

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

Himani Tandon, Alexandre de Brevern, Narayanaswamy Srinivasan

▶ **To cite this version:**

Himani Tandon, Alexandre de Brevern, Narayanaswamy Srinivasan. Transient association between proteins elicits alteration of dynamics at sites far away from interfaces. Structure, 2021, 29 (4), pp.371-384.e3. 10.1016/j.str.2020.11.015 . inserm-03543689

HAL Id: inserm-03543689

https://inserm.hal.science/inserm-03543689

Submitted on 26 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Published in Structure (2021) 29(4):371-384

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

Himani Tandon[1], Alexandre G. deBrevern[2,3,4,5] and Narayanaswamy Srinivasan[1*]

[1]Molecular Biophysics Unit

[2]INSERM, U 1134, DSIMB, F-75739 Paris, France.

[3]Univ Paris, UMR_S 1134, F-75739 Paris, France.

[4]Institut National de la Transfusion Sanguine (INTS), F-75739 Paris, France.

[5]Laboratoire d'Excellence GR-Ex, F-75739 Paris, France.

* Author for correspondence and Lead Contact: ns@iisc.ac.in, Tel: +91-80-22932837

Summary

Proteins are known to undergo structural changes upon binding to partner proteins. But the prevalence, extent, location, and function of change in protein dynamics due to transient protein-protein interactions is not well-documented. Here, we have analysed a dataset of 58 protein-protein complexes of known 3-D structure and structures of their corresponding unbound forms to evaluate dynamics changes induced by binding. 55% 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 for about 90% 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, results presented here strongly suggest that alteration of protein dynamics due to binding of a partner protein commonly occurs.

Keywords: Protein-protein interactions, protein structure, protein dynamics, normal mode analysis, functional analysis, allostery, bioinformatics

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 signalling process, recruit components of bigger complex, and inhibit or trigger molecular function (Tsai *et al*., 2009). Generally, a conformational change is induced by 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 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 *viz*., induced fit and conformational selection, were proposed as early as in 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. Further, Frauenfelder *et al* discussed the existence of static broad ensembles of states (Frauenfelder et al., 1991). Later, a dynamic landscape, which is the basis of 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 analysed 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). Further, Swapna *et al* 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 (Swapna *et al*., 2012a). The idea of allostery without a significant structural change was proposed in 1984 (Cooper and Dryden, 1984), but it is only until recently, it has been fully appreciated. In their seminal paper, Cooper and Dryden 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 (Cooper and Dryden, 1984). This changing view of allostery has been appreciated in the last decade with popularisation of NMR methods to study protein motions. Proteins like catabolite activator protein (CAP) have

been shown to exhibit dynamic allostery, arising due to changes in intrinsic dynamics of the structure upon cAMP-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 change 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 above-mentioned 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. 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 analyses 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 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, for 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, for most of them, the modes are re-ordered. We further highlight with examples how an alteration in dynamics is related to function. Taken together, results presented here strongly suggest that the alteration of dynamics, upon interaction of two proteins, occur more frequently than previously thought.

Results

Binding of proteins influences the conformations of associated proteins

In an earlier study, Swapna *et al* showed that proteins bound to other proteins undergo larger structural changes compared to proteins in the unbound form (Swapna *et al*., 2012a). With the availability of far more structures of proteins in bound and unbound forms in protein data bank (PDB), it is worthwhile exploring the extent of structural changes induced by PPIs. So, a comprehensive dataset of 120 protein complexes was prepared for this study as described in materials and methods section. The RMSD and 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 out 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, RMSD distribution for 120 complexes was compared with that of control dataset 1 (details in methods) (**Figure 1B**). The two distributions (mean values 0.45Å and 1.47Å for control dataset 1 and PPC dataset respectively) were found to be significantly different (two sample KS-test, p-value < 2.2×10^{-16}), suggesting that the observed differences in global protein conformation is mainly due to binding of another protein and not due to crystallisation artefacts.

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

Residue dynamics get altered upon protein-protein complex formation

To understand the extent of change in residue dynamics between the bound and unbound form of a protein, normalised square fluctuations obtained from normal mode analysis (NMA) were analysed for 58 complexes. These were obtained after filtering cases with missing residues in the structure in either bound or unbound form (see materials and methods section). 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 normalised square fluctuations for the proteins in bound and unbound forms were found to be significantly different (two-sample KS test, p-value < 2.2e-16) **(Figure 2A)**. A higher variance in the fluctuation distribution of bound form was observed, suggesting that many residues in the bound form show change in flexibility **(Figure 2A)**. To impose confidence on the results, two control studies were performed. First, the difference between normalised square fluctuations of the bound and unbound form of proteins from PPC dataset was compared with the difference between normalised 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-value < 2.2×10^{-16}). Second, the normalised square fluctuations of the unbound form were compared with those of the fictitious unbound dataset i.e. control dataset 2 (details in methods) **(Figure 2C)**. The fluctuation profiles of both the datasets were not significantly different from each other (two-sample KS test, p-value = 0.10). These results suggest that differences in fluctuations between bound and unbound form **(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, difference between residue fluctuations was calculated. 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 bound form than the unbound form. On the other hand, ~11% of the residues showed significantly higher fluctuations in unbound form when compared to the bound form. To ascertain that these differences are insensitive to distance cut-off of 15Å used for NMA calculations, normalised fluctuations of the bound and unbound proteins were calculated at 12Å and 10Å cut-offs as well. It was observed that differences in distributions of normalised square fluctuations were significant irrespective of the cut-offs used **(Figure S2A,B)**. For both cut-offs (i.e.12Å and 10Å), ~7% of the residues showed significantly higher fluctuations

in bound form than the unbound form and ~6% of the residues showed significantly higher fluctuations in unbound form than 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 analysed separately **(Figure 3)**. The distributions were found to be significantly different for the three regions (two-sample KS test, p-value < 2.2e-16). Intuitively, interface residues showed higher fluctuations in unbound form with ~28% of the residues showing significantly higher fluctuations **(Figure 3A)**. Rest 72% showed comparable fluctuations. The interface residues were further divided into "core" and "rim" **(Figure 3A)**. It must be noticed 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. It underlines 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 near interface region, ~5% 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% residues showed higher fluctuations in the bound form than in unbound form, whereas ~5% residues showed higher fluctuations in the unbound forms **(Figure 3C)**. This result counters the general idea that flexibility of interacting partners should decrease, when in complex. Our findings suggest that a loss of conformational entropy at the interface is likely compensated by 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 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, RMSD and $RMSD^f$ values were compared 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, the dataset (58*2=116 proteins) was divided into 4 categories: i) proteins that only show significant structural changes away from the interfaces, ii) proteins that only show significant change in dynamics, away from the interfaces, iii) proteins that show significant changes in structure and dynamics, away from the interfaces, and iv) proteins that show no significant changes either in structure or dynamics, away from the interfaces **(Figure 4B and Table S4)**. 15/116 cases (12.9%) belong to first category, 17/116 (14.7%) to second, 35/116 (30.2%) to third and 49/116 (42.2%) to fourth category. It is interesting to note that though 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 was 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). Further, residues coupled in motion are helpful in constructing a pathway between allosteric site and functional site (Gerek and Ozkan, 2011). So, 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 gets affected. R_v coefficient was calculated between matrices of bound and unbound forms **(Figure 5A)**. Coupling between the fluctuations was found to be unaffected for only 12/116 proteins (10.3%) after binding of the partner protein (R_v coefficient ≥ 0.7). Remaining proteins showed change in synchronisation 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 to the unbound form. An example from the dataset where cross-correlation gets affected by binding of a partner protein is shown in **Figure 5B** and one example where it does not get affected by binding of a partner protein is shown in **Figure 5C**. 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 gets affected by binding of a partner protein **(Figure 5D)**.

Low frequency global modes of unbound form get perturbed by binding of interacting partner

Low frequency global modes from NMA are known to be biologically relevant (Bahar *et al*., 1998, 2010a). Previously, Marcos *et al* showed that new modes of motions are acquired by enzymes form the amino acid kinase family upon oligomerisation, which regulate the substrate binding and allostery (Marcos *et al*., 2011). Also, an attempt was made to model conformational changes upon binding by re-ranking of normal modes (Oliwa and Shen, 2015). Hence, we asked the question, if this is a commonly observed phenomenon upon protein binding? In other words, how the low frequency modes of unbound protein get altered in the presence of an interacting partner? or, are the low frequency motions accessible to a protein unique or are they maintained in bound form too? To answer this question, we analysed the similarities/differences between the modes of motion accessible to a protein in its bound and unbound forms by calculating overlap between the top 10 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. Smaller the overlap, different are the two modes of motion. ~58% cases showed high overlap (> |0.7|) between bound and unbound forms for at least one mode (from top 10) **(Figure 6A)**. However, out of these, ~60% of the cases showed change in mode order/preference **(Figure S3)**. A change in mode preference means, 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 though some modes of motion are preserved between the bound and unbound forms, their frequency, size and shape changes as suggested by re-ordering of normal modes. For the remaining ~42% cases, very few global modes in unbound form were found maintained in the bound form, that too with weaker correlation (<|0.7|) and re-ordering of modes **(Figure 6A)**. This clearly suggests that dynamics of the unbound form gets affected by binding of partner protein. Example of a case where the mode order and shape were retained between the bound and unbound form is shown in **Figure 6B**, example of change in mode preference in a high overlap case is shown in **Figure 6C**, example of a case with low overlap value is shown in **Figure 6D.** To rule out the effect of crystal packing on the observed dissimilarity of intrinsic modes, overlap for all the pairs in control dataset 1 was calculated. Each pair showed an overlap value ≥ |0.9|, suggesting similarities in their modes **(Figure 6E)**. Hence, the global motions of 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, re-enforcing 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, 17/116 proteins showed significant change in dynamics at sites away from interfaces without a significant structural change. Out of the 17 proteins, 3 are antigen-antibody complexes, 4 are enzyme-inhibitor complexes and 10 are either enzyme-substrate or signalling complexes. This study proposes that these differences are likely involved in either the stability of the complexes or signalling a downstream protein or both. Role of differential dynamics is presented for two cases below and three other cases in supporting information (**Data S1**).

Differential dynamics of Cyclophilin A (CypA) likely plays a role in stabilising HIV-1 capsid assembly

The first example we describe is the human Cyclophilin A (CypA), a peptidyl-prolyl enzyme that catalyses the isomerisation of peptide bonds from trans to cis form and participates in various biological process such as protein folding, apoptosis and signalling (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)stabilisation 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). Though 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 cryo-EM structure of CypA-CA assembly was solved at 8Å (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 site, thus stabilising the CA assembly. Moreover, the non-canonical binding site, by itself, has weaker affinity for CA, but binds 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 analysed the fluctuation profiles, cross-correlation matrices and overlaps obtained from NMA of the free and bound-form of CypA. A comparison of the structure of free and CA-bound at canonical

binding site of CypA suggests no significant structural changes in CypA (RMSD=0.3Å) **(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 form 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 ~16Å, it suggests a long-range communication between the two sites **(Figure 7E)**. The residue Val29 (part of 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 unbound form with more global, collective motions in bound form. It presents 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 a manner like 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/differences between the dynamics of CypA bound to two evolutionarily conserved partners and our results propose a possibility of similar binding mode like HIV CA (**see also Data S1, Figure S4**).

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

DNAse-I is an endonuclease that cleaves dsDNA 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, 1991). It also interacts with actin monomer to form a 1:1 actin-DNAse-I complex (Hitchcock, 1980). Though 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 code: 2dnj) (Lahm and Suck, 1991) and bound to actin monomer (PDB code: 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 stearic hinderance 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 0.35Å) **(Figure 8B)**. But 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 synchronised motions of DNA-binding sites of DNAse-I. This is suggestive of existence of a communication pathway between actin-interface and DNA-interface **(Figure 8D)**. From these results, we propose that apart from steric hinderance for DNA-binding, actin molecule can affect the vibrational motions intrinsic to DNAse-I., which in turn can affect the dynamics of DNA-binding residues, contributing towards inactivity of DNAse-I by actin.

Apart from these two cases, we observed a redistribution of dynamics upon binding of partner protein in two anitgen-antibody complexes, *viz.* cytochrome-C (Cyt-C) & E8 antibody and sonic-hedgehog (Shh) protein & 5E1 antibody fragment complex. Such redistribution was found to contribute to the stability of the complex (**See also Data S1, Figure S5**). Another protein, beta-lactamase inhibitor protein-II (BLIP-II), 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 also Data S1, Figure S6**). The analysis was also performed for thioredoxin, a moonlighting protein. Since moonlighting proteins can bind diverse partners under different conditions, it is interesting to analyze if dynamics of a moonlighting protein gets altered upon binding of another protein without a significant change in structure. It was observed that a change in Thioredoxin (trx) flexibility likely helps in better packing of T7 bacteriophage replisome (**See also Data S1, Figure S7**).

Discussion

Proteins are dynamic systems that undergo post-translational modifications (PTMs), bind to small molecules or other proteins and elicit allosteric responses at sites which are 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 the recent years,

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). It 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). Further, it has been proposed that large-scale motions associated with proteins are important carriers of allosteric signal 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 dynamics change in them (especially at the sites away from interfaces) upon binding of 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 coarse-grained anisotropic network model based NMA (ANM-NMA). Before studying flexibility/dynamics we note that our observations from structural analysis, albeit on a bigger dataset, fall in line with the findings from previous studies on individual proteins or on 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 change in conformation.

The distributions of normalised 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 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 partner protein. We believe that a loss of conformational entropy at the interface is compensated by reorganisation of motions within a protein, resulting in higher flexibility at other regions. This result reinforces similar findings by Grunberg *et al* where, for a very small dataset of 17 PPCs, authors reported an increase in flexibility of many proteins in the complexed form (Grünberg *et al*., 2006).

