuations for interface residues (atoms within a distance $\leq$ 4.5 Å) in bound and unbound proteins in PPC dataset. The distributions are significantly different from each other (two-sample KS test, $p < 2.2 \times 10^{-16}$). Scatterplot displays fluctuations for the corresponding residues in bound and unbound form and is useful in identifying residues that show increase/decrease in fluctuations. The interface is further divided into core and rim, and their distributions are provided in the panel on the right.

(B) Box plot showing distribution of Cα fluctuations for near-interface residues (atoms within 4.5–10 Å) in bound and unbound proteins in PPC dataset. The distributions are significantly different from each other (two-sample KS test, $p < 2.2 \times 10^{-16}$). Scatterplot shows fluctuations for corresponding residues in bound and unbound forms.

(C) Box plot showing distribution of Cα fluctuations for residues far from interface (atoms at a distance >10 Å) in bound and unbound proteins in PPC dataset. The distributions are significantly different from each other (two-sample KS test, $p < 2.2 \times 10^{-16}$). Scatterplot shows fluctuations for corresponding residues in bound and unbound forms.

For all scatterplots, the x axis represents normalized square fluctuations in bound form and the y axis represents normalized square fluctuations in unbound form. See also Figures S2C and S2D; Table S1.

fluctuations was found to be unaffected for only 12 of 116 proteins (10.3%) after binding of the partner protein (R_v coefficient $\geq$0.7). The remaining proteins showed a change in synchronization of residue motions, suggesting a high influence of protein-protein binding on the residue couplings. It was further observed that residues of a protein become tightly coupled in the bound form as compared with the unbound form. An example from the dataset where cross-correlation is affected by binding of a partner protein is shown in Figure 5B, and an example where it does not get affected by binding of a partner protein is shown in Figure 5C. A Pearson correlation coefficient of 0.5 between the protein size and R_v coefficient suggests a slight dependence of protein size on the extent to which its residue communication is affected by binding of a partner protein (Figure 5D).

Low-frequency global modes of the unbound form are perturbed by binding of the interacting partner

Low-frequency global modes from NMA are known to be biologically relevant (Bahar et al., 1998, 2010a). Previously, Marcos et al. (2011) showed that new modes of motions are acquired by enzymes to form the amino acid kinase family upon oligomerization, which regulates the substrate binding and allostery. Also, an attempt was made to model conformational changes upon binding by reranking of normal modes

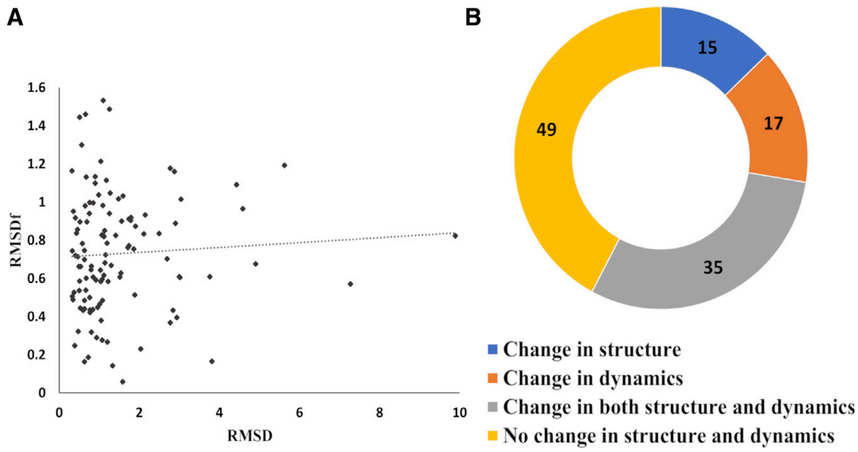


Figure 4. Change in dynamics need not always be accompanied by change in structure

(A) RMSD (in Å) of bound and free proteins in the PPC dataset is plotted on the x axis vis-à-vis the root-mean-square difference of fluctuations ($RMSD^f$) of unbound and free proteins on the y axis. The dotted line is the line of best fit. The plot suggests no correlation between the observed structural change and dynamics change.

(B) Doughnut chart shows the number of cases in each category as mentioned in text.

See also Table S4.

(Oliwa and Shen, 2015). Hence we asked the question, is this is a commonly observed phenomenon upon protein binding? In other words, how do the low-frequency modes of unbound protein become altered in the presence of an interacting partner? Or, are the low-frequency motions accessible to a unique protein or are they maintained in bound form too? To answer these questions, we analyzed the similarities/differences between the modes of motion accessible to a protein in its bound and unbound forms by calculating the overlap between the top ten low-frequency global motions obtained from NMA (Figure 6). The overlap value is an indicator of similarity between the modes in terms of its frequency, shape, and size: the smaller the overlap, the smaller the difference in the two modes of motion. About 58% of cases showed high overlap (>|0.7|) between bound and unbound forms for at least one mode (from the top ten) (Figure 6A). However, of these, ~60% of the cases showed change in mode order/preference (Figure S3). A change in mode preference means that, if mode ''m'' exists as a low-frequency mode in one form, the same or similar mode (defined by a high overlap value) exists with an altered frequency in another form. This suggests that although some modes of motion are preserved between the bound and unbound forms, their frequency, size, and shape changes as suggested by reordering of normal modes. For the remaining ~42% of cases, very few global modes in unbound form were found maintained in the bound form, along with weaker correlation (<|0.7|) and reordering of modes (Figure 6A). This clearly suggests that dynamics of the unbound form are affected by binding of the partner protein. Examples are shown of a case where the mode order and shape were retained between the bound and unbound form in Figure 6B, a change in mode preference in a high-overlap case in Figure 6C, and a case with low overlap value in Figure 6D. To rule out the effect of crystal packing on the observed dissimilarity of intrinsic modes, we calculated overlap for all pairs in control dataset 1. Each pair showed an overlap value $\geq|0.9|$, suggesting similarities in their modes (Figure 6E). Hence, the global motions of the unbound form are superseded by low-frequency global motions of the bound form. This becomes especially important when no visible conformational changes are observed at the macro level, reinforcing the notion that the absence of conformational changes does not mean that there is no allostery.

Analysis of cases with no observed structural changes reveals prevalence of dynamic allostery and its functional role

As mentioned earlier, 17 of 116 proteins showed significant change in dynamics at sites away from interfaces without a significant structural change. Of these 17 proteins, three are antigen-antibody complexes, four are enzyme-inhibitor complexes, and ten are either enzyme-substrate or signaling complexes. This study proposes that these differences are likely involved in either the stability of the complexes or signaling a downstream protein, or both. The role of differential dynamics is presented for two cases below and three other cases in Data S1.

Differential dynamics of cyclophilin A likely plays a role in stabilizing HIV-1 capsid assembly

The first example we describe is human cyclophilin A (CypA), a peptidyl-prolyl enzyme that catalyzes the isomerization of peptide bonds from *trans* to *cis* form and participates in various biological process such as protein folding, apoptosis, and signaling (Nigro et al., 2013). Many studies have reported dynamic allostery associated with CypA, which couples the active-site and distal residues, regulating the enzymatic activity (Agarwal, 2005; Rodriguez-Bussey et al., 2018; Wapeesittipan et al., 2019). In addition to the native functions, CypA plays an important role in (de)stabilization of HIV-1 capsid and hence is often recruited by HIV-1 during its life cycle in host cells (Lu et al., 2015; Thali et al., 1994).Although the structure of CypA and HIV-1 capsid (CA) have been available for a long time (Gamble et al., 1996), it was not clear how CypA modulates the CA stability until recently, when the cryoelectron microscopy (cryo-EM) structure of CypA-CA assembly was solved at 8 Å (Liu et al., 2016). It was reported that a single CypA molecule binds to two CA molecules at two different sites, one canonical and the other non-canonical, thus stabilizing the CA assembly. Moreover, the non-canonical binding site by itself has weaker affinity for CA, but binds the second CA molecule with strong avidity in presence of CA at the canonical site (Liu et al., 2016).

