frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 25 August 2020
doi: 10.3389/fmolb.2020.00210

A Coarse-Grained Methodology Identifies Intrinsic Mechanisms That Dissociate Interacting Protein Pairs

*Haleh Abdizadeh[1], Farzaneh Jalalypour[2], Ali Rana Atilgan[2] and Canan Atilgan[2]**

[1] Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, Netherlands, [2] Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey

We address the problem of triggering dissociation events between proteins that have formed a complex. We have collected a set of 25 non-redundant, functionally diverse protein complexes having high-resolution three-dimensional structures in both the unbound and bound forms. We unify elastic network models with perturbation response scanning (PRS) methodology as an efficient approach for predicting residues that have the propensity to trigger dissociation of an interacting protein pair, using the three-dimensional structures of the bound and unbound proteins as input. PRS reveals that while for a group of protein pairs, residues involved in the conformational shifts are confined to regions with large motions, there are others where they originate from parts of the protein unaffected structurally by binding. Strikingly, only a few of the complexes have interface residues responsible for dissociation. We find two main modes of response: In one mode, remote control of dissociation in which disruption of the electrostatic potential distribution along protein surfaces play the major role; in the alternative mode, mechanical control of dissociation by remote residues prevail. In the former, dissociation is triggered by changes in the local environment of the protein, e.g., pH or ionic strength, while in the latter, specific perturbations arriving at the controlling residues, e.g., via binding to a third interacting partner is required for decomplexation. We resolve the observations by relying on an electromechanical coupling model which reduces to the usual elastic network result in the limit of the lack of coupling. We validate the approach by illustrating the biological significance of top residues selected by PRS on select cases where we show that the residues whose perturbation leads to the observed conformational changes correspond to either functionally important or highly conserved residues in the complex.

Keywords: perturbation response scanning, elastic network model, protein complexes, structural motifs, electrostatic potential distribution, protein–protein dissociation, allostery, cooperative conformational change

OPEN ACCESS

Edited by:
Pemra Doruker, University of Pittsburgh, United States

Reviewed by:
Ruth Nussinov, National Cancer Institute (NCI), United States
Pavel Srb, Academy of Sciences of the Czech Republic (ASCR), Czechia

***Correspondence:**
Canan Atilgan
canan@sabanciuniv.edu

Specialty section:
This article was submitted to Biological Modeling and Simulation, a section of the journal Frontiers in Molecular Biosciences

Received: *24 June 2020*
Accepted: *03 August 2020*
Published: *25 August 2020*

Citation:
Abdizadeh H, Jalalypour F, Atilgan AR and Atilgan C (2020) A Coarse-Grained Methodology Identifies Intrinsic Mechanisms That Dissociate Interacting Protein Pairs. Front. Mol. Biosci. 7:210. doi: 10.3389/fmolb.2020.00210

INTRODUCTION

Chemical and physical processes within assemblies of proteins in the cellular environment are events often encompassing multiple time and length scales. Therefore, different modeling tools are commonly used to describe the network of interactions featuring the protein dynamics (Adcock and McCammon, 2006; Tozzini, 2010). In this vein, coarse-grained models (CG), with several

atoms of the protein grouped into one bead and in the absence of atomic details of the solvent molecules have been developed to supplement the extremely expensive atomistic modeling of large scale motions of biomolecular aggregates (Orozco, 2014; Atilgan, 2018). CG models have proved to aid sampling efficiency, predict allosteric regulations (Ming and Wall, 2005) and describe conformational transition pathways (Kim et al., 2002; Miyashita et al., 2003). One useful measure of large-scale protein mechanics in the context of CG models is the elastic network model (ENM) (Tirion, 1996; Bahar et al., 1997; Hinsen, 1998). ENMs are based on the assumption that the potential energy of the system is approximated harmonically about a single minimum energy conformation. Methodologically, in ENM, the atomic details of the biomolecule structure are reduced to a network of nodes (typically one site per residue) connected by harmonic springs. Since all springs are in a relaxed state in the network, no energy minimization is required, in comparison to normal mode analysis in which an expensive initial energy minimization is required prior to calculating the Hessian matrix. For large biomolecules and multi-protein complexes, ENM models with a resolution lower than standard have been used (Durand et al., 1994; Doruker et al., 2002; Chennubhotla et al., 2005; Ahmed and Gohlke, 2006; Kurkcuoglu et al., 2009; Ross C. et al., 2018). The gross representation of large assemblies has proven to predict dynamics of the rigid and flexible parts of the proteins (Ross C.J. et al., 2018). Anisotropic network model (ANM) and Gaussian network model (GNM) are amongst the most widely used ENM-based methods (Bahar et al., 1997; Atilgan et al., 2001). While GNM is applied to produce the amplitudes of isotropic thermal motions, ANM describes both amplitudes and directionality of anisotropic motions. Increased amount of data for proteins of different forms (free, liganded, or complexed), elucidates the correlation between protein function observed in experiments and the global motions predicted by ANM/GNM analyses. Numerous studies have employed ENM-based models to explore various aspects of protein structural dynamics. These include identifying and visualizing collective motions (Kong et al., 2003), predicting modes of motion underlying function (Baysal and Atilgan, 2001b; Keskin et al., 2002; Zheng and Doniach, 2003; Whitford et al., 2007), and explaining details of conformational changes of various types and amplitudes (Tama and Sanejouand, 2001; Krebs et al., 2002; Zheng et al., 2007). ENMs may be applied to the refinement of medium to low-resolution structures of electron microscopy density maps of large macromolecular complexes or molecular envelopes derived from small-angle x-ray scattering (SAXS) data (Delarue and Dumas, 2004; Hinsen et al., 2005). Within the concept of ENM, researchers have developed approaches to generate feasible pathways for conformational transitions between two end conformers (Kim et al., 2002; Orellana et al., 2019), removing the need for expensive molecular dynamics (MD) simulations and all-atom empirical force fields to set up intermediate conformations. ENMs are also applied to determine the main evolutionary transformations of structural changes among homologous proteins (Leo-Macias et al., 2005; Han et al., 2008). In such an approach, for a given set of proteins, evolutionary direction is argued to be a combination of a small subspace projected by a few low frequency modes imposed by inter-residue contact topology.

We have extended ENM to analyze allosterically significant residues and function-related motions of proteins via a technique named perturbation-response scanning (PRS) (Atilgan and Atilgan, 2009; Atilgan et al., 2010a). The methodology inserts fictitious forces on selected atoms and predicts the response within the realm of linear response theory (LRT). *In vivo*, the perturbation may arrive in the form of changing environmental conditions such as pH or ionic strength (Abdizadeh et al., 2015a; Sensoy et al., 2017), or it may act directly on the chain as in pulling (Carrion-Vazquez et al., 2003; Dietz et al., 2006) or other single-molecule experiments, as well as through mutations or ligand binding. PRS serves as a tool to gain insight into molecular origins of mechanical feedbacks of bimolecular structures through recording response to each inserted force on each residue of a protein (Atilgan and Atilgan, 2009; Atilgan et al., 2010a, 2011; Abdizadeh et al., 2015b; Abdizadeh and Atilgan, 2016). It is further capable of recognizing how directionality of the inserted force may coordinate the response of the protein in a functional motion (Jalalypour et al., 2020). PRS requires two distinct conformations of a protein, determined, e.g., by x-ray crystallography, as input; and relies on LRT to relate virtual external forces acting on a protein to the perturbed positions of the residues (Ikeguchi et al., 2005). In PRS, one scans a protein structure residue by residue through applying forces in many directions and records the subset of responses that encourage conversion to another known conformation of the protein. Thus, one can map the regions on the protein surface whose perturbation might lead to the expected conformational change. Besides mapping active site residues that are prime regions for invoking conformational transitions, this approach also has the potential of pointing out allosteric locations and drug target regions. For example, previously, we have studied the proteins calmodulin (Atilgan et al., 2011) and ferric binding protein (Atilgan and Atilgan, 2009) via PRS. By mutating those residues that were implicated in allosteric communication, we later verified through classical MD simulations that they affect the conformation distributions (Aykut et al., 2013; Guven et al., 2014). In a later study, we have performed PRS on subtilisin in complex with its inhibitor to pinpoint hot residues involved in catalytic mechanism and stability of the enzyme (Abdizadeh et al., 2015b). PRS has also been used in the conformation generation step of a flexible docking scheme for exploring protein-ligand interactions (Bolia et al., 2014). In a similar methodology, Mottonen et al. (2010) and Jacobs et al. (2012) have used a method based on distance constraint model to impose constraints on the torsional degrees of freedom of the protein to mimic a hypothetical ligand-binding situation.

In this manuscript, we utilize these CG approaches to address a major challenge for structural biology in providing a mechanistic view of the behavior of molecular complexes and their conformational changes. Protein-ligand and protein–protein interactions (PPI) govern most of the cellular processes (De Las Rivas and Fontanillo, 2010). Many studies investigate the protein-ligand complexes and look for functional regions, binding sites or druggable cavities (Lichtarge and Sowa, 2002;

Xie et al., 2009; Siragusa et al., 2015). On the other hand, PPI allow a protein to perform its biological function by interacting with another partner protein (Ozdemir et al., 2019). Therefore, the interface is usually considered as a candidate to be targeted by a potential drug such as orthosteric or allosteric PPI modulators (Xie et al., 2009; Ni et al., 2019). Studying PPIs and protein-interaction networks may provide insights into new opportunities in the medical, biotechnological, and pharmaceutical fields. Hence, several approaches have been proposed to study PPIs (Ozdemir et al., 2019). A number of bioinformatics techniques have been developed to predict PPI networks based on genomic-context, sequence homology and structural similarity (Shi et al., 2005). Most systematic studies involving protein–protein complexes focuses on the interaction interface to determine compatibility of the structures or attempts to study individual PPI and predicts residues, called hotspots, effective in recognition of partner proteins (Liu Q. et al., 2018; Qiao et al., 2018). Alanine scanning mutagenesis is the major experimental method to identify hotspots (Kenneth Morrow and Zhang, 2012). In one study, non-covalent interactions (hydrophobic, van der Waals, and hydrogen bonding) are found to account as the major forces operating at the PPI interfaces (Gao et al., 2004). Available computational techniques for PPI hotspot prediction are roughly divided to two groups whereby most use the complex structure and a few utilize unbound structures (Ozdemir et al., 2019). Generally, hotspots resulting from the computations are compared to those from alanine scanning mutagenesis experiments (Bradshaw et al., 2011; Ibarra et al., 2019). In addition, machine learning-based methods have been developed to predict hotspots, considering the amino-acid features and conservation information (Liu S. et al., 2018). Most recently, by ignoring the internal structures of the molecules and scanning the protein surface for the so-called "interaction fingerprints," geometric deep learning algorithms have been developed for predicting protein-protein complexes (Gainza et al., 2020). Although these approaches attempt to define a general interaction pattern based on parameters such as structure, hydrophobicity or polarity, there is no general rule to be used in PPI prediction due to their diversity (Ni et al., 2019).