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 show 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 get superseded by low-frequency global motions of the bound form. We believe that synchronised motions in 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 few proteins by 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). Further, 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, though the modes overlapped well, they were maintained with a re-ordering of modes. This seems to suggest that global motions are affected by binding of a partner protein to control the functions of complex. To the best of our knowledge this was not known before 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 hypothesised that global and local modes obtained from NMA can carry signals from 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 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 towards the stability of a complex by adding to the positive gain in entropy (in case of BLIP-II interactions with beta-lactamases and antigen-antibody complexes) or regulating a downstream function such as binding of another protein (in case of CypA binding to HIV CA) or altering the functional capacity of an enzyme (in case of actin binding to DNAse I). Binding of 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 analyses of the effect of binding of two or more proteins on the dynamics of individual proteins. Since the impact on dynamics is independent of sizes 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). But Cui and Karplus raised a valid question in their classic review if these perturbation induced by binding should be considered a manifestation of allostery or the term dynamic allostery should be used where the allosteric effect has a biological function? (Cui and Karplus, 2008) Whether or not our results provide an answer to that question, the analyses of the cases in this study do reveal functional relevance of these changes. We further believe that results presented here will lead to better appreciation of allostery mediated by PPIs. We propose two direct applications of our work. First, the lessons learnt are expected to be applicable to the growing knowledge on 3-D structures of large multi-protein assemblies. An understanding of allostery, dynamics and their relationship to the biological function of the proteins studied here may aid in understanding the organisation 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 towards affinity calculations. Two previous studies have shown the importance of vibrational entropy in 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 if including the free energy contributions from vibrational entropy helps in improving the accuracy of affinity calculations. Our observations are presented in **Data S1, Figures 2, S7, Table S7**. We find that regardless of proper weight assignment to vibrational entropy contributions, the accuracy of binding affinity calculation improves for many cases. Our study envisions a tremendous scope of improvement in this area.

Acknowledgments

NS and AdB acknowledge to Indo-French Centre for the Promotion of Advanced Research / CEFIPRA for collaborative grant (number 5302-2). NS 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, Bioinformatics and Computational Biology centre support from DBT and the DBT-IISc Partnership Program. N.S. is a J.C. Bose National Fellow. HT is an INSPIRE fellow.AdB 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 funded by the French National Research Agency, reference ANR-11-IDEX-0005-02. HT thank Dr. Sneha Vishwanath and other lab members for useful discussions. Authors also thank Ms. Seemadri, Ms. Yazhini and Ms. Sneha for technical assistance with remote data access during the pandemic lockdown.

Author contributions

Conceptualization:NS and ADB

Investigations:HT

Data Curation: HT

Methodology: HT, NS and ADB

Formal Analysis:HT

Validation: HT

Supervision: NS and ADB

Writing – Original Draft: HT

Writing – Review & Editing:HT, NS and ADB

Project Administration: NS and ADB

Funding Acquisition: NS and ADB

Declaration of interests

The authors declare no competing interests.

Main figure titles and legends

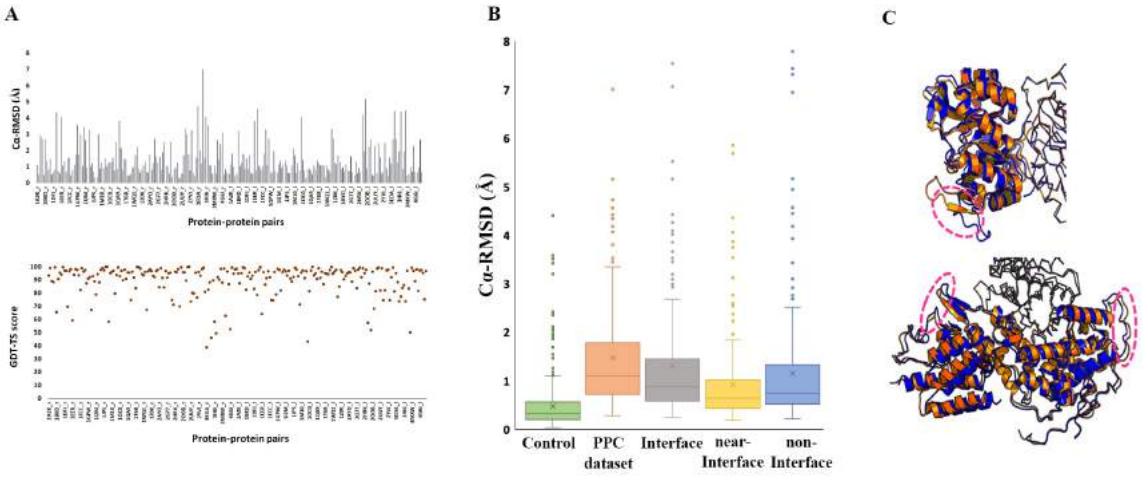


Figure 1| Structural analyses of the PPCs. (A) Cα-RMSD for protein pairs is plotted as bar plot (top panel). Higher the RMSD, greater is the structural differences between bound and unbound form. GDT-TS score is plotted in bottom panel. Here, higher the score, higher is the similarity between two structures. **(B)** Box plots showing distribution of Cα-RMSD (Å) for control dataset 1, PPC dataset, interface, near-interface & 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-value < 2.2×10^{-16}). **(C)** 2 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 code: 1gpw) and unbound (PDB code: 1thf) form. Bottom panel shows SOS (Son of sevenless) protein in its bound (PDB code: 1bkd) and unbound form (PDB code: 2iio) (See also Figure S1, Table S1,S2).

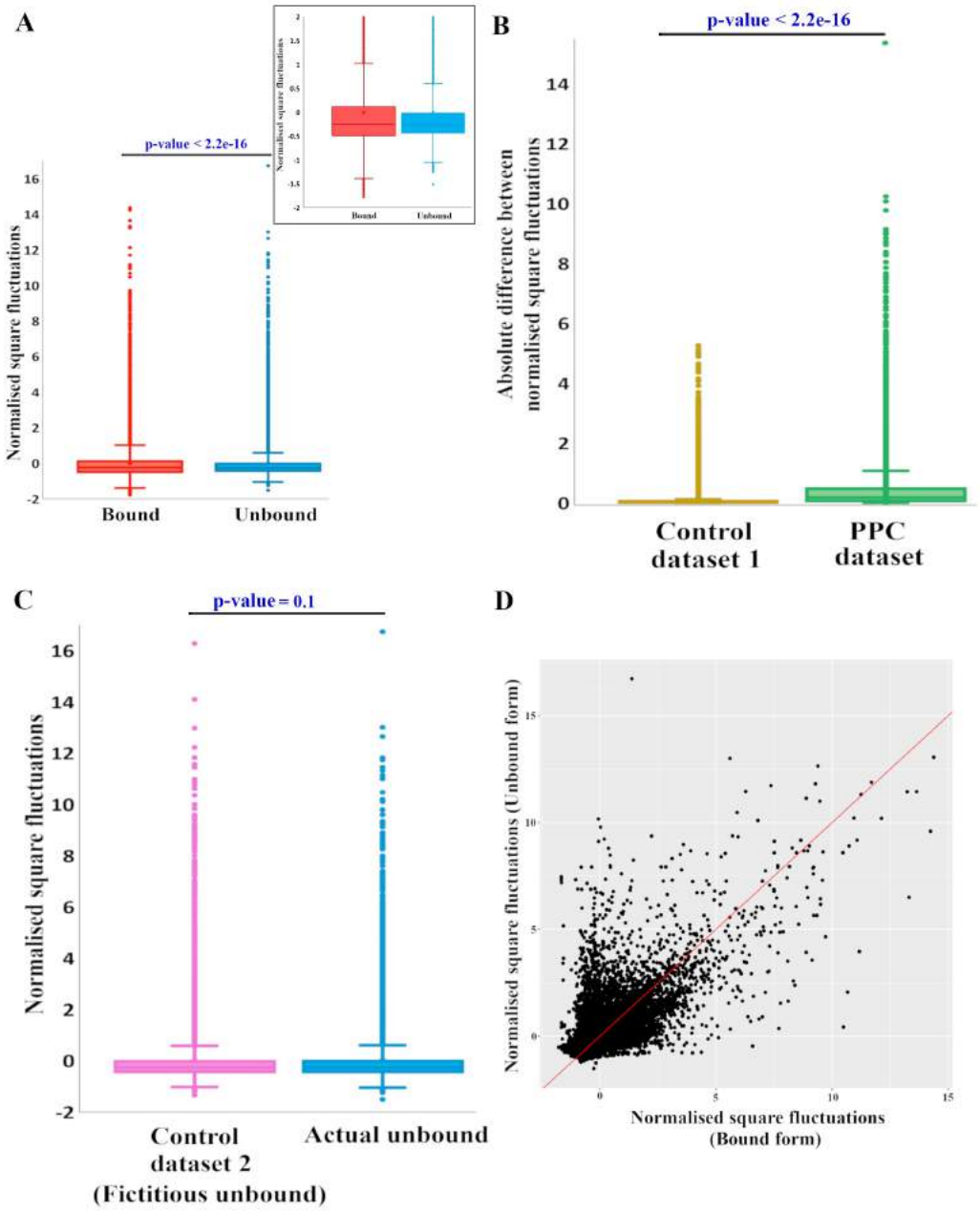


Figure 2| Analyses of dynamics of PPCs. (A) Box plots showing distribution of Cα fluctuations for all residues in bound and unbound forms. The two distributions are significantly different from each other (two-sample KS-test, p-value < $2.2x10^{-16}$), showing variation in flexibility in bound and free proteins. Box plots with fluctuations only between the range of -2 & 2 are shown in the inset for clarity. **(B)** Comparison of absolute difference between normalised fluctuations of control dataset 1 and PPC dataset. The distributions are significantly different (two-sample KS-test, p-value < 2.2x10-^{16}), suggesting no bias due to crystal packing. **(C)** Comparison of normalised 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-value=0.1), suggesting no effect of crystal packing. **(D)** Scatter plot of square fluctuations for all residues in PPC dataset. X-axis represents normalised square fluctuations for bound and Y-axis represents normalised 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 S2(A,B), S8, Data S1, Table S7).

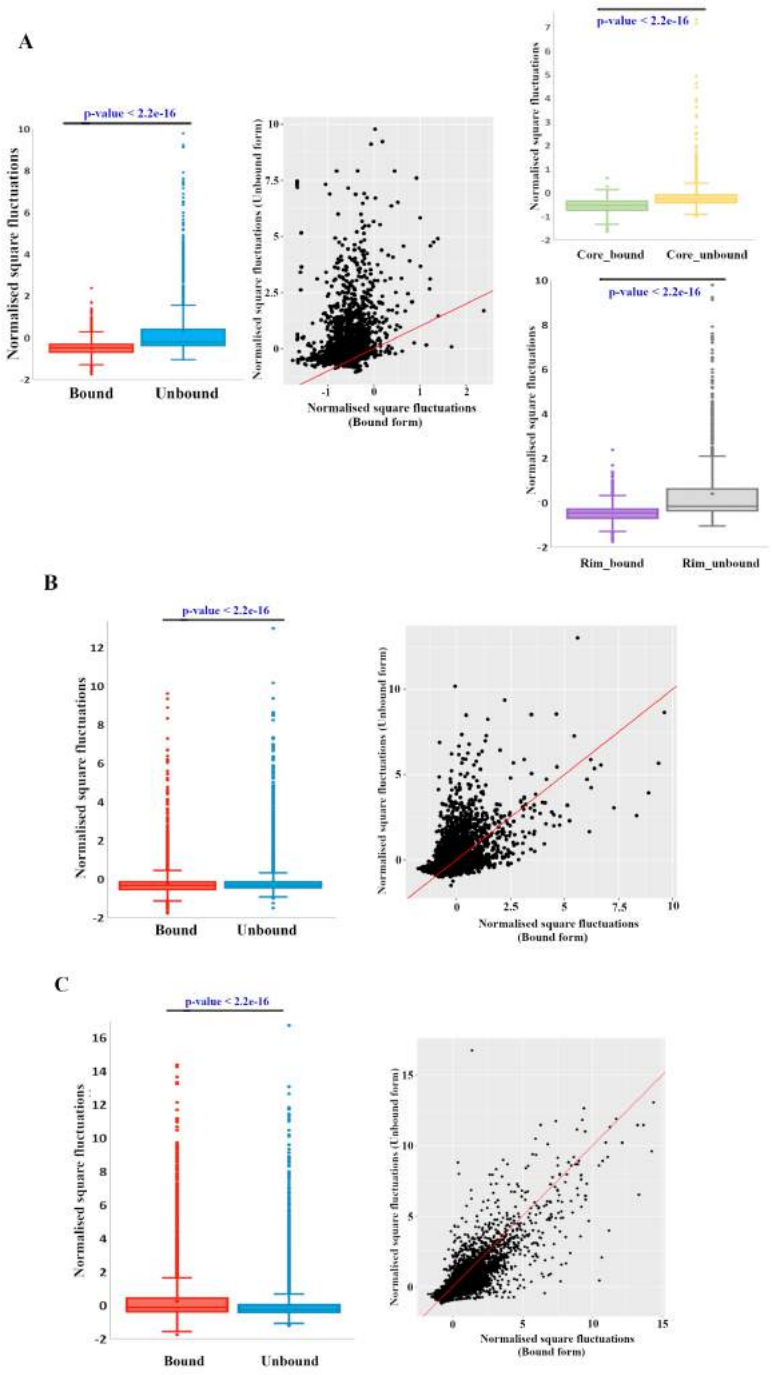


Figure 3| Comparison of fluctuations for different types of residues in proteins. (A) Box plot showing distribution of Cα fluctuations for interface residues (atoms within a distance ≤ 4.5Å) in bound and unbound proteins in PPC dataset. The distributions are significantly different from each other (two sample KS-test, p-value < $2.2x10^{-16}$). Scatter plot 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 right panel. **(B)** Box plot showing distribution of Cα fluctuations for near-interface residues (atoms within 4.5-10Å) in bound and unbound proteins in PPC dataset. The distributions are significantly different from each other (two sample KS-test, p-value < $2.2x10^{-16}$). Scatter plot 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 > 10Å) in bound and unbound proteins in PPC dataset. The distributions are significantly different from each other (two sample KS-test, p-value < $2.2x10^{-16}$). Scatter plot shows fluctuations for corresponding residues in bound and unbound forms. For all scatter plots, X-axis represents normalised square fluctuations in bound form and Y-axis represents normalised square fluctuations in unbound form (See also Figure S2(C,D), Table S3).