To understand the binding of CA at the non-canonical site of CypA, we analyzed the fluctuation profiles, cross-correlation matrices, and overlaps obtained from NMA of the free and bound forms of CypA. A comparison of the structure of free

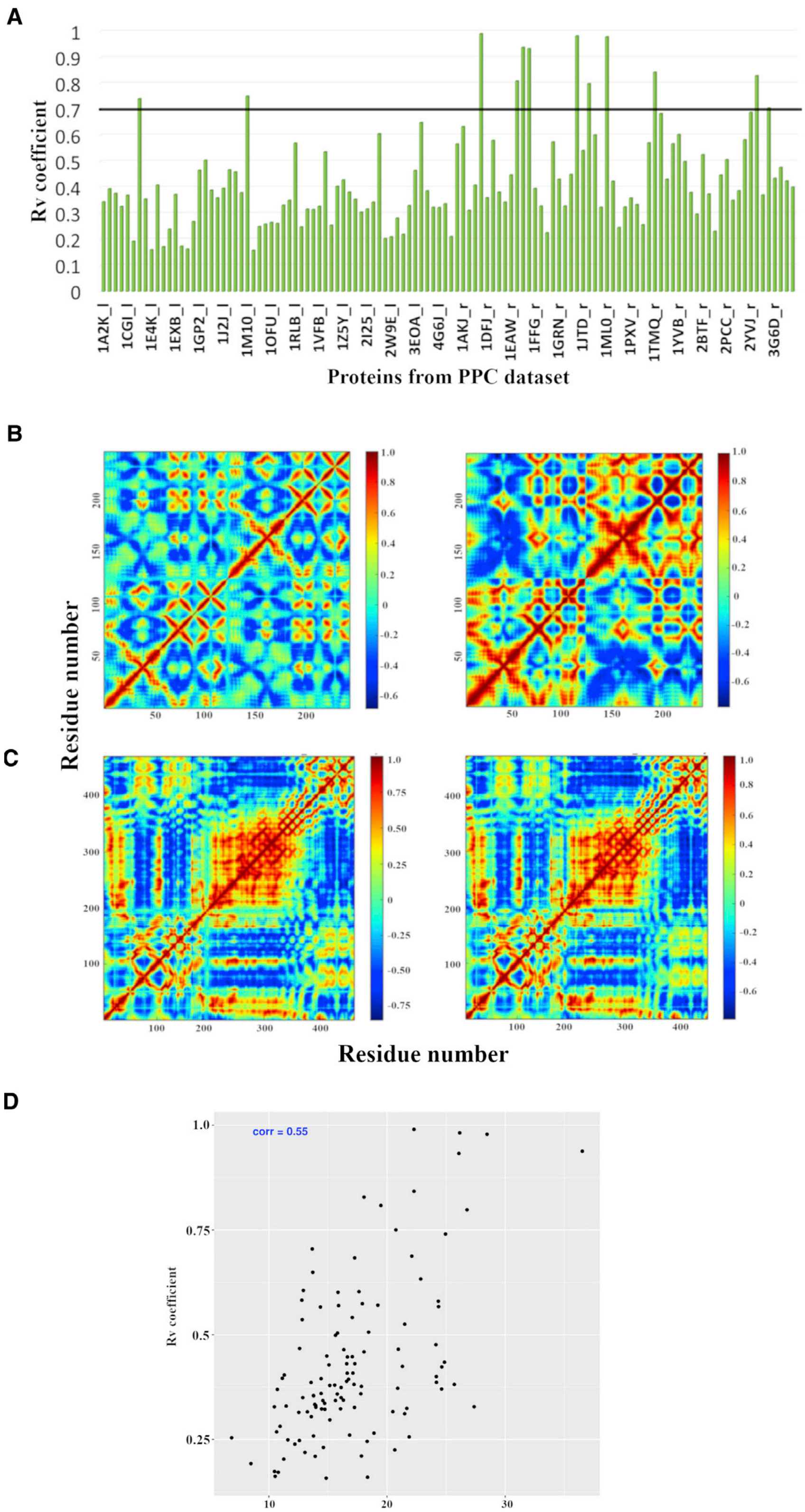


Figure 5. Cross-correlation analysis of proteins in PPC dataset

(A) Similarity between cross-correlation matrices of proteins in the free and bound form is plotted as bar plot. The x axis shows the proteins in the PPC dataset. Each entry code corresponds to the PDB code of the complex with "_l" and "_r" suffix representing two entries in the PPC dataset. The black horizontal line marks the cutoff for similarity.

(B) An example case from PPC dataset (PDB: 1a2k) where the binding of a partner protein changes the residue communication as shown by different cross-correlation matrices.

(C) An example case from PPC dataset (PDB: 1jiw) where the binding of a partner protein does not change the residue communication as shown by similar cross-correlation matrices.

(D) Radius of gyration (proxy for protein size) on the x axis is plotted vis-à-vis the R_v coefficient. The correlation coefficient of 0.5 suggests there is not a strong relation between the size of protein and its effect on residue communication.

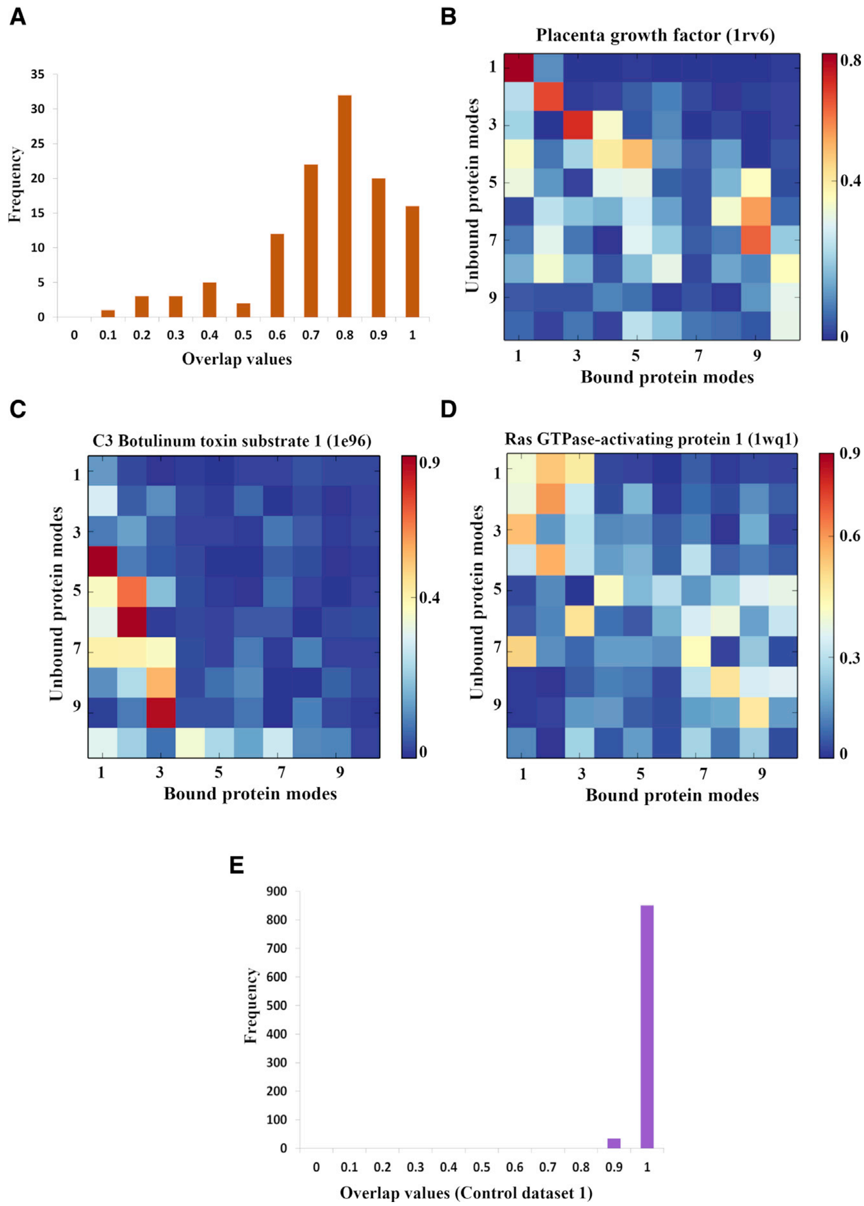


Figure 6. Overlap analysis of proteins in PPC dataset

(A) Absolute overlap score obtained for top ten low-frequency modes of bound and unbound proteins in the PPC dataset plotted as a histogram.

(B) Example case where the mode order and shape were retained between the bound and unbound forms.

(C) Example case showing change in mode preference.

(D) Example case showing low overlap values between bound and free forms. The PDB ID for the complex is mentioned in all three examples (B), (C), and (D).

(E) Absolute overlap score obtained for each pair in control dataset 1 is plotted as a histogram. The overlap is high between the pair of proteins in the control dataset 1.