In this work we address a reverse problem: How is it possible to trigger dissociation events between proteins which have already formed a complex? We study 25 sets of protein complexes utilizing PRS with the ENM harmonic potential to determine regions responsible for rendering known complexes incompatible. Elastic network construction helps one to probe conformational changes due to altered physical and chemical environment (Atilgan et al., 2012). Accordingly, the information needed for assessing protein–protein interactions can be derived from knowledge of inter-residue contact topology, buried in the Hessian matrix deduced from ENM (Bahar et al., 2010). Rather than focusing on the interface of the interacting subunits, we relate the physical effects of the internal dynamics of protein complexes to the motions involved in their dissociation. We show that PRS maps residues that may alone initiate the structural change between the bound and unbound forms during dissociation processes of the protein complexes.

MODELS AND METHODS

The conformational change was analyzed via PRS between two different conformers of a protein, one in its complexed form with another protein, and the other in its unbound form. The propensity to convert between conformations was examined for these two states of the protein by employing fictitious forces. The detailed theory of PRS has been laid-out in previous studies (Atilgan and Atilgan, 2009; Atilgan et al., 2010a). In brief, the unbound state of a protein may be described by a perturbation of the Hamiltonian of the bound state, or vice versa. Under LRT, the shift in the coordinates due to unbinding is approximated by Yilmaz and Atilgan (2000) and Atilgan and Atilgan (2009):

$$\Delta\mathbf{R}_1 = \langle\mathbf{R}\rangle_1 - \langle\mathbf{R}\rangle_0 \simeq \frac{1}{3k_BT}\langle\Delta\mathbf{R}\Delta\mathbf{R}^{\mathrm{T}}\rangle_0\Delta\mathbf{F} = \frac{1}{3k_BT}\mathbf{H}^{-1}\Delta\mathbf{F} \tag{1}$$

where the subscripts 1 and 0 denote perturbed and unperturbed configurations of the protein, k_B is the Boltzmann constant and T is temperature. $\Delta\mathbf{F}$ vector contains the components of the externally inserted force vectors on the selected residues; e.g., for the perturbation of a single residue i, $(\Delta\mathbf{F})^T = \left\{000\ldots\Delta F_x^i\Delta F_y^i\Delta F_z^i\ldots 000\right\}_{1\times 3N}$. $\mathbf{H}^{-1}$ is the variance-covariance matrix which may be obtained from either the atomic coordinate trajectories of MD simulations of suitable length (Atilgan et al., 2012), or by imposing the approximation of harmonic springs between pairs of interacting atoms. In this work, we generate the $\mathbf{H}^{-1}$ matrix from a coarse-grained approach, constructing a network of N nodes on the C_α atoms of the protein complexes whose coordinates are directly used from their protein data bank (PDB) entries. Any given pair of nodes interacts in accord with a conventional harmonic potential, if the nodes are within a cut-off distance, r_c, of each other. This leads to a total of M interactions. Within the scope of an elastic network of residues that are connected to their neighbors by linear-elastic springs, one gets the $3N \times M$ direction cosine matrix $\mathbf{B}$ (Yilmaz and Atilgan, 2000). $\mathbf{BB}^T$ is exactly the Hessian if harmonic interactions with uniform force constants for all M bonds in the network are assumed. $(\mathbf{BB}^T)^{-1}$ is the covariance matrix $\mathbf{H}^{-1}$ for a given configuration, which is also an $N \times N$ supermatrix whose ij^{th} element is the 3×3 matrix of correlations between the x-, y-, and z-components of the fluctuations $\Delta\mathbf{R}_i$ and $\Delta\mathbf{R}_j$ of residues i and j.

$\mathbf{H}$ of the system has at least six zero eigen-values corresponding to the purely translational and rotational motions. The eigen-value distribution of the Hessian of proteins is such that the low-frequency region is more crowded than expected of polymers or other condensed matter (Ben-Avraham, 1993). Thus, the choice of the cutoff distance, r_c, for the construction of the Hessian is critical for extracting protein-like properties from the systems studied (Atilgan et al., 2010b). For all the proteins studied in this work, we coarse-grain the crystal structure so that each residue is represented by the coordinates of its C_α atom. To account for the flexibility of proteins, we repeat the PRS analysis for a variety of cut-off distances in the range of 10.0–14.0 Å in increments of 1 Å; the lower limit of 10 Å agrees with the definition of first coordination shell of residues in proteins

(∼ 7 Å). For each network structure, we ensure that the system has six zero eigen-values corresponding to the translational and rotational degrees of freedom of the protein. The smallest common r_c at which we obtain six zero eigenvalues for all the proteins tested is 9 Å.

PRS technique relies on repeating the above LRT calculation (Equation 1) by scanning the residues of the protein one-by-one and focusing further on those perturbations that overlap with the conformational change, $\Delta\mathbf{R}_1 = \langle\mathbf{R}\rangle_1 - \langle\mathbf{R}\rangle_0$. There is no *a priori* assumption on how a force might be generated at a particular point. Conversely, after finding the force/residue pair that best leads to the conformational change of interest, we relate this finding to possible causes. In this study, PRS is applied by scanning each residue in 500 random directions.

To assess the quality of the predicted displacements of all residues resulting from a force applied on selected residue *i*, we use the correlation coefficient between the predicted and experimental displacements, averaged over all the affected residues, *k*:

$$C_i = \frac{\sum_{k=1}^{N}\left[(\Delta R_{\mathrm{k}})^{\mathrm{i}} - \left(\overline{\Delta R}\right)^{\mathrm{i}}\right]\left(\Delta S_{\mathrm{k}} - \overline{\Delta S}\right)}{(N-1)\,\sigma_R\sigma_S} \qquad (2)$$

In equation 2, the overbar indicates the average, $\Delta\mathbf{S}_k$ are the displacements between the initial and the target forms obtained from the PDB structures, σ_S and σ_R are the respective standard deviations for experimental and calculated displacements. A value close to 1 implies good agreement with the experimental conformational change, while a value of zero indicates lack of correlation between experimental and theoretical findings. Several approaches were taken to select the residues that are effective in directing the protein toward alternative conformations depending on the distribution of the maximum of the C_i values, C_i^{max}, obtained from the 500 perturbations and calculated through equation 2. We first list C_i^{max} in ascending order: (1) If there is a sharp decrease in the Pearson correlation values, we list the top residues until that gap. (2) If there is a smooth decrease in the Pearson correlation values, we list the residues that are common among top 10 residues of all cut-off values. We also check the location of the residues that do not survive these selection criteria. We have found that the variable residues observed among top residues in different cut-off values are spatial neighbors of the listed ones.

Protein Complex Selection

We analyzed a set of 25 protein complexes in their bound and unbound forms (**Table 1**, column 2). We collected protein pair structures from those reported in Benchmark 0.0 of ZDOCK (Chen and Weng, 2002). The complexes are non-redundant and they have X-ray structure solved at better than 2.90 Å resolution. They include a wide variety of function and affinities; they belong to different biological families. The constituent unbound forms of all the 25 complexes are available in the PDB with solution NMR or X-ray structures solved at better than 3.50 Å resolution. For the proteins resolved by solution NMR, we always use the first model for the PRS calculations. More specifically, we have chosen the protein complexes with no less than two missing residues along the protein chain, either in the bound or unbound form. Furthermore, if the number of residues in the bound and unbound components differ, we only analyze the common parts of the bound and unbound proteins.

Optimization of the Cut-Off Distance

For each pair of experimental structures, the unbound form is superimposed on the bound form, followed by the residue displacement vectors computation, $\Delta\mathbf{S}$. In this study, we perturb the bound form of each protein by applying a random force to the C_α atom of each residue in the complex. We select residues whose perturbation leads to the $\Delta\mathbf{R}$ vector (equation 1) that best resembles the dissociated proteins using equation 2 as the measure. For a given protein, we select the cut-off distance, r_{opt} that yields the closest agreement with the displacement vectors from experiments for at least one residue (**Table 1**, column 8). We verify that the residue indices that provide the best Pearson correlation value are always present within the range perceived as the highest correlation value for all cut-off distances studied. We note that the correlation values reported in **Table 2** (column 5 and 6) are not affected for the range of values $r_{opt} \pm 1$ Å for all of the 25 protein complexes. Strikingly, although many of these proteins display large Pearson correlation values, the numbers of residues yielding the highest values differ among proteins. For some proteins, there is not much specificity on the residue to be perturbed to reproduce the conformational change. For others, by perturbing a very specific location, the complete conformational change is obtained. The former category is exemplified by fibrinogen-binding protein (pdb code: 3D5S; Haspel et al., 2008) while serine protease and its proteinaceous inhibitor (pdb code: 1D6R; Koepke et al., 2000) is an example of a protein complex that we need to perturb in a specific location to mimic the dissociated conformation.

RESULTS AND DISCUSSION

The structural difference between the bound and the unbound forms, based on the C_α superposition of the binding partners, show that while for 22 of the cases the interface RMSD is less than 1.5 Å; for two cases interface RMSD is 1.5–2.2 Å and two cases have interface RMSD greater than 2.2 Å (Humphrey et al., 1996). For the pair of initial and target forms of the proteins present in bound and unbound form, we perform STAMP structural alignment, implemented in VMD 1.9.1 MultiSeq plugin (Humphrey et al., 1996). We record the RMSD between the structures of the target forms with the initial structure which vary between 0.4 and 4.3Å (**Table 1**, column 6). Although large motions of side chains and surface loops is always present as a local conformational change, we do not detect any discernible shape change on a global scale in proteins in their bound form compared to the respective unbound constituents. For all cases, the main chains essentially have the same conformation in bound and unbound forms and the binding partners interact as rigid bodies. However, the superposition of mobile segments of the proteins in the complex and corresponding unbound forms produces higher RMSDs (**Table 1**, column 7) and can be as large

TABLE 1 | General features of protein complexes studied.

Type protein complex		PDB codes (Bound: chain/Unbound: chain) (X-ray resolution, Å)	Amino acid length	*N*	Global RMSD (Å)	Local RMSD Regions of motion (residue:Å)	Cutoff, r_{opt} (Å)
Type I	**Alpha amylase: Alpha amylase inhibitor**	1CLV:A/1JAE:A (2.00/1.65)	2–471	470	0.5	81–90: 1.8 290–300: 2.5 440–450:1.7	12
		1CLV:I/1HTX:A (2.00/NMR)	501–532	32	0.9	511–520:1.0 520–525:1.3	
	MAP kinase-activated protein kinase 2:Mitogen-activated protein kinase 14	2OZA:A/1KWP:A (2.70/2.80)	51–215	165	2.0	65–70:6.2 70–75:8.0 75–80:3.0	14
		2OZA:B/1P38:A (2.70/2.10)	16–169	154	1.7	16–21:1.8 25–30:2.3 30–35:5.3 35–40:2.6 55–58:2.4 97–100:2.2 115–120:2.9 120–125:2.9	
	Trypsin:Trypsin inhibitor	1AVX:A/1QQU:A (1.90/1.63)	16–245	230	0.5	95–100:0.85 165–170:0.85 215–220:0.90 240–245:0.75	13
		1AVX:B/1BA7:A (1.90/2.50)	501–623	123	0.5	545–550:0.6 595–600:0.8	
	Ras-related protein Ral-A:Mono-ADP-ribosyltransferase C3	2A9K:A/1U90:A (1.73/2.00)	13–178	166	0.6	47–51:0.8 56–62:0.8 70–75:2.1	11
		2A9K:B/2C8B:X (1.73/1.70)	45–245	201	1.2	70–82:1.2 140–150:1.5 180–185:2.9 209–216:3.3	
	Ribonuclease SA:Barstar	1AY7:A/1RGH:A (1.70/1.20)	1–96	96	0.5	28–30:0.7 39–41:0.7 61–65:0.9	13
		1AY7:B/1A19:A (1.70/2.76)	1–89	89	0.6	5–15:0.95 55–66:0.95	
	Carboxypeptidase A:Metallocarboxypeptidase inhibitor	4CPA:A/1YME:A (2.50/1.53)	1–307	307	0.5	132–136:1.8 245–249:1.2	13
		4CPA:I/1H20:A (2.50/NMR)	4–37	34	1.1	10–16:1.6 17–21:2.1	
	Ribonuclease A:Ribonuclease inhibitor	1DFJ:E/9RSA:A (2.50/1.80)	1–124	124	0.7	15–17:0.8 90–95:1.2 110–114:1.0	12
		1DFJ:I/2BNH:A (2.50/2.30)	1–456	456	1.5	1–30:2.9 40–55:2.6 414–423:2.7 443–451:3.1	
	Superoxide dismutase:copper chaperone for superoxide dismutase	1JK9:A/2JCW:A (2.90/1.70)	1–153	153	0.8	51–62:2.5	13
		1JK9:B/1QUP:A (2.90/1.80)	3–222	220	4.3	3–10:5.4 30–40:5.3 50–70:6.9 160–170:5.7	
	Carboxypeptidase A1:Metallocarboxypeptidase inhibitor	2ABZ:A/1M4L:A (2.16/1.25)	5–305	301	0.4	245–250:1.3	13

