Published in final edited form as:
Curr Opin Pharmacol. 2010 December ; 10(6): 738–744. doi:10.1016/j.coph.2010.09.016.

Molecular Dynamics Simulations of Protein Dynamics and their relevance to drug discovery

Freddie R. Salsbury Jr.
Department of Physics, Wake Forest University, 1834 Wake Forest Road, Winston-Salem NC 27106 USA

Abstract

Molecular dynamics simulations have become increasingly useful in studying biological systems of biomedical interest, and not just in the study of model or toy systems. In this article, the methods and principles of all-atom molecular dynamics will be elucidated with several examples provided of their utility to investigators interested on drug discovery.

Introduction

Molecular dynamics simulations of proteins were initially developed in the early 1980's [1] to harness the emerging power of computers to study the motions of proteins and other biopolymers. Molecular dynamics simulations with a wide-variety of different approximations, have been particularly successful in studying the protein folding problem, for example [2,3], and the impact of protein motions on catalysis and ligand binding [4,5]. These latter studies have been particularly influential as they have required considerable discussion of the interplay of conformational change, such as changes in active site geometries in DHFR [4] or metallo-beta-lactamases [5], and coupled protein fluctuations [4,6], which show that within a single protein conformation, long-range coupling networks exist and are sensitive to interactions with different ligands.

These two types of protein dynamics, both of which are thermodynamic in nature, changes in protein conformations, and coupled fluctuations, are the two types of dynamics that are most amenable to study by molecular dynamics simulations, and have become particularly relevant to pharmacology with the development of the concept generalized allostery [7]. This seminal generalization of the classic notion of allostery suggests that virtually all proteins are allosteric in some sense; either due to coupling of conformational changes or due to long-range communication between parts of proteins or protein complexes, and that this allostery can be exploited in the drug-discovery process [7,8]

Basics of Molecular Dynamics Simulations: Structure, Forcefield and Simulation Suite

Classical all-atom molecular dynamics simulations start simple as simulations where Newtonian equation of motions are solved for each atom in the system, and the basics are well-covered in advanced textbooks [9]. These simulations require only three items : initial

coordinates, a potential, and algorithms for propagation. The initial coordinates can be obtained from experimental structures, or from models, or some combination thereof. Systematic studies of the effects of using different sources are rare; a typical solution is to use the best structure available and to compare predictions from simulations with predictions from experiments, such as by comparing NMR order parameters [6,10] or, especially when the initial coordinates involve some model-building, to determine if residues predicted to mediate some functionally important interaction actually affect function when those residues are mutated and the function assayed [11]. The potential is obtained from a force field along with the coordinates. Conceptually, a force field is simple; a parameterization of the energy surface of the protein. However, given the complexity of the details of protein structure, protein force fields are divided into multiple terms, built up from model-systems and transferability is assumed. Not surprisingly, given that there are many different ways to develop model-systems and to parameterize potential energy surfaces, there are many different force field models. The most common currently used are the include the latest CHARMM [12] AMBER, [13] and GROMACS [14] force fields. These force fields, which are associated with particular modeling suites, have been used to simulate a wide-variety of macromolecules. There is little consensus to suggest that one is preferable over the others for protein simulations, and often simulations performed on the same structure with different force fields generate consistent results, for example [15;5,6]

The choice of forcefield is therefore usually a combination of personal preference and choice of molecular simulation suite. The four most commonly used simulation suites are the CHARMM [16], AMBER [17], GROMACS [18] and NAMD [19] suites. These packages share common basic features, but vary in their capacities, and underlying philosophies. CHARMM at one extreme is a very complete modeling program that requires mastery of a fairly complex scripting language, but within which one can conduct a wide-variety of simulations and perform the widest variety of simulation analyses. However, this flexibility comes at a cost; a steep learning curve, although there is an ongoing attempt to alievate this with a GUI [28], and poorer parallel performance then, for example, NAMD. NAMD is in many ways the opposite of CHARMM in usability; it has a much simpler scripting language and reduced functionally. However, the reduction in functionally is in analysis and simulation methods, but does contain the simulation methods needed for classical all-atom simulations. Of the four packages, NAMD is the most capable of performing large classical all-atom simulations, and has been used to perform simulate particularly large proteins and protein-complexes, for examples [8,21,22,23]. GROMACS and AMBER are closer to NAMD in scope and complexity, with GROMACS not using a scripting language, and possessing a large number of external tools for trajectory analysis. Gromacs has the advantage of being the only one of these four that is open-source.

