Mechanical changes in proteins with large-scale motions highlight the formation of structural locks

Sophie Sacquin-Mora

Laboratoire de Biochimie Théorique, CNRS UPR9080, Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

E-mail: sacquin@ibpc.fr

Running title : Structural locks in protein closed forms

Keywords: protein domain motion; proteins mechanics; coarse-grain simulations; elastic network model.

A preliminary version of this work, [doi], was deposited in bioRxiv on november 17th 2017.

Abstract: Protein function depends just as much on flexibility as on structure, and in numerous cases, a protein's biological activity involves transitions that will impact both its conformation and its mechanical properties. Here, we use a coarse-grain approach to investigate the impact of structural changes on protein flexibility. More particularly, we focus our study on proteins presenting large-scale motions. We show how calculating directional force constants within residue pairs, and investigating their variation upon protein closure, can lead to the detection of a limited set of residues that form a *structural lock* in the protein's closed conformation. This lock, which is composed of residues whose side-chains are tightly interacting, highlights residues that are important for protein function by stabilizing the closed structure, and that cannot be detected using earlier tools like local rigidity profiles or distance variations maps, or alternative bioinformatics approaches, such as coevolution scores.

1. Introduction

Decades of protein studies have now made it clear that flexibility is just as important as structure for defining a protein's activity (1-4). Indeed, conformational diversity is an essential feature in a functional protein, wether it is involved in catalysis, signal transduction (5), recognition, or allostery (6). However, protein motions are difficult to observe directly using experimental approaches, as they can cover a wide range of timescales, from a few picoseconds to 10's of millisecond or more (7). From a theoretical perspective, Molecular Dynamics (MD) simulations now represent a serious alternative. Still, the use of all-atom MD remains a costly strategy to access rare events, such as slow, large-amplitude conformational transitions, and usually requires the development of enhanced sampling strategies (8), or access to remarkable computational power (9). Coarse-grain models, based on simplified protein representation and energy functions, are a second option to investigate structural changes that are not accessible to all-atom MD simulations (10). Elastic Network Models (ENM) combined with Normal Mode Analysis (NMA) in particular, have proved particularly useful in this regard, since a small number of low-frequency normal modes will provide information regarding large-amplitude conformational changes in a protein (11-18).

In that perspective, the ProPHet (for Probing Protein Heterogeneity) program, which combines a coarse-grain ENM with Brownian Dynamics (BD) simulations, was initially developed to investigate protein mechanical properties on the residue level (19). Later work on proteins undergoing functionally related structural transitions (such as enzymes) showed that the mechanical variations induced by these conformational changes target a limited number of specific residues occupying key locations for

enzymatic activity (20). In an previous, multi-scale, study on the conformational transition of guanylate kinase, we showed that the enzyme closure leads to the formation of a *structural lock* constituted of residues forming a dense set of interactions and stabilizing the protein's closed form (21). In this study, we used ProPHet to investigate the mechanical variations associated with structural changes for a set of 51 proteins, 11 of which exhibit large-scale motions (LSM) leading to root mean-square deviations (RMSDs) > 4 Å. Calculations of directional force constant variations within residue pairs show the formation of similar structural locks in the closed structures of 9 out of 11 proteins. Interestingly, neither residue specific force constants, nor distance variation maps permit to detect these functionally relevant residues. Additional calculations of coevolution scores for residue pairs also fail to predict the location of the structural lock within the tested proteins.

2. Material and Methods

Coarse-grain Simulation: Coarse-grain Brownian Dynamics (BD) simulations were performed using the ProPHet program (http://bioserv.rpbs.univ-paris-diderot.fr/services/ProPHet/) (22-24) on a set of 51 proteins listed in table SI-1 in the Supporting Material, for which the unbound/ligand-bound or open/closed structures are available in the Protein Data Bank (PDB). Note that 31 of these proteins are enzymes that were already investigated in more details in ref. (20). The root mean square deviations (RMSDs) between the unbound and ligand-bound (respectively the open and closed) structures ranges from 0.17 Å (for endo-xylanase) to 15.6 Å (for diphteria toxin), and 11 proteins are considered to exhibit large-scale conformational transitions, with RMSDs > 4 Å.

Diverging from most common CG models, where each residue is described by a single pseudo-atom (25), ProPHet uses a more detailed model enabling different residues to be distinguished, in which the amino acids are represented by one pseudo-atom located at the Cα position, and either one or two (for larger residues) pseudo-atoms replacing the side-chain (with the exception of Gly) (26). Interactions between the pseudo-atoms are treated according to the standard Elastic Network Model (ENM) (25), that is, pseudoatoms closer than the cutoff parameter, Rc = 9 Å, are joined by gaussian springs that all have identical spring constants of γ = 0.42 N m^{-1} (0.6 kcal mol^{-1} Å^{-2}). The springs are taken to be relaxed for the starting conformation of the protein, i. e. its crystallographic structure. Mechanical properties are obtained from 200,000 BD steps at 300 K. The simulations are analyzed in terms of the fluctuations of the mean distance between each pseudo-atom belonging to a given amino acid and the pseudo-atoms belonging to the remaining protein residues. The inverse of these fluctuations yields an effective force constant k_i that describes the ease of moving a pseudo-atom i with respect to the overall protein structure:

$$k_i = \frac{3k_B T}{\left\langle \left(d_i - \langle d_i \rangle \right)^2 \right\rangle}$$

where $\langle\rangle$ denotes an average taken over the whole simulation and $d_i = \langle d_{ij} \rangle_{j*}$ is the average distance from particle i to the other particles j in the protein (the sum over $j*$ implies the exclusion of the pseudo-atoms belonging to the same residue as i). The distances between the Cα pseudo-atom of residue k and the Cα pseudo-atoms of the adjacent residues k-1 and k+1 are excluded, since the corresponding distances are virtually constant. The force constant for each residue k in the protein is the average of the force constants for all its constituent pseudo-atoms i. We will use the term «rigidity profile»

to describe the ordered set of force constants for all the residues in a given protein. Note that, following earlier studies which showed how small ligands had little influence on calculated force constants (19, 23), we chose not to include ligands in the protein representations. This enables us to study the proteins intrinsic mechanical properties independently of the nature and position of any bound ligand.