See also Figure S3.

and CA-bound forms at the canonical binding site of CypA suggests no significant structural changes in CypA (RMSD = 0.3 Å) (Figure 7A). Nonetheless, differences in fluctuations and cross-correlations were observed between the bound and unbound form of CypA (Figures 7B and 7C). Interestingly, higher fluctuations were observed for the bound form of CypA. Regions away from the active site or non-canonical binding site specifically showed increased fluctuations (Figure 7B). A higher cross-correlation was observed between the active site (residues Arg55, Gln63, Asn102, Trp121, and His126) and non-canonical binding region (residues 25–31) of CypA in the bound form (Figure 7C). The lowest-frequency global modes for both bound and unbound forms showed that the individual residues in canonical and non-canonical binding sites move in different directions in the free form but show better coordination in the bound form (Figure 7D). Since the distance from the centroid of canonical binding site to non-canonical binding site is ~16 Å, this suggests a long-range communication between the two sites (Figure 7E). The residue Val29 (part of the non-canonical binding site) has already been shown to be involved in allosteric communication within CypA (Holliday et al., 2017). Results from this study suggest that binding of HIV-1 CA at the canonical site strongly affects the dynamics of the distant non-canonical site and supersedes the local motions of the unbound form with more global, collective motions in the bound form. These results present a clear case for dynamic allostery between the two sites, which promotes CA binding at non-canonical site after its binding at the canonical site.

Interestingly, CypA has also been shown to bind other prehistoric endogenous lentiviruses, e.g., from rabbits (RELIK) and lemurs (PSIV). The crystal structure of CypA with RELIK-capsid (rCA) shows that the active site of CypA binds rCA in like manner to CypA-HIV CA, but the orientation of CAs differ in the two crystal structures (Goldstone et al., 2010). Since CypA shows a conserved binding mode with lentiviral capsids, its interaction with CA and rCA was further compared to find the similarity or differences between the dynamics of CypA bound to two evolutionarily conserved partners, the results of which propose a possibility of a binding mode similar to that of HIV CA (see Data S1 and Figure S4).

DNA-binding site of DNAse-I is affected upon binding of actin molecule

DNAse-I is an endonuclease that cleaves double-stranded DNA in a sequence-specific manner at phosphodiester linkages. Many studies have reported the residues important for binding and cleaving the DNA molecule (Lahm and Suck,

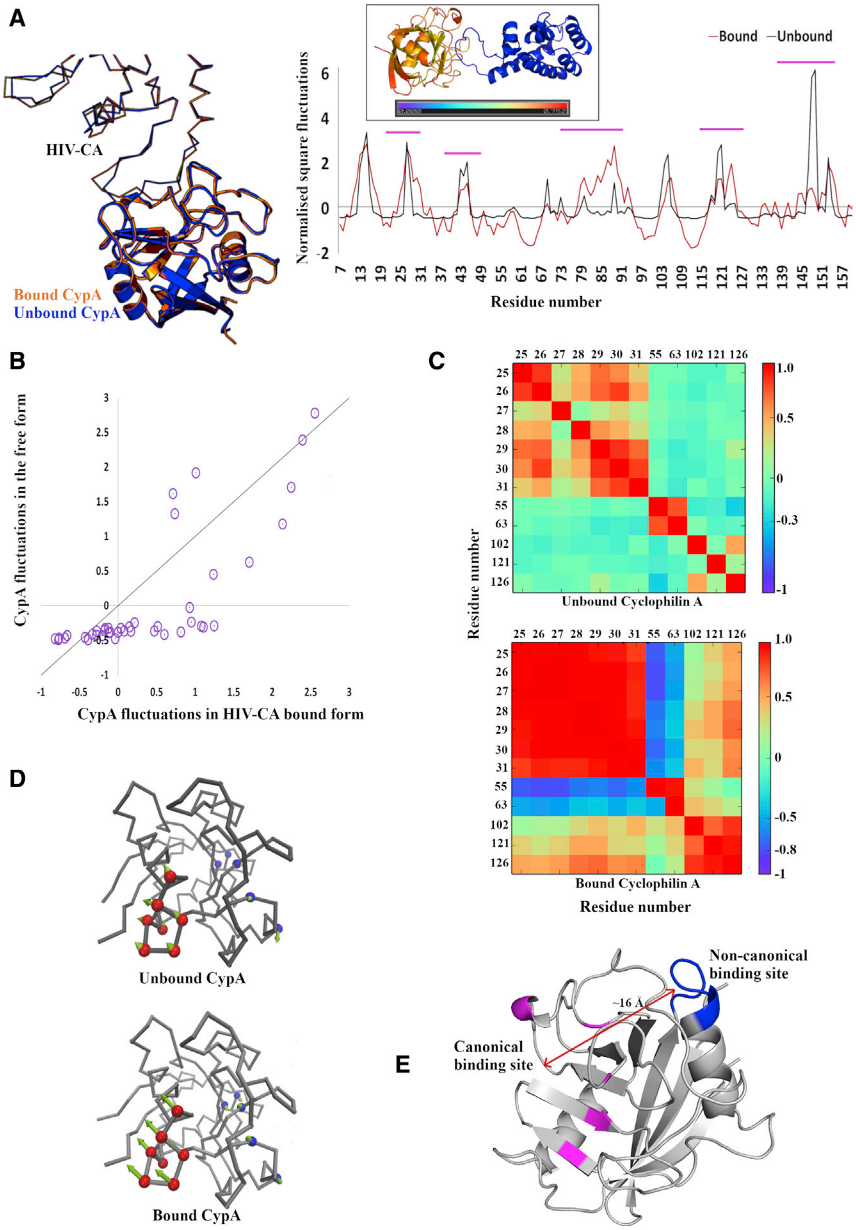


Figure 7. Understanding dynamic allostery in cyclophilin A

(A) Free and HIV CA bound CypA is superposed with an RMSD of 0.3 Å. CypA in both the forms is rendered as cartoon and the HIV CA as ribbon. Bound CypA is shown in orange and unbound in blue color. Normalized square fluctuations of CypA in CA-bound and unbound form are shown in the right panel. The x axis represents residue number and the y axis represents normalized square fluctuations. Pink horizontal lines signify regions away from interface. Inset shows the absolute difference between square fluctuations mapped onto the CypA structure.

(B) Scatterplot of normalized square fluctuations for non-interface residues showing generally higher fluctuations in the bound form.

(C) Matrices show weak cross-correlation between the canonical and non-canonical binding sites in the unbound (top) CypA and a tighter coupling in bound (bottom) CypA.

(D) Backbone trace of CypA in the unbound and bound form, with canonical binding site in blue spheres and non-canonical binding site in red spheres. The green arrows show the direction of motion, and their length is proportional to the magnitude of fluctuation.

(E) Cartoon representation of CypA showing the spatial distance between canonical (pink) and non-canonical (blue) binding site.

See also Data S1 and Figures S4–S7.

1991). It also interacts with actin monomer to form a 1:1 actin-DNAse-I complex (Hitchcock, 1980). Although the function of this interaction is not clear, it renders DNAse-I inactive. An inspection of the crystal structures of DNAse-I bound to octamer DNA (PDB: 2dnj) (Lahm and Suck, 1991) and bound to actin monomer (PDB: 1atn) (Kabsch et al., 1990) shows that binding sites of DNA and actin are proximal but do not overlap (Figure 8A). However, it has been proposed that actin monomer provides steric hindrance to binding of DNA (Kabsch et al., 1990).

A comparison of the actin-bound and free forms of DNAse-I suggested high similarity between them (RMSD = 0.35 Å) (Figure 8B). However, differences were observed in the flexibility of DNAse-I between the bound and unbound forms (Figure 8C). The presence of actin molecule was found to modulate the synchronized motions of DNA-binding sites of DNAse-I. This is suggestive of the existence of a communication pathway between the actin interface and DNA interface (Figure 8D). From these results, we propose that apart from steric hindrance for DNA binding, the actin molecule can affect the vibrational motions intrinsic to DNAse-I, which in turn can affect the dynamics of DNA-binding residues, contributing toward inactivity of DNAse-I by actin.