(Continued)

TABLE 1 | Continued

Type protein complex		PDB codes (Bound: chain/Unbound: chain) (X-ray resolution, Å)	Amino acid length	*N*	Global RMSD (Å)	Local RMSD Regions of motion (residue:Å)	Cutoff, r_{opt} (Å)
		2ABZ:C/1DTV:A (2.16/NMR)	5–65	61	1.3	12–15:2.1 52–56:1.5	
	Ferredoxin-NADP reductase:Ferredoxin I	1EWY:A/1GJR:A (2.38/2.10)	9–303	295	1.1	50–54:1.4 67–74:3.3 102–112:2.0 221–236:2.3 262–270:1.4 280–287:1.2	12
		1EWY:C/1CZP:A (2.38/1.17)	1–98	98	0.8	10–15:1.5 50–67:1.5	
	Cysteine protease:cysteine protease Inhibitor	1PXV:A/1X9Y:A (1.80/2.50)	223–392	170	2.5	330–339:9.1 375–382:3.9	11
		1PXV:C/1NYC:A (1.80/1.40)	0–109	110	0.9	0–5:1.6 91–96:1.3 104–109:2.7	
Type II	**Chemotaxis protein CHEY:Chemotaxis protein CHEA**	1FFW:A/3CHY:A (2.70/1.66)	2–129	128	0.5	90–92:1.1 110–112:0.9 124–130:1.3	12
		1FFW:B/1FWP:A (2.70/NMR)	160–226	67	1.8	165–170:2.6 200–205:2.6 210–215:2.3	
	Cell division protein FTSZ:Hypothetical protein PA3008	1OFU:A/2VAW:A (2.10/2.90)	11–316	306	0.5	70–72:0.9 122–125:0.9 202–208:2.7 231–235:1.2 268–273:1.8 288–293:0.9 299–306:2.0	12
		1OFU:X/1OFT:A (2.10/2.90)	45–160	116	0.8	70–72:1.1 87–90:1.0	
	Complement C3:Fibrinogen-binding protein	3D5S:A/1C3D:A (2.30/1.80)	8–298	291	0.4	44–51:1.7 165–171:1.0	12
		3D5S:C/2GOM:A (2.30/1.25)	15–75	61	0.4	15–19:0.7 70–75:0.9	
	Bovine hymotrypsinogen A:human pancreatic secretory trypsin inhibitor (Kazal-type)	1CGI:E/2CGA:A (2.30/1.80)	1–245	245	1.5	31–39:1.2 140–154:4.5 183–195:3.5 216–223:1.6	13
		1CGI:I/1HPT:A (2.30/2.30)	(1–56)	56	1.8	1–5:4.4 10–14:3.2 17–20:1.5	
	Elastase:Elafin	1FLE:E/1QNJ:A (1.90/1.10)	16–245	229	0.9	59–62:1.4 96–106:1.6 166–176:1.7 186–196:1.3 216–226:1.7	13 Å
		1FLE:I/2REL:A (1.90/NMR)	11–57	47	2.8	11–43:3.5	
Type III	**Interleukin-6 receptor beta chain Leukemia inhibitory factor**	1PVH:A/1BQU:B (2.50/2.00)	101–301	201	0.9	130–135:1.7 210–215:1.3 240–245:1.3 295–301:1.8	12
		1PVH:B/1EMR:A (2.50/3.50)	22–180	159	0.9	135–140:1.3 140–150:2.0 150–160:1.6	

(Continued)

TABLE 1 | Continued

Type protein complex		PDB codes (Bound: chain/Unbound: chain) (X-ray resolution, Å)	Amino acid length	*N*	Global RMSD (Å)	Local RMSD Regions of motion (residue:Å)	Cutoff, r_{opt} (Å)
	Alkaline metalloproteinase:Proteinase inhibitor	1JIW:P/1AKL:A (1.74/2.00)	1–470	470	1.2	20–22:1.5 185–195:7.3	10
		1JIW:I/2RN4:A (1.74/NMR)	8–105	98	1.4	18–25:3.4 50–55:1.2 95–98:2.2	
	Heat shock protein HSP82:HSP90 Co-chaperone CDC37	1US7:A/1AH6:A (2.30/1.80)	2–207	206	0.8	54–57:1.3 91–99:3.5 100–117:1.7 198–203:1.3	10
		1US7:B/2W0G:A (2.30/1.88)	148–276	129	1.0	224–228:1.4 235–240:1.3 240–255:2.7 270–276:1.9	
	Trypsinogen:Bowman-Birk proteinase inhibitor precursor	1D6R:A/2TGT:A (2.30/1.70)	19–245	217	0.6	96–99:0.7 144–146:0.8 188–194:2.6 217–220:1.6	11
		1D6R:I/1K9B:A (2.30/2.80)	7–63	57	1.0	30–35:1.3 42–46:1.4 57–59:1.1	
Type IV	**Falcipain 2:Chagasin**	2OUL:A/2GHU:A (2.20/3.10)	15 to 224	240	0.6	1–5:2.1 115–120:1.0 153–158:1.2 185–195:2.6	12
		2OUL:B/2H7W:A (2.20/1.70)	4–110	107	0.6	60–65:1.7	
	ATP-dependent Clp protease ATP-binding subunit clpA: ATP-dependent Clp protease adaptor protein clpS	1R6Q:A/1R6B:X (2.35/2.25)	1–141	141	1.0	15–20:1.2 68–76:2.8 91–102:1.6	13
		1R6Q:C/3O1F:A (2.35/1.40)	26–106	81	0.4	37–40:0.6	
	NAD(P) transhydrogenase subunit alpha part 1: NAD(P) transhydrogenase subunit beta	2OOR:A/1L7D:A (2.32/1.81)	1–220	220	0.6	45–50:0.9 80–85:1.3 164–170:1.6 204–207:1.2 214–220:1.1	12
		2OOR:C/1E3T:A (2.32/NMR)	30–201	172	2.2	30–50:2.8 140–150:2.9 170–190:2.4	
	Alpha-Chymotrypsin:Protease inhibitor LCMI II	1GL1:A/1MTN:F (2.10/2.80)	16–146	131	0.5	71–82:1.3 143–146:1.0	12
		1GL1:I/1PMC:A (2.10/NMR)	2–33	32	1.6	7–23:2.1	
	Alpha amylase:Tendamistat	1BVN:P/1PIG:A (2.50/2.20)	2–496	495	0.7	51–55:0.8 108–112:2.3 140–150:1.5 238–243:1.9 303–308:2.7 347–351:1.4	13
		1BVN:T/1HOE:A (2.50/2.00)	804–874	71	0.5	827–830:0.6 838–840:0.7 860–862:0.7	

as 9.1 Å (see the case of cysteine protease and its inhibitor; Filipek et al., 2003), the largest deviations being mostly confined to flexible unstructured stretches, i.e., turns and bends.

Recording the residues identified by PRS (**Table 2**; those with C_i values exceeding the value listed in column 6), displays those encouraging the conformational change from the bound form

TABLE 2 | PRS results and classification of the protein complexes.

Type protein complex		PDB codes (Bound:chain/ Unbound:chain) (X-ray resolution, Å)	Correlation Bound/Unbound	C_i^{max}	$>C_i$	Interface residues	First shell residues	Remote site residues
Type I	**Alpha amylase:Alpha amylase inhibitor**	1CLV:A/1JAE:A (2.00/1.65)	0.92	0.87	0.87	G292:A	N/A	N/A
		1CLV:I/1HTX:A (2.00/NMR)	0.75	0.76	0.66	W57:A, N138:A, V151:A, G152:A	N/A	N/A
	MAP kinase-activated protein kinase 2:Mitogen-activated protein kinase 14	2OZA:A/1KWP:A (2.70/2.80)	0.97	0.68	0.65	L70:A	G71:A	N/A
		2OZA:B/1P38:A (2.70/2.10)	0.96	0.80	0.80	N/A	N/A	G171:A, Q175:A, Y176:A
	Trypsin:Trypsin inhibitor	1AVX:A/1QQU:A (1.90/1.63)	0.9	0.62	0.59	N/A	N/A	L520:B, K552:B
		1AVX:B/1BA7:A (1.90/2.50)	0.87	0.42	0.40	N/A	N/A	S579:B, F580:B, A581:B, D598:B, K611:B
	Ras-related protein Ral-A:Mono-ADP-ribosyltransferase C3	2A9K:A/1U90:A (1.73/2.00)	0.78	0.8	0.79	N/A	N/A	Y66:B, G67:B, L68:B, S69:B, D112:B
		2A9K:B/2C8B:X (1.73/1.70)	0.73	0.68	0.62	N/A	N/A	S181:B, F209:B, A210:B, G211:B
	Ribonuclease SA:Barstar	1AY7:A/1RGH:A (1.70/1.20)	0.93	0.72	0.69	N/A	N/A	L41:B, T42:B, G43:B, W44:B
		1AY7:B/1A19:A (1.70/2.76)	0.75	0.52	0.50	N/A	N/A	E8:B, E57:B, Q58:B
	Carboxypeptidase A:Metallocarboxypeptidase inhibitor	4CPA:A/1YME:A (2.50/1.53)	0.97	0.70	0.70	N/A	N/A	S134:A
		4CPA:I/1H20:A (2.50/NMR)	0.50	0.70	0.66	N/A	N/A	K177:A, S199:A, I274:A
	Ribonuclease A:Ribonuclease inhibitor	1DFJ:E/9RSA:A (2.50/1.80)	0.76	0.59	0.58	N/A	N/A	G186:I, D213:I, P450:I, G451:I
		1DFJ:I/2BNH:A (2.50/2.30)	0.80	0.91	0.89	N/A	N/A	L22:I, A46:I, L47:I, R48:I, A49:I
	Superoxide dismutase:copper chaperone for superoxide dismutase	1JK9:A/2JCW:A (2.90/1.70)	0.70	0.78	0.76	N/A	I69:B	D67:B, A68:B
		1JK9:B/1QUP:A (2.90/1.80)	0.70	0.60	0.57	N/A	N/A	C27:B, P54:B, S55:B
	Carboxypeptidase A1:Metallocarboxypeptidase inhibitor	2ABZ:A/1M4L:A (2.16/1.25)	0.96	0.75	0.73	Q16:C, V17:C	C18:C	N/A
		2ABZ:C/1DTV:A (2.16/NMR)	0.56	0.75	0.75	N/A	E31:C	N/A
	Ferredoxin-NADP reductase:Ferredoxin I	1EWY:A/1GJR:A (2.38/2.10)	1.00	0.70	0.67	N/A	N/A	V67:A, D68:A, K69:A
		1EWY:C/1CZP:A (2.38/1.17)	0.80	0.74	0.72	N/A	I62:A	T164:A, F183:A