Parameters and Solvation in Molecular Dynamics Simulations

The next crucial step in using all-atom molecular dynamics is to decide upon a solvent model; two typical choices are between a Generalized Born implicit solvent model (GB) and explicit solvent model. The second-generation of the GB models [24,25,26], which can be found in both CHARMM and AMBER, have shown great promise and application over the past several years [27-31], and are under active development [32,33]. These models have two key features: 1) the use of a semi-empirical potential which approximates the interaction between two charges within a protein (or more precisely two charges in a single dielectric material embedded within another dielectric) by extrapolating between two exact forms: the infinity-separated ion pair and the unified ion, and 2) approximating the Born radius, i.e., the energy of placing a single charge in a dielectric-embedded in water. Both approximations have been under active development by attempting to best approximate solutions, at least numerically to the Possion-Boltzmann equation [24,26,33-35]. The use of these approximations allows for longer

simulations due to reduced computational need; there is no explicit computations of the dynamics of water molecules.

Despite the great promise of the latest generation GB models, there has been no systematic study of the influence of using an implicit solvent model on protein dynamics and there have been indicates that the GB models are still lacking the accuracy needed for precise analysis of conformations, especially of protein complexes [36-38].

The simplest and most common solvent model used is explicit solvation; in which the water molecules and counterions are explicitly-represented in the forcefield. This model can be built, for exampke, using VMD's [http://www.ks.uiuc.edu/Research/vmd/plugins/] solvate and autoionize packages. The current “gold-standard” simulation protocol would be to then use a constant-pressure and temperature algorithm, particle mesh ewald and images; all of which are defaults in standard publically-avaliable NAMD scripts (http://www.ks.uiuc.edu/Training/Tutorials/), for example. This combinations models the protein-water-ion systems a periodic system in which the long-range electrostatics are calculated on a grid with fourier transforms and where the pressure and temperature are kept approximately constant with a thermostat and a barostat.

Analysis of Molecular Dynamics Simulations

With modern computers and modern algorithms, moderately-large complexes can be simulated for tens of nanoseconds in full atomic detail; for example, ten million time steps (20ns) of a 151940 atom system consisting of two proteins with 1829 amino acids, a strand of DNA with water molecules and counterions can readily be simulated in five weeks on 6 quad-core Intel 2.4GHz processors using NAMD and the CHARMM27 forcefield [unpublished work]. Such a simulation would result in up to 10 million distinct protein conformations for analysis; though since saving all those conformations would require ~17TB of disk space, a common approach is to save every 100th conformation and then to remove the water and counterions as uninteresting; so that the number of conformations (100,000) is more manageable and the disk space demands are more reasonable (~34GB). The sheer number of conformations that can be sampled with modern computers and algorithms require increasingly sophisticated methods for analysis. These methods can generally be divided up into four types: 1) gross measures of protein and simulation stability, 2) clustering analysis, 3) quasiharmonic and principal component analysis, and 4) correlation function analysis.

The first, gross measures of protein and simulation stability, are used to check the simulation integrity and estimate the equilibration timescale of the simulation; such as 1) calculations of root-mean-square-deviation (RMSD) to quantitate how much the protein has changed, or how much the protein conformations vary during the course of the simulation, 2) calculations of average temperature and pressure and their fluctuations during the simulation to determine if the simulation is physically-realistic or if some error has occurred during the simulations, likely in the setup. These measures are standard and otherwise uninteresting for dynamics analysis.

Clustering analysis is commonly used to discover protein conformations in protein simulations, for example [39]. Clustering analysis is most commonly done with respect to coordinate RMSD, though can be done with other measures to such as in phi/psi space, The RMSD of each conformation relative to another is calculated, so that for a full simulation as discussed above would require the calculation of nearly 50 million RMSDs, which could easily require more computer time than the original simulation. Though not all degrees of freedom need be including a clustering analysis, for example, only a binding site could be included and the rest of the protein disregarded, or only a subset of conformations could be examined, say every 1000th structure. The resulting matrix of RMSD is then used to divide the conformations into

disjoint groups using either a k-means or hierarchical process. In a k-means approach, a specific radius is specified so that, for example, all conformations within 1 Angstrom are clustered together. Whereas, in the hierarchical approach, the entire set of conformations is divided into groups, and then each group is subdivided and so on until a specified stopping criteria is met; such as when the number of clusters reaches a certain predefined number [40,41].