Apart from these two cases, we observed a redistribution of dynamics upon binding of partner protein in two antigen-antibody complexes, namely cytochrome *c* (Cyt-C) and E8 antibody (Mylvaganam et al., 1998), and sonic-hedgehog (Shh) protein and 5E1 antibody fragment complex (Maun et al., 2010). Such redistribution was found to contribute to the stability of the complex (see Data S1 and Figure S5). Another protein, β-lactamase inhibitor protein-II (BLIP-II) (Brown et al., 2011, 2013), showed subtle differences in residue-residue communication when bound to two homologous partners, and these differences were found to be in concordance with gain in entropy (see Data S1 and Figure S6). The analysis was also performed for thioredoxin, a moonlighting protein (Akabayov et al., 2010, Ghosh et al., 2008, Kulczyk and Richardson, 2016, Lee et al., 2018, Zeller and Klug, 2006). Since moonlighting proteins can bind diverse partners under different conditions (Gancedo et al., 2016; Huberts and van der Klei,

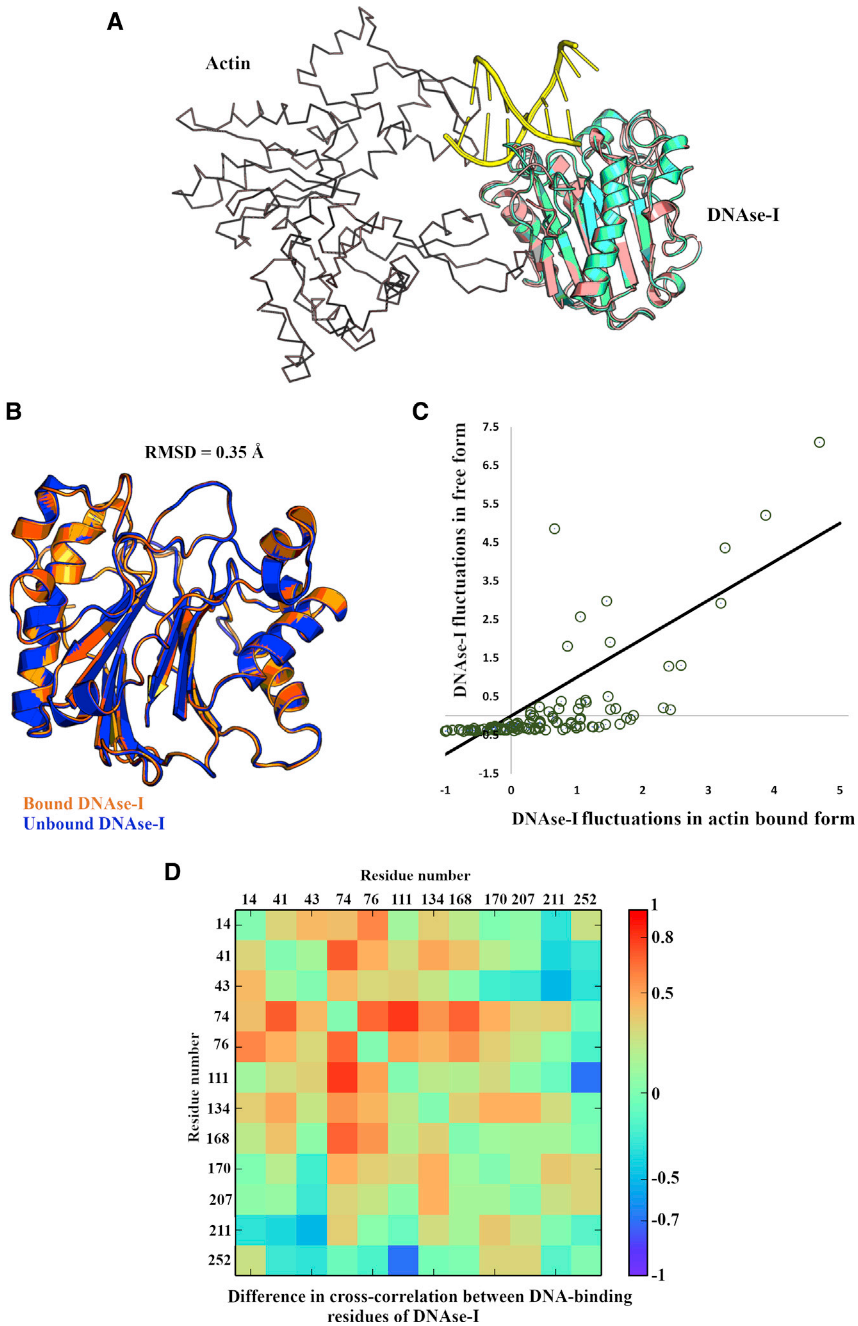


Figure 8. Analysis of DNAse-1 activity upon actin binding

(A) Superposition of structures of DNAse-1 in complex with DNA and in complex with actin shows non-overlapping DNA and actin binding sites. DNAse-1 is represented as cartoon and actin is shown as ribbon. DNA is shown in green color.

(B) Superposition of DNAse-1 in the actin-bound (orange) and free (blue) form shows an RMSD of 0.35 Å.

(C) Scatterplot showing normalized square fluctuations of DNAse-1 in actin-bound form (x axis) and free form (y axis). Solid line represents the unity line.

(D) Difference between cross-correlation of DNA-binding residues in DNAse-1 in bound and unbound form shows effect of actin binding on DNA-binding site.

See also Data S1 and Figures S4–S7.

2010; Jeffery, 2009), it was of interest to analyze whether dynamics of a moonlighting protein are altered upon binding of another protein without a significant change in structure (Marsh and Teichmann, 2014). We observed that a change in thioredoxin flexibility likely helps in better packing of T7 bacteriophage replisome (see Data S1 and Figure S7).

DISCUSSION

Proteins are dynamic systems that undergo post-translational modifications, bind to small molecules or other proteins, and elicit allosteric responses at sites implicated in function. Historically, allostery was suggested to be mediated by a change in the mean conformation of a protein, contributing to enthalpy gain. However, in recent years the definition of allostery has been broadened by including the alteration in nature and extent of dynamics at sites away from the site of perturbation, and few proteins have been reported to show such dynamic allostery (Arumugam et al., 2003; Fayos et al., 2003; Grünberg et al., 2006; Louet et al., 2015; Popovych et al., 2006). Allostery is achieved through altered entropy of the protein side chain or backbone (and/or other entropic effects) (Popovych et al., 2006; Tzeng and Kalodimos, 2015). Furthermore, it has been proposed that large-scale motions associated with proteins are important carriers of allosteric signals without requiring a conformational change (Rodgers et al., 2013). Motivated by these observations on isolated examples of proteins, we performed systematic analyses on a dataset of proteins in their bound and free forms to understand their change in dynamics (especially at the sites away from interfaces) upon binding of the partner protein. To achieve this, we used fundamental and widely used metrics such as squared fluctuations, residue couplings, and overlap of intrinsic dynamics obtained from a coarse-grained anisotropic network model (ANM)-based NMA. Before studying flexibility/dynamics, we note that our observations from structural analysis, albeit with a bigger dataset, fall in line with the findings from previous studies of individual proteins or smaller datasets (Agarwal et al., 2010; Martin et al., 2008b; Smith et al., 2005; Swapna et al., 2012a) and suggest prevalence of allostery mediated by PPIs via a change in conformation.

The distributions of normalized square fluctuations were found to be significantly different for proteins in bound and unbound forms, with ~10% of residues showing higher fluctuations in bound form and 11% in unbound form. This suggested a strong effect of partner binding on the vibrational entropy of protein residues. We show by using various controls that the observed changes are not due to crystal packing defects. Intuitively, the flexibility of proteins should decrease upon forming a complex. While the proteins showed decreased atomic fluctuations at the interface in the bound form, interestingly, non-interface residues showed a general increase in fluctuations and ~11% of the residues showed significantly higher fluctuations in bound forms, suggesting redistribution of motions within a protein upon binding of the partner protein. We believe that a loss of conformational entropy at the interface is compensated by reorganization of motions within a protein, resulting in higher flexibility at other regions. This result reinforces similar findings by Grünberg et al. (2006) who reported, for a very small dataset of 17 PPCs, an increase in flexibility of many proteins in the complexed form.

Binding of two or more proteins was observed to strongly affect the residue-residue couplings (for ~90% of the cases), and proteins in the bound form showed highly correlated motions (either positive or negative). Such changes were commonly observed for the proteins in our dataset, suggesting that the low-frequency local motions are superseded by low-frequency global motions of the bound form. We believe that synchronized motions in the bound form can create allosteric communication between the interface and non-interface regions, which can lead to increase/decrease in the flexibility of these regions. Such changes are likely to have implications on the stability or function of the complex. Indeed, allosteric regulation of a few proteins by a change in flexibility and/or change in correlated motions has been reported earlier (DuBay et al., 2011; Kern and Zuiderweg, 2003; Zhang et al., 2014). Furthermore, binding of a partner protein was observed to perturb the modes of motions in 42% of cases, and new modes of motions were acquired. Interestingly, for 52% of the cases, although the modes overlapped well, they were maintained with a reordering of modes. This seems to suggest that global motions are affected by binding of a partner protein to control the functions of the complex. To the best of our knowledge, this was not previously known as a common feature in many PPCs.