(Continued)

TABLE 2 | Continued

Type protein complex		PDB codes (Bound:chain/ Unbound:chain) (X-ray resolution, Å)	Correlation Bound/Unbound	C_i^{max}	$>C_i$	Interface residues	First shell residues	Remote site residues
	Cysteine protease:cysteine protease Inhibitor	1PXV:A/1X9Y:A (1.80/2.50)	0.55	0.9	0.85	N/A	H54:C	V10:C, Y11:C, H44:C
		1PXV:C/1NYC:A (1.80/1.40)	0.72	0.80	0.80	N/A	N/A	V109:C
Type II	**Chemotaxis protein CHEY:Chemotaxis protein CHEA**	1FFW:A/3CHY:A (2.70/1.66)	0.83	0.75	0.74	N/A	N/A	K190:B, G191:B, L195:B, A197:B
		1FFW:B/1FWP:A (2.70/NMR)	0.77	0.61	0.54	N/A	N/A	G52:A, V54:A, D57:A, N59:A
	Cell division protein FTSZ:Hypothetical protein PA3008	1OFU:A/2VAW:A (2.10/2.90)	0.92	0.83	0.82	H89:X, R93:X	L87:X, T88:X	N/A
		1OFU:X/1OFT:A (2.10/2.90)	0.90	0.58	0.56	I207:A	D210:A, L271:A, S272:A	N/A
	Complement C3:Fibrinogen-binding protein	3D5S:A/1C3D:A (2.30/1.80)	0.92	0.70	0.69	N/A	N67:C, K70:C, Q71:C	N/A
		3D5S:C/2GOM:A (2.30/1.25)	0.78	0.80	0.78	N/A	N/A	R10:A, L11:A, K12:A, H13:A, L14:A, I15:A, V16:A, T17:A
	Bovine hymotrypsinogen A:human pancreatic secretory trypsin inhibitor (Kazal-type)	1CGI:E/2CGA:A (2.30/1.80)	0.60	0.70	0.68	T30:I, Y31:I, P32:I	N/A	N/A
		1CGI:I/1HPT:A (2.30/2.30)	0.48	0.83	0.79	G197:E	A179:E	N/A
	Elastase:Elafin	1FLE:E/1QNJ:A (1.90/1.10)	0.96	0.75	0.65	N/A	N/A	T11:I, K12:I, P13:I, L33:I, K34:I
		1FLE:I/2REL:A (1.90/NMR)	-0.30	0.71	0.66	N/A	N/A	L123:E, A208:E, V209:E
Type III	**Interleukin-6 receptor beta chain Leukemia inhibitory factor**	1PVH:A/1BQU:B (2.50/2.00)	0.89	0.85	0.83	N/A	N/A	R276:A, I277:A, E294:A, A295:A, S296:A, G297:A
		1PVH:B/1EMR:A (2.50/3.50)	0.84	0.58	0.54	N/A	N/A	G147:B, P148:B, D149:B, T150:B
	Alkaline metalloproteinase: Proteinase inhibitor	1JIW:P/1AKL:A (1.74/2.00)	0.98	0.85	0.78	N191:P, A192:P	G193:P	N/A
		1JIW:I/2RN4:A (1.74/NMR)	0.80	0.79	0.73	N/A	N/A	E21:I, A22:I
	Heat shock protein HSP82:HSP90 Co-chaperone CDC37	1US7:A/1AH6:A (2.30/1.80)	0.80	0.79	0.79	N/A	A97:A	N/A
		1US7:B/2W0G:A (2.30/1.88)	0.92	0.79	0.79	N/A	N/A	A244:B
	Trypsinogen:Bowman-Birk proteinase inhibitor precursor	1D6R:A/2TGT:A (2.30/1.70)	0.96	0.78	0.78	N/A	N/A	N/A

(Continued)

TABLE 2 | Continued

Type protein complex		PDB codes (Bound:chain/ Unbound:chain) (X-ray resolution, Å)	Correlation Bound/Unbound	C_i^{max}	$>C_i$	Interface residues	First shell residues	Remote site residues
		1D6R:I/1K9B:A (2.30/2.80)	0.50	0.70	0.70	N/A	N/A	H33:I, S34:I
Type IV	**Falcipain 2:Chagasin**	2OUL:A/2GHU:A (2.20/3.10)	1.00	0.87	0.83	N/A	D148:A, Y37:B, G41:B	G40:B
		2OUL:B/2H7W:A (2.20/1.70)	0.25	0.78	0.76	N/A	I68:A	S113:A, V114:A, D148:A, F219:A
	ATP-dependent Clp protease ATP-binding subunit clpA:ATP-dependent Clp protease adaptor protein clpS	1R6Q:A/1R6B:X (2.35/2.25)	0.85	0.82	0.77	N/A	E73:A, K49:C	D45:C, L61:C
		1R6Q:C/3O1F:A (2.35/1.40)	0.76	0.76	0.74	S118:A	N/A	E7:A, Y122:A
	NAD(P) transhydrogenase subunit alpha part 1:NAD(P) transhydrogenase subunit beta	2OOR:A/1L7D:A (2.32/1.81)	0.23	0.68	0.68	N/A	N103:C, P105:C	L214:A, T220:A
		2OOR:C/1E3T:A (2.32/NMR)	0.52	0.50	0.49	A166:A	N/A	M167:C
	Alpha-Chymotrypsin:Protease inhibitor LCMI II	1GL1:A/1MTN:F (2.10/2.80)	1.00	0.76	0.52	N/A	N/A	S76:A, S77:A
		1GL1:I/1PMC:A (2.10/NMR)	−0.40	0.87	0.76	C58:A, C14:I	K13:I	N/A
	Alpha amylase:Tendamistat	1BVN:P/1PIG:A (2.50/2.20)	0.89	0.64	0.64	A823:T	C811:T	N/A
		1BVN:T/1HOE:A (2.50/2.00)	0.68	0.57	0.53	N/A	N/A	Q5:P, T6:P, Q7:P, S8:P, R10:P, V804:T

to unbound form do not necessarily reside on these flexible structures. We have collected 161 effector residues from PRS calculations (**Table 2**, columns 7–9). While 63 of them are located on flexible loops with large motions and high RMSD values, we identify 30 residues residing on α-helices and 35 on β-strands. The remaining are on loops that do not display any large structural change upon binding.

The average motions of the proteins, quantified by the root mean square fluctuations (RMSFs), are usually expected to dampen upon binding, especially at the binding interface residues, even when the protein conformation is unaltered (Baysal and Atilgan, 2001a). RMSFs of each protein complex constituents in their bound and unbound form are derived from auto-correlation of the residues in each protein pair. By treating the $\mathbf{H}^{-1}$ matrix as an $N \times N$ supermatrix, whose ij^{th} element is the 3 × 3 second moment matrix of correlations between the x-, y-, and z-components of the fluctuations $\Delta\mathbf{R}_i$ and $\Delta\mathbf{R}_j$ of residues i and j (Baysal and Atilgan, 2001b) are calculated, whose diagonal elements predict the RMSFs (Atilgan et al., 2001). The cut-off distances, r_{opt}, optimized for building the Hessian matrix of each protein complex have the same values as r_c chosen for PRS analysis. We report the correlation values between proteins in their bound and unbound form for all the protein pairs; the similarity between the RMSF profiles of a protein in its bound and unbound form is expressed as a Pearson correlation and is listed in column 4 of **Table 2**. We observe that there is a significant change in RMSF of binding region residues in one of the constituents in each pair, so far as to have a negative correlation in some cases; e.g., fluctuation patterns in some regions of the protein is reversed upon complexation. This means that the local fluctuations of the interface area vary in at least one protein upon complex formation, while local fluctuations of their binding partners display the same pattern as the unbound form. However, in an exceptional case of transhydrogenase complex (pdb code: 2OOR; Bhakta et al., 2007), we observe low correlations, 0.23 and 0.52, between the binding proteins and their unbound form. For this protein complex, the RMSF curves of both the constituents display a significant transformation of fluctuation pattern upon protein binding.

Features of the Amino Acids Involved in Dissociation Event

From the 8828 residues, 161 of them are selected by PRS. These residues, whose perturbation encourages the unbound over

TABLE 3 | Types of amino acids selected by PRS compared to all residues in the study.

	Non-polar				Charged									Polar amino						
Residue type	**A**	**V**	**L**	**I**	**M**	**F**	**G**	**W**	**P**	**D**	**E**	**R**	**K**	**N**	**Q**	**T**	**Y**	**S**	**H**	**C**
PRS residues (%)	10	7	9	4	1	3	13	1	3	6	4	3	6	4	4	5	4	9	2	3
All residues (%)	8	7	9	5	2	4	8	2	4	6	6	4	6	5	4	6	4	7	2	3

TABLE 4 | Secondary structure attributes of amino acids selected by PRS compared to all residues in the study.

	α-helix	**π-helix**	**3-10 helix**	**β-strand**	**Isolated bridge**	**Turn**	**Coil**	**Total**
PRS residues (%)	19	0.0	2	20	1	29	29	100.0
All residues (%)	26	0.0	4	24	2	26	18	100.0

the bound form, are distributed among all amino acid types. The percentage of each amino acid type in our analysis pool and their contributions in PRS analysis are listed in **Table 3**. PRS does not display any preference over amino acid types and any contribution to PRS selection is corroborated to the population of the amino acid type in the total analysis pool. For example, methionines and tryptophans, each with 1% frequency are the least detected residues by PRS. They are also less frequently seen in the analysis pool (2% of the population). The only residue type that is observed significantly more that in the average pool is glycine which constitutes 13% of all PRS selected residues, compared to its 8% abundance in the residue data set.

In **Table 4**, we report the local secondary structure attributes of residues detected by PRS compared to all residues. The secondary structures are assigned by the "Timeline" plugin of VMD and are calculated based on the STRIDE algorithm (Humphrey et al., 1996). Among all the protein complexes, we do not find any π-helix type of structure. Residues in the total analysis pool are mostly populated by turns, α-helices and β-strands with 26, 26, and 24% distribution, respectively. However, we find that most preferred regions by PRS defined residues are on coils and turns, each with 29% of all PRS defined residues, although they populate only 44% of the analysis pool. In particular, the enhancement of coil residues in the PRS selection is statistically significant, as these are represented by 29% in the PRS sub-pool, up from 18% of all residues in the original pool of residues.