In ProPHet, we can also use the fluctuations of the inter-residue distances d_{ij} to compute force constants similar to those obtained by Eyal and Bahar (27) with a one-point-per-residue ENM. The resulting directional force constants (DFCs) are simply calculated as

$$k_{ij} = \frac{3k_B T}{\left\langle \left(d_{ij} - \langle d_{ij} \rangle^2 \right)^2 \right\rangle}$$

and can be used to construct a complete map of the mechanical resistance of a protein in response to all possible pulling directions.

Calculation of coevolution scores: Coevolution scores for residue pairs were calculated using the CoeViz tool (28) on the Polyview web-server (http:// polyview.cchmc.org/) (29). Given a target protein sequence, a multiple sequence alignment (MSA) is generated using PSI-BLAST(30), and CoeViz computes coevolution scores from the MSA using three different covariance metrics : chi-square statistics, χ^2 (31), mutual information MI (32), and Pearson correlation (see equations 1-3 in ref. (28)). The tool also computes one conservation metric, which is the joint Shannon entropy (JE), see equation 4 in ref. (28).

3. Results and Discussion

The rigidity profiles and mechanical maps of the open/unbound and closed/ligand-bound structures of the proteins listed in table SI-1 were systematically compared in order to investigate the mechanical changes induced by conformational variations.

General considerations

There is a general positive correlation between the amplitude of the conformational variation in a protein and its average mechanical variation, as can be seen on Figure1a-b (with correlation coefficients of 0,79 and 0.67 for the average force constant variation $\langle \Delta k \rangle$ and the root mean square force constant variation RMSΔk, respectively). As shown in a previous work made specifically on proteins undergoing domain-domain motions (24), large hinge-bending-type conformational transitions involve the closure of one domain on the other, usually resulting in a larger buried surface in the interdomain interface and leading to a more rigid structure. This is for example the case of the T4 lysozyme mutant (see Figure 2a-c), where protein closure induced by hinge motion of the two domains leads to a large rigidity increase for catalytic reside Glu11, ligand binding residues Gly30 and Met102 (33-35), and channel lining residues Met106 and Trp138 (36). On the contrary, cavity lining residues Met6 and Cys97, which form hydrogen bonds with water molecules (37), see a decrease in their force constant upon proteins closure, thus suggesting a disruption of this interaction network. More generally, one must keep in mind that the force constants variations resulting from conformational change are highly heterogeneous along the protein sequence and cannot simply be predicted from the RMSD between the initial and final structures. As shown in earlier work (38), even conformational changes upon ligand binding with small

RMSDs (under 0.6 Å) can both lead to minor mechanical changes (see the case of ribonuclease MC1 in Figure 2d-f), or major mechanical variations for a limited set of specific residues, as can be seen in Figure 2g-i for ligand bound galactose mutarotase, where hinge residues Asn175, Ser177 and Glu307 undergo an important rigidity loss). In addition, a small average mechanical variation over the whole protein can also result from a mixed mechanical response to structural change, with some residues (usually located in the active site) becoming more rigid upon ligand binding, while others (frequently lying on hinges and domain interfaces) become more flexible (see the case of carbonic anhydrase in ref. (20)).

Another example of a mixed mechanical response can be seen for the N-terminal lobe human serum transferrin (HST) in Figure 3. HST binds ferric ions in the bloodstream and transports them to cells. Its N-lobe (residues 1-337 of the native protein) folds in two domains (N1, residues 1-93+247-315, and N2, residues 94-246) separated by a hinge region. Upon ligand biding, the protein presents a rigid body rotation of the N2 domain relative to the N1 domain leading to a RMSD of 6.7 Å, while the RMSD calculated over individual domain are below 0.6 Å (39). This closure motion leads to an important increase in the rigidity of a limited set of residues (see Figure 3b), that are all involved either in the iron binding site, Asp63, Tyr188, His249 and Lys296, or the carbonate anion binding site, Thr120 and Arg124 (see Figure 3c). Meanwhile, hinge residue Ala244, Val246 and Pro247, which lie on the domain-domain interface, become more flexible upon closure.

Interestingly, we can also observe a negative correlation (with a correlation coefficient of 0.67) between the proteins average rigidity and the RMSD between their open/

unbound and closed/ligand-bound structures (see Figure 1c). Proteins that are, on average, more flexible, are also more likely to undergo large conformational transitions. More specifically, if we define *soft* proteins as those with an average rigidity below 20 kcal $mol^{-1}Å^{-2}$, 8 out of 11 elements from the large-scale motion group can be considered as soft proteins, while only 3 out of the 40 remaining proteins in the dataset belong to the same category. A noticeable outlier is the trp-repressor from E. coli, a small protein that appears to be remarkably flexible (with $\langle k \rangle$ = 4.7 kcal mol^{-1} $Å^{-2}$) despite the small RMSD (1.70 Å) between its open and closed forms. However, after searching the PDB for additional conformations of this protein, we found an NMR structure of the trp-repressor complexed with DNA (PDB entry 1rcs). The RMSDs between the 15 conformers for this DNA-bound form and our reference open form of the trp-repressor are comprised between 5 and 5.6 Å (which corresponds to the green point in Figure 1c), so eventually the low average flexibility calculated for this protein accurately reflects its ability to undergo large conformational changes. In addition, this correlation between $\langle k \rangle$ and the RMSD means that one should be able to estimate the amplitude of the conformational change undergone by a protein based on its general mechanical properties and without having any prior information regarding its final state.

Proteins undergoing large conformational transitions

Variations in the mechanical map: For the 11 proteins presenting large (with RMSDs > 4 Å) conformational transitions between their open/unbound and closed/ligand bound states, we also investigated the proteins DFC matrices and the variation in these mechanical maps resulting from the structural change.