It is difficult to appreciate the presence of allostery without a change in conformation, but absence of a structural change does not imply that allostery is not in play (Nussinov and Tsai, 2015). It has been hypothesized that global and local modes obtained from NMA can carry signals from the binding site to other regions of proteins without requiring structural changes (Hawkins and McLeish, 2004, 2006). From our analyses, we found that at least 17 cases from the dataset showed differences in fluctuations for non-interface regions without an observed structural change. We believe that this should be carefully considered in future studies, otherwise allostery due to PPIs may go unnoticed in several cases. A close inspection of these cases revealed that such changes contribute to the stability of a complex by adding to the positive gain in entropy (in the case of BLIP-II interactions with β-lactamases and antigen-antibody complexes), regulating a downstream function such as binding of another protein (in the case of CypA binding to HIV CA), or altering the functional capacity of an enzyme (in the case of actin binding to DNAse-I). Binding of a partner protein was also proposed as one of the contributing factors to functional switching of bacterial thioredoxin and better packing of T7 replisome.

Taken together, this study provides a comprehensive analysis of the effect of binding of two or more proteins on the dynamics of individual proteins. Since the impact on dynamics is independent of size and interface area of two interacting proteins, changes mediated by PPI can be thought of as an intrinsic property of these interactions. Furthermore, allostery has been proposed to be an intrinsic property of monomeric proteins (Gunasekaran et al., 2004). However, Cui and Karplus (2008) raised a valid question in their classic review as to whether these perturbations induced by binding should be considered a manifestation of allostery or whether the term dynamic allostery should be used when the allosteric effect has a biological function. Whether or not our results provide an answer to that question, the analyses of the cases in this study do reveal the functional relevance of these changes. We further believe that the results presented here will lead to better appreciation of allostery mediated by PPIs. We propose two direct applications of our work. First, the lessons learned are expected to be applicable to the growing knowledge on three-dimensional structures of large multi-protein assemblies. An understanding of allostery, dynamics, and their relationship with the biological function of the proteins studied here may aid in understanding the organization of subunits within the multi-protein assemblies determined often by cryo-EM. Second, the structural fluctuations have a critical impact on thermodynamics of PPIs and hence are likely to be an essential contributor to binding affinities. Results from this analysis strongly suggest modulation of dynamics upon binding of two proteins, and hence support inclusion of contributions from vibrational entropy toward affinity calculations. Two previous studies have shown the importance of vibrational entropy in the calculation of binding affinities and thus act as a proof of principle for the results we have discussed in our analyses (Moal et al., 2011; Skrbic et al., 2018). Furthermore, encouraged by our findings, we also performed a pilot study to understand whether including the free energy contributions from vibrational entropy helps in improving the accuracy of affinity calculations (Gilson and Zhou, 2007, Kollman, 1993, Xue et al., 2016). Our observations are presented in Data S1, Figure S7, and Table S7. We find that regardless of proper weight assignment to vibrational entropy contributions, the accuracy of binding affinity calculation improves in many cases. Our study envisions a tremendous scope of improvement in this area.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABLITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS

- Dataset preparation
- PPC dataset for dynamics analyses
- Control datasets
- Structural analyses
- Dynamics analyses using normal mode analysis (NMA)
- Classification of interface & non-interface residues

- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2020.11.015.

ACKNOWLEDGMENTS

N.S. and A.G.d.B. acknowledge the Indo-French Center for the Promotion of Advanced Research/CEFIPRA for collaborative grant number 5302-2. N.S. acknowledges funding for infrastructural support from the following agencies or programs of the Government of India: DBT-COE, Ministry of Human Resource Development, DST-FIST, UGC Center for Advanced Study, and the DBT-IISc Partnership Program. N.S. is a J.C. Bose National Fellow. H.T. is an INSPIRE fellow. A.G.d.B. acknowledges grants from the Ministry of Research (France), University de Paris, University Paris Diderot, Sorbonne, Paris Cité (France), National Institute for Blood Transfusion (INTS, France), National Institute for Health and Medical Research (Inserm, France), IdEx ANR-18-IDEX-0001, and labex GR-Ex. The labex GR-Ex, reference ANR-11- LABX-0051 is funded by the French National Research Agency, reference ANR-11-IDEX-0005-02. H.T. thank Dr. Sneha Vishwanath and other lab members for useful discussions. The authors also thank Ms. Seemadri, Ms. Yazhini, and Ms. Sneha for technical assistance with remote data access during the pandemic lockdown.

AUTHOR CONTRIBUTIONS

Conceptualization, N.S. and A.G.d.B.; Investigations, H.T.; Data Curation, H.T.; Methodology, H.T., N.S., and A.G.d.B.; Formal Analysis, H.T.; Validation, H.T.; Supervision, N.S. and A.G.d.B.; Writing – Original Draft, H.T.; Writing – Review & Editing, H.T., N.S., and A.G.d.B.; Project Administration, N.S. and A.G.d.B.; Funding Acquisition, N.S. and A.G.d.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 8, 2020
Revised: October 1, 2020
Accepted: November 17, 2020
Published: December 10, 2020

REFERENCES

Acuner Ozbabacan, S.E., Engin, H.B., Gursoy, A., and Keskin, O. (2011). Transient protein-protein interactions. Protein Eng. Des. Sel. *24*, 635–648.

Agarwal, P.K. (2005). Role of protein dynamics in reaction rate enhancement by enzymes. J. Am. Chem. Soc. *127*, 15248–15256.

Agarwal, G., Dinesh, D.C., Srinivasan, N., and de Brevern, A.G. (2010). Characterization of conformational patterns in active and inactive forms of kinases using protein blocks approach. In Computational Intelligence and Pattern Analysis in Biological Informatics, U. Maulik, S. Bandyopadhyay, and J.T.L. Wang, eds., pp. 169–187.

Akabayov, B., Akabayov, S.R., Lee, S.-J., Tabor, S., Kulczyk, A.W., and Richardson, C.C. (2010). Conformational dynamics of bacteriophage T7 DNA polymerase and its processivity factor, *Escherichia coli* thioredoxin. Proc. Natl. Acad. Sci. U S A *107*, 15033–15038.

Arumugam, S., Gao, G., Patton, B.L., Semenchenko, V., Brew, K., and Van Doren, S.R. (2003). Increased backbone mobility in β-barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. J. Mol. Biol. *327*, 719–734.

Bahar, I., and Rader, A.J. (2005). Coarse-grained normal mode analysis in structural biology. Curr. Opin. Struct. Biol. *15*, 586–592.

Bahar, I., Atilgan, A.R., Demirel, M.C., and Erman, B. (1998). Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. Phys. Rev. Lett. *80*, 2733–2736.

Bahar, I., Lezon, T.R., Yang, L.-W., and Eyal, E. (2010a). Global dynamics of proteins: bridging between structure and function. Annu. Rev. Biophys. *39*, 23–42.

Bahar, I., Lezon, T.R., Bakan, A., and Shrivastava, I.H. (2010b). Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. Chem. Rev. *110*, 1463–1497.

Bakan, A., Meireles, L.M., and Bahar, I. (2011). ProDy: protein dynamics inferred from theory and experiments. Bioinformatics *27*, 1575–1577.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Betts, M.J., and Sternberg, M.J. (1999). An analysis of conformational changes on protein-protein association: implications for predictive docking. Protein Eng. *12*, 271–283.

Brooks, B., and Karplus, M. (1983). Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proc. Natl. Acad. Sci. U S A *80*, 6571–6575.

Brown, N.G., Chow, D.C., Sankaran, B., Zwart, P., Venkataram Prasad, B.V., and Palzkill, T. (2011). Analysis of the binding forces driving the tight interactions between β-lactamase inhibitory protein-II (BLIP-II) and class A β-lactamases. J. Biol. Chem. *286*, 32723–32735.

Brown, N.G., Chow, D.C., Ruprecht, K.E., and Palzkill, T. (2013). Identification of the β-Lactamase Inhibitor Protein-II (BLIP-II) interface residues essential for binding affinity and specificity for class A β-lactamases. J. Biol. Chem. *288*, 17156–17166.