We divide the protein structure into three zones; interface, first coordination and remote, so as to categorize the location of the PRS selected residues. The interface of the two proteins present in the complex is defined by the C_α atoms of the residues from the two sides of the pair residing within 7 Å cut-off distance of each other. We define first coordination shell residues as those located within 7 Å cut-off distance from any interface residue. All remaining residues are classified as remote, defined as those residing beyond the first coordination shell of the interface. We observe that except for the case of alpha amylase and its inhibitor (pdb code:1CLV; Pereira et al., 1999), PRS selects for remote residues (**Table 2**). In fact, in 9 cases PRS selected only residues away from the interface. The remaining 16 protein complexes display residues from different parts of the protein in their PRS analysis, including, but not limited to the interface. In fact, these residues are overwhelmingly located on or near the outer surface, as indicated by their depth values from the surface as calculated by the DEPTH server (Tan et al., 2013) and listed in **Supplementary Table S1**. In fact, those few that are deeply located (depth greater than 5 Å; shown in bold) are part of a network of interactions whose one end is located on the surface. Thus, the interface of a protein complex is not the controlling region for dissociation of the two proteins.

Remarkably, residues signaling the dissociation of each protein in a given complex are not located on the same protein in all cases. Accordingly, effective sites involved in the dissociation of various protein complexes found in PRS analysis are categorized into four groups based on their responses to the perturbations on the protein. Proteins in which dissociation is signaled through remote residues of the complex are labeled as Type I (11 cases). In this group, PRS top rated residues are all confined to one of the binding proteins. Thus, residues on this protein also control the conformational changes of the binding partner. Type II are the proteins in which residues confined to one of the proteins control the other binding protein and vice versa. Type III are the proteins in which each constituent controls its own dissociation event; therefore, essential residues are clustered on the "self" protein. Finally, in Type IV both partners are essential for the dissociation to occur, as residues signaling the dissociation are scattered on both binding partners. **Tables 1**, **2** are organized according to these four distinct groups (I–IV).

Long Range Control of Dissociation Is Coupled to Electrostatic Effects

In a subset of the cases, perturbation of specific sites on only one of the constituents in the protein complex modulates dissociation. We label these as Type I group of protein complexes. The functional amino acids defined by PRS which are involved in disintegrating the contact network displays no specific perturbation location in Type I; they may be located on the interface, first shell or remote locations of the protein tertiary structure. Thus, the local perturbations which lead to global conformational shifts between bound and unbound states are not bundled in a specific region.

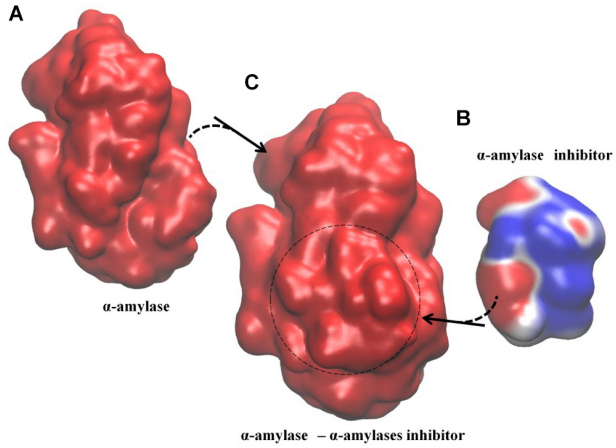

FIGURE 1 | Example of Type I electrostatic isocontour shifts upon binding; α-amylase/α-amylase inhibitor drawn at ±0.5 k_BT/e; blue is positive and red is negative. The signaling protein, α-amylase, where all PRS determined residues reside (**Table 2**), maintains its electrostatic potential distribution while α-amylase inhibitor displays altered electrostatic potential distribution. **(A)** α-amylase in the unbound form, with overall negative electrostatic potential distribution along the surface. **(B)** α-amylase inhibitor in its unbound form with a mixed pattern of negative and positive negative electrostatic potential distribution along the surface. **(C)** α-amylase/α-amylase inhibitor complex with overall negative electrostatic potential distribution. The spatial orientation of the proteins in the complex is kept the same as the respective unbound forms. Dashed circle indicates positioning of the inhibitor in the complex.

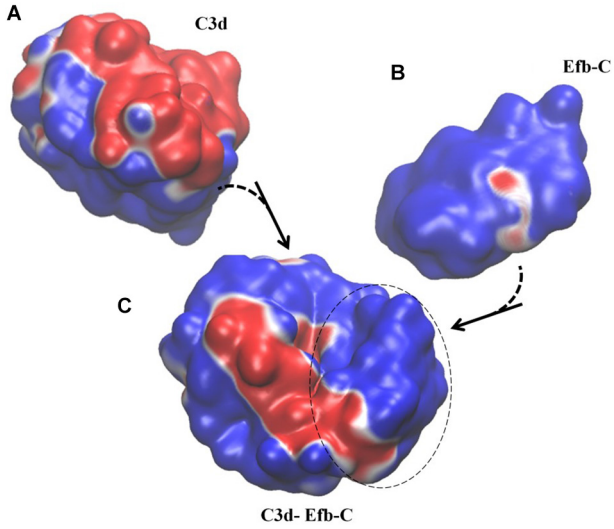

FIGURE 2 | Example of Type II electrostatic isocontour shifts upon binding; Efb-C and its complement target C3d drawn at ±0.5 k_BT/e. Blue is positive and red is negative. **(A)** C3d in its unbound form, with a mix of negative/positive electrostatic potential distribution along the surface. **(B)** Efb-C in its unbound form with predominantly positive electrostatic potential distribution along the surface. **(C)** Efb-C/C3d complex. Dashed circle displays Efb-C protein in the complex and the rest of the surface belongs to C3d protein. Both proteins display a mixture of negative/positive electrostatic potential distribution along the surface. In particular, positive surface of the Efb-C displays increased negative areas upon complex formation, while C3d loses negative patches. The spatial orientation of the proteins in the complex is kept the same as that presented in unbound forms.

To determine if these long range effects are controlled by electrostatics, we obtain the electrostatic potential distributions on the biomolecular surface using the APBS package (Baker et al., 2001). In APBS calculations, parameters are set to their default values; i.e., biomolecular and solvent dielectric constant are set to 2 and 78.54, respectively, the radius of the solvent molecules is 1.4 Å and the temperature is 298.15 K; finally, the cubic B-spline discretization method is used to grid biomolecular point charges. Electrostatic effects play a major role in the functionality of this group of protein complexes. Monitoring the electrostatic potential distribution along the protein surfaces in their bound and unbound forms reveals that any given protein in the complex whose electrostatic potential distribution state is stable in their bound and unbound form is also the protein controlling dissociation. Conversely, the binding partner that does not have any PRS determined residue displays considerable change in its charge distribution. In **Figure 1A**, we exemplify how the electrostatic potential distribution on the surface of the protein changes from the free to the bound form.

Type II group represents another set of protein complexes with remotely controlled conformation changes from the bound to the unbound form. In this group, there is cross-controlled dissociation; i.e., residues that lead to the conformational change upon dissociation on one protein are located on the partner in the complex. We observe that perturbations in a stretch of consecutive resdues is required to trigger the interconversion between two conformational endpoints (see **Table 2**, Type II). In addition, analysis of the electrostatic potential distribution shows that the proteins interacting with each other possess a similar state of charge distribution. Thus, if the unbound forms had a different electrostatic potential distribution, they reorient themselves to the same state upon protein complex formation. In **Figure 2**, we illustrate electrostatic potential surfaces before and after complex formation in Efb-C and its complement target C3d (pdb code: 3D5S; Haspel et al., 2008). In fact, it has been reported that formation and stability of Efb-C binding to C3d is electrostatic in nature (Haspel et al., 2008). Kinetic experiments in salt concentrations of 75–600 mM indicate the sensitivity of association/dissociation phases of wild type and various mutants to ionicity (Haspel et al., 2008). This suggests overall electrostatic contribution to be of importance in the initial complex formation and further in stabilizing the complex under the prevailing conditions. Our PRS analysis of Types I and II is further improved by these observations such that under various environmental perturbations, disruption of long-range and short-range electrostatic complementarity seem to impair stability and affect complex formation of binding partners.

Remarkably, electrostatic effects provide an excellent description for the observed pattern in both Type I and Type II complexes. Electrostatic interactions are the primary factors of pH dependent processes in biochemical reactions. Particularly, we find that among the 16 protein complexes

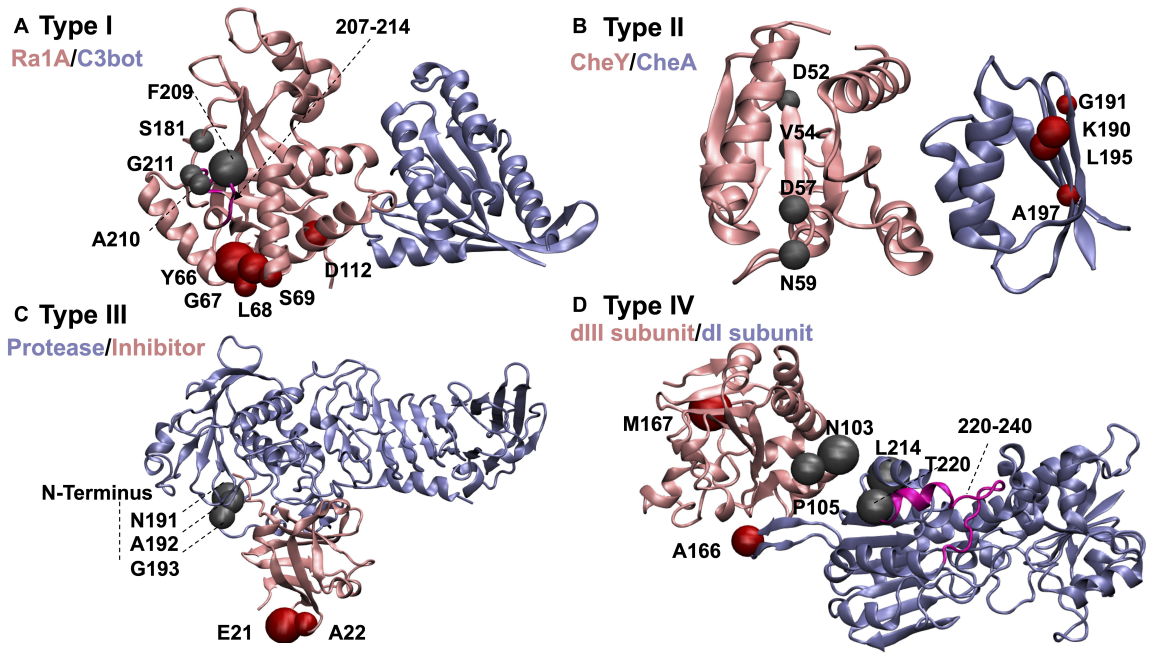

FIGURE 3 | Example complexes demonstrating dissociation scenarios discussed in the text. The color of each subunit is associated with its name. Residues controlling dissociation of the pink colored proteins are displayed as red beads, and those for ice blue ones as gray beads: **(A)** C3bot-RalA complex as an example of Type I proteins. PRS selected residues that control the dissociation are confined to RalA compartment. Substrate recognition site of C3bot (residues 207–214) is shown in magenta. **(B)** Chemotaxis CheA:CheY complex as an example of Type II proteins. **(C)** Alkaline protease and its cognate inhibitor as an example of Type III proteins. **(D)** dI_2dIII_1 complex of proton-translocating transhydrogenase as an example of Type IV proteins. Loop residues that become less mobile upon ligand binding due to surface closing of the protein are displayed in magenta.

included into Type I and II groups, 13 of them belong to enzymes. Enzymatic activities are known to be pH dependent and protonation state of catalytic and active site residues are effective in potential distribution of the binding region. Consequently, charge distribution of these regions will modulate the interactions between the proteins and the reaction products. It has been reported that, enzymes make use of their preoriented environment to stabilize the transition state and the reduction in catalytic energy is accomplished by electrostatic stabilization of the active site of the enzyme (Warshel, 1998; Warshel et al., 2006).

