Biophysical Journal Volume 105 November 2013 2412–2417

Molecular Simulations of a Dynamic Protein Complex: Role of Salt-Bridges and Polar Interactions in Configurational Transitions

Liqun Zhang[†] and Matthias Buck[†‡*]
[†]Department of Physiology and Biophysics; and [‡]Department of Neurosciences, Department of Pharmacology, Case Comprehensive Cancer Center and Center for Proteomics and Bioinformatics, Case Western Reserve University, School of Medicine, Cleveland, Ohio

ABSTRACT Ion charge pairs and hydrogen bonds have been extensively studied for their roles in stabilizing protein complexes and in steering the process of protein association. Recently, it has become clear that some protein complexes are dynamic in that they interconvert between several alternate configurations. We have previously characterized one such system: the EphA2:SHIP2 SAM-SAM heterodimer by solution NMR. Here we carried out extensive all-atom molecular-dynamics simulations on a microsecond time-scale starting with different NMR-derived structures for the complex. Transitions are observed between several discernible configurations at average time intervals of 50–100 ns. The domains reorient relative to one another by substantial rotation and a slight shifting of the interfaces. Bifurcated and intermediary salt-bridge and hydrogen-bond interactions play a role in the transitions in a process that can be described as moving along a "monkey-bar". We notice an increased density of salt bridges near protein interaction surfaces that appear to enable these transitions, also suggesting why the trajectories can become kinetically hindered in regions where fewer of such interactions are possible. In this context, even microsecond molecular-dynamics simulations are not sufficient to sample the energy landscape unless the structures remain close to their experimentally derived low-energy configurations.

INTRODUCTION

Accurate prediction of protein complexes remains a challenge when no structural or sequence homology is available, or in cases where the proteins undergo considerable conformational changes upon association (1,2). Recently, it has also become apparent that some protein complexes are inherently dynamic, fluctuating locally or even at the (sub-) domain level, between different configurations (3–6). We have been characterizing such a protein complex by solution NMR measurements (6), and here by molecular-dynamics simulations. Previous studies have shown that the free-energy landscape of protein-protein complexes can range from simple funnel-like to highly complicated terrains (7,8). How residue-residue and residue-solvent interactions determine the nature of the landscape, and which interactions allow transitions between low energy configurations, are important features of protein interactions that remain to be understood in detail (9,10). In particular, the roles of ion-pair and hydrogen-bond interactions for the stability of the protein complex and for the process of protein association have been heavily studied, indicating sometimes complex and often context-dependent contributions (11,12).

Here we study two features of the ensemble of structures that comprise a dynamic protein complex, as explored by extensive molecular dynamics simulations.

First, the 2.4-μs all-atom molecular dynamics simulations of the complex are used to sample the energy landscape; cluster centers are validated with respect to the experimental NMR restraints. However, the results also suggest that simulations can go off-course, leaving a simple landscape to become trapped in a more restrictive terrain. Second, the trajectories are used to examine transitions between different configurations of the protein-protein complex, especially with respect to the role of ion pairs (salt bridges) and hydrogen-bond interactions in enabling the transitions. Finally, these two findings are proposed to be related, considering the distribution of surface groups.

COMPUTATIONAL METHODS

Simulations were performed starting from NMR-derived configurations of the EphA2:SHIP2 SAM-SAM complex, using both the software NAMD Vers.2.8 (13) and the ANTON Supercomputer (14) simulation programs. The structures have been determined previously in the laboratory (6) from solution NMR restraints.

Molecular-dynamics simulations

Six simulations were run, each started with different structures. Three simulations began with the lowest energy structures of three different configurations that were derived (Clusters1–3). A fourth structure was the next lowest energy structure of the Cluster3 configuration. The fifth simulation was started with a Cluster1-like configuration that was derived by NMR restraints in a calculation omitting several unambiguous distance restraints. The sixth simulation began with a structure far away from the others (arising as a minor population in the second, control calculation as described in Lee et al. (6)). The starting structures were solvated in a rectangular box of explicitly represented water (TIP3P), minimized and equilibrated for 20 ns using unrestrained all-atom molecular-dynamics (MD) simulation at constant temperature and pressure (300 K and 1 atm). The simulation box contained ~16,500 TIP3P water molecules, and the box size was around 90 × 70 × 70 Å^3. The standard particle-mesh Ewald method was applied to calculate the long-range electrostatic interactions. The system was neutralized by adding ions to a concentration of 0.15 M

Submitted July 30, 2013, and accepted for publication September 24, 2013.
*Correspondence: matthias.buck@case.edu
Editor: Michael Feig.

0006-3495/13/11/2412/6 $2.00
http://dx.doi.org/10.1016/j.bpj.2013.09.052

NaCl. The CHARMM27 all-atom potential function was used including the CMAP correction (15). For nonbonded calculations, a cutoff of 12 Å was used. All bonds, involving hydrogen, were kept rigid using the SHAKE algorithm, allowing 2-fs timesteps. After 20-ns NAMD simulations, we continued the calculations with the ANTON Supercomputer for 2.4 μs. Coordinates were saved every 50 ps.

Trajectory analysis

K-class clustering of the trajectories was carried out using the software WORDOM (16). Solvent-accessible surface calculation used the Lee and Richards algorithm in the CHARMM program (Academic edition from Harvard University) with a probe radius of 1.4 Å. Distances of all protein side-chain hydrogen-bond donors and acceptors to the nearest group across the interface and to water were calculated. The groups involved were identified and intergroup distances were plotted (see Fig. 4). For Table 1, any residue involved in interactions in the four main interface configurations (derived by clustering the four trajectories showing transitions) was counted as an interface residue.