Cooper, A., and Dryden, D.T.F. (1984). Allostery without conformational change—a plausible model. Eur. Biophys. J. *11*, 103–109.

Cui, Q., and Karplus, M. (2008). Allostery and cooperativity revisited. Protein Sci. *17*, 1295–1307.

Delarue, M., and Sanejouand, Y.H. (2002). Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the Elastic Network Model. J. Mol. Biol. *320*, 1011–1024.

DuBay, K.H., Bothma, J.P., and Geissler, P.L. (2011). Long-range intra-protein communication can be transmitted by correlated side-chain fluctuations alone. PLoS Comput. Biol. *7*, e1002168.

Fayos, R., Melacini, G., Newlon, M.G., Burns, L., Scott, J.D., and Jennings, P.A. (2003). Induction of flexibility through protein-protein interactions. J. Biol. Chem. *278*, 18581–18587.

Frauenfelder, H., Sligar, S.G., and Wolynes, P.G. (1991). The energy landscapes and motions of proteins. Science *254*, 1598–1603.

Fox, N.K., Brenner, S.E., and Chandonia, J.M. (2014). SCOPe: structural Classification of Proteins-extended, integrating SCOP and ASTRAL data and classification of new structures. Nucleic Acids Res. *42*, D304–9.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics *28*, 3150–3152.

Fuglebakk, E., Tiwari, S.P., and Reuter, N. (2015). Comparing the intrinsic dynamics of multiple protein structures using elastic network models. Biochim. Biophys. Acta *1850*, 911–922.

Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I., and Hill, C.P. (1996). Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell *87*, 1285–1294.

Gancedo, C., Flores, C.-L., and Gancedo, J.M. (2016). The expanding landscape of moonlighting proteins in yeasts. Microbiol. Mol. Biol. Rev. *80*, 765–777.

General, I.J., Liu, Y., Blackburn, M.E., Mao, W., Gierasch, L.M., and Bahar, I. (2014). ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones. PLoS Comput. Biol. *10*, e1003624.

Gerek, Z.N., and Ozkan, S.B. (2011). Change in allosteric network affects binding affinities of PDZ domains: analysis through perturbation response scanning. PLoS Comput. Biol. *7*, e1002154.

Ghosh, S., Hamdan, S.M., Cook, T.E., and Richardson, C.C. (2008). Interactions of *Escherichia coli* thioredoxin, the processivity factor, with bacteriophage T7 DNA polymerase and helicase. J. Biol. Chem. *283*, 32077–32084.

Gilson, M.K., and Zhou, H.X. (2007). Calculation of protein-ligand binding affinities. Annu. Rev. Biophys. Biomol. Struct. *36*, 21–42.

Go, N., Noguti, T., and Nishikawa, T. (1983). Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc. Natl. Acad. Sci. U S A *80*, 3696–3700.

Goldstone, D.C., Yap, M.W., Robertson, L.E., Haire, L.F., Taylor, W.R., Katzourakis, A., Stoye, J.P., and Taylor, I.A. (2010). Structural and functional analysis of prehistoric lentiviruses uncovers an ancient molecular interface. Cell Host Microbe *8*, 248–259.

Goodey, N.M., and Benkovic, S.J. (2008). Allosteric regulation and catalysis emerge via a common route. Nat. Chem. Biol. *4*, 474–482.

Grant, B.J., Gorfe, A.A., and McCammon, J.A. (2010). Large conformational changes in proteins: signaling and other functions. Curr. Opin. Struct. Biol. *20*, 142–147.

Grünberg, R., Nilges, M., and Leckner, J. (2006). Flexibility and conformational entropy in protein-protein binding. Structure *14*, 683–693.

Gunasekaran, K., Ma, B., and Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? Proteins Struct. Funct. Genet. *57*, 433–443.

Hawkins, R.J., and McLeish, T.C.B. (2004). Coarse-grained model of entropic allostery. Phys. Rev. Lett. *93*, 098104.

Hawkins, R.J., and McLeish, T.C.B. (2006). Coupling of global and local vibrational modes in dynamic allostery of proteins. Biophys. J. *91*, 2055–2062.

Hitchcock, S.E. (1980). Actin deoxyribonuclease I interaction. J. Biol. Chem. *255*, 5668–5673.

Holliday, M.J., Camilloni, C., Armstrong, G.S., Vendruscolo, M., and Eisenmesser, E.Z. (2017). Networks of dynamic allostery regulate enzyme function. Structure *25*, 276–286.

Huberts, D.H.E.W., and van der Klei, I.J. (2010). Moonlighting proteins: an intriguing mode of multitasking. Biochim. Biophys. Acta *1803*, 520–525.

Jeffery, C.J. (2009). Moonlighting proteins—an update. Mol. Biosyst. *5*, 345–350.

Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., and Holmes, K.C. (1990). Atomic structure of the actin:DNase I complex. Nature *347*, 37–44.

Kern, D., and Zuiderweg, E.R.P. (2003). The role of dynamics in allosteric regulation. Curr. Opin. Struct. Biol. *13*, 748–757.

Kollman, P.A. (1993). Free energy calculations: applications to chemical and biochemical phenomena. Chem. Rev. *93*, 2395–2417.

Koshland, D.E. (1959). Enzyme flexibility and enzyme action. J. Cell. Comp. Physiol. *54*, 245–258.

Krull, F., Korff, G., Elghobashi-Meinhardt, N., and Knapp, E.W. (2015). ProPairs: a data set for protein-protein docking. J. Chem. Inf. Model. *55*, 1495–1507.

Kulczyk, A.W., and Richardson, C.C. (2016). The replication system of bacteriophage T7. Enzymes *39*, 89–136.

Kumar, S., Ma, B., Tsai, C.J., Sinha, N., and Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. Protein Sci. *9*, 10–19.

Lahm, A., and Suck, D. (1991). DNase I-induced DNA conformation. 2 Å structure of a DNase I-octamer complex. J. Mol. Biol. *222*, 645–667.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., Mcgettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947–2948.

Lee, S.J., Tran, N.Q., Lee, J., and Richardson, C.C. (2018). Hydrophobic residue in *Escherichia coli* thioredoxin critical for the processivity of T7 DNA polymerase. Biochemistry *57*, 587–5017.

Levy, E.D., and Pereira-Leal, J.B. (2008). Evolution and dynamics of protein interactions and networks. Curr. Opin. Struct. Biol. *18*, 349–357.

Liu, C., Perilla, J.R., Ning, J., Lu, M., Hou, G., Ramalho, R., Himes, B.A., Zhao, G., Bedwell, G.J., Byeon, I.J., et al. (2016). Cyclophilin A stabilizes the HIV-1 capsid through a novel non-canonical binding site. Nat. Commun. *7*, 10714.

Louet, M., Seifert, C., Hensen, U., and Gräter, F. (2015). Dynamic allostery of the catabolite activator protein revealed by interatomic forces. PLoS Comput. Biol. *11*, e1004358.

Lu, M., Hou, G., Zhang, H., Suiter, C.L., Ahn, J., Byeon, I.-J.L., Perilla, J.R., Langmead, C.J., Hung, I., Gor'kov, P.L., et al. (2015). Dynamic allostery governs cyclophilin A–HIV capsid interplay. Proc. Natl. Acad. Sci. U S A *112*, 14617–14622.

Marcos, E., Crehuet, R., and Bahar, I. (2011). Changes in dynamics upon oligomerization regulate substrate binding and allostery in amino acid kinase family members. PLoS Comput. Biol. *7*, e1002201.

Marsh, J.A., and Teichmann, S.A. (2014). Protein flexibility facilitates quaternary structure assembly and evolution. PLoS Biol. *12*, e1001870.

Martin, J., Regad, L., Lecornet, H., and Camproux, A.-C. (2008a). Structural deformation upon protein-protein interaction: a structural alphabet approach. BMC Struct. Biol. *8*, 12.

Martin, J., Regad, L., Etchebest, C., and Camproux, A.C. (2008b). Taking advantage of local structure descriptors to analyze interresidue contacts in protein structures and protein complexes. Proteins Struct. Funct. Genet. *73*, 672–689.

Maun, H.R., Wen, X., Lingel, A., De Sauvage, F.J., Lazarus, R.A., Scales, S.J., and Hymowitz, S.G. (2010). Hedgehog pathway antagonist 5E1 binds hedgehog at the pseudo-active site. J. Biol. Chem. *285*, 26570–26580.