In the same vein, we propose that the electrostatic characteristics of residues top-rated by PRS might be found essential to specificity and ligand binding properties in enzymatic reactions. For example, in the Type I complex C3bot-RalA (Ras-related protein Ral-A:Mono-ADP-ribosyltransferase C3, pdb code:2A9K; Pautsch et al., 2005), displayed in **Figure 3A**, residues 207–214 of C3bot (displayed in magenta in **Figure 3A**) are part of the substrate recognition site and important in catalytic activity. PRS selects residues 209–211 as candidate positions to provoke the dissociation of the complex. Another key region involved in dissociation upon force application is residue 112 of C3bot, which also contributes to complex formation. In addition, residue 181 is located on a loop responsible in substrate recognition. Accordingly, perturbation of specific sites found by PRS analysis and experimentally verified to be functionally important might reorient the enzyme/substrate dipoles that organize the catalysis and destabilize the charged transition state (Thomas et al., 1985). These changes might thus prevent the binding event or promote dissociation after the chemical reaction has terminated.

We also find that the effect of the local charge distribution on enzyme functions is not limited to the active site and that the remote locations on the protein are effectively involved in the dissociation process. Mutations of charged patches remote from either the protein or ligand binding site might alter the binding kinetic rates, shift pK_as and lead to weak molecular recognition (Thomas et al., 1985). In response to a particular perturbation, exposure to a different environment, reorganization of charged atomic groups and dielectric relaxation of the protein affects the electrostatic potential distribution of the interface or active site region considerably, creating the steering forces that guide the dissociation reaction. Thus, a local change of interactions at a remote site leads to a global structural change that modifies the organization of the interface contact network and leads to dissociation of the two proteins. The contribution of this distal perturbation on enzyme/inhibitor activity basically may be viewed as leading to a cooperative conformational transition. In both Type I and Type II, the information transmission between remote functional sites on one protein and the entire structure of the interacting partner naturally occurs via the interface linkage. In such proteins, the network of contacts in the interface could form a so-called "conductive" region so that the signal from

one protein is transferred through the interface to control the functionality of the second binding protein.

As one of the test beds, we have studied chemotaxis CheA:CheY complex (pdb code:1FFW; Gouet et al., 2001). PRS finds residues 52, 54, 57, and 59 of CheY to play a role in selecting the unbound over bound conformation in the presence of an external force (**Figure 3B**). This complex is an example of remote communication between a two-component signal-transducer. pH dependent catalytic activity has also been observed in chemotaxis CheA:CheY, in which ligand binding on CheA is conducted to CheY. Phosphorylation reaction on H48 of CheA subsequently transfers to D57 of CheY and the complex dissociates (Silversmith et al., 1997). CheY itself is incapable of providing an acidic residue during the phosphorylation event and the complex formation with CheA results in a conformational change on CheY as an acidic residue (D57) donates a proton to a phosphodonor in an optimal orientation and the protein–protein phosphotransfer occurs (Silversmith et al., 1997). The pH dependence of the phosphotransfer kinetics in the pH range of 7.5–10, studied through two mutants of CheY active site residues shows simply a moderate decrease in rate constants compared to the wild type CheY. This observation suggests that conserved active site residues do not have an essential and direct role in catalysis. Thus, the loss of activity throughout this range for phosphotransfer to wild type CheY is attributed not to deprotonation events on CheY; rather it is likely due to deprotonations in CheA (Silversmith et al., 1997). Variability of position 59 of CheY as a non-conserved and indirect active site residue in modulating the autophosphorylation of CheY with small molecule phosphodonors shows no detectable binding between the phosphodonor and CheY, validating the significant impact of position 59 on autophosphorylation kinetics (Thomas et al., 2013). Mutation of N59 to R, K, M, L, A, D, and E results in both increase and decrease in autophosphorylation rate constants. Substitution with positively charged residues increases the kinetic rates while substitution with negatively charged residues decreases the rate, implying how local electrostatic interactions at position 59 modulate the CheY autophosphorylation (Thomas et al., 2013).

Auto-Controlled Conformational Transitions Are Correlated to Mechanical Organization of the Protein in Type III Complexes

In Type III group of protein complexes, dissociations of either constituent of the complexes are governed by a local perturbation on the respective protein. We label this behavior as auto-controlled dissociation. In this category, all the residues involved in global transformation of bound to unbound forms are located out of the interface region, trypsinogen and its Bowman-Birk proteinase inhibitor precursor (pdb code:1D6R; Koepke et al., 2000) being an exception. We observe that in Type III complexes, the global RMSD between bound and unbound forms of each protein varies between 0.6 and 1.4 Å. However, regions of high mobility exist in which the local RMSD may take values as high as 7.3 Å (**Table 1**). Such regions belong to unstructured surface exposed loops where they display large deviations upon complex formation. The amino acids identified by PRS in this group belong to these regions and very specific locations may be perturbed so that the local fluctuations of the amino acids choose the conformational switch to the target structure. Thus, mechanical motions of the loops produce essential conformational transitions under such point perturbations.

Contrary to Type I and II complexes, the electrostatic potential distribution along the protein surfaces of Type III reveals that each protein maintains the same electrostatic potential distribution state in their respective bound and unbound forms. This means that for any protein complex in this group, orientation of the negatively and positively charged surfaces of the constituents has the same distribution in their respective free forms. Thus, the analysis identifies an "insulating" interface area that prevents the allosteric communication between interacting proteins and each protein functions independently under various perturbations. We note that while there are also stretches with large RMSDs in Type II protein complexes, those regions cross-control the dissociation of the binding proteins remotely and their own shape changes are not mechanically controlled.

As an example case, for the alkaline protease and its cognate inhibitor (pdb code: 1JIW; Hege et al., 2001) classified as Type III, PRS finds residues 191–193 of protease which are in direct contact with the N-terminus of the inhibitor (**Figure 3C**). The latter has been shown to coordinate the catalytic zinc anion (Hege et al., 2001). Such an interaction adds to the structural stability and leads to a low dissociation constant. Upon any kind of perturbation applied to residues 191–193 of the protease, disruption of the interactions in this region would change the extended conformation of the N-terminus and modify the zinc coordination and thus might facilitate the dissociation process.

Combined Perturbations on Both Partners Is Essential for Dissociation in Type IV Complexes

In Type IV group of protein complexes, dissociation of the two constituents might be triggered by point perturbations on either subunit. This contrasts Type III complexes where perturbation of any subunit will mediate the disruption of interactions in the protein pair, facilitating dissociation. Different functional subdomains may exist on these proteins which contribute to binding to diverse set of ligands in their functional pathway and promote dissociation. Additionally, complexes in our test set involving transmembrane proteins also fall into this group. Presence of each fragment of the protein pair in different compartments of the cell environment assist their exposure to different perturbation scenarios and support the idea of simultaneous perturbation of both partners of the protein complex in the dissociation process. As an example, we focus on dI_2dIII_1 complex of proton-translocating transhydrogenase (pdb code:2OOR; Bhakta et al., 2007). The complex is found in the membrane compartment of the bacteria or animal cells. Proton transfer across the membranes is facilitated by conformational changes of transhydrogenase. PRS identifies residues A166, L214 and T220 of dI subunit and residues N103, P105 and M167

TABLE 5 | Summary of types of complexes and properties of their dissociation mechanisms.

Complex Type (# of complexes out of total 25)	Control	Range	Electrostatic change upon binding
I (11)	Centralized on one protein; residues on one protein controls conformational changes on both proteins	Long-range	Yes
II (5)	Cross-control; residues on one protein controls the conformational change in the other partner	Long-range	Yes
III (4)	Auto-control; residues on the same protein control their own conformational change	Local	No
IV (5)	Scattered control; residues on both proteins control the conformational changes	Local	No

of dIII subunit to be effective in dissociation of the complex (**Figure 3D**). dI and dIII protrude from the membrane while a third compartment dII spans the membrane. Thus, each part of the protein is exposed to a different part of the cell, making the protein susceptible to alternative perturbation scenarios. Residue T220 of dI is part of a loop (residues 220–240; **Figure 3D** magenta) that becomes less mobile upon ligand binding due to surface closing of the protein (Bhakta et al., 2007). Furthermore, M167 of dIII is in the neighboring site of H_2NADH and any perturbation on this site might alter the proton pump reactions due to changes made to interaction network of ligand binding region. Accordingly, any perturbation on H_2NADH binding region in dIII subdomain may alter the structural features of dI subdomain through remote communication.

CONCLUSION

There is plethora of work addressing association of proteins partaking in complexes, and the consensus is to focus on the interface to determine the major features of binding events (Gainza et al., 2020), concentrating on, e.g., pockets formed upon complexation (Li et al., 2004), prediction of binding energies based on the interface (Moreira et al., 2007), close-range electrostatic interactions (Kumar and Nussinov, 2002), and conserved residues along the interface (Kumar and Nussinov, 2002; Li et al., 2004). However, to alter protein functions, e.g., for therapeutic applications, it is also essential to understand the mechanisms affecting their dissociation, a question that has not been thoroughly explored, to the best of our knowledge. In this work, we have studied the characteristics of residues responsible for the dissociation of a set of 25 non-redundant protein complexes, using PRS as the predictor of residues whose perturbation encourages the unbound conformations. Significance of the residues identified by PRS are discussed in detail for four sample cases (**Figure 3**).

In a statistical analysis of the PRS identified residues, we find that in terms of residue types, the only significant enhancement is in glycine residues, up from 8% of all residues found in the protein set to 13% in the subset of residues implicated in protein dissociation. This is in contrast to the studies reporting on hotspots on the interaction surface, whereby tryptophan (21%), arginine (13.3%), and tyrosine (12.3%) have the highest probabilities of occurrence (Moreira et al., 2007). Moreover, PRS identifies residues labeled as controlling dissociation are also significantly enhanced on loops, and are predominantly located on the complex surface, remote from the interface. This finding is plausible, since in contrast to an association event whereby interface compatibility is a major determinant, exposed residues are expected to partake in disrupting the complex.

We find that dissociation events disclosed by PRS analysis may be classified into four main groups as summarized in **Table 5**. The nature of the events leading to dissociation are either expected to be due to mechanical perturbations arriving at certain locations on the surface, or due to environmental triggers that interfere with the electrostatic potential distribution of the complex. In the latter case, a signature event is in the change of the electrostatic potential distributions of one of the binding partners (see Type I in **Tables 1**, **2** and **Figure 1**) or both of them (Type II complexes; **Figure 2**). The physics of these observations are resolved by an electro-mechanical coupled ENM proposed by our group (Sensoy et al., 2017). Accordingly, we find that even in the absence of an external force, positional displacements may still be obtained, provided there is electromechanical coupling. The role of PRS is to identify residues where such equivalent forces are focused on and are relevant to the observed conformational change. In the absence of coupling, the conformational change may indeed be triggered by an external force, e.g., upon binding of a ligand. Therefore, PRS either identifies the residues that facilitate displacements by mechanical perturbations (Type III and IV), or those which are mechanical mimics to the effects expedited by changes in electric displacement (Type I and II). In the latter situation, a coupling term that links electrical drive to mechanical response survives.