RESULTS AND DISCUSSION

We previously reported the refined structure of the EphA2:SHIP2 SAM-SAM complex (6). While 80% of the NMR-refined structures populated one cluster of conformations (referred to as Cluster1), two additional clusters of structures were observed (Clusters2 and 3; Fig. 1). These structures required further investigation because one of them (Cluster2) was close to the earlier model that had been derived by another group from fewer, and only ambiguous, NMR restraints (17). However, our study also indicated fluctuations toward these alternative states in a 20-ns MD simulation (6). Here we computed lengthy 2.4 μs trajectories on the MD-optimized supercomputer ANTON (15), starting with a structure representative of each of these three clusters as well as other configurations. The simulations were run with standard protocols (see Computational Methods). A typical analysis carried out with the trajectories is shown in Fig. 2. It is apparent that interconversions between Clusters1 and 2 and more rarely with Cluster3-like structures take place on a timescale of 50–100 ns (see Movie S1 in the Supporting Material).

Interestingly, the other two trajectories (started with Cluster2 and 3 structures) also sample states near Cluster1, as did a fourth trajectory started with a Cluster1-like structure (see Fig. S1 in the Supporting Material). The four trajectories were combined and clustered (Fig. 3 *a*), yielding four main clusters of configurations, represented by cluster centers, with occupancies of 53, 17, 15, and 9%, respectively, as well as two minor clusters (<5% occupancy).

TABLE 1 Distribution of charged and H-bonding groups on SAM domains

	EphA2 charged	H-bond	SHIP2 charged	H-bond
Interface	38%	46%	47%	21%
Noninterface	17%	23%	27%	15%

When these configurations are compared with the experimental NMR data, it is clear that four of six clusters (populating 82% of structures) satisfy the NMR data reasonably well, whereas two other clusters satisfy the NMR data less well (see Table S1 *a* in the Supporting Material). A similar picture emerges when structures from the two additional trajectories, started with minor configurations (see Computational Methods), are added to the clustering. It is apparent that these two trajectories do not move closer to any of the main configurations, but are unconverged and appear kinetically trapped in a region of the energy landscape (Fig. 3 *b*, and see Table S1 *b*, Fig. S2, and Movie S2).

Clustering that includes these latter two trajectories results in 14 cluster centers (only three with populations >10%). Most do not satisfy the experimental data as compared to the clusters from the four initial trajectories (see Table S1). This raises the following question: what features of a protein surface generate a smooth landscape that provides facile but contained transitions, compared to surfaces and interactions that allow protein structures to wander into complex and trapped terrains?

The principal clusters, which are sampled in the initial four well-behaved trajectories, correspond to the experimentally derived structures and show interconversion, as noted above. Geometrically, the structures differ by the alignment of the two SAM domains, both in terms of rotation (measured by the angle between helix 5 of each domain, Fig. 1, *a–c*) and by a slight shift in the interface (not shown). An explanation for these different configurations are different pairings of charged residues across the protein-protein interfaces, as shown in Fig. 1 *d*. Several of the pairings overlap, in part, participating in bifurcated bonding with the same acceptor group. The evolution of key interactions in the trajectory started with the Cluster1 configuration is summarized in Fig. 4. It is apparent that, although some of the residue-residue contacts, such as K917-D1224 and K917-D1230, are mutually exclusive for large parts of the trajectory, others display considerable degeneracy; that is, they occur in several of the clusters, as pointed out for the experimental structure determination. For example, in the NMR-derived Cluster2 configuration, K956-D1235 and K956K-E1238, can form simultaneously and these interactions persist for considerable parts of the trajectory, because the two residues (D1235 and E1238) are located close to one another. This interaction is also seen as an intermediate in transitioning to Cluster1-like configurations, where only K956-D1235 can form.

When the closest separation distances are plotted as a function of simulation time, it becomes apparent that many transitions are associated with making and breaking of contacts but in a manner that resembles movement along a monkey bar, i.e., a new contact is formed before the previous one is released. This way the proteins always make some interactions (i.e., in the vernacular sense, at least one of the hands is on the bar). This is true for the other three

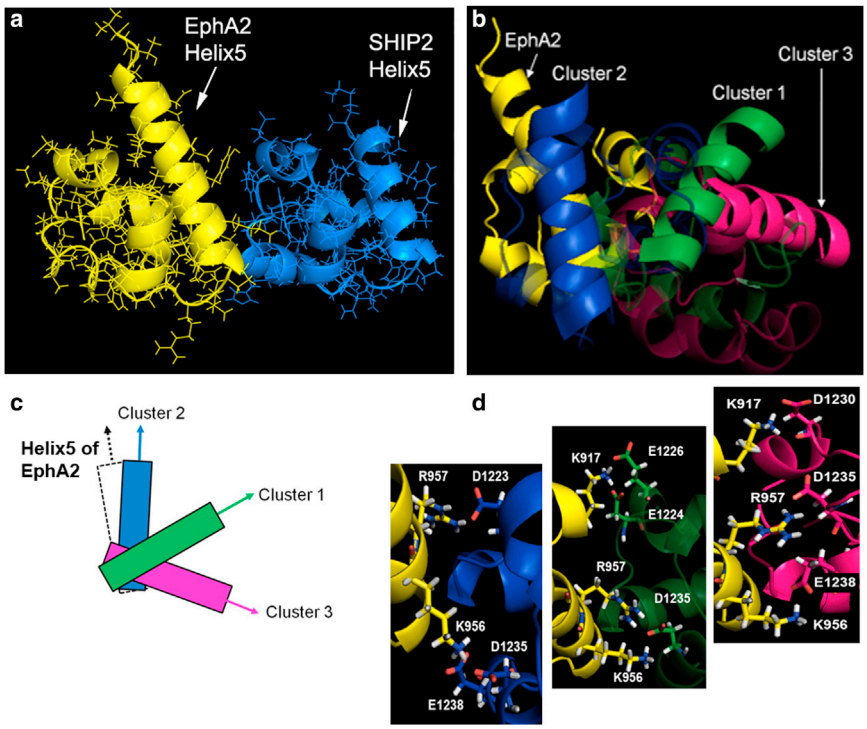


FIGURE 1 SAM-SAM heterodimer configurations derived from NMR (6) (*a*) Cluster 2 with fifth helix indicated in each domain; (*b*) Clusters 1–3, with fifth helix in the SHIP2 SAM domain highlighted; (*c*) schematic showing different domain orientations; and (*d*) ion-pairs across the interface in the different configurations.