Mercier, P., Spyracopoulos, L., and Sykes, B.D. (2001). Structure, dynamics, and thermodynamics of the structural domain of troponin C in complex with the regulatory peptide 1-40 of troponin I. Biochemistry *40*, 10063–10077.

Moal, I.H., Agius, R., and Bates, P.A. (2011). Protein-protein binding affinity prediction on a diverse set of structures. Bioinformatics *27*, 3002–3009.

Monod, J., Wyman, J., and Changeux, J.P. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88–118.

Mylvaganam, S.E., Paterson, Y., and Getzoff, E.D. (1998). Structural basis for the binding of an anti-cytochrome c antibody to its antigen: crystal structures of FabE8-cytochrome c complex to 1.8 Å resolution and FabE8 to 2.26 Å resolution. J. Mol. Biol. *281*, 301–322.

Nigro, P., Pompilio, G., and Capogrossi, M.C. (2013). Cyclophilin A: a key player for human disease. Cell Death Dis. *4*, e888.

Nussinov, R., and Tsai, C.J. (2013). Allostery in disease and in drug discovery. Cell *153*, 293–305.

Nussinov, R., and Tsai, C.J. (2015). Allostery without a conformational change? Revisiting the paradigm. Curr. Opin. Struct. Biol. *30*, 17–24.

Oliwa, T., and Shen, Y. (2015). cNMA: a framework of encounter complex-based normal mode analysis to model conformational changes in protein interactions. Bioinformatics *31*, i151–i160.

Olejniczak, E.T., Zhou, M.M., and Fesik, S.W. (1997). Changes in the NMR-derived motional parameters of the insulin receptor substrate 1 phosphotyrosine binding domain upon binding to an interleukin 4 receptor phosphopeptide. Biochemistry *36*, 4118–4124.

Popovych, N., Sun, S., Ebright, R.H., and Kalodimos, C.G. (2006). Dynamically driven protein allostery. Nat. Struct. Mol. Biol. *13*, 831–838.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing (R Foundation).

Robert, P., and Escoufier, Y. (2006). A unifying tool for linear multivariate statistical methods: the RV- coefficient. Appl. Stat. *25*, 257.

Rodgers, T.L., Townsend, P.D., Burnell, D., Jones, M.L., Richards, S.A., McLeish, T.C.B., Pohl, E., Wilson, M.R., and Cann, M.J. (2013). Modulation of global low-frequency motions underlies allosteric regulation: demonstration in CRP/FNR family transcription factors. PLoS Biol. *11*, e1001651.

Rodriguez-Bussey, I., Yao, X.Q., Shouaib, A.D., Lopez, J., and Hamelberg, D. (2018). Decoding allosteric communication pathways in cyclophilin A with a comparative analysis of perturbed conformational ensembles. J. Phys. Chem. B *122*, 6528–6535.

Schreiber, G., and Keating, A.E. (2011). Protein binding specificity versus promiscuity. Curr. Opin. Struct. Biol. *21*, 50–61.

Siew, N., Elofsson, A., Rychlewski, L., and Fischer, D. (2000). MaxSub: an automated measure for the assessment of protein structure prediction quality. Bioinformatics *16*, 776–785.

Skrbic, T., Zamuner, S., Hong, R., Seno, F., Laio, A., and Trovato, A. (2018). Vibrational entropy estimation can improve binding affinity prediction for non-obligatory protein complexes. Proteins *86*, 393–404.

Smith, G.R., Sternberg, M.J.E., and Bates, P.A. (2005). The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking. J. Mol. Biol. *347*, 1077–1101.

Swapna, L.S., Mahajan, S., de Brevern, A.G., and Srinivasan, N. (2012a). Comparison of tertiary structures of proteins in protein-protein complexes with unbound forms suggests prevalence of allostery in signalling proteins. BMC Struct. Biol. *12*, 6.

Tama, F., and Sanejouand, Y.-H. (2002). Conformational change of proteins arising from normal mode calculations. Protein Eng. Des. Sel. *14*, 1–6.

Thali, M., Bukovsky, A., Kondo, E., Rosenwlrth, B., Walsh, C.T., Sodroski, J., and Göttlinger, H.G. (1994). Functional association of cyclophilin A with HIV-1 virions. Nature *372*, 363–365.

Tirion, M.M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys. Rev. Lett. *77*, 1905–1908.

Tsai, C.J., and Nussinov, R. (2014). A unified view of "how allostery works. PLoS Comput. Biol. *10*, e1003394.

Tsai, C.J., Del Sol, A., and Nussinov, R. (2009). Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. Mol. Biosyst. *5*, 207–216.

Tzeng, S.R., and Kalodimos, C.G. (2015). The role of slow and fast protein motions in allosteric interactions. Biophys. Rev. *7*, 251–255.

Valadié, H., Lacapčre, J.J., Sanejouand, Y.H., and Etchebest, C. (2003). Dynamical properties of the MscL of *Escherichia coli*: a normal mode analysis. J. Mol. Biol. *332*, 657–674.

Vishwanath, S., de Brevern, A.G., and Srinivasan, N. (2018). Same but not alike: structure, flexibility and energetics of domains in multi-domain proteins are influenced by the presence of other domains. PLoS Comput. Biol. *14*, e1006008.

Vreven, T., Moal, I.H., Vangone, A., Pierce, B.G., Kastritis, P.L., Torchala, M., Chaleil, R., Jiménez-García, B., Bates, P.A., Fernandez-Recio, J., et al. (2015). Updates to the integrated protein-protein interaction benchmarks: docking benchmark version 5 and affinity benchmark version 2. J. Mol. Biol. *427*, 3031–3041.

Wang, C., Pawley, N.H., and Nicholson, L.K. (2001). The role of backbone motions in ligand binding to the c-Src SH3 domain. J. Mol. Biol. *313*, 873–887.

Wapeesittipan, P., Mey, A., Walkinshaw, M., and Michel, J. (2019). Allosteric effects in catalytic impaired variants of the enzyme cyclophilin A may be explained by changes in nano-microsecond time scale motions. Commun. Chem. *2*, 41.

Xue, L.C., Rodrigues, J.P., Kastritis, P.L., Bonvin, A.M., and Vangone, A. (2016). PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. Bioinformatics *32*, 3676–3678.

Zemla, A. (2003). LGA: a method for finding 3D similarities in protein structures. Nucleic Acids Res. *31*, 3370–3374.

Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. *33*, 2302–2309.

Zhang, L., Centa, T., and Buck, M. (2014). Structure and dynamics analysis on plexin-B1 Rho GTPase binding domain as a monomer and dimer. J. Phys. Chem. B *118*, 7302–7311.

Zidek, L., Novotny, M.V., and Stone, M.J. (1999). Increased protein backbone conformational entropy upon hydrophobic ligand binding. Nat. Struct. Biol. *6*, 1118–1121.

Zeller, T., and Klug, G. (2006). Thioredoxins in bacteria: functions in oxidative stress response and regulation of thioredoxin genes. Naturwissenschaften *93*, 259–266.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Protein data bank	Berman et al., 2000	www.rscb.org
Protein-protein benchmark dataset 5.0	Vreven et al., 2015	https://zlab.umassmed.edu/benchmark/
ProPairs	Krull et al., 2015	http://propairs.github.io
Software and Algorithms		
MaxCluster	Siew et al., 2000	http://www.sbg.bio.ic.ac.uk/maxcluster/
ProDy	Bakan et al., 2011	http://prody.csb.pitt.edu/
PyMol molecular graphics system	Schrödinger, LLC	https://pymol.org/2/

RESOURCE AVAILABLITY

Lead contact

Further information and requests for information on method, dataset or computational resources should be directed to and will be fulfilled by the Lead Contact, Prof. N. Srinivasan (ns@iisc.ac.in).

Materials availability

No new unique reagents or methods were produced in this study.