DFC matrices for the proteins in their open/unbound state highlight the connections between pairs of residues belonging to the same protein domains, which will present higher directional force constants than pairs of residues belonging to different domains. For example, in Figure 4a one can clearly see how the high DFC areas correspond to the N1 and N2 domains in human serum transferrin. After closure, we can observe a shift of the DFC distribution toward larger values (see the histograms in Figure 4c), with the disappearance of many residue pairs with low force constants (below 2 kcal mol^{-1} $Å^{-2}$). The evolution of the mechanical map shown in Figure 4b indicates that the increase in DFC values upon protein closure is restricted to a small number of inter-domain residue pairs. By selecting the residue pairs that present the largest relative increase (i. e, with at least 80% of the maximum relative increase observed for all pairs) in their DFC induced by conformational change, three structural groups are detected (see Figure 5). These residues come in close contact upon protein closure, which might explain the strong increase in the associated DFCs, with their side-chains tightly interacting, thus forming a *structural lock* that would help maintaining the protein in its closed conformation. Experimentally, the interactions between residue Val11 and Ser12 from the N1 domain, and Ser180 and Thr181 from the N2 domain involve hydrogen bonds through two buried, highly ordered water molecules (40). Interestingly, when considered on their own, the residues involved in this structural lock are not rigid in the initial open protein conformation (their positions along the sequence are shown as red bars on Figure 3a), neither do they become rigid upon closure. This means that they could not be detected using only simple rigidity profiles. Nevertheless, after closure they belong to rigid pairs and their detection thus necessitates the use of mechanical maps that will highlight their functional role in the protein. In addition, one can note that DFC maps are qualitatively

similar to the contact maps based on the inter-residue distances only (see for exemple the case of human serum transferrin in Figure SI-1a), with both matrices showing the protein domain composition along the sequence. DFC variations however, will highlight a much more limited set of residue pairs than simply looking at variation in the residue distances upon protein closure (cf. Figure SI-1b). Residues forming the protein structural lock are not only residues getting nearer in the protein closed form. They present specific mechanical properties, that are related to the protein dynamics, and cannot be identified using structural information only.

Similar structural locks are observed for all the proteins in our large conformational change set (see Figures SI-2-11) with the exceptions of adenylate kinase (AK) and GroEl, where the residues forming rigid pair upon protein closure seem to be more distributed over the protein structure. This came as a surprise to us, as our first observation of structural locks using ProPHet had been made on a relatively similar enzyme, guanylate kinase from *M. tuberculosis* (GK_{MT}), where residues from the LID and GMP binding domains form a tight lock after ligand binding (21). For AK, the rigid pairs in the closed form also involve residues from the LID and NMP binding domains, but these residues remain distant and do not form direct contacts upon closure (see Figure 6f). Interestingly, while the structural transition in GK_{MT} corresponds to a simple hinge-bending motion, that is essentially defined by the enzyme's first mode of motion (41), the conformational change in AK seems to be more complex, involving additional modes of motion and inducing a torsion of the LID domain over the NMP-BD (14, 42, 43). This might explain why rigid pairs residues no longer end up facing each other in the closed form of AK compared to the GK case.

The second protein with rigid pair residues remaining distant in the closed conformation is the GroEl monomer, where pairs undergoing a large increase in their directional force constants upon closure involve residues from the apical and equatorial domains (which correspond respectively to the upper and lower parts of the cartoon representations in Figure SI-9). Remarkably, the conformational transition of GroEL was also investigated using normal modes by Uyar et al. in Ref. (14), and turned out to be one of the most difficult case for predicting the protein closed conformation when starting from an open structure. Once again, these results suggest a complex conformational pathway between the two forms, which would result in the distribution of rigid residues pairs all over the protein structure.

Coevolution analysis of the residues: In our earlier work on guanylate kinase we noticed that the residues forming the structural lock seemed to occupy similar positions to the set of coevolved residues that were detected by Armenta-Medina et al.(44) using statistical coupling analysis (SCA) in their work on adenylate kinase. This qualitative agreement also appears with the structural lock residues detected for AK in the present work, which are distributed on the LID, NMP and CORE domains (see Figure 6e) and seem to overlap with the coevolved residues of Armenta-Medina et al. shown on Figure 1 in ref. (44). In addition, coevolution has also been used to determine alternative conformational states in proteins by Sfriso et al. (45). In order to get a more quantitative view of this phenomenon and to determine wether structural lock residues are in fact coevolved, we performed a systematic investigation of the residues co-evolution for 10 proteins presenting large conformational transitions in our dataset (the endonuclease homodimer was not included in this test) using the CoeViz tool. The resulting

coevolution (or conservation) scores for each residue pair were plotted as a function of the DFC variation for this pair upon closure. Figure 7 shows the resulting plots in the case of adenylate kinase. While coevolution and important DFC variations may target the same areas in the protein, there is no significative correlation between the two quantities independently of the metric used to compute the coevolution scores (χ^2 , MI or Pearson correlation). Moreover, the different coevolution scores also correlate badly with each other. Similar results were obtained for all the tested proteins, as can be seen in Figure SI-12.

4. Conclusion

We combine a coarse-grain elastic network protein representation and Brownian Dynamics simulations to investigate mechanical variations in proteins undergoing conformational changes, with a particular focus on proteins showing large scale motions with RMDSs between their open and closed forms larger than 4 Å. The variations in the residues force constants upon protein closure are highly heterogeneous along the sequence and difficult to predict. They are however tightly related to a protein's biological activity, and residues undergoing the larger mechanical variation during the conformational changes are usually important for function. Interestingly, a protein's average rigidity in its open form also provides a good estimate on wether it is likely to undergo large-scale motions or not.

For proteins presenting important conformational changes upon closure, we also calculated the directional force constants between all residue pairs. The resulting maps highlight how the proteins are structured in adjacent domains. Investigating the DFCs

variation upon protein closure leads to the identification of a subset of residues whose side-chains are tightly interacting in the closed structure, and forming what we call a *structural lock*. Therefore, the calculation of DFCs brings additional information regarding functionally relevant residues in proteins with large-scale motions. From a protein design perspective, the structural lock residues represent potential mutation targets to change the closed state stability in order to enhance or modulate protein function. On can also note that these pairs could not be identified using alternative methods such as rigidity profiles, since the individual residues do not correspond to rigid spots in the protein, and distance variation matrices will highlight a much larger subset of residue pairs than the mechanical approach. Additional calculations of coevolution and conservation scores also failed to predict the specific residues that will form a structural lock within the protein closed conformation.