The result of either clustering analysis is a set of clusters, perhaps arranged hierarchically, with populations associated to each cluster and with each conformation identified to a specific cluster. In principal, the specific populations could be used to estimate free-energy differences between different conformations by estimating Boltzmann weights, but this is rarely done as in both approaches clustering analysis is somewhat ad hoc, and different, equally valid, choices of clustering can result in different clusters. Furthermore, there is also the question of how to best represent the cluster; the center structure is commonly used to as a best approximation of the conformations contained within a specific cluster as opposed to the average structure in the cluster to ensure that the cluster representative is a physically-reasonable structure. While clustering analysis represents the current standard method of identifying specific protein conformations, it is indeed lacking, and development of more rigorous and less ad-hoc methods of conformational analysis remain a challenge for the field.

Clustering is one method of reducing the size of the problem of conformational analysis by partitioning the conformations sampled into disjoint sets, other methods can address this problem not by reducing the conformations sampled but by reducing the number of degrees of freedom. Two such methods include principal components analysis and quasiharmonic analysis, both of which are used to select the “most-important” dynamical modes of the system [42,43]. While these methods do not address the issue of identification of specific conformations, instead they address of issue of what are the most important motions of the protein, a useful complement.

In principal-component analysis, a matrix of atomic fluctuations is first constructed, and then this matrix is diagonalized to find obtain modes; linear-combinations of atomic fluctuations that account for the fluctuations observed in the system. While consideration of all the modes are necessary to account for all the fluctuations, the lowest modes typically account for most of the fluctuations, allowing for a possible reduction in the number of degrees of freedom that need be considered. [42] Quasiharmonic analysis is virtually identical except for mass-weighting, which results in obtaining exact normal modes when the system is purely harmonic. [43] Although since force-fields are highly non-harmonic, there appears to be no real reason to prefer quasiharmonic analysis over principal component analysis.

Correlations functions, in general, are simply a measure of how correlated two fluctuating quantities are over time, see for example [9]. While there are many different types of correlation functions that can be constructed, the most common types are bond-vector correlations and correlations of atomic fluctuations. Bond-vector correlations can be used to calculate NMR order parameters for quantitative comparison to experiments and characterizations of internal motions of proteins [10], whereas correlations of atomic fluctuations are used to quantify the correlations of motions of different residues in a protein, such as in [4,6]. A correlation function is a time-series and a full set of correlation functions would have a time-series for each unique residue pair in a protein. This is a wealth of detailed information about the time-evolution of the protein, however, due to difficulty in analyizing such a large quantity of data, usually plots are made of the first time-point of each atomic correlation function, resulting in a correlated motion (or more properly a correlated fluctuation) matrix. [4,6] Such a matrix will demonstrate whether distant parts of proteins, or protein complexes, move together, and increasingly more and more proteins have been shown to exhibit such highly non-local coupled motions. [4,6, 23,39]

Examples relevant to Pharamcology

Depending on the questions posed via different types of analysis and the simulations performed molecular dynamics simulations may be useful in pharmacology in a variety of cases. One case would be enabling the analysis of communication within proteins and protein-complexes [6, 10,23,44-46] with a future aim of enabling the development of drugs that might disrupt such communication. Another case lies in analysis of the effects of known ligands on protein dynamics; these studies help to elucidate the effects of ligand binding and can also provide a cautionary note for structure-based drug design [10,47]. A third case is when molecular dynamics simulations of can model protein conformations associated with specific cellular function to enable the discovery of conformationally-selective ligands [11,48,49]

Analysis of these communication with proteins, but several studies have attempted to analyze protein communication in detail by either linking correlated-motion matrices to functionally-relevant fluctuating properties [45], or by mapping changes in correlated - motion matrices, due to ligand-binding for example, onto structures [10,46], or by combining correlated motions with information from sequence alignment analysis [23]. These are all potentially powerful methods of merging functional or structural information with measures of dynamics, but have yet to be studied systematically. Although there has been a recent attempt to classify communication within proteins [44], which has the potential to lay the foundation for a general theory.

In one example of the combining functionally-relevant information with measures of protein dynamics [45], molecular dynamics simulations were combined with pK_a calculations of an activated cysteine in the active sites of a protein homodimer. The shifting of the pK_a was known to be critical to activity of this protein. The combined calculations showed that despite the presence of two active sites in the homodimer, that the cysteine residues were dynamically asymmetric; so that only one cysteine pK_a was lowered at any point in time. Although over a time-average, the monomers were nearly identical. The question of how the two active sites were coupled arose immediately; what communicated the states of the active sites to one another?