trajectories that show frequent transitions (see Fig. S3, *a–c*). Waters may participate in this process by forming bifurcated hydrogen bonding. A detailed trajectory analysis shows that amide and oxygen side-chain groups are close to water molecules throughout the trajectories (data not shown); the only residue side-chain group that becomes fully desolvated in many of the trajectories is SHIP2 W1223, but even this group is desolvated only temporarily (*last column* of Fig. 4, and similarly in Fig. S3, *a–c*). Solvation is likely to be important for the transitions, as well as for protein-dissociation and association processes (18). A dissociation of proteins is not seen in these simulations, but detailed in other simulations with a SAM-SAM complex where two key interface residues have been mutated (unpublished data).

The trajectories that move far away from the starting structures show few interconversion events involving ion pairs and hydrogen bonds (see Fig. S4, *a* and *b*). Examining the surfaces in contact, there are fewer charged residues (and H-bond donors and acceptors; see Table 1, above). We find the protein-protein complex configurations that are more highly populated also have more ion-pair and hydrogen-bonding residue contacts than others. However, having more interactions does not necessarily mean that the total protein-protein interaction is stronger, and a conclusion in this respect will require detailed free-energy calculations. The suggestion that native protein-protein interfaces have characteristics that are different from other protein surfaces has found support in several structural analyses (9–12). In a followup study, we find that the dissociation of mutant SAM-SAM complexes takes place from configurations in which at least one of the interface regions differs from those utilized by the dynamic and stable configurations near the experimental starting structures.

It should be noted that although the solvent-accessible surface area, which is buried between the proteins, can fluctuate significantly (e.g., Fig. S2 *a*), the proteins simulated here never truly separate. Interactions, such as those of R957-D1235 and R957-E1238, stay within contact distance (<5 Å) (see Fig. S4, *a* and *b*). It is intriguing to note that the latter interaction is not predominant in the four trajectories that show configurational interconversion. Considering the process of protein association, it is possible that such anchoring interactions (19,20) may form easily but might also trap the protein complex in a certain region of the configurational landscape.

Given resources, it would be desirable to run even longer simulations (hundreds of microseconds to milliseconds), but our observations also need to be viewed against a background of studies that have shown that sampling can be influenced by inadequacies in the potential function used in the simulations (e.g., Trbovic et al. (21)), especially when the energy barriers are high, such as in protein refinement (22). However, along with recent reports (e.g., Zeiske et al. (23)), this study suggests that if the protein complex is maintained near experimentally derived configurations, a realistic sampling and transitions may be obtained. Remarkably, we note that surfaces with fewer possible groups for protein interactions appear to lead to slower transition kinetics or even trapped states.

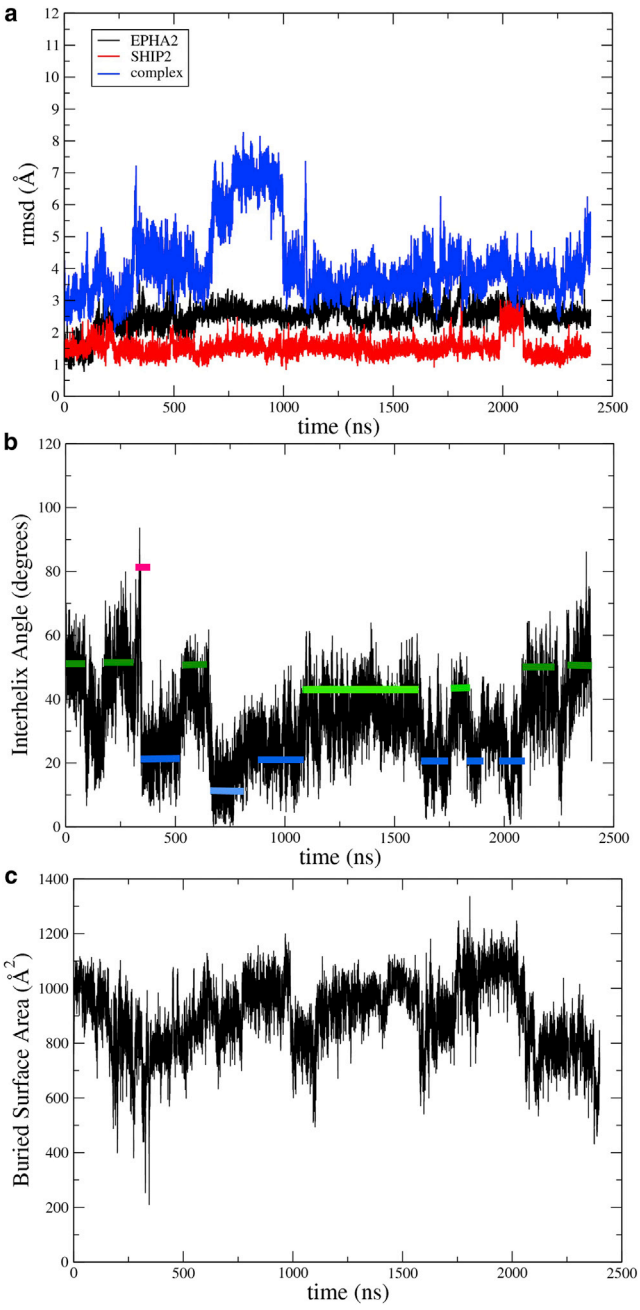


FIGURE 2 Trajectory analysis for a simulation started from a Cluster-1 conformation. (*a*) Root-mean-squared deviation from starting structures for EphA2 and SHIP2 SAM domains as well as for the complex; (*b*) angle between the fifth helix of the two domains (different configurations sampled are indicated by *color bars* with reference to Fig. 1 *c*); and (*c*) buried solvent-accessible surface area at the protein-protein interface. To see this figure in color, go online.