Acknowledgments: This work was supported by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01).

Supporting Material: A table listing all proteins studied in this work and 12 figures are available as supporting information.

Figure captions :

Figure 1. (a) Average force constant variation (in kcal mol^{-1} $Å^{-2}$) upon closure/ligand binding as a function of the RMSD for the 51 proteins in the dataset. (b) Root mean square force constant variation (in kcal mol^{-1} $Å^{-2}$) as a function of the RMSD for the 51 proteins in the dataset. (c) Average force constant (in kcal mol^{-1} $Å^{-2}$) as a function of the RMSD (in Å) for the 51 proteins in the dataset in their open/unbound form. The solid black lines show the best linear fit to the data.

Figure 2. (a) Rigidity profile for T4 lysozyme mutant in its open form (all the rigidity curves are in kcal mol^{-1} $Å^{-2}$) (b) Force constant variations upon closure (c) Cartoon representation of T4 lysozyme mutant in its closed form. Residues undergoing a large increase of their force constant upon closure are shown as red sticks, and residues undergoing an important rigidity decrease upon closure are shown as green van der Waals spheres. This color code is valid for all figures in the manuscript and supporting materials.

(d) Rigidity profile for ribonuclease MC1 in its unbound form (e) Force constant variations upon ligand binding (f) Cartoon representation of ribonuclease MC1 in its ligand bound form with the rigid protein core shown in blue.

(g) Rigidity profile for galactose mutarotase in its unbound form (h) Force constant variations upon ligand binding (i) Cartoon representation of galactose mutarotase in its ligand bound form.

Figures 2, 3, 5, 6 and figures in the supporting material were prepared using VMD (46).

Figure 3. (a) Rigidity profile for human serum transferrin in its open form. The red bars highlight the residues involved in a structural lock after closure (b) Force constant variations upon closure (c) Cartoon representation of human serum transferrin in its closed form, the ionic Fe^{3+} ligand is shown as a green dot in the center.

Figure 4. (a) Directional force constants matrix for human serum transferrin in its open form with the domains distribution along the sequence (b) Relative variation of the DFCs upon closure for human serum transferrin (c) DFCs distribution of human serum transferrin for the open (blue bars) and closed (transparent bars with black contours) conformations.

Figure 5. Cartoon representations of human serum transferrin with the residues forming rigid pairs upon closure shown as van der Waals spheres: Val11, Ser12, Glu13, His14, Glu15 and Ala16 (blue), Lys42, Ala43 and Ser44 (cyan), Cys179, Ser180, Thr181 and Asn183 (purple). (a) Open structure. (b) Closed structure.

Figure 6. Adenylate kinase (a) Upper panel: Rigidity profile for the protein in its open form. Lower panel: Force constant variations upon closure. (b) Directional force constants matrix for the protein in its open form. (c) Relative variation of the DFCs upon closure. (d) Cartoon representation of the protein in its closed form, the AP5 ligand is shown in the center.

Cartoon representations with the residues forming rigid pairs upon closure shown as van der Waals spheres: (e) Open structure (f) Closed structure.

Figure 7 Distribution of coevolution scores from the CoeViz webserver as a function of the relative variation in the DFCs upon closure for adenylate kinase.

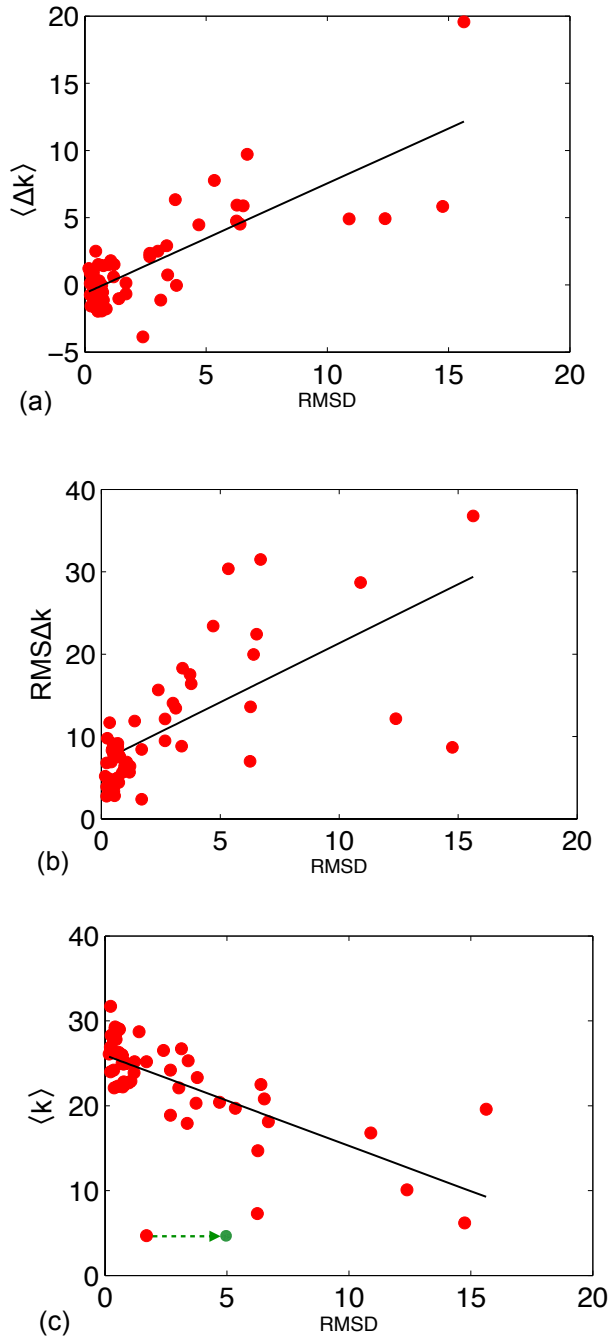

20
15
10
5
0
−5
⟨Δk⟩
0
5
10
15
20
RMSD
(a)
40
30
20
10
0
RMSΔk
0
5
10
15
20
RMSD
(b)
40
30
20
10
0
⟨k⟩
0
5
10
15
20
RMSD
(c)