Data and code availability

The PDB codes used for structural and dynamics analyses can be accessed from PDB. The codes are provided in the Supplementary Information as Tables S1-S4, S5, and S6. The program used to perform structural and dynamics analysis can be obtained from Maxcluster software and ProDy website (http://prody.csb.pitt.edu/) respectively.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Not applicable

METHOD DETAILS

Dataset preparation

PPC dataset for structural comparison

A dataset of transient PPCs was prepared using ProPairs program (Krull et al., 2015) and Benchmark dataset 5.0 (Vreven et al., 2015). ProPairs program compiles a dataset of proteins in their bound and unbound forms from PDB in an automated manner. Benchmark dataset 5.0 is a docking benchmark consisting of non-redundant structures of PPCs and unbound structures of their components. 2,943 redundant complexes were obtained using ProPairs and 230 complexes were collected from benchmark dataset. The two datasets were pruned to obtain cases that pass through the following criteria:

1. 3-D structures of all proteins in a complex should be available in their unbound forms and there should be no missing residues in the interface region.
2. Resolution of the structure of the complex and individual unbound forms should be better than 3.2Å.
3. Bound and unbound forms should have the same uniport identifier.
4. Unbound and bound forms should have same oligomeric state. PDB biological unit information, PISA as well as relevant literature were reviewed to enforce the condition of same oligomeric state of the proteins in complexed and non-complexed forms. This condition was imposed to ensure that differences obtained, if any, between the bound and unbound forms are not due to different oligomeric states.
5. Both bound and unbound forms should either have similar or no ligands bound to it. This condition was employed to minimise the bias due to presence of a ligand.
6. There should be no occurrence of disordered regions in the proteins. Benchmark dataset 5.0 already takes care of this condition.

The filtered dataset was then made non-redundant at the SCOP (Fox et al., 2014) family level. For the ProPairs entries, SCOP domains were assigned to each entry and clustered according to SCOP families, wherever available. The Benchmark dataset 5.0 is already non-redundant at SCOP family level. Individual cases were also inspected manually to check for any unexpected discrepancy that might have crept in during automated handling of the data. Finally, after applying these stringent filters, a dataset of high-quality, non-redundant, 120 complexes (i.e. 120*3 = 360 protein structures) were curated for structural analyses (see Table S5). The dataset comprises of 18 antigen-antibody, 45 enzyme-inhibitor and 57 other complexes (Nomenclature based on Benchmark dataset 5.0), representing the fourSCOP classes. This dataset has only 50% overlap with the dataset from (Swapna et al., 2012a)due to more strict conditions and stringent definition of oligomeric states and ligands.

PPC dataset for dynamics analyses

To understand the modulation of dynamics upon binding of two proteins, the above dataset was further filtered to remove cases with missing residues in the structure in either bound or unbound form. This condition was enforced to avoid introducing bias due to modelling of the missing regions in the structure. 58 complexes (i.e. 58 * 3 = 174 structures) matched the criteria and were selected for the analyses (see Table S6). This dataset consists of 11 antigen-antibody, 21 enzyme-inhibitor and 26 other complexes.

Control datasets

Two types of control datasets were used:

1. *Monomeric proteins solved under different crystal conditions (Control dataset 1)*: A collection of protein crystal structures with a single chain in both asymmetric unit (ASU) and biological unit (BU) were curated from PDB (Berman et al., 2000). Care was taken to assure that the structures do not have any other biological entity such as peptide, RNA or DNA in the ASU and BU. This set was further filtered using a resolution cut-off of 2.5 Å and was subjected to clustering at 100% sequence identity using CD-HIT (Fu et al., 2012). Additionally, structures with missing residues were removed and 883 pairs, in total, were chosen for analyses. This dataset was curated to understand the influence of crystal packing on protein conformation and was treated as a background noise while selecting cut-offs for significant difference between bound and unbound forms.
2. *Fictitious-unbound protein dataset (Control dataset 2)*: To analyse the effect of crystal packing on the dynamics of the proteins, this dataset was created by *in-silico* deletion of the partner protein from the PPC dataset for dynamics analyses. The deletion resulted into artificial unbound proteins which are equivalent to the sequence and length of actual unbound proteins.

Structural analyses

Proteins in the bound and their respective unbound forms were compared after identifying the residue equivalences using CLUSTALW (Larkin et al., 2007). TM-align (Zhang and Skolnick, 2005) was used to structurally align the bound and unbound forms. Two measures, namely, root mean square deviation (RMSD) and global distance test – total score (GDT-TS) (Zemla, 2003), were used to calculate the global similarities between the structures. While RMSD provides an estimate of distance between pairs of atoms, GDT is used to calculate similarities between structures of proteins with identical sequences. GDT is independent of protein length unlike RMSD and hence considered as a better similarity criterion to assess global similarity. RMSD is calculated as;

$$\sqrt{\frac{1}{N}\sum di^2},$$

where *d* is the distance between N pairs of equivalent atoms and *i* ranges from residue 1 to N. Lower the RMSD value, higher is the similarity. RMSD values greater than the standard deviation from the mean of the RMSDs for proteins in the control dataset 1 were considered significant. GDT is calculated as:

GDT-TS = 100 * (C1 + C2 + C3 + C4) / 4N
Where, C1 = Count of number of residues superposed below (threshold/4)
C2 = Count of number of residues superposed below (threshold/2)
C3 = Count of number of residues superposed below (threshold)
C4 = Count of number of residues superposed below (2*threshold)
N = Total number of residues

MAXCLUSTER (Siew et al., 2000) algorithm with a distance cut-off of 4 Å was used to calculate GDT-TS. Higher the GDT-TS value, higher is the similarity. Both RMSD and GDT-TS were calculated for all Cα positions.

To calculate local structure variations, individual Cα deviations between equivalent residue positions were obtained. All the residues that showed a Cα deviation greater than the standard deviation from the mean of the Cα deviations for residues in the control dataset 1 were considered as showing significant structural change.

Dynamics analyses using normal mode analysis (NMA)

Normal mode analysis (NMA) is one of the methods of choice to study long-timescale motions associated with proteins (Bahar and Rader, 2005; Brooks and Karplus, 1983; Go et al., 1983). NMA requires a set of cartesian coordinates from protein structure and a force-field that defines interactions between the atoms. A ''Hessian'' matrix is then generated from second derivate of the potential energy and is diagonalised to yield eigen vectors and eigen values. Low frequency collective motions, termed as, global motions of proteins have been shown to signify a biologically relevant function (Bahar et al., 1998; 2010b; General et al., 2014). An all-atom NMA poses a significant computational problem due to massive calculations to solve complete spectra of motions. Hence, for the current study, coarse-grained (Cα-level) anisotropic network model (ANM)-based NMA was used (Tirion, 1996). The low-frequency modes from Cα-level NMA have been shown to corroborate well with the experimental as well as molecular dynamics data in the past (Delarue and Sanejouand, 2002; Valadié et al., 2003; Vishwanath et al., 2018). Hence, coarse-grained models can successfully provide idea of dynamics for longer time-scales. All the calculations pertaining to normal mode analyses were performed using Prody package (Bakan et al., 2011).

Normal modes were calculated for the dataset of 58 PPCs both in bound and unbound states (i.e. 58*3=174 calculations), the control dataset 1 and control dataset 2. A distance cut-off of 15 Å was employed and normal modes pertaining to 80% of the variance were analysed. Contributions from 5 N-terminal and 5 C-terminal residues were removed, unless they were part of the interface. Mean squared fluctuations were scaled using a z-score normalisation. A difference greater than 1, between the normalised square fluctuations of bound and unbound forms was considered significant. The cut-off was derived based on the difference between the normalised square fluctuations of control dataset 1. To obtain an equivalent of RMSD, root mean squared difference of fluctuations ($RMSD^f$) was calculated for dynamics analyses. It was calculated like RMSD, but instead of deviations, difference between the normalised fluctuations of bound and unbound form was used. Correlation between the fluctuations, namely cross-correlation was also calculated for all the structures. R_v coefficient (Robert and Escoufier, 2006), which is a multivariate generalisation of Pearson's coefficient was used to quantify the similarity between cross-correlation matrices. Similarity between the conformational space accessible to a subset of modes, called as overlap (Fuglebakk et al., 2015; Tama and Sanejouand, 2002), was further calculated for 10 lowest frequency modes using Prody package. Overlap gives an estimate of the extent to which the intrinsic motions of the protein in bound form are accessible to the unbound form and is calculated as the inner product of the eigenvectors calculated using NMA.

Classification of interface & non-interface residues

Protein residues were classified in this study into three types, viz. interface, near-interface and far from interface as per the following distance cut-offs. Atoms of the residues from the two proteins which lie at a distance $\leq$ 4.5Å were classified as interface residues. Atoms from the two proteins which lie at a distance > 4.5Å but $\leq$10 Å, were called as near-interface residues and the rest were classified as far from interface. Interface residues were further categorised into "core" and "rim" residues depending upon their solvent accessibility. Interface residues having solvent accessibility $\leq$7% in bound form were considered buried and hence termed "core" and the remaining interface residues were termed as "rim".

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed using R free software environment (version 3.3.0) (R Development Core Team, 2011).

ADDITIONAL RESOURCES

Not applicable