This study extends observations and concepts developed for protein folding and single-domain structural transitions to protein-protein interactions. Specifically, hydrogen-bonding intermediary states are known to play a role in helix-coil transitions (24–26) and in a cell-signaling protein (e.g., Lei et al. (27)). The kinetic role of waters in these transitions is thought

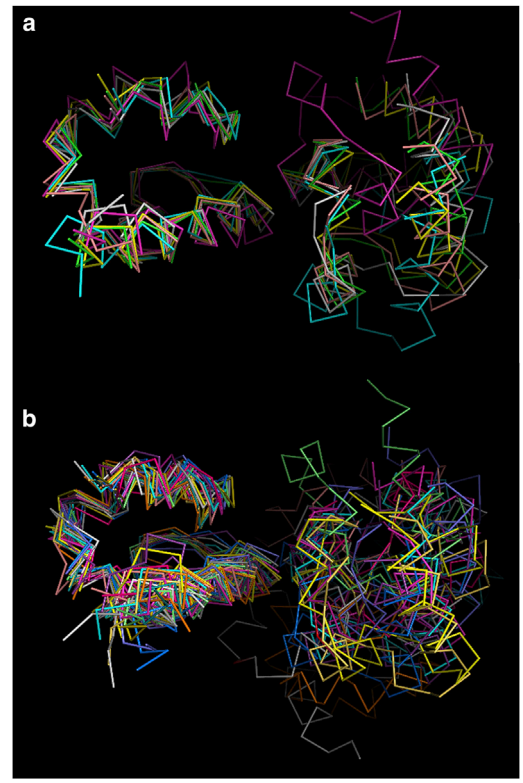


FIGURE 3 Cluster centers from (*a*) four and (*b*) six 2.4-μs trajectories. The clustering was done by side-chain, root-mean-squared deviation (5.5 Å cutoff) of the residues at the Clusters 1–3 SAM-SAM interface. Here the structures are superimposed on the EphA2 SAM domain (*left*) for visual display. To see this figure in color, go online.

to be small: Water moves around the protein much more quickly than even its side-chain transitions (28). More importantly, as seen in the configurational transitions described here, charged/polar side chains that line the protein-protein interface remain at least partially solvated, as found in the great majority of protein interfaces (18,29,30). By contrast, larger-scale desolvation may be expected to slow protein association events (e.g., Camacho et al. (31)).

Finally, experimental studies, examining possible functional differences between the different configurations of the EphA2:SHIP2 SAM-SAM complexes, are needed to show that the different configuration have a biological significance. This will be a considerable undertaking, because the studies need to be carried out for the whole-length receptor/SHIP2 protein as well as in proximity of the plasma membrane, which may influence the equilibrium. For now, however, our study delineates the behavior of the interaction at the SAM-SAM domain level, and the dynamic nature of the interface indicated by our previous NMR study can be seen directly in microsecond simulations.

SUPPORTING MATERIAL

Two movies are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)01129-6.

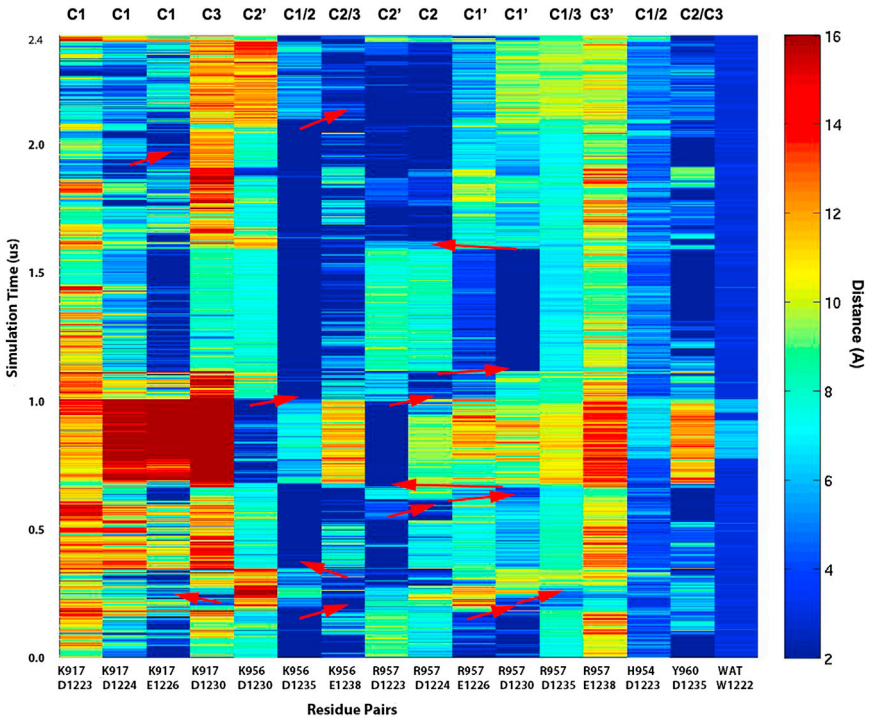


FIGURE 4 Transitions between residue pairs across the interface as a function of simulation time. (*Dark blue*) Close contacts; (*yellow* to *red*) distant interactions (*scale on right*). Contact pairs are identified below and associated with cluster/cluster-like configuration above (*C#/C#′*); see also Fig. S3 in the Supporting Material. It is apparent that in many transitions additional contacts form before old ones are broken (indicated by *red arrows*). For example, K956-D1238 forms before K956-D1235 is broken at ~0.25 μs. Similarly, R957 transitions from E1226 to E1230 and then D1235 as partners, before returning to D1224 and then D1223, around the same time. This is conceptually similar to holding onto and traveling along a monkey bar, as depicted below.