Figure 1

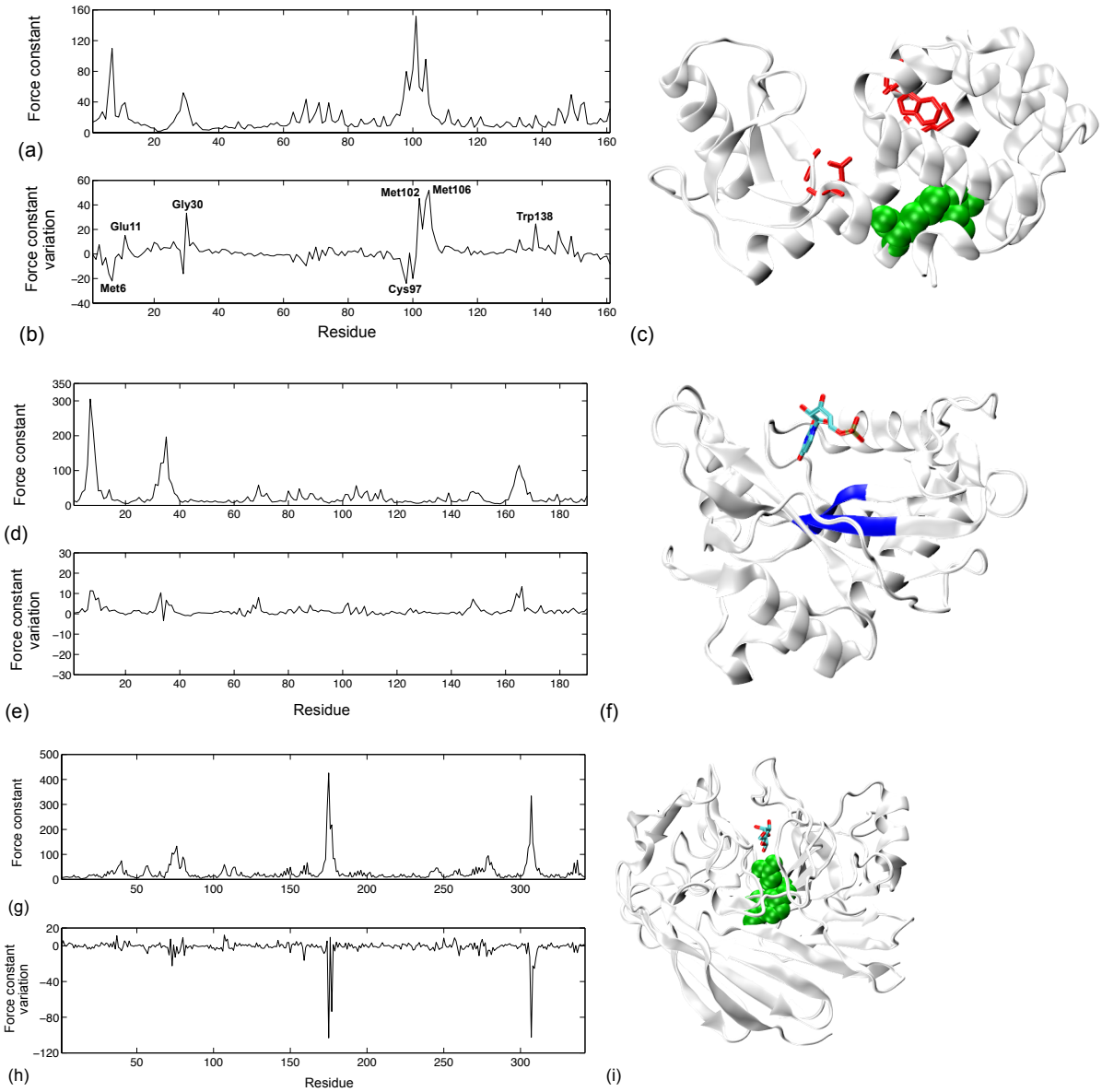

Force constant
(a)
Force constant variation
Glu11
Gly30
Met6
Met102
Met106
Cys97
Trp138
Residue
(b)
(c)
Force constant
(d)
Force constant variation
Residue
(e)
(f)
Force constant
(g)
Force constant variation
Residue
(h)
(i)

Figure 2

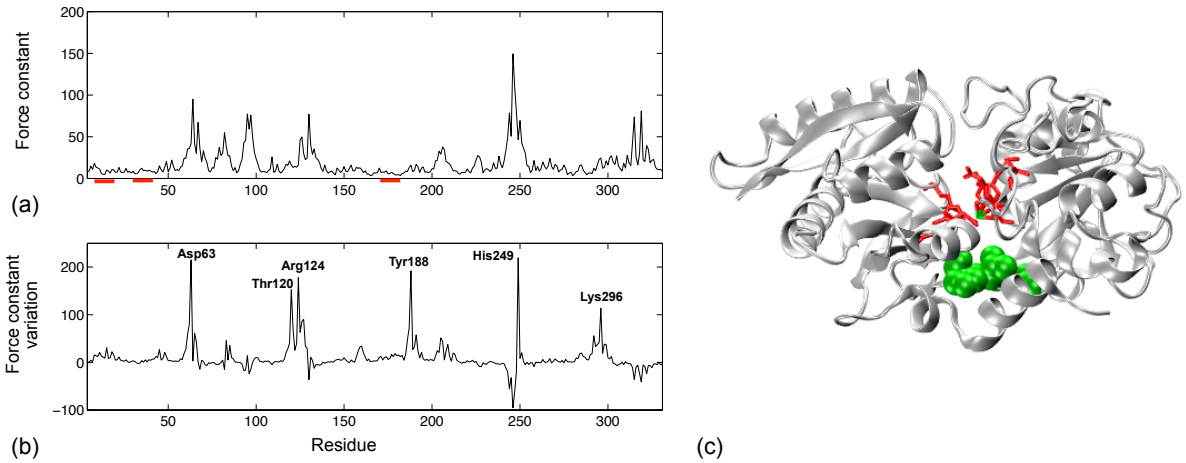

Force constant
200
150
100
50
0
50
100
150
200
250
300
(a)
Force constant variation
Asp63
Thr120
Arg124
Tyr188
His249
Lys296
200
100
0
−100
50
100
150
200
250
300
Residue
(b)
(c)