A pathway analysis method was developed to answer this question, and to find pathways of connected and coupled residues more generally. First, paths of connected residues -- residues close enough to interact -- were identified that connected the two activated cysteine; many such paths exist. Second, paths were found were each residue's psi or phi angle fluctuations correlated significantly with the difference in the active site pKas. The result was that two paths were found that coupled the two active sites together; pathways where the backbone dihedrals fluctuated with the different pKa states of the active sites predicting what residues controlled the communication between the different active sites.

In another study [23], simulations where combined in the study of a large protein/tRNA complex; the MetRS/tRNA complex. In this protein complex, communication occurs between two different protein domains that are well-separated in space (~7nm). Molecular dynamics simulations enabled the determination of the correlated motions between the residues in the MetRS protein. When combined with a network analysis on the simulated structures pathways of communication between the activation site and the anticodon recognition site were found. The detailed paths of communication were found to be consistent with experimental results on mutation in the MetRS/tRNA complex.

Several different studies have mapped correlated motions, and their perturbations due to mutation [50] or ligand-binding onto structures to link structural and thermodynamic changes [6,47]. The ligand-binding studies are more relevant here, and these two recent exhaustive

studies [6,47] have addressed questions relevant to the emergence of drug resistance. The first [6] study examined the effects of a suicide-inhibitor on the dynamics and structure of a metallo-beta-lactamase. These enzymes are involved enabling certain bacteria to become resistant to penicillins, and so have attracted interest as drug targets. Multiple NMR and molecular dynamics studies have been conducted on the enzyme from *B. fragillis* in particular, which is particularly interesting in that it has a active site with two zincs, a flexible loop that has been shown to fold onto the active site, and a deformable second loop that appears to change conformation upon binding. A consistent picture of structural changes upon binding has emerged from the NMR and molecular dynamics simulations [5,6,51] despite the simulations using different program suites, and different force fields including rather different models for the zinc coordination. In this picture, the flexible loop adopts variable conformations in the apo form and distinct different conformations in the bound form, and the deformable loop deforms slightly upon binding. There are also more subtle and slight rearrangements in the zinc coordination spheres. A recent molecular dynamics study [6] examined both these rearrangements in more details and also studied the effects of ligand-binding on protein dynamics as probed by changes in correlated motions. The pairs of residues that exhibited the greatest changes in correlated motion due to ligand binding were visualized on the protein structure and paths of residues with changed correlations were discovered that radiated outward from the two zinc atoms, where the inhibitor bound, through the rest of the protein. Intriguingly some of the most drastic changes in correlated motions occurred in the absence of conformational changes, indicating that dynamical changes can occur even in the absence of structural changes; an observation that has been noted by the Nussinov group as well [52]. This picture of a relatively simple active site that is plastic and which exhibits dynamic coupling to the rest of the protein does provide a cautionary note for structure-based drug-discovery.

A less cautionary study [47] involved the study of multiple drug, Tamiflu and Relenza, to multiple realted proteins; swine influenza A/H1N1, Spanish H1N1, and avian H5N1 flu N1 neuraminidases to also study possible mechanims of drug resistance, in a study which also combined molecular dynamics with electrostatic analysis. In this study a molecular model of the swine influenza A/H1N1 type-I neuraminidase was built based on the avian H5N1 type-I neuraminidase and then all three neuraminidases were simulated in their apo form and with the antivirals, Tamiflu or Relenza separately bound. When considered together the simulations allowed for the identification of conserved and unique drug-protein interactions across all three proteins mediated by hydrogen bonds. These hydrogen bond networks were analyzed to show how mutations could lead to drug resistance by disrupting the protein-drug networks. Furthermore, by examining an electrostatic pathway which is hypothesized to play a role in controlling drug access to the binding pocket amechanism by which another mutation acquires drug resistance was proposed.

A third example in which dynamics, especially conformational changes and changes in coupling, can affect and even be exploited in drug discovery, is in the developing story of the development of the MSH2/MSH6 protein complex as a drug target for chemotherapeutic development [8,11,46,48,49]. This protein complex was found via a combined computational-experimental study [8,11] to be multifunctional, acting as a sensor for both DNA damage and mismatches; the complex then recruits additional proteins which lead to either DNA repair or cell-death depending on the protein conformation. These initial studies lead to the thought that it might be possible to screen for small molecules which bind specifically to the “death” conformation of the protein complex. This was found to be the case [48,49]. The identification of a conformationally-selective ligand required the identification of specific protein conformations via molecular dynamics followed by structure-based screening; in this case the molecular dynamics was required to specifically select only the “death” conformation.