The work was funded by a National Institutes of Health grant to M.B. (No. R01GM092851) and a postdoctoral fellowship to L.Z. (No. T32DK007470). The use of the ANTON Supercomputer at the Pittsburgh Supercomputing Center is supported by a National Institutes of Health award (No. RC2GM093307) to Carnegie Mellon University. Part of this work also used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575.

REFERENCES

1. Janin, J. 2010. Protein-protein docking tested in blind predictions: the CAPRI experiment. *Mol. Biosyst.* 6:2351–2362.
2. de Vries, S. J., A. S. Melquiond, …, A. M. Bonvin. 2010. Strengths and weaknesses of data-driven docking in critical assessment of prediction of interactions. *Proteins.* 78:3242–3249.
3. Hu, J., K. Hu, …, G. M. Clore. 2008. Solution NMR structures of productive and non-productive complexes between the A and B domains of the cytoplasmic subunit of the mannose transporter of the *Escherichia coli* phosphotransferase system. *J. Biol. Chem.* 283:11024–11037.
4. Boehr, D. D., R. Nussinov, and P. E. Wright. 2009. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* 5:789–796.
5. Lindfors, H. E., J. W. Drijfhout, and M. Ubbink. 2012. The Src SH2 domain interacts dynamically with the focal adhesion kinase binding site as demonstrated by paramagnetic NMR spectroscopy. *IUBMB Life.* 64:538–544.
6. Lee, H. J., P. K. Hota, …, M. Buck. 2012. NMR structure of a heterodimeric SAM:SAM complex: characterization and manipulation of EphA2 binding reveal new cellular functions of SHIP2. *Structure.* 20:41–55.
7. Tsai, C. J., S. Kumar, …, R. Nussinov. 1999. Folding funnels, binding funnels, and protein function. *Protein Sci.* 8:1181–1190.
8. Schreiber, G., G. Haran, and H.-X. Zhou. 2009. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* 109:839–860.

9. Jones, S., and J. M. Thornton. 1996. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA.* 93:13–20.
10. Lo Conte, L., C. Chothia, and J. Janin. 1999. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* 285:2177–2198.
11. Xu, D., C. J. Tsai, and R. Nussinov. 1997. Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng.* 10:999–1012.
12. Sheinerman, F. B., R. Norel, and B. Honig. 2000. Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* 10:153–159.
13. Phillips, J. C., R. Braun, …, K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
14. Shaw, D. E., M. M. Deneroff, …, S. C. Wang. 2008. Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM.* 51:91–97.
15. Buck, M., S. Bouguet-Bonnet, …, A. D. MacKerell, Jr. 2006. Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys. J.* 90:L36–L38.
16. Seeber, M., M. Cecchini, …, A. Caflisch. 2007. WORDOM: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics.* 23:2625–2627.
17. Leone, M., J. Cellitti, and M. Pellecchia. 2008. NMR studies of a heterotypic SAM-SAM domain association: the interaction between the lipid phosphatase SHIP2 and the EPHA2 receptor. *Biochemistry.* 47:12721–12728.
18. Rodier, F., R. P. Bahadur, …, J. Janin. 2005. Hydration of protein-protein interfaces. *Proteins.* 60:36–45.
19. Kortemme, T., and D. Baker. 2002. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc. Natl. Acad. Sci. USA.* 99:14116–14121.
20. Keskin, O., B. Y. Ma, and R. Nussinov. 2005. Hot regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol.* 345:1281–1294.
21. Trbovic, N., B. Kim, …, A. G. Palmer, 3rd. 2008. Structural analysis of protein dynamics by MD simulations and NMR spin-relaxation. *Proteins.* 71:684–694.
22. Raval, A., S. Piana, …, D. E. Shaw. 2012. Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins.* 80:2071–2079.
23. Zeiske, T., K. A. Stafford, …, A. G. Palmer, 3rd. 2013. Starting-structure dependence of nanosecond timescale intersubstate transitions and reproducibility of MD-derived order parameters. *Proteins.* 81:499–509.
24. Tarek, M., and D. J. Tobias. 2002. Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys. Rev. Lett.* 88:138101.
25. Armen, R., D. O. Alonso, and V. Daggett. 2003. The role of α-, 3(10)-, and π-helix in helix$\rightarrow$coil transitions. *Protein Sci.* 12:1145–1157.
26. Monticelli, L., D. P. Tieleman, and G. Colombo. 2005. Mechanism of helix nucleation and propagation: microscopic view from microsecond time scale MD simulations. *J. Phys. Chem. B.* 109:20064–20067.
27. Lei, M., J. Velos, …, D. Kern. 2009. Segmented transition pathway of the signaling protein nitrogen regulatory protein C. *J. Mol. Biol.* 392:823–836.
28. Hamaneh, M. B., L. Zhang, and M. Buck. 2011. A direct coupling between global and internal motions in a single domain protein? MD investigation of extreme scenarios. *Biophys. J.* 101:196–204.
29. Bueno, M., N. A. Temiz, and C. J. Camacho. 2010. Novel modulation factor quantifies the role of water molecules in protein interactions. *Proteins.* 78:3226–3234.
30. Reichmann, D., Y. Phillip, …, G. Schreiber. 2008. On the contribution of water-mediated interactions to protein-complex stability. *Biochemistry.* 47:1051–1060.
31. Camacho, C. J., S. R. Kimura, …, S. Vajda. 2000. Kinetics of desolvation-mediated protein-protein binding. *Biophys. J.* 78:1094–1105.