Figure 3

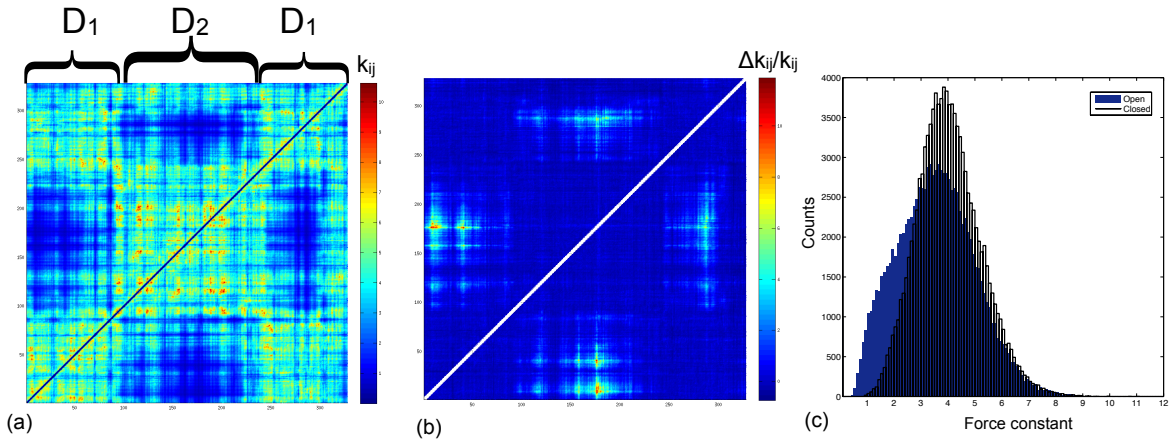
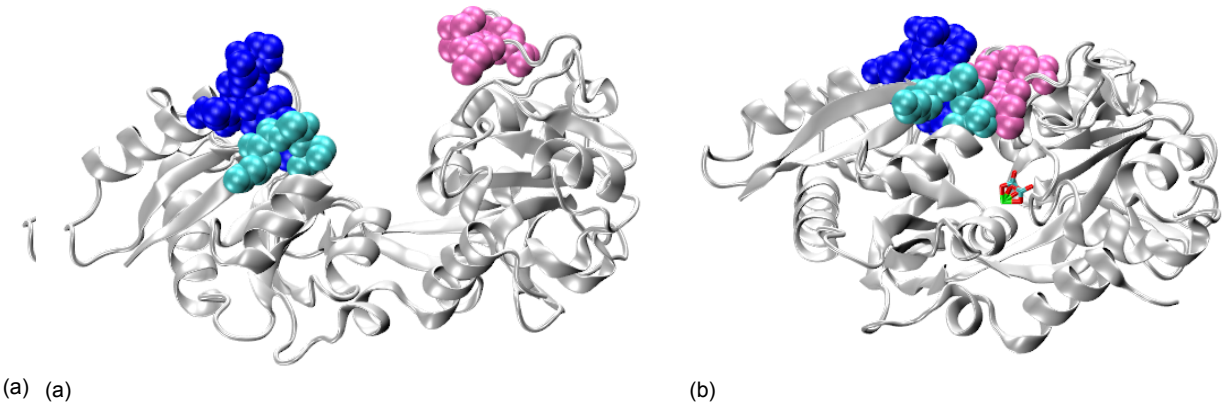


Figure 4



(a) (a) (b)

Figure 5

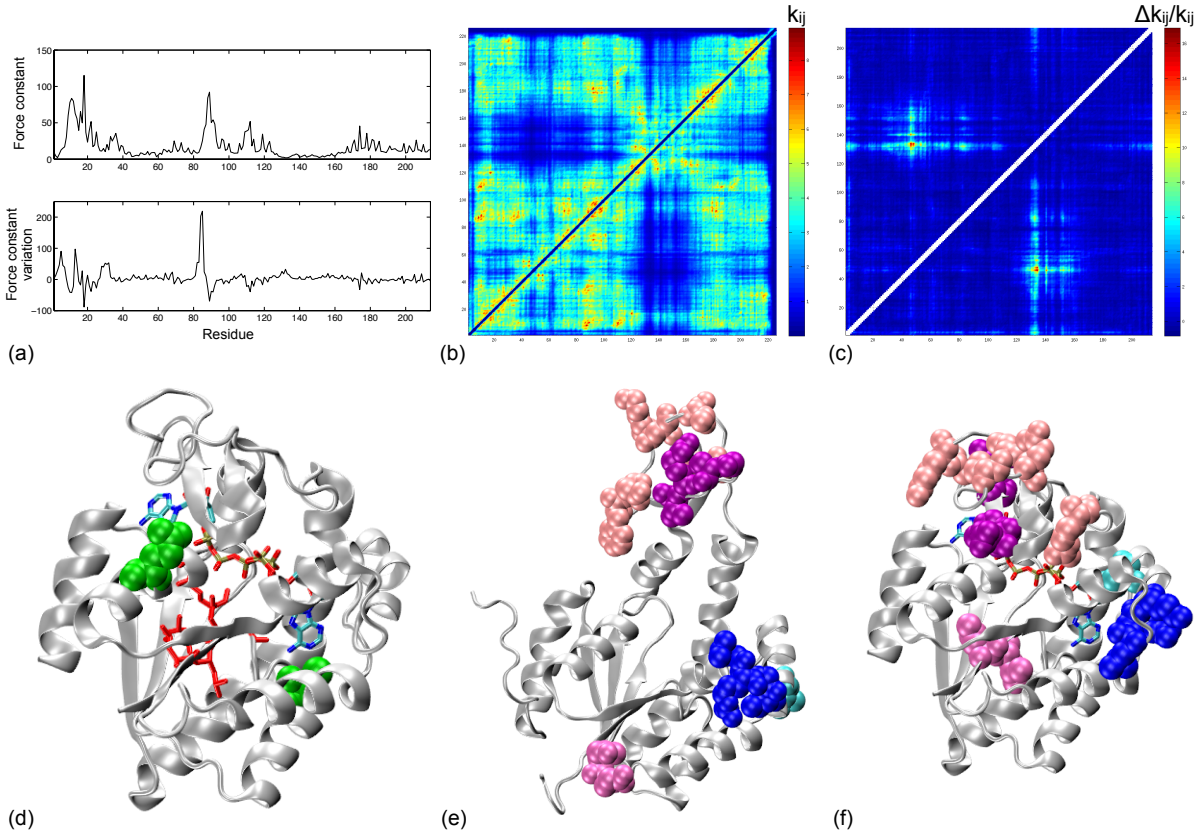

Force constant
Force constant variation
Residue
(a)
k_{ij}
(b)
$\Delta k_{ij}/k_{ij}$
(c)
(d)
(e)
(f)