This study is a step toward developing descriptors aimed at disrupting protein complexes with the aim of developing therapeutic approaches to alter the function of proteins working in tandem. In particular, targeting remote sites to destabilize interacting proteins using unique approaches will aid in the emerging field of allosteric drug design (Guarnera and Berezovsky, 2020).

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

HA designed the research, conducted the computer experiments, analyzed the results, interpreted the data, wrote the manuscript, and constructed the figures. FJ wrote the manuscript, interpreted the data, and constructed the figures. AA and CA designed the research, guided the computer experiments and analyses, interpreted the data, guided the structure and contents of the manuscript, and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Scientific and Technological Research Council of Turkey (Grant Number 116F229).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.00210/full#supplementary-material

REFERENCES

Abdizadeh, H., Atilgan, A. R., and Atilgan, C. (2015a). Detailed molecular dynamics simulations of human transferrin provide insights into iron release dynamics at serum and endosomal pH. *JBIC J. Biol. Inorgan. Chem.* 20, 705–718. doi: 10.1007/s00775-015-1256-4

Abdizadeh, H., Guven, G., Atilgan, A. R., and Atilgan, C. (2015b). Perturbation response scanning specifies key regions in subtilisin serine protease for both function and stability. *J. Enzyme Inhibit. Med. Chem.* 30, 867–873. doi: 10.3109/14756366.2014.979345

Abdizadeh, H., and Atilgan, C. (2016). Predicting long term cooperativity and specific modulators of receptor interactions in human transferrin from dynamics within a single microstate. *Phys. Chem. Chem. Phys.* 18, 7916–7926. doi: 10.1039/c5cp05107j

Adcock, S. A., and McCammon, J. A. (2006). Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* 106, 1589–1615. doi: 10.1002/chin.200630297

Ahmed, A., and Gohlke, H. (2006). Multiscale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory. *Proteins Struct. Funct. Bioinform.* 63, 1038–1051. doi: 10.1002/prot.20907

Atilgan, A. R., Aykut, A. O., and Atilgan, C. (2011). Subtle p H differences trigger single residue motions for moderating conformations of calmodulin. *J. Chem. Phys.* 135:10B613. doi: 10.1063/1.3651807

Atilgan, A. R., Durell, S., Jernigan, R. L., Demirel, M. C., Keskin, O., and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80, 505–515. doi: 10.1016/s0006-3495(01)76033-x

Atilgan, C. (2018). Computational methods for efficient sampling of protein landscapes and disclosing allosteric regions. *Adv. Protein Chem. Struct. Biol.* 113, 33–63. doi: 10.1016/bs.apcsb.2018.06.001

Atilgan, C., and Atilgan, A. R. (2009). Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. *PLoS Comput. Biol.* 5:e000544. doi: 10.1371/journal.pcbi.1000544

Atilgan, C., Gerek, Z., Ozkan, S., and Atilgan, A. (2010a). Manipulation of conformational change in proteins by single-residue perturbations. *Biophys. J.* 99, 933–943. doi: 10.1016/j.bpj.2010.05.020

Atilgan, C., Okan, O. B., and Atilgan, A. R. (2010b). How orientational order governs collectivity of folded proteins. *Proteins Struct. Funct. Bioinform.* 78, 3363–3375. doi: 10.1002/prot.22843

Atilgan, C., Okan, O. B., and Atilgan, A. R. (2012). Network-based models as tools hinting at nonevident protein functionality. *Annu. Rev. Biophys.* 41, 205–225. doi: 10.1146/annurev-biophys-050511-102305

Aykut, A. O., Atilgan, A. R., and Atilgan, C. (2013). Designing molecular dynamics simulations to shift populations of the conformational states of calmodulin. *PLoS Comput. Biol.* 9:1003366. doi: 10.1371/journal.pcbi.1003366

Bahar, I., Atilgan, A. R., and Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold. Design* 2, 173–181. doi: 10.1016/s1359-0278(97)00024-2

Bahar, I., Lezon, T. R., Yang, L.-W., and Eyal, E. (2010). Global dynamics of proteins: bridging between structure and function. *Annu. Rev. Biophys.* 39, 23–42. doi: 10.1146/annurev.biophys.093008.131258

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10037–10041. doi: 10.1073/pnas.181342398

Baysal, C., and Atilgan, A. R. (2001a). Coordination topology and stability for the native and binding conformers of chymotrypsin inhibitor 2. *Proteins Struct. Funct. Bioinform.* 45, 62–70. doi: 10.1002/prot.1124

Baysal, C., and Atilgan, A. R. (2001b). Elucidating the structural mechanisms for biological activity of the chemokine family. *Proteins Struct. Funct. Bioinform.* 43, 150–160. doi: 10.1002/1097-0134(20010501)43:2<150::aid-prot1027>3.0.co;2-m

Ben-Avraham, D. (1993). Vibrational normal-mode spectrum of globular proteins. *Phys. Rev. B* 47:14559. doi: 10.1103/physrevb.47.14559

Bhakta, T., Whitehead, S. J., Snaith, J. S., Dafforn, T. R., Wilkie, J., Rajesh, S., et al. (2007). Structures of the dI2dIII1 complex of proton-translocating transhydrogenase with bound, inactive analogues of NADH and NADPH reveal active site geometries. *Biochemistry* 46, 3304–3318. doi: 10.1021/bi061843r.s001

Bolia, A., Gerek, Z. N., and Ozkan, S. B. (2014). BP-Dock: a flexible docking scheme for exploring protein–ligand interactions based on unbound structures. *J. Chem. Inform. Model.* 54, 913–925. doi: 10.1021/ci4004927

Bradshaw, R. T., Patel, B. H., Tate, E. W., Leatherbarrow, R. J., and Gould, I. R. (2011). Comparing experimental and computational alanine scanning techniques for probing a prototypical protein–protein interaction. *Protein Eng. Design Select.* 24, 197–207. doi: 10.1093/protein/gzq047

Carrion-Vazquez, M., Li, H., Lu, H., Marszalek, P. E., Oberhauser, A. F., and Fernandez, J. M. (2003). The mechanical stability of ubiquitin is linkage dependent. *Nat. Struct. Mol. Biol.* 10, 738–743. doi: 10.1038/nsb965

Chen, R., and Weng, Z. (2002). Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins Struct. Funct. Bioinform.* 47, 281–294. doi: 10.1002/prot.10092

Chennubhotla, C., Rader, A., Yang, L.-W., and Bahar, I. (2005). Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. *Phys. Biol.* 2:S173. doi: 10.1088/1478-3975/2/4/s12

De Las Rivas, J., and Fontanillo, C. (2010). Protein–protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Comput. Biol.* 6:807. doi: 10.1371/journal.pcbi.1000807

Delarue, M., and Dumas, P. (2004). On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6957–6962. doi: 10.1073/pnas.0400301101

Dietz, H., Berkemeier, F., Bertz, M., and Rief, M. (2006). Anisotropic deformation response of single protein molecules. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12724–12728. doi: 10.1073/pnas.0602995103

Doruker, P., Jernigan, R. L., and Bahar, I. (2002). Dynamics of large proteins through hierarchical levels of coarse-grained structures. *J. Comput. Chem.* 23, 119–127. doi: 10.1002/jcc.1160

Durand, P., Trinquier, G., and Sanejouand, Y. H. (1994). A new approach for determining low-frequency normal modes in macromolecules. *Biopolym. Orig. Res. Biomol.* 34, 759–771. doi: 10.1002/bip.360340608

Filipek, R., Rzychon, M., Oleksy, A., Gruca, M., Dubin, A., Potempa, J., et al. (2003). The Staphostatin-Staphopain complex a forward binding inhibitor in complex with its target cysteine protease. *J. Biol. Chem.* 278, 40959–40966. doi: 10.2210/pdb1pxv/pdb

Gainza, P., Sverrisson, F., Monti, F., Rodola, E., Boscaini, D., Bronstein, M., et al. (2020). Deciphering interaction fingerprints from protein molecular surfaces

using geometric deep learning. *Nat. Methods* 17, 184–192. doi: 10.1038/s41592-019-0666-6

Gao, Y., Wang, R., and Lai, L. (2004). Structure-based method for analyzing protein–protein interfaces. *J. Mol. Model.* 10, 44–54. doi: 10.1007/s00894-003-0168-3

Gouet, P., Chinardet, N., Welch, M., Guillet, V., Cabantous, S., Birck, C., et al. (2001). Further insights into the mechanism of function of the response regulator CheY from crystallographic studies of the CheY–CheA124–257 complex. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 57, 44–51. doi: 10.1107/s090744490001492x

Guarnera, E., and Berezovsky, I. N. (2020). Allosteric drugs and mutations: chances, challenges, and necessity. *Curr. Opin. Struct. Biol.* 62, 149–157. doi: 10.1016/j.sbi.2020.01.010

Guven, G., Atilgan, A. R., and Atilgan, C. (2014). Protonation states of remote residues affect binding–release dynamics of the ligand but not the conformation of apo ferric binding protein. *J. Phys. Chem. B* 118, 11677–11687. doi: 10.1021/jp5079218

Han, R., Leo-Macias, A., Zerbino, D., Bastolla, U., Contreras-Moreira, B., and Ortiz, A. R. (2008). An efficient conformational sampling method for homology modeling. *Proteins Struct. Funct. Bioinform.* 71, 175–188. doi: 10.1002/prot.21672

Haspel, N., Ricklin, D., Geisbrecht, B. V., Kavraki, L. E., and Lambris, J. D. (2008). Electrostatic contributions drive the interaction between Staphylococcus aureus protein Efb-C and its complement target C3d. *Protein Sci.* 17, 1894–1906. doi: 10.1110/ps.036624.108

Hege, T., Feltzer, R. E., Gray, R. D., and Baumann, U. (2001). Crystal structure of a complex between *Pseudomonas aeruginosa* alkaline protease and its cognate inhibitor inhibition by a zinc–NH2 coordinative boND. *J. Biol. Chem.* 276, 35087–35092. doi: 10.1074/jbc.m104020200

Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. *Proteins Struct. Funct. Bioinform.* 33, 417–429. doi: 10.1002/(sici)1097-0134(19981115)33:3<417::aid-prot10<3.0.co;2-8

Hinsen, K., Reuter, N., Navaza, J., Stokes, D. L., and Lacapère, J.-J. (2005). Normal mode-based fitting of atomic structure into electron density maps: application to sarcoplasmic reticulum Ca-ATPase. *Biophys. J.* 88, 818–827. doi: 10.1529/biophysj.104.050716

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38. doi: 10.1016/0263-7855(96)00018-5

Ibarra, A. A., Bartlett, G. J., Hegedüs, Z. F., Dutt, S., Hobor, F., Horner, K. A., et al. (2019). Predicting and experimentally validating hot-spot residues at protein–protein interfaces. *ACS Chem. Biol.* 14, 2252–2263. doi: 10.1021/acschembio.9b00560.s001