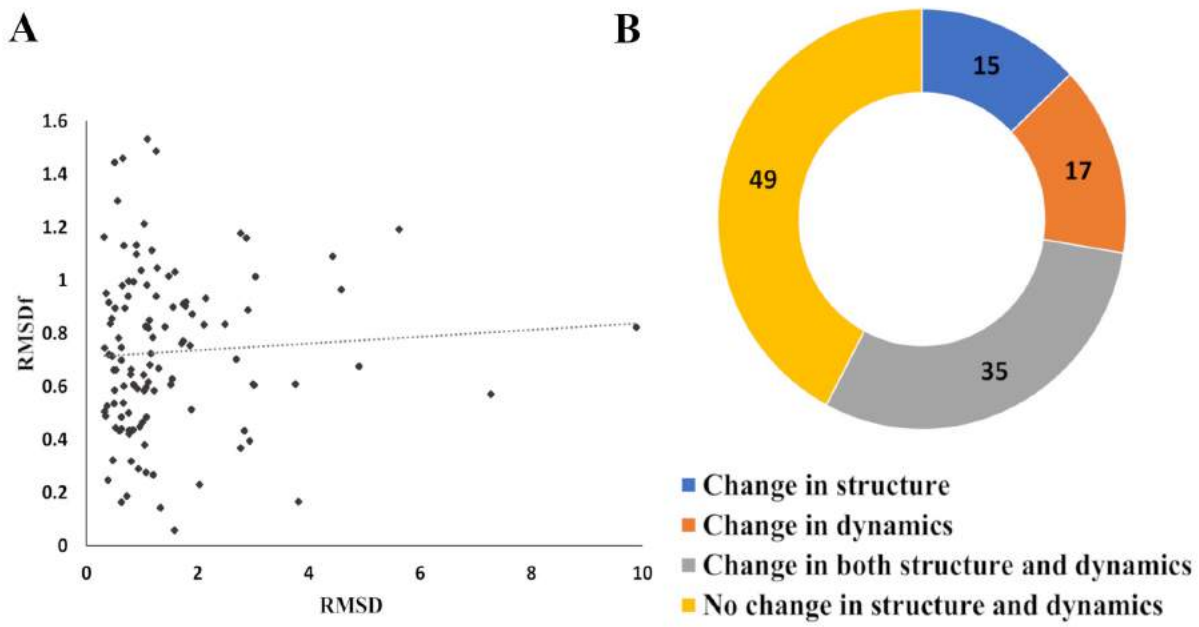


Figure 4| Change in dynamics need not always be accompanied by change in structure. **(A)** RMSD (in Å) of bound and free proteins in the PPC dataset is plotted on X-axis vis-à-vis the root mean square difference of fluctuations ($RMSD^f$) of unbound and free proteins on Y-axis. The dotted line is the 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).

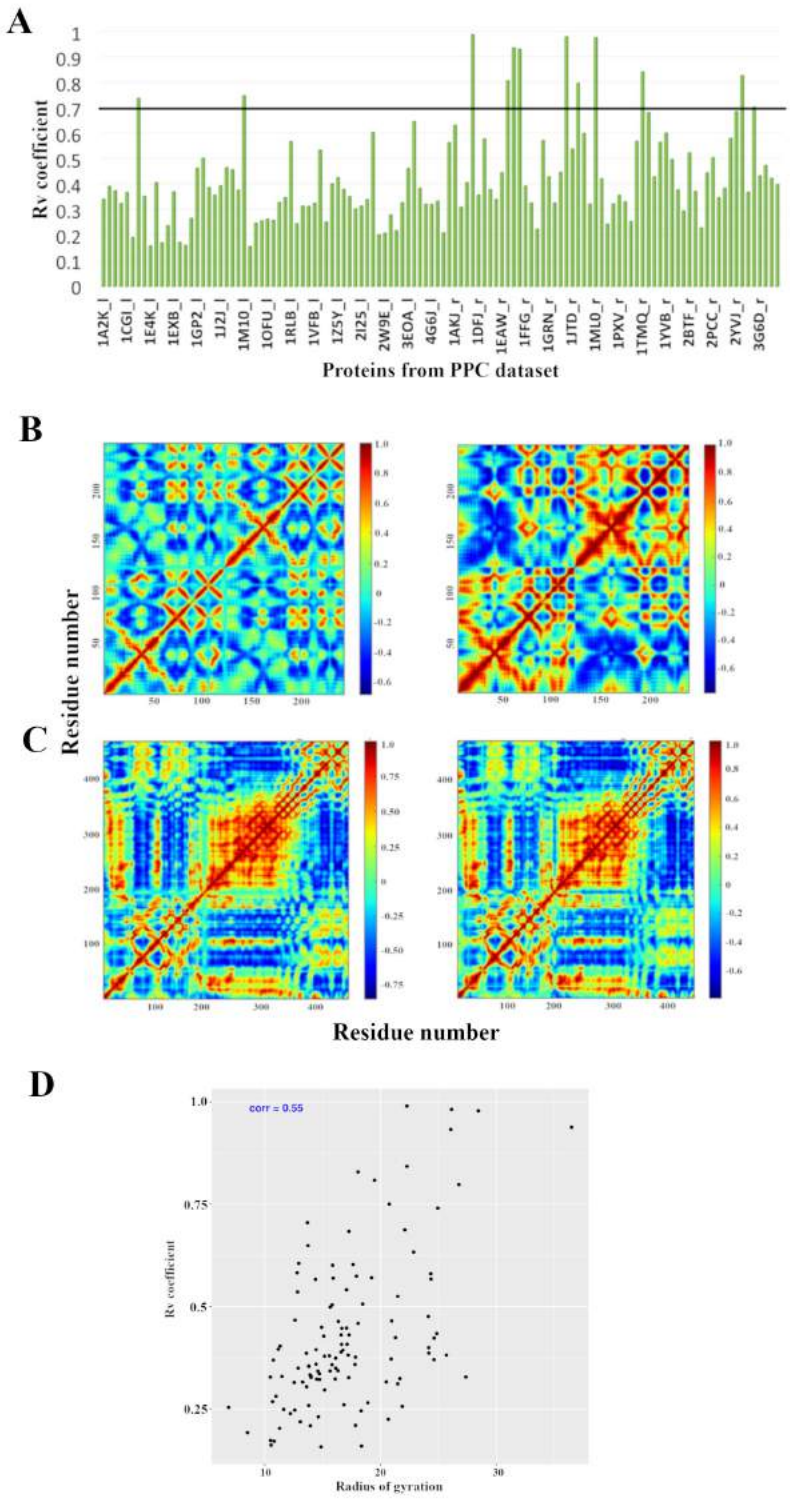


Figure 5| Cross-correlation analysis of proteins in PPC dataset. (A) Similarity between cross-correlation matrices of proteins in the free and bound form is plotted as bar plot. The X-axis shows the proteins in the PPC dataset. Each entry code corresponds to the PDB code of the complex with "_l" and "_r" suffix representing two entries in the PPC dataset. The black horizontal line marks the cut-off for similarity. **(B)** An example case from PPC dataset (PDB code: 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 code: 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 X-axis is plotted vis-à-vis the R_v coefficient. The correlation coefficient of 0.5 suggests not a strong relation between the size of protein and its effect on residue communication.

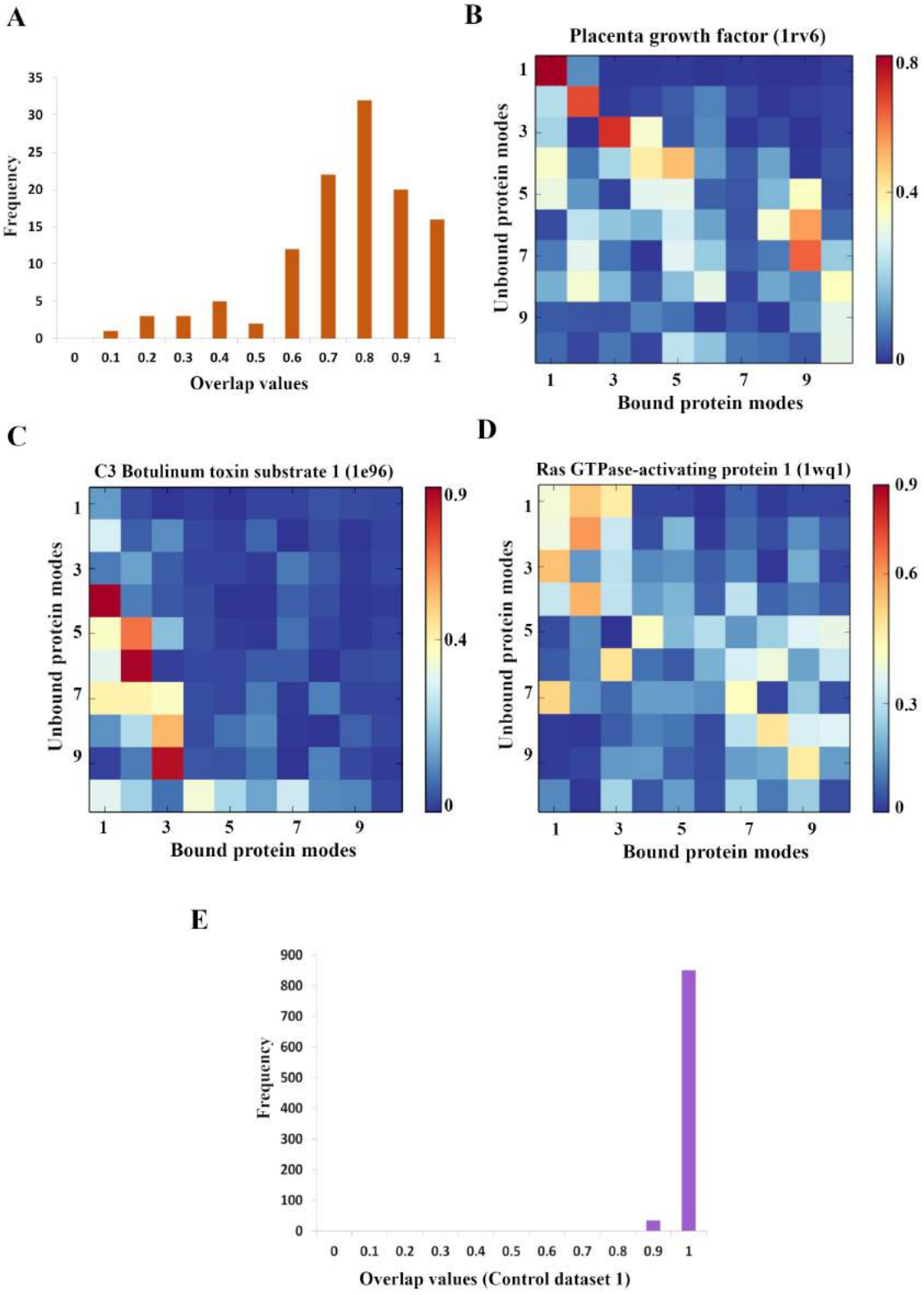


Figure 6| Overlap analysis of proteins in PPC dataset. (A) Absolute overlap score obtained for top 10 low-frequency modes of bound and unbound proteins in the PPC dataset plotted as histogram. **(B)** Example case where the mode order and shape were retained between the bound and unbound form. **(C)** Example case showing change in mode preference. **(D)** Example case showing low overlap values between bound and free form. PDB id for the complex is mentioned in all three examples. **(E)** Absolute overlap score obtained for each pair in control dataset 1 is plotted as histogram. The overlap is high between pair of proteins in the control dataset 1 (See also Figure S3).