Figure 6

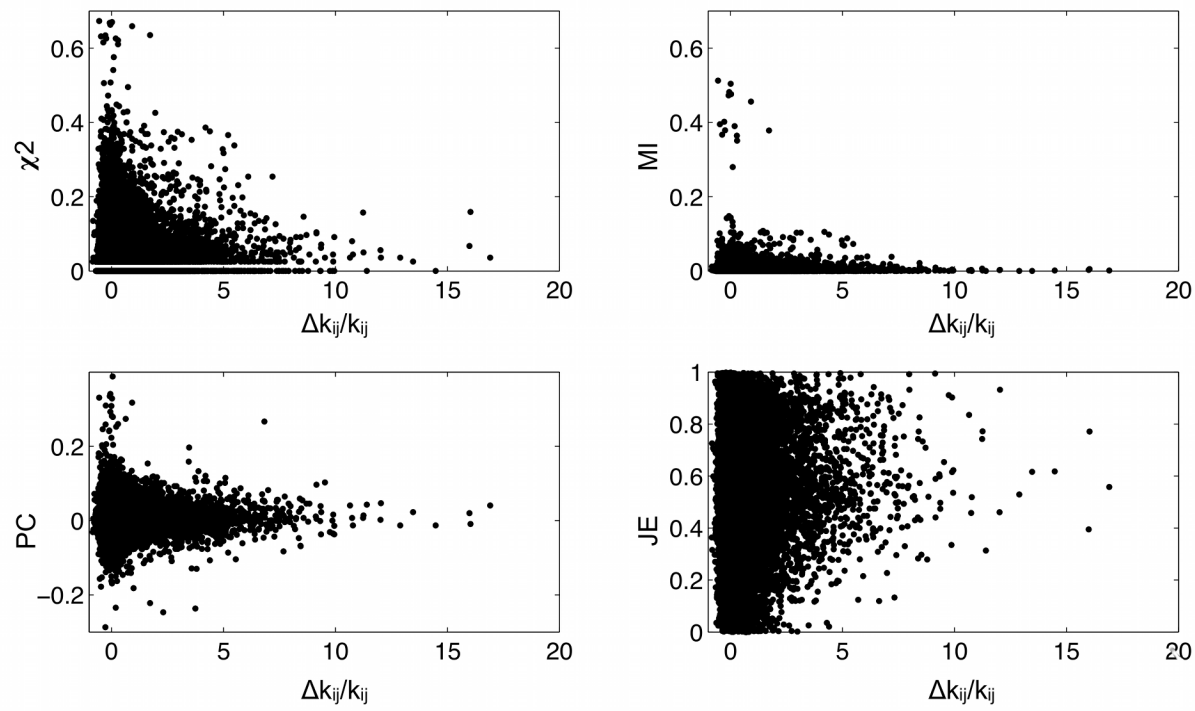


Figure 7

References

1. Henzler-Wildman, K., and D. Kern. 2007. Dynamic personalities of proteins. Nature 450:964-972.
2. Teilum, K., J. G. Olsen, and B. B. Kragelund. 2009. Functional aspects of protein flexibility. Cellular and Molecular Life Sciences 66:2231-2247.
3. Micheletti, C. 2013. Comparing proteins by their internal dynamics: exploring structure-function relationships beyond static structural alignments. Phys Life Rev 10:1-26.
4. Orozco, M. 2014. A theoretical view of protein dynamics. Chem. Soc. Rev. 43:5051-5066.
5. Nussinov, R., and B. Ma. 2012. Protein dynamics and conformational selection in bidirectional signal transduction. BMC Biol 10:2.
6. Gunasekaran, K., B. Ma, and R. Nussinov. 2004. Is allostery an intrinsic property of all dynamic proteins? Proteins 57:433-443.
7. Dror, R. O., R. M. Dirks, J. P. Grossman, H. Xu, and D. E. Shaw. 2012. Biomolecular simulation: a computational microscope for molecular biology. Annu Rev Biophys 41:429-452.
8. Maximova, T., R. Moffatt, B. Ma, R. Nussinov, and A. Shehu. 2016. Principles and Overview of Sampling Methods for Modeling Macromolecular Structure and Dynamics. PLoS Comput Biol 12:e1004619.
9. Shaw, D. E., P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers. 2010. Atomic-level characterization of the structural dynamics of proteins. Science 330:341-346.
10. Zheng, W., and H. Wen. 2017. A survey of coarse-grained methods for modeling protein conformational transitions. Curr. Opin. Struct. Biol. 42:24-30.
11. Tama, F., and Y. H. Sanejouand. 2001. Conformational change of proteins arising from normal mode calculations. Protein Eng. 14:1-6.
12. Yang, L., G. Song, and R. L. Jernigan. 2007. How well can we understand large-scale protein motions using normal modes of elastic network models? Biophys. J. 93:920-929.
13. Al-Bluwi, I., M. Vaisset, T. Simeon, and J. Cortes. 2013. Modeling protein conformational transitions by a combination of coarse-grained normal mode analysis and robotics-inspired methods. BMC Struct Biol 13 Suppl 1:S2.
14. Uyar, A., N. Kantarci-Carsibasi, T. Haliloglu, and P. Doruker. 2014. Features of large hinge-bending conformational transitions. Prediction of closed structure from open state. Biophys. J. 106:2656-2666.
15. Fuglebakk, E., S. P. Tiwari, and N. Reuter. 2015. Comparing the intrinsic dynamics of multiple protein structures using elastic network models. Biochim. Biophys. Acta 1850:911-922.
16. Mahajan, S., and Y. H. Sanejouand. 2015. On the relationship between low-frequency normal modes and the large-scale conformational changes of proteins. Arch. Biochem. Biophys. 567:59-65.
17. Kurkcuoglu, Z., I. Bahar, and P. Doruker. 2016. ClustENM: ENM-Based Sampling of Essential Conformational Space at Full Atomic Resolution. J Chem Theory Comput 12:4549-4562.