Molecular Simulations of a Dynamic Protein Complex: Role of Salt-Bridges and Polar Interactions in Configurational Transitions

Liqun Zhang+ and Matthias Buck*,+,$

SUPPORTING MATERIAL

Supplemental Movies

Transitions are evident in the trajectory started with the cluster1 configuration (movie1), but not in the trajectory started with a structure far away from the major configurations (movie2, see methods). In both movies, the SAM domain mainchain is shown in ribbon representation, and is aligned on the EphA2 SAM domain. Note that in the second movie, the configuration of the complex is with helix5 nearly antiparallel. This does not change/fluctuate over the course of the simulation (see Fig. S2b).

movie1.mpg

movie2.mpg

Supplemental Tables

Table S1. Analysis of cluster centers. Coordinate frames from the four and six trajectories were k-clustered by interface sidechain RMSD as described (cut-off 5.5 Å). Cluster center population in the trajectories and RMSD to starting structures is given (underlined value is the closest). The extent to which the experimental restraints are satisfied in these unrestrained calculations is given; Q-factors for RDCs from the two alignment media and RMS deviation from ambiguous and unambigious NOEs (experimental data from reference 5) are given. Q-factors ≥ 31% and RMS NOE ≥ 1.0 A are in bold. In clustering all trajectories, the 8 of 14 cluster center structures no longer satisfy the NMR data well, amounting to 26% of structures (another 26% of structures in cluster12 are near the cut-off). * Cluster centers are structures from the two non-converged trajectories.

Clustering of 4 trajectories and rel. population	**Q-factor RDC1**	**Q-factor RDC2**	**RMS_NOE (Å)**	**RMSD to cluster1 (Å)**	**RMSD to cluster2 (Å)**	**RMSD to cluster3 (Å)**
Cluster1 15%	28.0	29.9	0.65	1.74	4.03	4.74
Cluster2 4%	29.5	28.1	0.67	2.68	4.61	4.72
Cluster3 1%	**42.2**	**35.6**	0.44	4.14	7.50	2.46
Cluster4 17%	29.1	**31.9**	0.67	2.29	3.04	5.89
Cluster5 9%	28.1	28.7	0.67	2.07	3.89	5.10
Cluster6 53%	28.8	30.5	0.62	2.54	5.11	4.20
Clustering of 6 trajectories and rel. population	**Q-factor RDC1 (%)**	**Q-factor RDC2 (%)**	**RMS_NOE (Å)**	**RMSD to cluster1 (Å)**	**RMSD to cluster2 (Å)**	**RMSD to cluster3(Å)**
Cluster1 4%	**39.0**	**31.7**	0.70	3.20	5.00	5.00
Cluster2 6%	27.9	29.1	0.63	2.90	3.44	6.28
Cluster3 7%	29.4	30.5	0.47	3.93	4.75	6.07
Cluster4 17%	27.7	28.8	0.66	2.51	3.81	5.70
Cluster5 1%	28.2	30.3	**1.08**	7.12	5.38	9.77
Cluster6 3%	29.2	29.5	**1.20**	2.76	5.06	4.61
Cluster7* 2%	**41.9**	**43.1**	**2.01**	11.38	11.53	8.74
Cluster8* 1%	27.2	29.9	0.94	5.66	8.83	2.98
Cluster9* 5%	**31.0**	28.9	**2.50**	10.85	11.81	8.22
Cluster10* 4%	29.8	30.6	**1.26**	10.40	11.57	7.53
Cluster11 17%	29.3	28.5	0.48	3.91	5.49	5.10
Cluster12 26%	30.6	**31.0**	0.54	3.61	4.11	6.39
Cluster13 4%	**38.0**	**32.8**	0.83	5.48	3.41	8.83
Cluster14 3%	**35.1**	29.1	0.77	3.42	4.47	5.76

Supplemental Figures

Figure S1. Analysis of three trajectories (see Fig. 2 for description of parameters). Trajectories were started with a) a cluster1-like configuration (structure refined without unambiguous NOEs from reference 6), with b) cluster2 and c) the cluster3 structure. The RMSD is calculated from the starting structure (blue) and from the NMR derived cluster1 (green) as the reference.