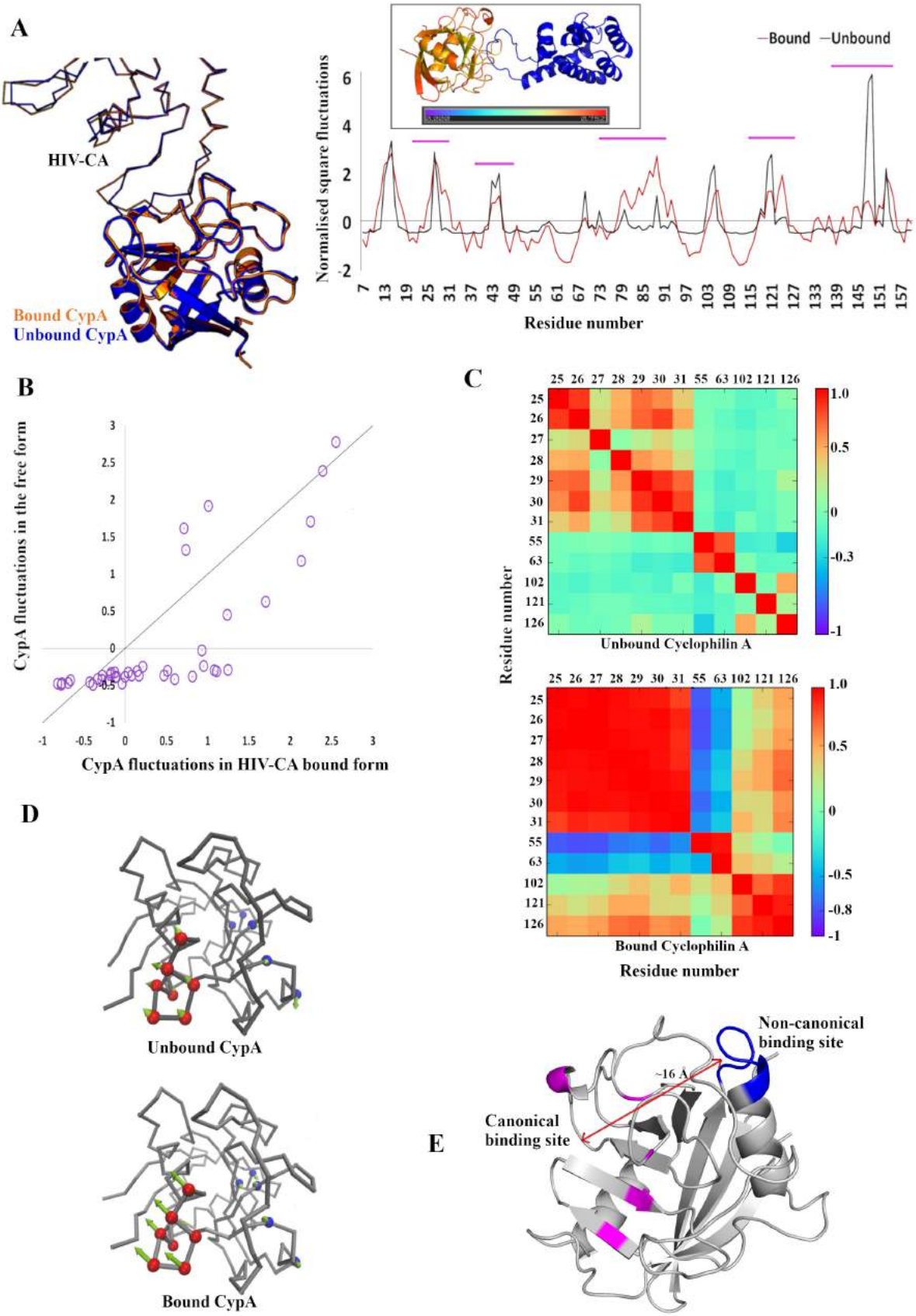


Figure 7| Understanding dynamic allostery in Cyclophilin A. (A) Free and HIV-CA bound CypA is superposed with an RMSD of 0.3Å. CypA in both the forms is rendered as

cartoon and the HIV-CA as ribbon. Bound CypA is shown in orange and unbound in blue colour. Normalised square fluctuations of CypA in CA-bound and unbound form are shown in right panel. X-axis represents residue number and Y-axis represents normalised square fluctuations. Pink horizontal lines signify regions away from interface. Inset shows the absolute difference between square fluctuations mapped ono the CypA structure. **(B)** Scatter plot of normalised 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, Figures S4-S7).

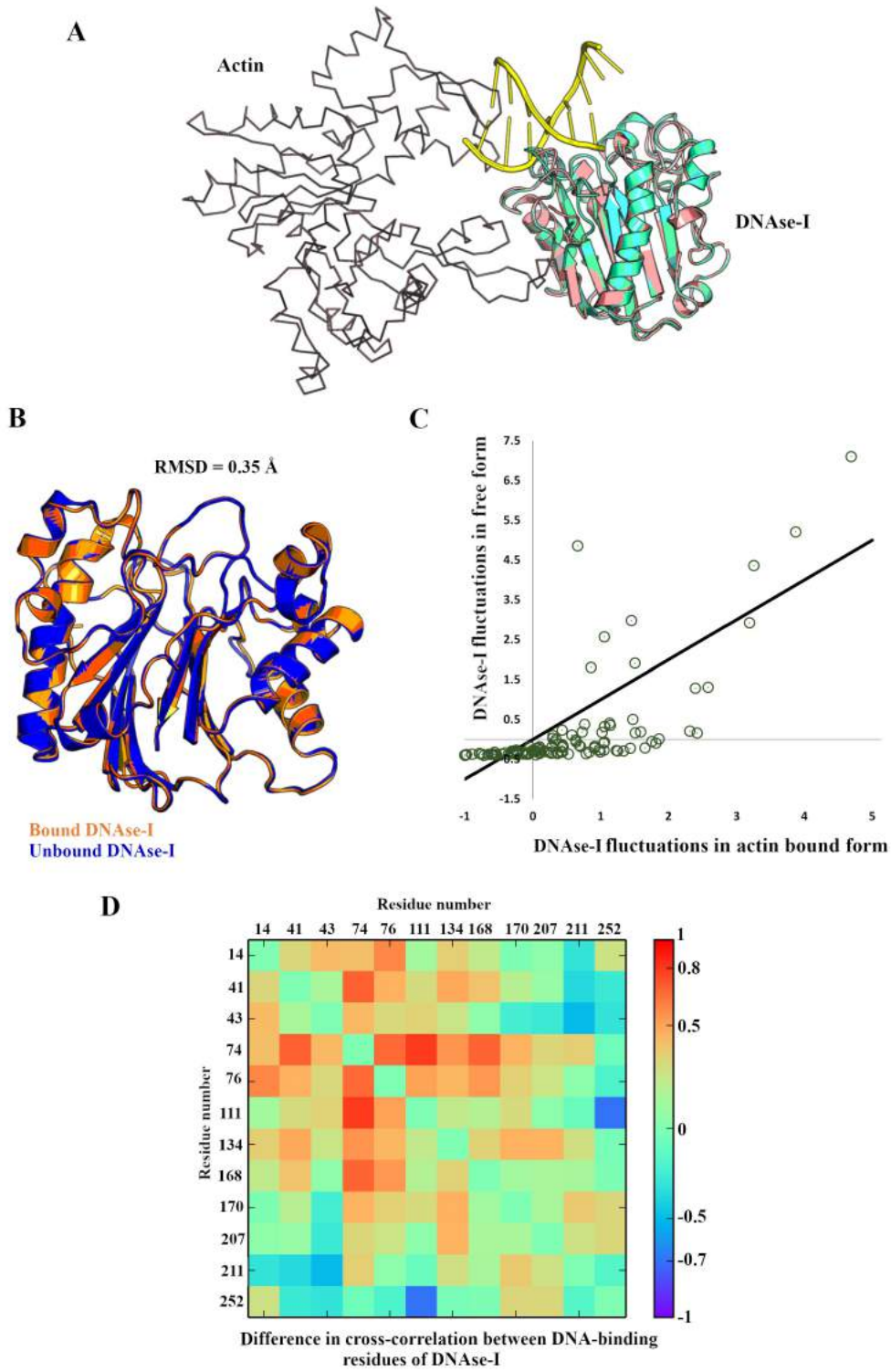


Figure 8| Analysis of DNAse-1 activity upon actin binding. (A) Superposition of structures of DNAse-1 in complex with DNA and in complex with actin shows non-overlapping DNA and actin binding sites. DNAse-1 is represented as cartoon and actin is shown as ribbon. DNA is shown in green colour. **(B)** Superposition of DNAse-1 in the actin bound (orange) and free form (blue) shows an RMSD of 0.35Å. **(C)** Scatter plot showing normalised 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, Figures S4-S7).

STAR METHODS

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 Supplemental information as Tables S1-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 four SCOP 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*., 2005). 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 ≤ 4.5Å were classified as interface residues. Atoms from the two proteins which lie at a distance > 4.5Å but ≤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 ≤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

Supplemental item titles

Table S5| Protein-protein complex dataset used for structural analyses (Related to STAR methods). PDB codes along with chain identifiers for 120 complexes and their interacting partners are provided. Resolution of crystal structure entries is mentioned in Å. Key for type of complex is as follows: EI: enzyme-inhibitor complex, ES: enzyme-substrate complex, ER: enzyme complex with a regulatory chain, OG: others, G-protein containing, OR: others, receptor-containing and OX: others, miscellaneous and A: antigen-antibody complex.

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, 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. *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. *27*, 1575–1577.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2005). The Protein Data Bank. Struct. Bioinforma. 181–198.

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. *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*.

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*.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). BIOINFORMATICS APPLICATIONS NOTE Sequence analysis CD-HIT : accelerated for clustering the next-generation sequencing data. *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 - Gen. Subj. *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. *10*.

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*.

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. *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*.

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 - Mol. Cell Res. *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. In Enzymes, pp. 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*.

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*.

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. *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*.

Marsh, J.A., and Teichmann, S.A. (2014). Protein Flexibility Facilitates Quaternary Structure Assembly and Evolution. PLoS Biol. *12*.

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.

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*.

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*, i151-160.

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 Found. Stat. Comput. Vienna Austria *0*, {ISBN} 3-900051-07-0.

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*.

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*.

Swapna, L.S., Bhaskara, R.M., Sharma, J., and Srinivasan, N. (2012b). Roles of residues in the interface of transient protein-protein complexes before complexation. Sci. Rep. *2*, 334.

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*.

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*.

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*.

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. *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.

Supplementary material (tables, figures and text):

Table S1|List of proteins that show structural changes away from interface, Related to Figure 1. PDB codes for the complexed form is mentioned. Individual PDB codes for the interacting partners can be mapped from **Table S5**. Throughout the study, the PDB codes of the complexed structures have been used for reference and their individual interacting partners have been referred to as either receptor (using suffix, "_r"), for the bigger of the two protein or ligand (using suffix, "_l") for the smaller protein, e.g. if the protein complex 1A2K has been referred, then its interacting partner 1 has been referred as 1A2K_r and the interacting partner 2 as 1A2K_l.

1AK4_l	1GPW_r	1OC0_r	1ZHI_r	2IDO_r	3K75_r
1ATN_r	1HE1_l	1RV6_r	2A5T_l	2O3B_l	3L89_r
1BUH_r	1HE8_r	1SYX_r	2A78_l	2OT3_r	3S9D_r
1DFJ_l	1I4D_r	1TMQ_l	2AYO_r	2UUY_l	3VLB_r
1DQJ_r	1IB1_r	1WEJ_r	2B4J_r	2VDB_r	1FFG_l
1E96_r	1J2J_r	1XD3_r	2BTF_r	2YVJ_r	1JTD_l
1F51_r	1JIW_l	1XUI_r	2GAF_r	3EOA_r	1OPH_r
1FCC_r	1M10_r	1YVB_L	2GTP_r	3FN1_r	3MXW_r
1FQ1_l	1Ml0_r	1Z5Y_r	2HQS_r	3H2V_r	

Table S2| List of complexes where both partners show structural changes away from interface, Related to Figure 1. PDB codes for the complexes is mentioned.

1A2K	1GXD	1XQS	3HI6
1AKJ	1I2M	2FJU	3HMX
1BKD	1JPS	2G77	3L5W
1BVK	1JZD	2H7V	4FZA
1E4K	1KXP	2NZ8	4G6J
1EER	1MLC	2OOR	4G6M
1EWY	1N2C	2W9E	4H03
1EXB	1PVH	2Z0E	4L4W
1GP2	1RLB	3DAW	IJPS
1GRN	1T6B	3G6D	

Table S3| Interface area and protein length for structures used for dynamics analyses, Related to Figure 3. PDB codes for the complex is mentioned. Individual PDB codes for interacting proteins can be mapped from **Table S5**. Contribution to interface area for partners proteins was calculated by finding difference between the accessible surface area of the protein in its unbound and bound form.

Entry name	*Surface area for interface of protein 1 (Å²)*	*Protein length (size of protein)*	*Entry name*	*Surface area for interface of protein 2 (Å²)*	*Protein length (size of protein)*
1A2K_l	1424.1	196	1A2K_r	1334.2	244
1AK4_l	532.1	145	1AK4_r	312.7	164
1AKJ_l	483	228	1AKJ_r	491.2	373
1ATN_l	1281.8	258	1ATN_r	885.6	363
1CGI_l	1099.3	56	1CGI_r	1155.6	245
1CLV_l	1231.2	32	1CLV_r	570.2	470
1DFJ_l	1080.3	456	1DFJ_r	1189.1	124
1DQJ_l	685.5	129	1DQJ_r	1130.2	424
1E4K_l	811.9	172	1E4K_r	279.9	414
1E96_l	1005.7	185	1E96_r	593.8	178
1EAW_l	966.9	56	1EAW_r	930.6	241
1EWY_l	560.8	98	1EWY_r	79.1	295
1EXB_l	2612.9	360	1EXB_r	1234	1300
1FCC_l	698.3	56	1FCC_r	642.8	412
1FFG_l	682.3	68	1FFG_r	550.7	128
1GCQ_l	466.7	66	1GCQ_r	9369.1	57
1GP2_l	1375	388	1GP2_r	2567.9	310
1GPW_l	1093.2	200	1GPW_r	807.6	246
1GRN_l	839.6	185	1GRN_r	1073.3	190
1GXD_l	837.2	181	1GXD_r	242.5	616
1J2J_l	432.8	38	1J2J_r	858.4	165
1JIW_l	1115.8	98	1JIW_r	1152.2	469
1JTD_l	1827.1	262	1JTD_r	723.3	268
1JZD_l	895.7	116	1JZD_r	786.3	428
1M10_l	821	263	1M10_r	1595.9	199
1M27_l	54.3	57	1M27_r	372.1	115
1ML0_l	1825.6	63	1ML0_r	936.6	742
1MLC_l	601	129	1MLC_r	563.6	432
1OFU_l	776.7	306	1OFU_r	968.4	237
1OPH_l	503.8	223	1OPH_r	1685.8	371
1PXV_l	1294	111	1PXV_r	57.4	167
1QA9_l	4582.8	95	1QA9_r	4682.5	102
1RLB_l	937.2	174	1RLB_r	1238.5	453
1RV6_l	509.2	92	1RV6_r	1877.5	189
1TMQ_l	1181.3	115	1TMQ_r	946.6	470
1UDI_l	1261	83	1UDI_r	1296.3	227
1VFB_l	834.1	129	1VFB_r	657.9	223

1WEJ_l	663.5	104	1WEJ_r	6.3	433
1YVB_l	695	105	1YVB_r	235.6	241
1Z0K_l	109.1	46	1Z0K_r	1880.5	169
1Z5Y_l	106.7	117	1Z5Y_r	1035.1	136
2A78_l	896.4	201	2A78_r	543.1	166
2BTF_l	1080.3	139	2BTF_r	1004.2	365
2HRK_l	592.4	103	2HRK_r	1269.4	177
2I25_l	425.1	129	2I25_r	837.9	113
2O3B_l	29.4	122	2O3B_r	291.5	239
2PCC_l	807.1	108	2PCC_r	11.4	293
2UUY_l	975.7	51	2UUY_r	489.3	223
2W9E_l	2154.4	99	2W9E_r	1029.4	424
2X9A_l	969.1	61	2X9A_r	1902.9	90
2YVJ_l	914.5	106	2YVJ_r	652.6	401
3D5S_l	607.1	61	3D5S_r	669.6	294
3EOA_l	271.1	179	3EOA_r	951.8	410
3F1P_l	1183.9	109	3F1P_r	922	111
3G6D_l	1270.6	105	3G6D_r	1606.4	394
3MXW_l	2349.2	150	3MXW_r	867.4	432
4G6J_l	988.5	149	4G6J_r	1511.4	428
4G6M_l	708	149	4G6M_r	741.8	432

Table S4| List of proteins that show changes at non-interface residues, Related to Figure 4. PDB codes for complexed form is mentioned. Four columns are as follows: i) proteins that only show significant structural changes away from the interfaces, ii) proteins that only show significant change in dynamics, away from the interfaces, iii) proteins that show significant changes in structure as well as dynamics, away from the interfaces and iv) proteins that show no significant changes either in structure or dynamics, away from the interfaces.