18. Lopez-Blanco, J. R., and P. Chacon. 2016. New generation of elastic network models. Curr. Opin. Struct. Biol. 37:46-53.
19. Sacquin-Mora, S., and R. Lavery. 2006. Investigating the local flexibility of functional residues in hemoproteins. Biophys. J. 90:2706-2717.
20. Sacquin-Mora, S. 2016. Bridging Enzymatic Structure Function via Mechanics: A Coarse-Grain Approach. Methods Enzymol. 578:227-248.
21. Delalande, O., S. Sacquin-Mora, and M. Baaden. 2011. Enzyme closure and nucleotide binding structurally lock guanylate kinase. Biophys. J. 101:1440-1449.
22. Lavery, R., and S. Sacquin-Mora. 2007. Protein mechanics: a route from structure to function. J. Biosci. 32:891-898.
23. Sacquin-Mora, S., E. Laforet, and R. Lavery. 2007. Locating the active sites of enzymes using mechanical properties. Proteins 67:350-359.
24. Sacquin-Mora, S. 2014. Motions and mechanics: investigating conformational transitions in multi-domain proteins with coarse-grain simulations. Mol. Simul. 40:229-236.
25. Tozzini, V. 2005. Coarse-grained models for proteins. Curr. Opin. Struct. Biol. 15:144-150.
26. Zacharias, M. 2003. Protein-protein docking with a reduced protein model accounting for side-chain flexibility. Protein Sci. 12:1271-1282.
27. Eyal, E., and I. Bahar. 2008. Toward a molecular understanding of the anisotropic response of proteins to external forces: Insights from elastic network models. Biophys. J 94:3424-3435.
28. Baker, F. N., and A. Porollo. 2016. CoeViz: a web-based tool for coevolution analysis of protein residues. BMC Bioinformatics 17:119.
29. Porollo, A. A., R. Adamczak, and J. Meller. 2004. POLYVIEW: a flexible visualization tool for structural and functional annotations of proteins. Bioinformatics 20:2460-2462.
30. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.
31. Larson, S. M., A. A. Di Nardo, and A. R. Davidson. 2000. Analysis of covariation in an SH3 domain sequence alignment: applications in tertiary contact prediction and the design of compensating hydrophobic core substitutions. J. Mol. Biol. 303:433-446.
32. Clarke, N. D. 1995. Covariation of residues in the homeodomain sequence family. Protein science : a publication of the Protein Society 4:2269-2278.
33. Weaver, L. H., and B. W. Matthews. 1987. Structure of bacteriophage T4 lysozyme refined at 1.7 A resolution. Journal of Molecular Biology 193:189-199.
34. Shoichet, B. K., W. A. Baase, R. Kuroki, and B. W. Matthews. 1995. A relationship between protein stability and protein function. Proceedings of the National Academy of Sciences of the United States of America 92:452-456.
35. Merski, M., and B. K. Shoichet. 2012. Engineering a model protein cavity to catalyze the Kemp elimination. Proceedings of the National Academy of Sciences of the United States of America 109:16179-16183.

36. Collins, M. D., M. L. Quillin, G. Hummer, B. W. Matthews, and S. M. Gruner. 2007. Structural rigidity of a large cavity-containing protein revealed by high-pressure crystallography. Journal of Molecular Biology 367:752-763.
37. Dixon, M. M., H. Nicholson, L. Shewchuk, W. A. Baase, and B. W. Matthews. 1992. Structure of a hinge-bending bacteriophage T4 lysozyme mutant, Ile3-->Pro. Journal of Molecular Biology 227:917-933.
38. Sacquin-Mora, S., P. Sebban, V. Derrien, B. Frick, R. Lavery, and C. Alba-Simionesco. 2007. Probing the flexibility of the bacterial reaction center: the wild-type protein is more rigid than two site-specific mutants. Biochemistry 46:14960-14968.
39. Bocahut, A., S. Bernad, P. Sebban, and S. Sacquin-Mora. 2011. Frontier residues lining globin internal cavities present specific mechanical properties. J. Am. Chem. Soc. 133:8753-8761.
40. Jeffrey, P. D., M. C. Bewley, R. T. MacGillivray, A. B. Mason, R. C. Woodworth, and E. N. Baker. 1998. Ligand-induced conformational change in transferrins: crystal structure of the open form of the N-terminal half-molecule of human transferrin. Biochemistry 37:13978-13986.
41. Sacquin-Mora, S., O. Delalande, and M. Baaden. 2010. Functional modes and residue flexibility control the anisotropic response of guanylate kinase to mechanical stress. Biophys. J. 99:3412-3419.
42. Seyler, S. L., and O. Beckstein. 2014. Sampling large conformational transitions: adenylate kinase as a testing ground. Molecular Simulation 40:855-877.
43. Unan, H., A. Yildirim, and M. Tekpinar. 2015. Opening mechanism of adenylate kinase can vary according to selected molecular dynamics force field. J Comput Aided Mol Des 29:655-665.
44. Armenta-Medina, D., E. Perez-Rueda, and L. Segovia. 2011. Identification of functional motions in the adenylate kinase (ADK) protein family by computational hybrid approaches. Proteins 79:1662-1671.
45. Sfriso, P., M. Duran-Frigola, R. Mosca, A. Emperador, P. Aloy, and M. Orozco. 2016. Residues Coevolution Guides the Systematic Identification of Alternative Functional Conformations in Proteins. Structure 24:116-126.
46. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14:33-38, 27-38.