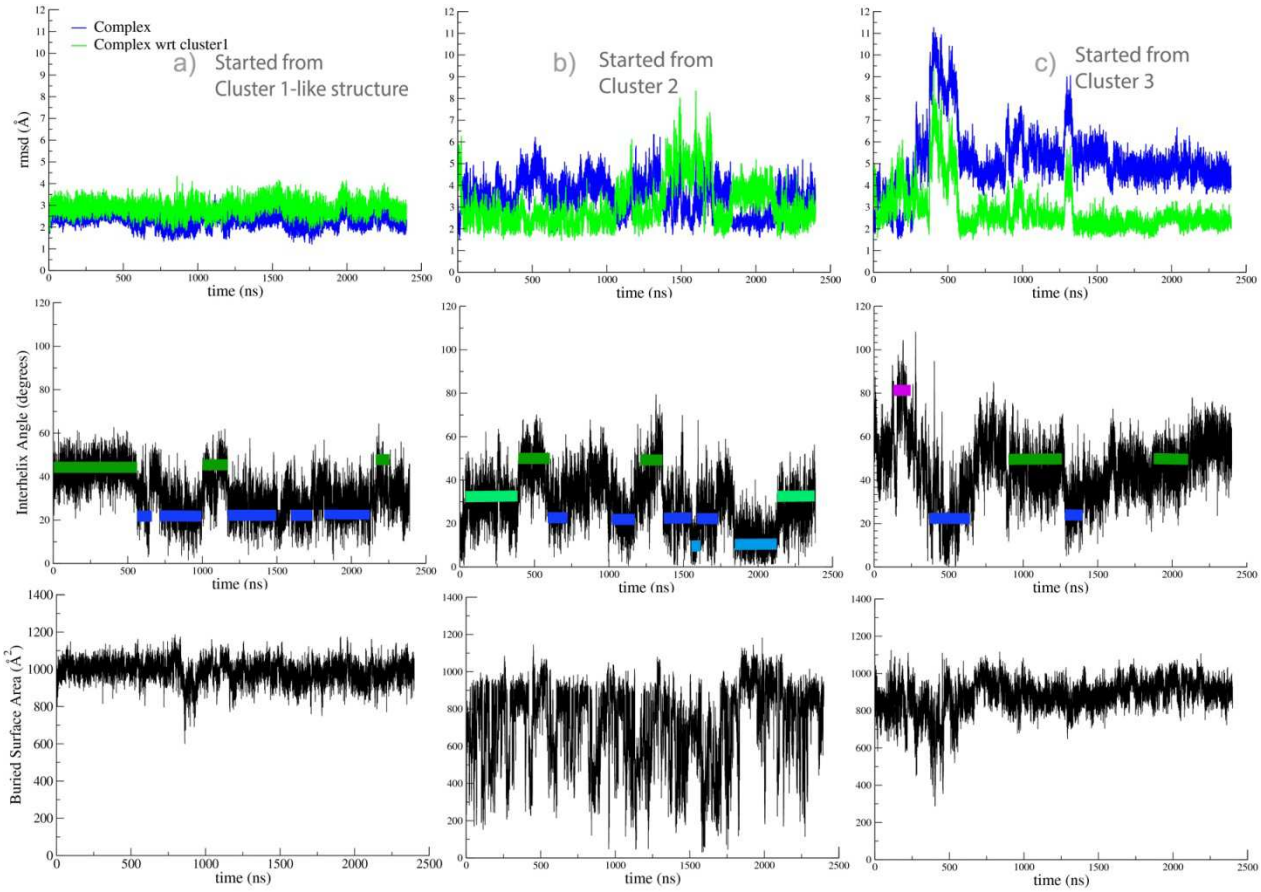


Figure S2. Analysis of two trajectories (see description for Fig. 2 and S1, above). Two trajectories that moved/stayed away from the NMR derived structures a) started with a cluster3 structure and b) started with a structure away from the three NMR derived configurations (a minor configuration in the structure determination without unambiguous restraints, ref. 6). Note the expanded scale for the inter-helix angle compared to Fig. 2 and S1.

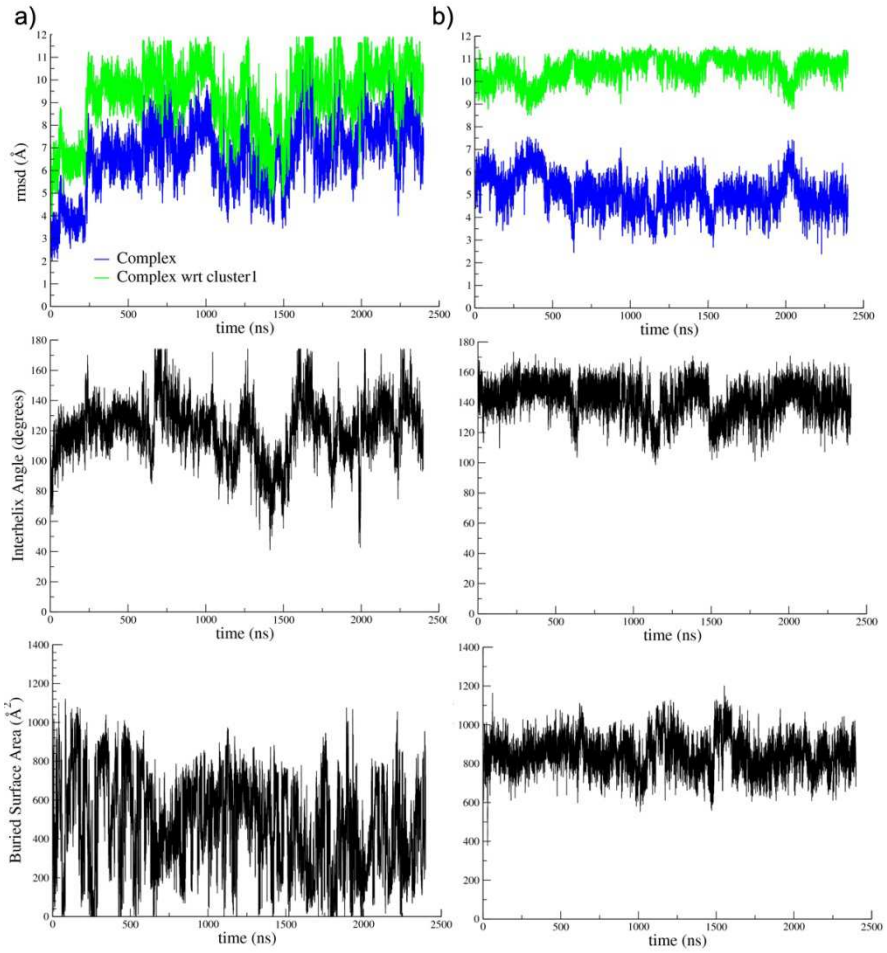
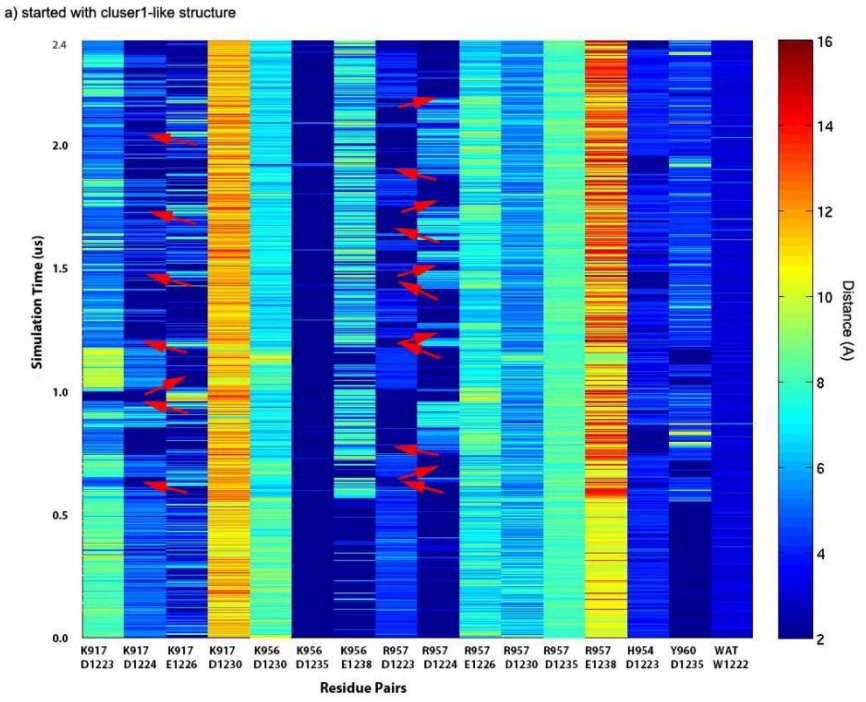
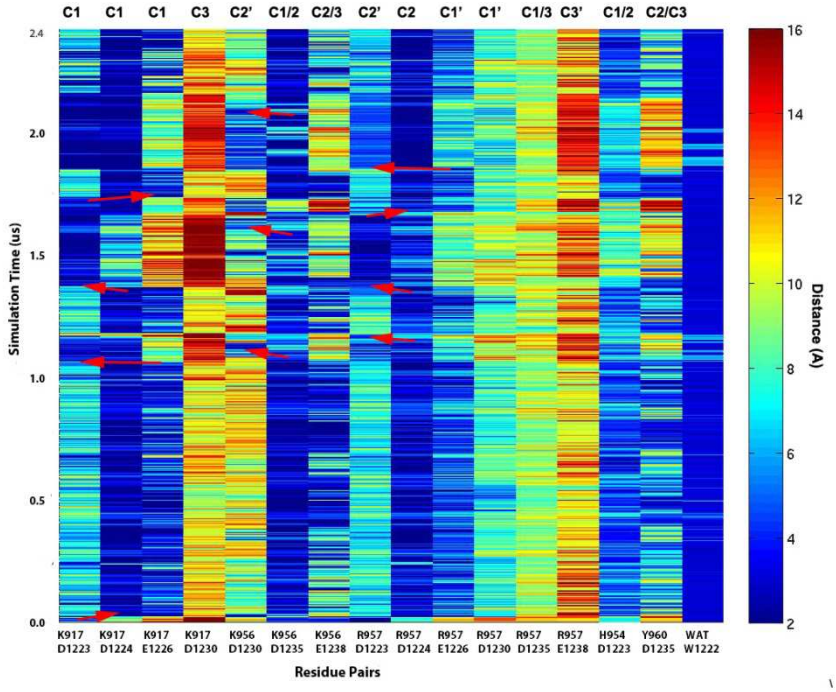