Only change in structure	*Only change in dynamics*	*Change in both structure and dynamics*	*No change in structure and dynamics*	
1DFJ_l	1AK4_r	1AKJ_l	1CGI_l	2I25_r
1E96_r	1ATN_l	1AKJ_r	1CGI_r	2O3B_r
1EWY_l	1DFJ_r	1A2K_l	1CLV_r	2PCC_l
1EWY_r	1E96_l	1A2K_r	1CLV_l	2PCC_r
1FCC_r	1J2J_r	1AK4_l	1DQJ_l	2UUY_r
1FFG_l	1JTD_r	1ATN_r	1EAW_l	2X9A_l
1GRN_l	1M10_r	1DQJ_r	1EAW_r	2X9A_r
1GXD_l	1OFU_r	1E4K_l	1FCC_l	2YVJ_r
1RLB_r	1OPH_l	1E4K_r	1FFG_r	3D5S_l
1JTD_l	1PXV_l	1EXB_l	1GCQ_r	3D5S_r
1JZD_l	1QA9_l	1EXB_r	1GCQ_l	3F1P_l

1RLB_l	1QA9_r	1GP2_l	1GPW_l	3F1P_r
1Z5Y_r	1WEJ_l	1GP2_r	1JIW_r	4G6J_l
2UUY_l	1Z5Y_l	1GPW_r	1J2J_l	4G6M_l
4G6M_r	2YVJ_l	1GRN_r	1M10_l	
	3EOA_l	1GXD_r	1M27_l	
	3MXW_l	1JIW_l	1M27_r	
		1JZD_r	1ML0_l	
		1ML0_r	1OFU_l	
		1MLC_l	1PXV_r	
		1MLC_r	1RV6_l	
		1OPH_r	1TMQ_r	
		1RV6_r	1UDI_l	
		1TMQ_l	1UDI_r	
		1YVB_l	1VFB_l	
		2A78_l	1VFB_r	
		2BTF_r	1WEJ_r	
		2O3B_l	1YVB_r	
		2W9E_l	1Z0K_l	
		2W9E_r	1Z0K_r	
		3EOA_r	2A78_l	
		3G6D_l	2BTF_l	
		3G6D_r	2HRK_l	
		3MXW_r	2HRK_r	
		4G6J_r	2I25_l	

Table S6| Protein-protein complex dataset used for dynamics analyses, Related to STAR methods. PDB codes along with chain identifiers for 58 complexes and their interacting partners are provided. This dataset is a subset of dataset in Table S5 and has been manually curated to remove entries which have missing residues anywhere in the structures. Resolution of crystal structure entries is provided in Å. Key for type of complex is as follows: EI: enzyme-inhibitor complex, ES: enzyme-substrate complex, ER: enzyme complex with a regulatory chain, OG: others, G-protein containing, OR: others, receptor-containing and OX: others, miscellaneous and A: antigen-antibody complex.

Bound (PDB code)	*Type of complex*	*Resolution (Å)*	*Interacting partner 1 (Receptor) (PDB code)*	*Resolution (Å)*	*Interacting partner 2 (Ligand) (PDB code)*	*Resolution (Å)*
1A2K_C:AB	OG	2.5	1QG4_A	2.5	1OUN_AB	2.3
1AK4_A:D	OX	2.3	2CPL_	1.63	4J93_A/2PXR_A	1.74
1AKJ_AB:DE	OX	2.6	2CLR_DE	2	1CD8_AB	2.6
1ATN_A:D	OX	2.8	1IJJ_B	2.85	3DNI_	2
1CGI_E:I	EI	2.3	2CGA_B	1.8	1HPT_	2.3
1CLV_A:I	EI	2.0	1JAE_A	1.65	1QFD_A	NMR
1DFJ_I:E	EI	2.5	9RSA_B	1.8	2BNH_	2.3
1DQJ_AB:C	A	2.0	1DQQ_CD	1.8	3LZT_	0.93

1E4K_AB:C	OR	3.2	3AVE_AB	2	1FNL_A	1.8
1E96_A:B	OG	2.4	1MH1_	1.38	1HH8_A	1.8
1EAW_A:B	EI	2.9	1EAX_A	1.3	9PTI_	1.22
1EWY_A:C	ES	2.3	1GJR_A	2.1	1CZP_A	1.17
1EXB_ABDC:EGFH	OX	2.1	1QRQ_ABCD	2.8	1QDV_ABCD	1.6
1FCC_AB:C	OX	3.2	1FC1_AB	2.9	2IGG_A	NMR
1FFG_A:B	OX	2.7	3CHY_A	1.66	1FWP_A	NMR
1GCQ_B:C	OX	1.6	1GRI_B	3.1	1GCP_B	2.1
1GP2_A:BG	OG	2.3	1GIA_	2	1TBG_DH	2.1
1GPW_A:B	OX	2.4	1THF_D	1.45	1K9V_F	2.4
1GRN_A:B *	OG	2.1	1A4R_A	2.5	1RGP_	2
1GXD_A:C	EI	3.1	1CK7_A	2.8	1BR9_A	2.1
1J2J_A:B	OG	1.6	1O3Y_A	1.5	1OXZ_A	2.8
1JIW_P:I	EI	1.7	1AKL_A	2.0	2RN4_A	NMR
1JTD_B:A	EI	2.3	3QI0_A	2.8	1BTL_A	1.8
1JZD_AB:C	ER	2.3	1JZO_AB	1.9	1JPE_A	1.9
1M10_A:B	ER	3.1	1AUQ_	2.3	1M0Z_B	1.85
1M27_AB:C	OX	2.5	1D4T_AB	1.1	3UA6_A	1.85
1ML0_AB:D	OR	2.8	1MKF_AB	2.1	1DOL_	2.4
1MLC_AB:E	A	2.5	1MLB_AB	2.1	3LZT_	0.93
1OFU_XY:A	OX	2.1	1OFT_AB	2.9	2VAW_A	2.9
1OPH_A:B	EI	2.3	1QLP_A	2.0	1UTQ_A	1.15
1PXV_A:C	EI	1.8	1X9Y_A	2.5	1NYC_A	1.4
1QA9_A:B	OX	3.2	1HNF_	2.5	1CCZ_A	1.8
1RLB_ABCD:E	OX	3.1	2PAB_ABCD	1.8	1HBP_	1.9
1RV6_VW:X	OR	2.4	1FZV_AB	2.0	1QSZ_A	NMR
1TMQ_A:B	EI	2.5	1JAE_	1.6	1B1U_A	2.2
1UDI_E:I	EI	2.7	1UDH_	1.7	2UGI_B	2.2
1VFB_AB:C	A	1.8	1VFA_AB	1.8	8LYZ_	2.5
1WEJ_HL:F	A	1.8	1QBL_HL	2.2	1HRC_	1.9
1YVB_A:I	EI	2.7	2GHU_A	3.1	1CEW_I	2
1Z0K_A:B	OG	1.9	2BME_A	1.5	1YZM_A	1.5
1Z5Y_D:E	ES	1.9	1L6P	1.6	2B1K_A	1.9
2A78_A:B	ES	1.8	1U90_A	2.0	2C8B_X	1.7
2BTF_A:P	OX	2.5	1IJJ_B	2.8	1PNE_	2
2HRK_A:B	OX	2.0	2HRA_A	1.9	2HQT_A	1.9
2I25_N:L	A	1.8	2I24_N	1.3	3LZT	0.93
2O3B_A:B	EI	2.3	1ZM8_A	1.9	1J57_A	NMR
2PCC_A:B	ES	2.3	1CCP_	2.2	1YCC_	1.23
2UUY_A:B	EI	1.1	1HJ9_A	0.9	2UUX_A	1.4
2W9E_HL:A	A	2.9	2W9D_HL	1.5	1QM1_A	NMR
2X9A_D:C	OR	2.4	1S62_A	NMR	2X9B_A	2.92
2YVJ_A:B	ER	1.9	2YVF_A	1.6	2E4P_A	2
3D5S_A:C	OX	2.3	1C3D_A	1.8	2GOM_A	1.25
3EOA_LH:I	A	2.8	3EO9_LH	1.8	3F74_A	1.7

3F1P_A:B	OX	1.1	1P97_A	NMR	1X0O_A	NMR
3G6D_LH:A	A	3.2	3G6A_LH	2.1	1IK0_A	NMR
3MXW_LH:A	A	1.8	3MXV_LH	1.9	3M1N_A	1.85
4G6J_HL:A	A	2.0	4G5Z_HL	1.8	4I1B_A	2
4G6M_HL:A	A	1.8	4G6K_HL	1.9	4I1B_A	2

Table S7| Experimental ΔG, predicted ΔG from PRODIGY and the new ΔG after adding the ΔG_{vib} term to it, Related to Figures 2 and S8.

PDB-dataset	*Experimental ΔG (Kcal/mol)*	*Prodigy ΔG (Kcal/mol)*	*Prodigy ΔG_new (Kcal/mol)*	*Experimental Kd (M)*	*Kd-PRODIGY (M)*	*Kd-PRODIGY-new (M)*
1DQJ	-11.7	-12.7	-12.07	2.80E-09	4.54E-10	1.32E-09
1PXV	-12.97	-14	-13.36	3.10E-10	5.02E-11	1.47E-10
1DFJ	-18.05	-13.6	-13.36	5.90E-14	9.88E-11	1.48E-10
1GXD	-11.7	-12.6	-12.05	5.20E-09	5.38E-10	1.37E-09
1AKJ	-5.3	-7.3	-6.76	0.000126	4.26E-06	1.07E-05
1ATN	-12.07	-10.9	-10.05	2E-09	9.58E-09	4.06E-08
1Z0K	-7	-11.7	-11.14	0.0000077	2.47E-09	6.41E-09
1MLC	-9.6	-10.3	-9.58	9.1E-08	2.65E-08	8.99E-08
1A2K	-9.3	-9	-8.22	0.00000015	2.39E-07	8.92E-07
2I25	-12.3	-12	-11.23	1E-09	1.49E-09	5.51E-09
2HRK	-11	-8.7	-8.11	9E-09	3.98E-07	1.09E-06
1WEJ	-12.48	-9.4	-8.63	7.14E-10	1.22E-07	4.49E-07
1E4K	-7.9	-9.4	-9.03	0.0000017	1.22E-07	2.26E-07
1GCQ	-6.5	-9	-8.44	0.000017	2.39E-07	6.15E-07
1OPH	-11.32	-11.6	-11.79	5E-09	2.93E-09	2.13E-09
1J2J	-8.13	-7	-6.44	0.0000011	7.08E-06	1.84E-05
1VFB	-11.5	-11.4	-10.68	3.70E-09	4.11E-09	1.39E-08
1E96	-7.42	-8	-7.32	0.0000027	1.30E-06	4.15E-06
1EWY	-7.4	-8.3	-7.78	0.00000357	7.83E-07	1.89E-06
2PCC	-7.9	-8.1	-7.38	0.0000016	1.10E-06	3.70E-06
1AK4	-6.43	-7.2	-6.32	0.000016	5.05E-06	2.26E-05
1QA9	-7.16	-9.6	-8.92	0.000009	8.66E-08	2.75E-07
1JTD	-14.41	-10.2	-9.5	2.72E-11	3.13E-08	1.03E-07
3EOA	-11.81	-9.2	-8.46	2.20E-09	1.71E-07	5.95E-07
3MXW	-11.31	-12.3	-11.54	7E-09	8.94E-10	3.22E-09
1M10	-11.24	-10.5	-9.98	5.80E-09	1.89E-08	4.59E-08

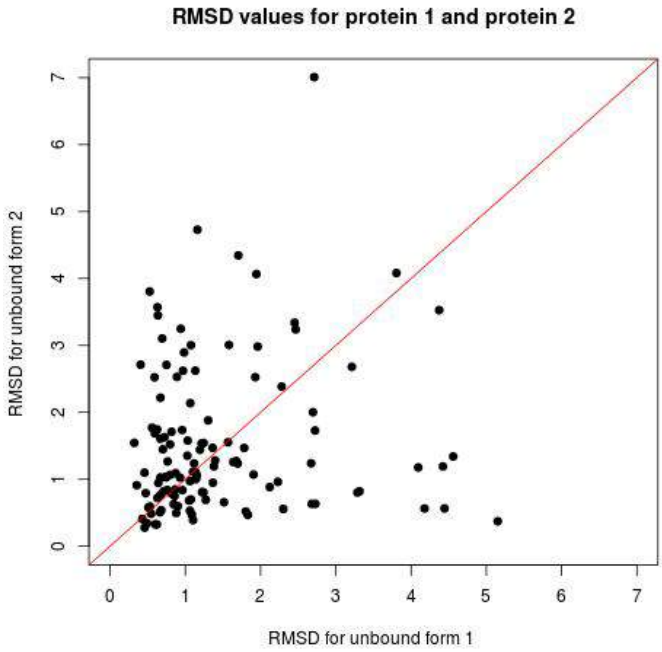


Figure S1|RMSD between the interacting partners, **Related to Figure 1.** The RMSD between bound and unbound form of interacting partner 1 (X-axis) is compared to the RMSD between bound and free form of interacting partner 2 (Y-axis). The red solid line is the unity line. For 53 complexes, significant change in structures is observed only for one partner.