Ikeguchi, M., Ueno, J., Sato, M., and Kidera, A. (2005). Protein structural change upon ligand binding: linear response theory. *Phys. Rev. Lett.* 94:078102. doi: 10.1103/physrevlett.94.078102

Jacobs, D. J., Livesay, D. R., Mottonen, J. M., Vorov, O. K., Istomin, A. Y., and Verma, D. (2012). Ensemble properties of network rigidity reveal allosteric mechanisms. *Methods Mol. Biol.* 796, 279–304. doi: 10.1007/978-1-61779-334-9_15

Jalalypour, F., Sensoy, O., and Atilgan, C. (2020). Perturb-Scan-Pull: a novel method facilitating conformational transitions in proteins. *J. Chem. Theory Comput.* 16, 3825–3841. doi: 10.1021/acs.jctc.9b01222

Kenneth Morrow, J., and Zhang, S. (2012). Computational prediction of protein hot spot residues. *Curr. Pharm. Design* 18, 1255–1265. doi: 10.2174/138161212799436412

Keskin, O., Bahar, I., Flatow, D., Covell, D., and Jernigan, R. (2002). Molecular mechanisms of chaperonin GroEL- GroES function. *Biochemistry* 41, 491–501. doi: 10.1021/bi011393x

Kim, M. K., Jernigan, R. L., and Chirikjian, G. S. (2002). Efficient generation of feasible pathways for protein conformational transitions. *Biophys. J.* 83, 1620–1630. doi: 10.1016/s0006-3495(02)73931-3

Koepke, J., Ermler, U., Warkentin, E., Wenzl, G., and Flecker, P. (2000). Crystal structure of cancer chemopreventive Bowman-Birk inhibitor in ternary complex with bovine trypsin at 2.3 Å resolution. Structural basis of Janus-faced serine protease inhibitor specificity. *J. Mol. Biol.* 298, 477–491. doi: 10.2210/pdb1d6r/pdb

Kong, Y., Ming, D., Wu, Y., Stoops, J. K., Zhou, Z. H., and Ma, J. (2003). Conformational flexibility of pyruvate dehydrogenase complexes: a computational analysis by quantized elastic deformational model. *J. Mol. Biol.* 330, 129–135. doi: 10.1016/s0022-2836(03)00555-2

Krebs, W. G., Alexandrov, V., Wilson, C. A., Echols, N., Yu, H., and Gerstein, M. (2002). Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic. *Proteins Struct. Funct. Bioinform.* 48, 682–695. doi: 10.1002/prot.10168

Kumar, S., and Nussinov, R. (2002). Close-range electrostatic interactions in proteins. *Chembiochem* 3, 604–617.

Kurkcuoglu, O., Turgut, O. T., Cansu, S., Jernigan, R. L., and Doruker, P. (2009). Focused functional dynamics of supramolecules by use of a mixed-resolution elastic network model. *Biophys. J.* 97, 1178–1187. doi: 10.1016/j.bpj.2009.06.009

Leo-Macias, A., Lopez-Romero, P., Lupyan, D., Zerbino, D., and Ortiz, A. R. (2005). An analysis of core deformations in protein superfamilies. *Biophys. J.* 88, 1291–1299. doi: 10.1529/biophysj.104.052449

Li, X., Keskin, O., Ma, B., Nussinov, R., and Liang, J. (2004). Protein–protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *J. Mol. Biol.* 344, 781–795. doi: 10.1016/j.jmb.2004.09.051

Lichtarge, O., and Sowa, M. E. (2002). Evolutionary predictions of binding surfaces and interactions. *Curr. Opin. Struct. Biol.* 12, 21–27. doi: 10.1016/s0959-440x(02)00284-1

Liu, Q., Chen, P., Wang, B., Zhang, J., and Li, J. (2018). Hot spot prediction in protein-protein interactions by an ensemble system. *BMC Syst. Biol.* 12:132. doi: 10.1186/s12918-018-0665-8

Liu, S., Liu, C., and Deng, L. (2018). Machine learning approaches for protein–protein interaction hot spot prediction: progress and comparative assessment. *Molecules* 23:2535. doi: 10.3390/molecules23102535

Ming, D., and Wall, M. E. (2005). Quantifying allosteric effects in proteins. *Proteins Struct. Funct. Bioinform.* 59, 697–707. doi: 10.1002/prot.20440

Miyashita, O., Onuchic, J. N., and Wolynes, P. G. (2003). Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. U.S.A.* 100, 12570–12575. doi: 10.3410/f.1016237.200300

Moreira, I. S., Fernandes, P. A., and Ramos, M. J. (2007). Hot spots—A review of the protein–protein interface determinant amino-acid residues. *Proteins Struct. Funct. Bioinform.* 68, 803–812. doi: 10.1002/prot.21396

Mottonen, J. M., Jacobs, D. J., and Livesay, D. R. (2010). Allosteric response is both conserved and variable across three CheY orthologs. *Biophys. J.* 99, 2245–2254. doi: 10.1016/j.bpj.2010.07.043

Ni, D., Liu, N., and Sheng, C. (2019). Allosteric modulators of protein–protein interactions (PPIs). *Protein Allostery Drug Discov.* 1163, 313–334. doi: 10.1007/978-981-13-8719-7_13

Orellana, L., Gustavsson, J., Bergh, C., Yoluk, O., and Lindahl, E. (2019). eBDIMS server: protein transition pathways with ensemble analysis in 2D-motion spaces. *Bioinformatics* 35, 3505–3507. doi: 10.1093/bioinformatics/btz104

Orozco, M. (2014). A theoretical view of protein dynamics. *Chem. Soc. Rev.* 43, 5051–5066. doi: 10.1002/chin.201437300

Ozdemir, E. S., Halakou, F., Nussinov, R., Gursoy, A., and Keskin, O. (2019). Methods for discovering and targeting druggable protein-protein interfaces and their application to repurposing. *Computat. Methods Drug Repurp.* 1903, 1–21. doi: 10.1007/978-1-4939-8955-3_1

Pautsch, A., Vogelsgesang, M., Tränkle, J., Herrmann, C., and Aktories, K. (2005). Crystal structure of the C3bot–RalA complex reveals a novel type of action of a bacterial exoenzyme. *EMBO J.* 24, 3670–3680. doi: 10.2210/pdb2a78/pdb

Pereira, P. J. B., Lozanov, V., Patthy, A., Huber, R., Bode, W., Pongor, S., et al. (1999). Specific inhibition of insect α-amylases: yellow meal worm α-amylase in complex with the Amaranth α-amylase inhibitor at 2.0 *Å resolution*. *Structure* 7, 1079–1088. doi: 10.1016/s0969-2126(99)80175-0

Qiao, Y., Xiong, Y., Gao, H., Zhu, X., and Chen, P. (2018). Protein-protein interface hot spots prediction based on a hybrid feature selection strategy. *BMC Bioinform.* 19:14. doi: 10.1186/s12859-018-2009-5

Ross, C., Nizami, B., Glenister, M., Sheik Amamuddy, O., Atilgan, A. R., Atilgan, C., et al. (2018). MODE-TASK: large-scale protein motion tools. *Bioinformatics* 34, 3759–3763. doi: 10.1093/bioinformatics/bty427

Ross, C. J., Atilgan, A. R., Bishop, ÖT., and Atilgan, C. (2018). Unraveling the motions behind enterovirus 71 uncoating. *Biophys. J.* 114, 822–838. doi: 10.1016/j.bpj.2017.12.021

Sensoy, O., Atilgan, A. R., and Atilgan, C. (2017). FbpA iron storage and release are governed by periplasmic microenvironments. *Phys. Chem. Chem. Phys.* 19, 6064–6075. doi: 10.1039/c6cp06961d

Shi, T.-L., Li, Y.-X., Cai, Y.-D., and Chou, K.-C. (2005). Computational methods for protein-protein interaction and their application. *Curr. Protein Peptide Sci.* 6, 443–449. doi: 10.5772/36716

Silversmith, R. E., Appleby, J. L., and Bourret, R. B. (1997). Catalytic mechanism of phosphorylation and dephosphorylation of CheY: kinetic characterization of imidazole phosphates as phosphodonors and the role of acid catalysis. *Biochemistry* 36, 14965–14974. doi: 10.1021/bi9715573

Siragusa, L., Cross, S., Baroni, M., Goracci, L., and Cruciani, G. (2015). BioGPS: navigating biological space to predict polypharmacology, off-targeting, and selectivity. *Proteins Struct. Funct. Bioinform.* 83, 517–532. doi: 10.1002/prot.24753

Tama, F., and Sanejouand, Y.-H. (2001). Conformational change of proteins arising from normal mode calculations. *Protein Eng.* 14, 1–6. doi: 10.1093/protein/14.1.1

Tan, K. P., Nguyen, T. B., Patel, S., Varadarajan, R., and Madhusudhan, M. S. (2013). Depth: a web server to compute depth, cavity sizes, detect potential small-molecule ligand-binding cavities and predict the pKa of ionizable residues in proteins. *Nucleic Acids Res.* 41, W314–W321. doi: 10.1093/nar/gkt503

Thomas, P. G., Russell, A. J., and Fersht, A. R. (1985). Tailoring the p H dependence of enzyme catalysis using protein engineering. *Nature* 318, 375–376. doi: 10.1038/318375a0

Thomas, S. A., Immormino, R. M., Bourret, R. B., and Silversmith, R. E. (2013). Nonconserved active site residues modulate CheY autophosphorylation kinetics and phosphodonor preference. *Biochemistry* 52, 2262–2273. doi: 10.1021/bi301654m

Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905. doi: 10.1103/physrevlett.77.1905

Tozzini, V. (2010). Multiscale modeling of proteins. *Acc. Chem. Res.* 43, 220–230. doi: 10.1021/ar9001476

Warshel, A. (1998). Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J. Biol. Chem.* 273, 27035–27038. doi: 10.1074/jbc.273.42.27035

Warshel, A., Sharma, P. K., Kato, M., Xiang, Y., Liu, H., and Olsson, M. H. (2006). Electrostatic basis for enzyme catalysis. *Chem. Rev.* 106, 3210–3235. doi: 10.1007/978-94-009-5574-5_2

Whitford, P. C., Miyashita, O., Levy, Y., and Onuchic, J. N. (2007). Conformational transitions of adenylate kinase: switching by cracking. *J. Mol. Biol.* 366, 1661–1671. doi: 10.1016/j.jmb.2006.11.085

Xie, L., Li, J., Xie, L., and Bourne, P. E. (2009). Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. *PLoS Comput. Biol.* 5:387. doi: 10.1371/journal.pcbi.1000387

Yilmaz, L. S., and Atilgan, A. R. (2000). Identifying the adaptive mechanism in globular proteins: fluctuations in densely packed regions manipulate flexible parts. *J. Chem. Phys.* 113, 4454–4464. doi: 10.1063/1.1288512

Zheng, W., Brooks, B. R., and Hummer, G. (2007). Protein conformational transitions explored by mixed elastic network models. *Proteins Struct. Funct. Bioinform.* 69, 43–57. doi: 10.1002/prot.21465

Zheng, W., and Doniach, S. (2003). A comparative study of motor-protein motions by using a simple elastic-network model. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13253–13258. doi: 10.1073/pnas.2235686100

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.