Figure S3. Distances between ion pairs/hydrogen bond donor-acceptor groups across the interface, plotted as a function of simulation time. Distances are indicated by colors (dark blue 2 Å to dark red 16 Å). The trajectories analyzed are those of Fig. S1. a) a cluster1-like configuration (see legend of Fig. S1), b) cluster2 and c) the cluster3 structure. For those contacts that match/are close to those the NMR derived clusters we use the label C1, C2, C3, whereas for those with contacts that partially match or a little away from NMR derived contact pairings we use C1'.



b) started with cluster2 structure



c) started with cluster3 structure

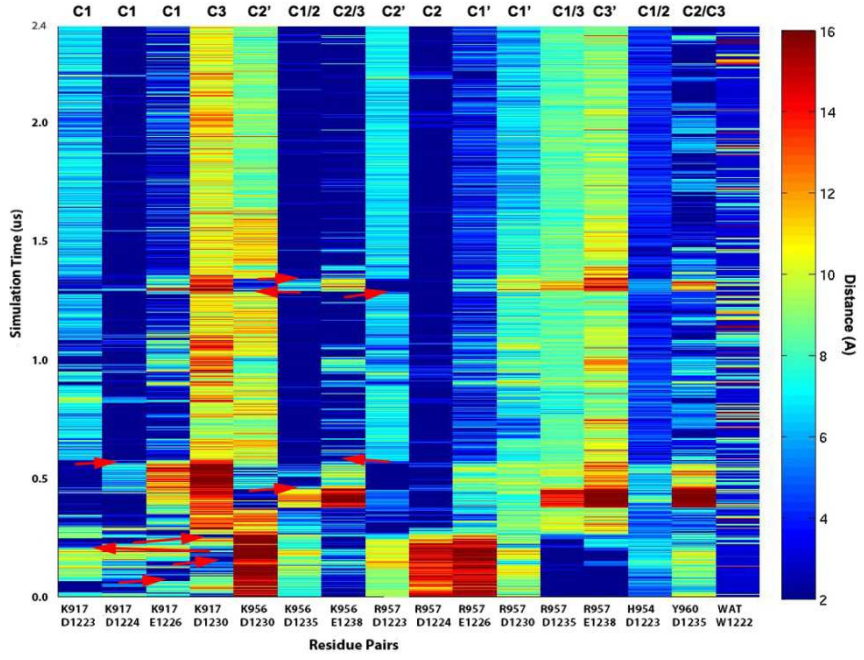


Figure S4. Same as Fig. S3, but for the two trajectories that showed few transitions (see Fig. S2).