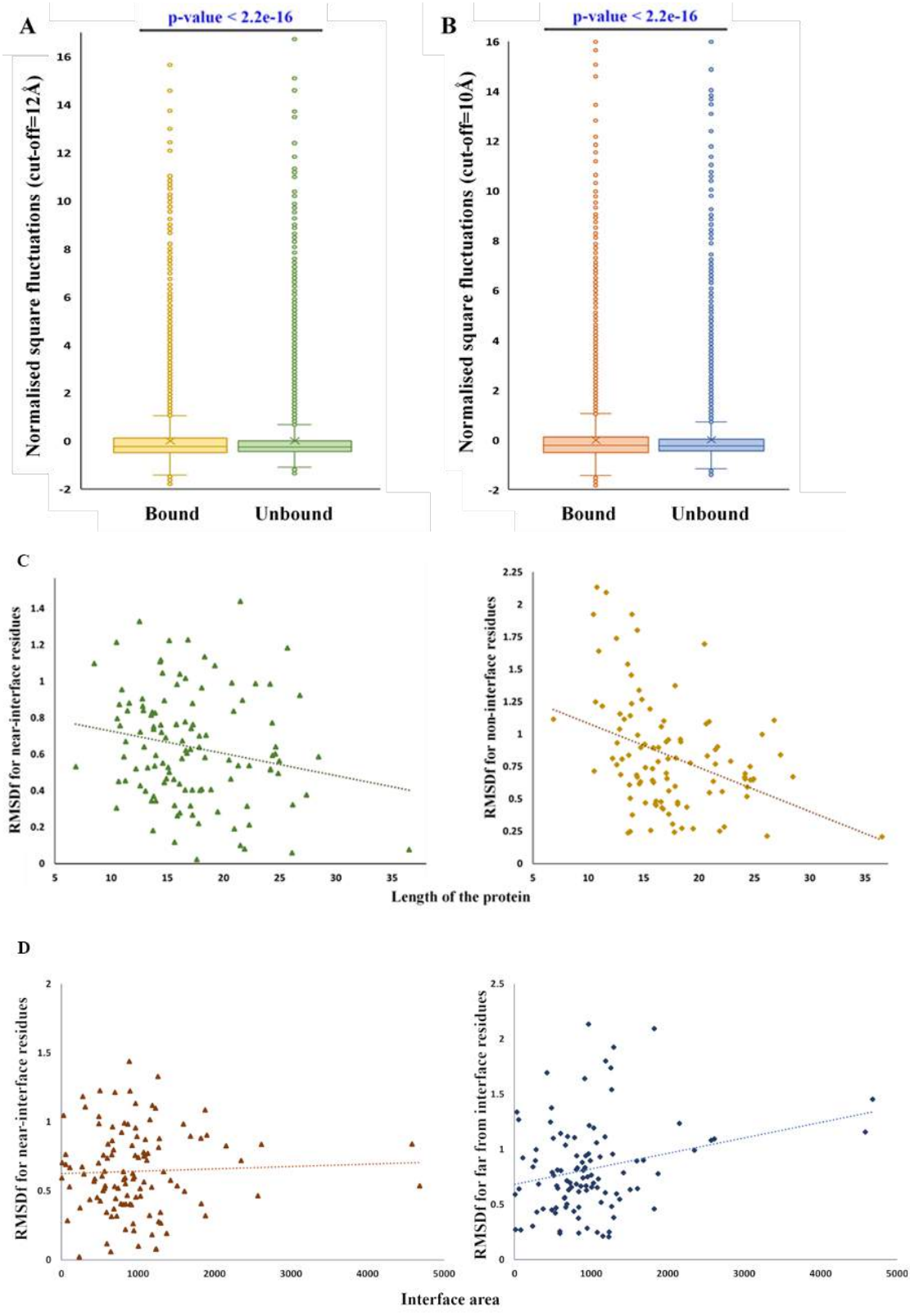


Figure S2|Analyses of dynamics of PPCs and Influence of size of interacting proteins and their interface area on observed fluctuation differences, Related to Figures 2 and 3. Box

plots showing distribution of Cα fluctuations for all residues in bound and unbound forms at a distance cut-off of **(A)** 12Å and **(B)** 10Å. The two distributions are significantly different from each other (two-sample KS-test, p-value < $2.2x10^{-16}$), showing variation in flexibility of bound and free proteins. **(C)**Length of protein (proxy for protein size) on X-axis is plotted vis-à-vis the root mean square fluctuations ($RMSD^f$) on Y-axis for near interface residues (left panel) and residues far from interface (right panel). The dotted line shows the best line of fit. **(D)** Interface area on X-axis is plotted vis-à-vis the root mean square difference of fluctuations ($RMSD^f$) on Y-axis for near interface residues (left panel) and residues far from interface (right panel). The dotted line shows the line of best fit. The results suggest no significant effect of size or interface area of the protein on observed fluctuation differences as the Pearson's correlation coefficient of -0.16 and -0.39 was obtained.

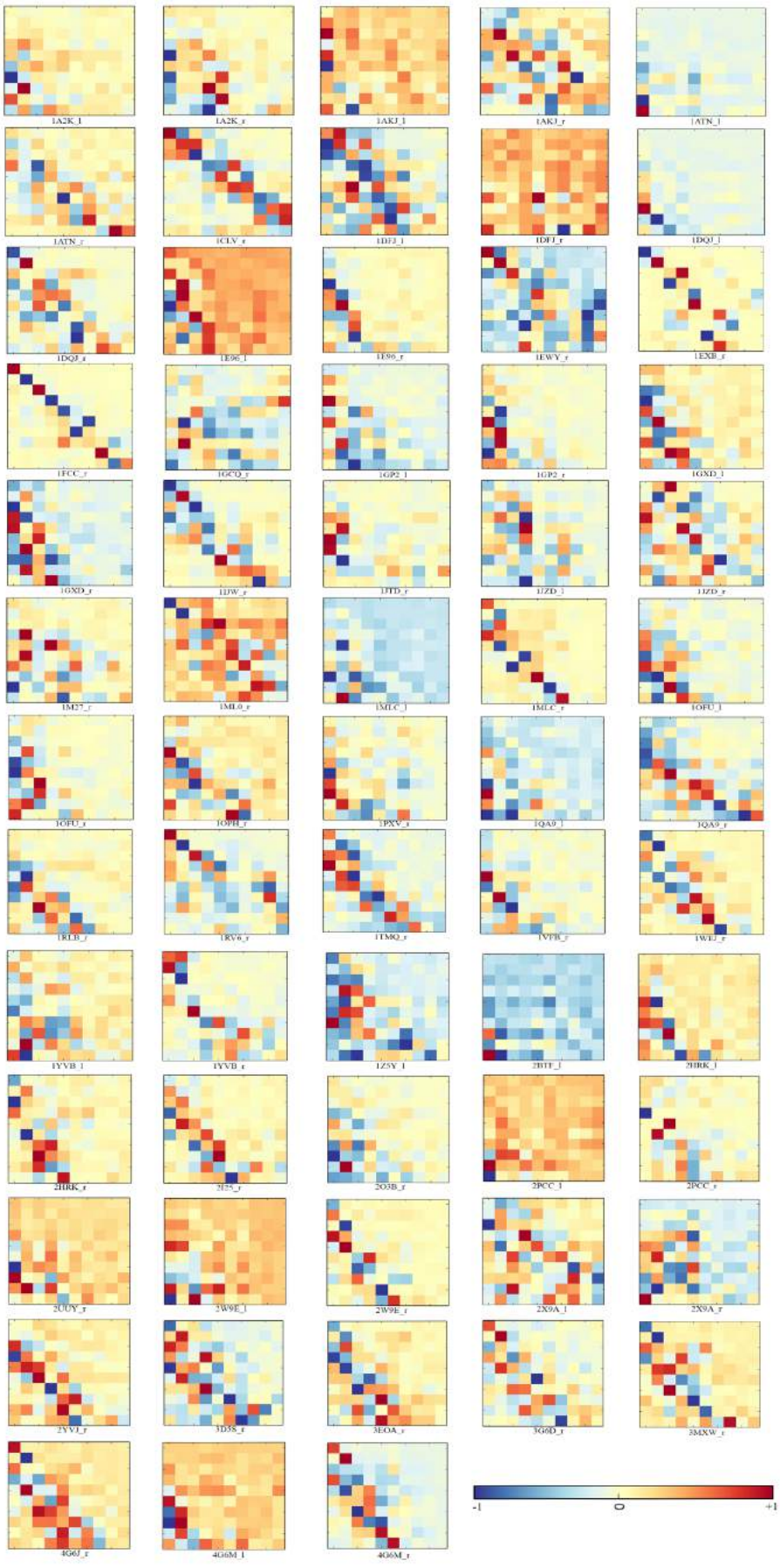

1A2K_l
1A2K_r
1AKJ_l
1AKJ_r
1ATN_l
1ATN_r
1CLV_r
1DFJ_l
1DFJ_r
1DQJ_l
1DQJ_r
1E96_l
1E96_r
1EWY_r
1EXB_r
1FCC_r
1GCQ_r
1GP2_l
1GP2_r
1GXD_l
1GXD_r
1IJW_r
1ITD_r
1IZD_l
1IZD_r
1M27_r
1ML0_r
1MLC_l
1MLC_r
1OFU_l
1OFU_r
1OPH_r
1PXV_r
1QA9_l
1QA9_r
1RLB_r
1RV6_r
1TM9_r
1VPB_r
1WEJ_r
1YVB_l
1YVB_r
1Z5Y_l
2BTF_l
2HRK_l
2HRK_r
2I25_r
2O3B_r
2PCC_l
2PCC_r
2UUY_r
2W9E_l
2W9E_r
2X9A_l
2X9A_r
2YVI_r
3D5S_r
3EOA_r
3G6D_r
3MXW_r
4G6J_r
4G6M_l
4G6M_r
-1
0
+1

Figure S3| Overlap heatmaps for the proteins that showed at least one mode with high score (>|0.7|), Related to Figure 6. The X-axis in all the maps show the 10 lowest frequency modes and Y-axis shows the overlap score. The color scale for these maps is shown at the bottom and is scaled between -1 and +1. In all the cases – and + signs do not have a physical meaning and are considered equivalent. These maps were analysed to understand the conservation of modes of motion for high overlap cases. It is evident from these maps that despite high correlation, many proteins show reordering of modes in the bound form when compared to the unbound form.

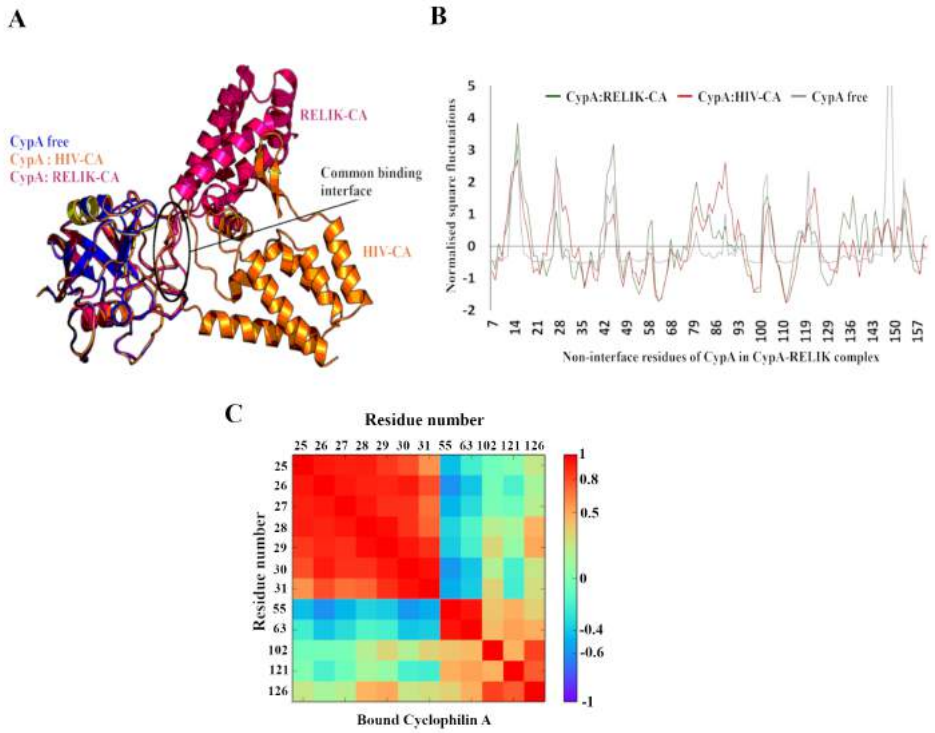


Figure S4| Role of dynamics in CypA and RELIK binding, Related to Figures 7. (A) Superposition of CypA structure in free form (blue), bound to HIV-CA (CA) (orange) and RELIK-CA (rCA) (pink) shows no significant difference in structure. The common binding site of CA and rCA on CypA is encircled in black. **(B)** Normalised square fluctuations for non-interface residues of CypA in free, CA-bound and rCA-bound form. **(C)** Cross-correlation between canonical and non-canonical binding sites on CypA in RELIK-bound form.

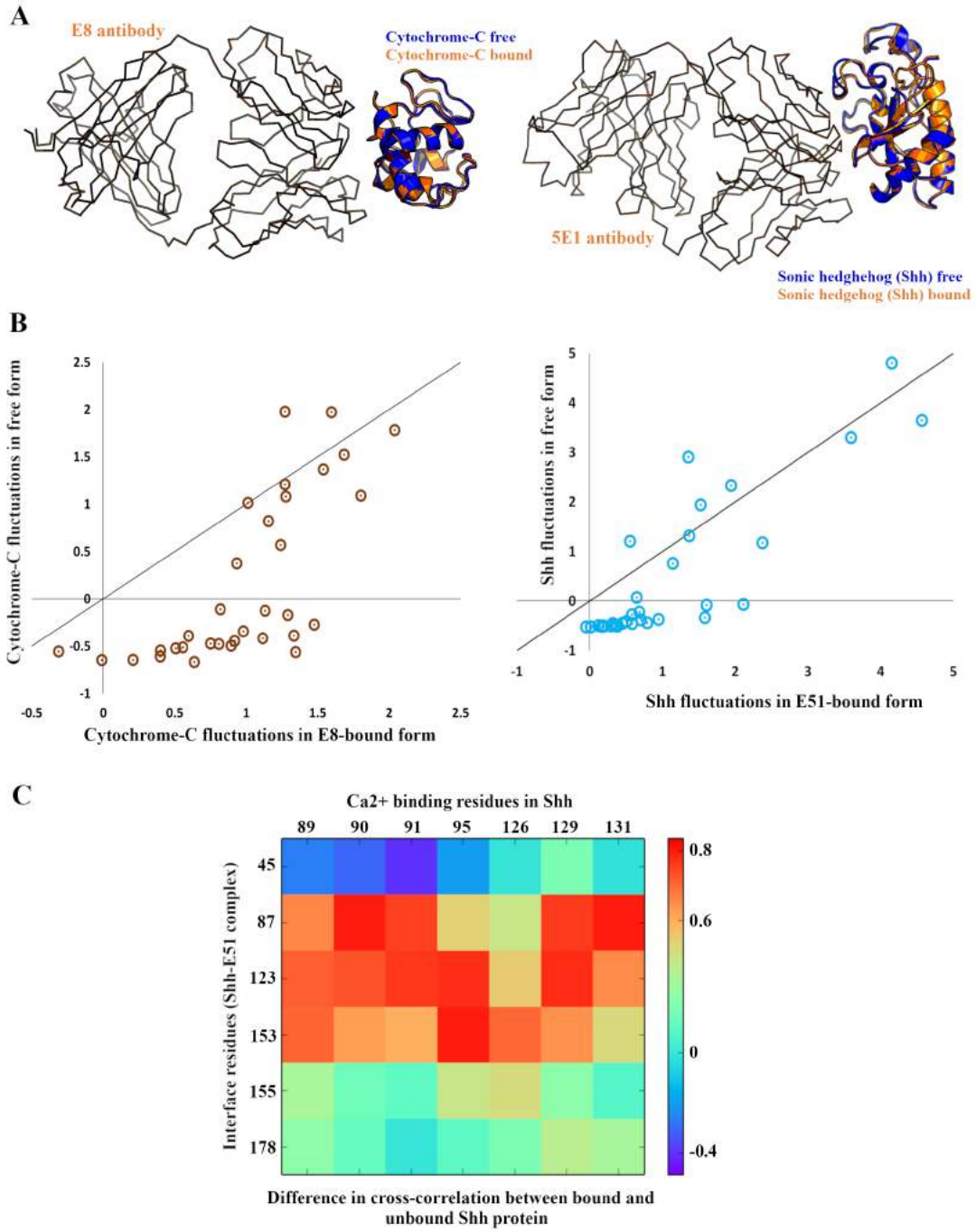


Figure S5| Role of dynamic allostery in stability of antigen-antibody complexes, Related to Figure 7 and 8. (A) Superposed structures of E8-antibody (ribbon representation) and Cytochrome-C (Cyt-C) protein (cartoon representation) in left panel and 5E1 antibody (ribbon representation) and Sonic hedgehog (Shh) protein (cartoon representation) in right panel. Bound protein is shown in orange and unbound protein in blue colour. **(B)** Scatter plots of non-interface residues in bound and unbound forms for Cyt-C (left) and Shh (right)

show higher fluctuations in bound form. X-axis shows normalised square fluctuations in bound forms and Y-axis shows normalised square fluctuations in unbound forms. **(C)** Difference in cross-correlation of Ca^{2+} binding residues between bound and unbound Shh protein suggests an allosteric communication between interface and Ca^{2+} binding residues.

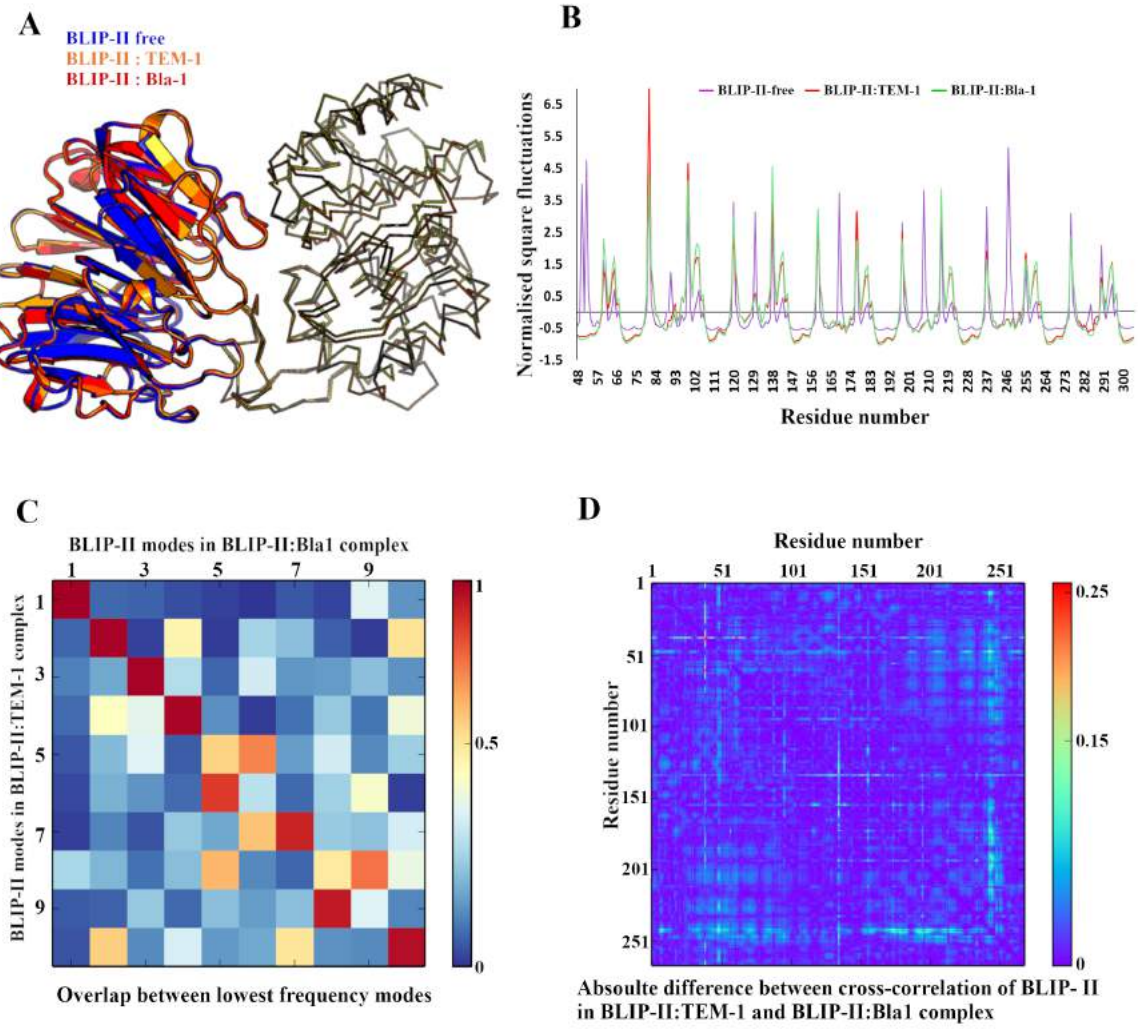


Figure S6| Bla-1 binding to BLIP-II induces subtle differences in residue-communication, Related to Figure 7 and 8. (A) Superposed BLIP-II structures in free (blue), TEM-1 bound (orange) and Bla-1 bound (red) show no significant structural changes. BLIP-II protein is shown as cartoon and TEM-1 and Bla-1 proteins are shown as ribbons. **(B)** Normalised square fluctuations for residues in BLIP-II in free (pink), TEM-1 bound (red) and Bla-1 bound (green) form. Profiles show significant differences in fluctuations between free and bound forms. **(C)** Overlap between 10 lowest frequency modes of TEM-1 and Bla-1 bound BLIP-II. **(D)** Absolute differences between cross-correlation in BLIP-II residues in TEM-1 and Bla-1 bound forms show alteration in residue coupling in Bla-1 bound form.

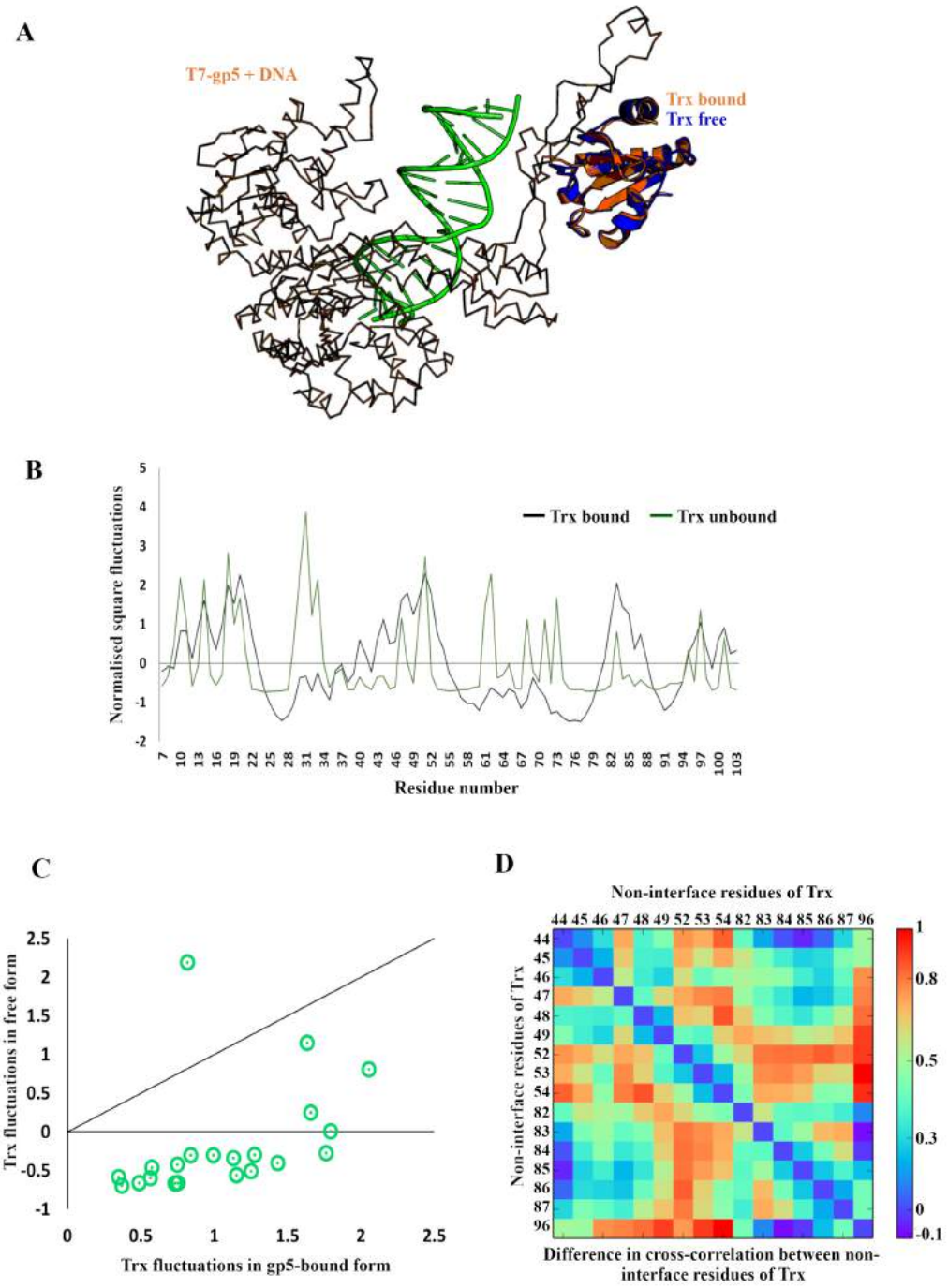


Figure S7| Putative role of dynamics in packing of replisome complex, Related to Figure 7 and 8. (A) Superposed structures of thioredoxin (trx) protein in T7 gp5+DNA bound (orange) and unbound form (blue). Trx is represented as cartoon and gp5 as ribbon. DNA is shown in green. **(B)** Normalised square fluctuations of trx residues in gp5+DNA bound form (black) and free form (green). **(C)** Scatter plot showing normalised square fluctuations for non-interface residues of trx in gp5+DNA bound form on X-axis and free form on Y-axis. **(D)**

Cross-correlation differences between non-interface residues of trx in gp5+DNA bound form and unbound form.

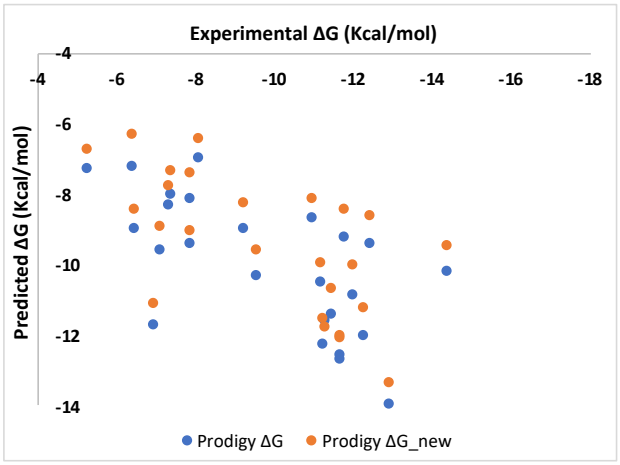


Figure S8| Graphical representation of predicted and experimental ΔG values, Related to Figure 2 and Table S7.

Data S1| Includes the analysis of three more cases with change in dynamics but no observed structural changes and the pilot study to improve binding affinity prediction, Related to Figures 2, 7, 8 and Discussion.

Differential dynamics of Cyclophilin A (CypA) may promote RELIK capsid stabilisation

Interestingly, CypA has also been shown to bind other prehistoric endogenous lentiviruses, e.g., from rabbits (RELIK) and lemurs (PSIV), etc. The crystal structure of CypA with RELIK-capsid (rCA), shows that the active site of CypA binds rCA in a manner like 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 the lentiviral capsids, its interaction with CA and rCA was compared to find the similarity or differences between the dynamics of CypA bound to two evolutionarily conserved partners. Here again, CypA does not show differences in the structure between rCA bound and free form **(Figure S4A)**, but the two CA molecules are positioned differently in crystal structures. These differences affect the dynamics in a slightly different way. Comparison of normalised fluctuations for CypA-rCA and CypA-CA revealed regions of similar and differential dynamics between the two. Interestingly, the regions away from the interface show change in dynamics in CypA-rCA structure also **(Figure S4B)**. Despite only subtle differences between the normalised square fluctuations of non-canonical binding site in the free and rCA bound structures, residue communication within the site gets altered in CypA-rCA structure too **(Figure S4C)**. Though it is unclear if CypA binds two rCA molecules, the current results propose a possibility of similar binding mode like HIV CA.

Redistribution of dynamics upon binding of partner protein contributes to the stability of the complex

Apart from participating in a downstream signalling process, recruiting a partner protein or modulating the activity of the enzymes, binding of two or more proteins can also lead to redistribution of dynamics (Grünberg *et al*., 2006). Reorganisation of motions within a protein contributes favourably to the entropy and hence stabilises the complex. For many cases in **Table S4**, increase in fluctuations was observed in the bound form. When the relevant literature was reviewed, it was found that many a times, an increase in flexibility in regions other than interface has been reported to positively contribute towards the stability of the complex. For example, the interactions between cytochrome-C (Cyt-C) and E8 antibody

(PDB code: 1wej) (Mylvaganam *et al*., 1998) as well as sonic-hedgehog (Shh) protein and 5E1 antibody fragment complex (PDB code: 3mxw) (Maun *et al*., 2010), does not induce any structural changes in Cyt-C and Shh protein respectively **(Figure S5A)**. Antibodies generally induce long-range conformational changes in the antigen. These cases raise special interest since a change in dynamics was observed in the antigens, despite no change in structure. Many residues, away from the interfaces, showed higher fluctuations in the bound form of antigens than the unbound forms **(Figure S5B)**.

The Cyt-C : E8 interaction has been shown to be both enthalpically and entropically favoured (Mylvaganam *et al*., 1998). Hence, the increase in fluctuation of non-interface residues in both the antigens could imply a positive gain in entropy upon complex formation, thereby compensating for the loss of conformational freedom at the interface. Additionally, Shh-E51 complex is shown to have better affinity in the presence of Ca^{2+}. Cross-correlation analysis performed in this study suggested a better correlation between interface (Lys45, Lys87, Arg123, Arg153, Arg155, and Lys178) and Ca^{2+}-binding residues (Glu89, Glu90, Asp95, Asp126, Asp129, and Asp131) of Shh in the E51 bound form than the unbound form **(Figure S5C)**. This suggests the existence of communication between the two regions, which likely leads to a positive effect of Ca^{2+} binding.

Beta-lactamase inhibitor protein-II shows subtle differences in residue-residue communication when bound to two homologous partners

Beta-lactamase inhibitor protein-II (BLIP-II) binds beta-lactamases and inhibit their activities. It is a highly potent inhibitor and binds to the active site of beta-lactamases with femtomolar to picomolar affinity (Brown *et al*., 2013). Two such beta-lactamases that BLIP-II binds to are TEM-1 and Bla-1. Previous studies have shown that though BLIP-II binds to beta-lactamases with varying strengths, its affinity for TEM-1 and Bla-1 is comparable (K_d of 0.79 pM and 1.1 pM respectively) (Brown *et al*., 2011). But there exists a difference between the pre-steady state kinetics, where association rate for BLIP-II and TEM-1 is faster as compared to BLIP-II and Bla-1, but once bound, BLIP-II dissociates from Bla-1 with a slow dissociation rate (Brown *et al*., 2011). TEM-1 and Bla-1 are homologous proteins with 38% sequence identity. BLIP-II binds to both with similar interface with no backbone or side-chain movement observed in free and the two bound forms of BLIP-II **(Figure S6A)**. This raises a possibility that dynamics may play a role in regulating the pre-steady state kinetics.

No differences between normalised square fluctuations of BLIP-II were observed **(Figure S6B)**. Overlap score was also observed to be high between the low frequency normal modes of BLIP-II bound to TEM-1 and Bla-1 **(Figure S6C)**. However, the cross-correlation matrices, interestingly, showed subtle differences between the two forms **(Figure S6D)**.Inter-residue communication gets marginally altered within BLIP-II bound to Bla-1 as compared to TEM-1 which can contribute to slow dissociation rates of BLIP-II from Bla-1 complex. Further, in previous studies, it was observed that BLIP-II : Bla-1 complex shows a more positive gain in entropy when compared to BLIP-II : TEM-1 complex (Brown et al., 2011). Results here are in concordance with gain in entropy, as the fluctuation profiles indicate the residues which have higher fluctuations in two bound forms of BLIP-II vis-à-vis the unbound form. This, in turn, seems to affect the synchronised motions within BLIP-II. It is to be noted that regions in BLIP-II, which show differences in cross-correlation are situated far away from the interface suggesting the possibility of a communication pathway between them. These residues can be further mutated to confirm their role in allosteric communication.

Change in Thioredoxin (trx) flexibility likely helps in better packing of T7 bacteriophage replisome

Some proteins are known to perform multiple functions depending on their cellular location, cell-signal or binding partners and are known as moonlighting proteins (Jeffery, 2009). A large number of reviews are available on such proteins, listing their form and functions (Gancedo *et al*., 2016; Huberts and van der Klei, 2010; Jeffery, 2009). Since these proteins can bind diverse partners under different conditions, it is interesting to analyze if the dynamics of any moonlighting protein gets altered upon binding of another protein without a significant change in structure. Proteins from T7-bacteriophage replisome complex were selected as a case study. T7 replisome complex consists of four proteins *viz*. T7 DNA polymerase (gp5), host bacterial thioredoxin (trx), ssDNA-binding protein (gp2.5) and DNA primase-helicase (gp4) (Kulczyk and Richardson, 2016). T7 gp5 is a non-processive polymerase in the absence of bacterial trx (Akabayov *et al*., 2010). Trx, in its functional form in host bacteria, maintains the redox environment in cell whenever required (Zeller and Klug, 2006). Proteins, gp2.5 and gp4, bind more strongly to gp5/trx (k_d = 130 nM and 90 nM respectively) as compared to either gp5 (K_d = 1600nM and 370nM respectively) or trx alone (K_d = 500 μM and 130 μM respectively) (Ghosh *et al*., 2008). It was believed for a long time that trx itself doesn't have any role in recruiting gp2.5 and gp4 and it only stabilizes the gp5 loops for gp2.5 and gp4 binding. However, it was later shown that though in the absence of

DNA, gp5/trx complex recruits gp2.5 and gp2.4 via two loops in gp5 but in the presence of DNA, gp4 binds to gp5 via trx residue Arg36 (Ghosh *et al*., 2008).

Thioredoxin in its gp5-bound and free form does not show significant change in structure **(Figure S7A)**. To understand the role of trx dynamics in recruiting gp4, normal modes were calculated for trx in free and gp5-DNA bound form. A general increase in flexibility was observed for trx in its bound form except for the residues at the interface **(Figure S7B)**. Arg36 showed comparable flexibility in the bound and unbound form suggesting its role in gp4 binding. Further, it has been experimentally shown that residues 60-91 of *E. coli* trx are critical because replacement of this region with human trx leads to complete inactivity of T7-polymerase (Lee *et al*., 2018). Many residues in region 60-91 are away from gp5 interface and were observed to show higher fluctuations **(Figure S7C)** and showed better synchronisation of motions **(Figure S7D)**. Though there are no other trx residues currently known to bind directly to gp4, these results warrant for the role of increased flexibility in recruitment of gp4. Moreover, it has been earlier suggested that for many quaternary complexes, flexibility helps in better packing of distinct heteromeric subunits (Marsh and Teichmann, 2014).

Pilot study to understand the vibrational entropy contributions towards binding affinities

The current methods to predict binding affinity or the binding free energy fall into two categories: simulation-based numerical methods and statistics-based theoretical methods. Simulation-based methods like MM/GBSA or MM/PBSA are time-consuming and computationally expensive procedures (Gilson and Zhou, 2007; Kollman, 1993). The statistics-based methods employ heuristics and thus are fast but always have a scope of improvement. We started with an aim to include the contribution of vibrational entropy to the calculation of binding affinity. The frequency values to calculate vibrational entropy was obtained from the standard Cα-based normal mode analysis performed in this study and its contribution to free energy was calculated using the formula:

$$G_{vib}^{Str} = -RT \sum_{k=1}^{N} ln(\frac{K_b T}{h\square^k})$$

Where, *R* is the universal gas constant, *T* is the temperature (here taken as 298.15K), *N* is the total number of modes corresponding to 80% variance in motion, *h* is the Planck's constant, λ

is the natural frequency of the k^{th} mode, and K_b is the Boltzmann constant. All units converted from SI to Kcal/mol.

Out of the 58 complexes in our dataset, binding affinity values have been experimentally reported for 27 complexes (Vreven et al., 2015). So, we calculated the G_{vib}^{Str} for the 27 complexes in their bound and respective unbound forms. ΔG_{vib} was calculated as:

$$\Delta G_{vib} = G_{vib}^{comp} - G_{vib}^{ub1} - G_{vib}^{ub2}$$

Where, *comp* is the bound complex, ub^1 is the unbound protein 1 and ub^2 is the unbound protein 2. We added the calculated ΔG_{vib} term to the predicted ΔG from an existing software, PRODIGY (Xue et al., 2016), to predict binding free energy/binding affinity. This algorithm is based on a simple linear regression of interfacial contacts (ICs) and some of the properties of the non-interacting surfaces (NIS), which have been shown to influence the binding affinity. It is to be noted that this method is currently one of the best performing statistical methods, but does not take into account any entropic term. So, it seems worthwhile adding the calculated ΔG_{vib} to the predicted ΔG and analyse for any improvements.

Table S7 consists of the experimental ΔG, predicted ΔG from PRODIGY and the new ΔG after adding the ΔG_{vib} term to it. The corresponding K_d values were calculated using the formula:

$$\Delta G = RTlnK_d$$

Figure S8 shows the graphical representation of predicted and experimental ΔG values. The Pearson's correlation coefficient between the experimental ΔG and the new PRODIGY ΔG was found to be ~0.7 and the RMSE was 2.15 Kcal/mol. The Pearson's correlation coefficient between the experimental K_d and K_d Prodigy-new was found to be ~0.5 and RMSE was 2.3E-05 M.

We reckon that at this stage of the analysis, it is not a drastic improvement (as compared to PRODIGY) in prediction efficacy but it is evident from this pilot study that the PRODIGY-new values show marginal improvement and moved closer to the experimental values, especially in cases where PRODIGY over predicted the free energy of binding (entries highlighted in Bold). This, we believe, is moving one step closer to improving the accuracy of binding affinity predictions. Moreover, it is to be noted that ΔG_{vib} values at this stage are "raw", "uncalibrated" values added to the ΔG from PRODIGY. The vibrational entropy

obtained from NMA may be an under-estimate and hence require calibration using weights from MD simulations or experiments available in literature. We strongly believe that calibrating these values by assigning weights will have a positive impact and is expected to improve the accuracy. In a previous study from 2018 by Skrbic *et al*. (Skrbic et al., 2018) authors have incorporated vibrational entropy term in a similar fashion while calculating binding energies/affinities and they report that it slightly improves the prediction. We believe that there are many instances within the current methodology to improve the efficiency. But it demands a full-fledged, dedicated attention and analysis of its own which will include understanding of different methods and if required, application of machine learning methods to train the model with high efficiency.