Conclusion

It has become increasingly evident that molecular dynamics simulations have become valuable tools in probing details of protein structure and dynamics. The programs and methods underlying simulations have become increasingly mature, leaving the much of the future in the application and analysis of simulations. Increasingly large and proteins and protein-complexes are amenable to simulation, and more attention has been paid recently to the development of force-fields appropriate for simulations with a wide variety of drug molecules [53].

The promise is great for using molecular dynamics in drug-discovery and development, but challenges remain. First, the development of a general method for finding conformations in and determining the free energies from simulations, without the need for somewhat arbitrary clustering methods, while not ignoring the multi-dimensional configuration space that protein move in remains a challenge. This challenge is particularly important to address in order to develop conformationally-selective drugs. Second, the analysis of protein dynamics is often semi-quantitative.; While correlated motion matrices, or any desired correlation function, can readily be calculated quantitatively, the results are often looked at visually and qualitatively. The major difficulty here is the amount of information that can be obtained from simulations,. Hopefully continued progress will be made in the application of statistical methods, and analyses that integrate structural and dynamical information to make the analysis of simulation data more quantitative and routine.

References and Recommended Readings

* of special interest

* * of outstanding interest

1. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature 1980;(267):585–590.
2. Zhang J, Li WF, Wang J, Qin M, Wu L, Yan ZQ, Xu WX, Zuo GH, Wang W. Protein folding simulations: from coarse-grained Model to all-atom model. IUBMB LIFE 2009;61:627–643. [PubMed: 19472192]
3. Cho SS, Weinkam P, Wolynes PG. Origins of barriers and barrierless folding in BBL. PNAS 2009;105:118–123. [PubMed: 18172203]
4. Radkiewicz JL, Brooks CL. Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase. JACS 2000;122:225–231.
5. Salsbury FR, Crowley MF, Brooks CL. Modeling of the metallo-beta-lactamase from B-fragilis: Structural and dynamic effects of inhibitor binding. Proteins 2001;44:448–459. [PubMed: 11484222]
6. Salsbury FR, Crowder MW, Kingsmore SF, Huntley JJA. Molecular dynamic simulations of the metallo-beta-lactamase from Bacteroides fragilis in the presence and absence of a tight-binding inhibitor. J Mol Model 2009;15:133–145. [PubMed: 19039608]
7. Gunasekaran K, Ma BY, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Proteins 2004;57:433–443. [PubMed: 15382234]
8. Vasilyeva A, Clodfelter JE, Rector B, Hollis T, Scarpinato KD, Salsbury FR. Small molecule induction of MSH2-dependent cell death suggests a vital role of mismatch repair proteins in cell death. DNA Repair 2009;8:103–113. [PubMed: 18955167] **An account of combining computational modeling with cell biology to discover a small molecule that selectively targets a specific protein conformation associated with a specific cellular function.
9. Schlick, T. Molecular Modeling and Simulation: An Interdisciplinary Guide. Springer; 2010.
10. Wrabl JO, Shortle D, Woolf TB. Correlation between changes in nuclear magnetic resonance order parameters and conformational entropy: Molecular dynamics simulations of native and denatured staphylococcal nuclease. Proteins 2000;38:123–133. [PubMed: 10656260]

11. Salsbury FR, Clodfelter JE, Gentry MB, Hollis T, Scarpinato KD. The molecular mechanism of DNA damage recognition by MutS homologs and its consequences for cell death response. Nuc Acids Res 2006;34:2173–2185.
12. MacKerell AD Jr. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
13. Ponder JW, Case DA. Force fields for protein simulations. Adv Prot Chem 2003;66:27–85.
14. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 2004;25:1656–1676. [PubMed: 15264259]
15. Price DJ, Brooks CL. Modern protein force fields behave comparably in molecular dynamics simulations. J Comput Chem 2002;23:1045–1057. [PubMed: 12116391]
16. Brooks BR, et al. CHARMM: The Biomolecular simulation Program. J Comp Chem 2009;30:1545–1615. [PubMed: 19444816] *An overview of a particularly flexible simulation program.
17. Case DA, et al. The Amber biomolecular simulation programs. J Comp Chem 2005;26:1668–1688. [PubMed: 16200636]
18. Hess B, Kutzner C, van der Spoel D, et al. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4:435–447.
19. Phillips JC, et al. Scalable molecular dynamics with NAMD. J Comp Chem 2009;26:1781–1802. [PubMed: 16222654]
20. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. J Comput Chem 2008;29:1859–1865. [PubMed: 18351591]
21. Gumbart J, Trabuco LG, Schreiner E, Villa E, Schulten K. Regulation of the protein-conducting channel by a bound ribosome. Structure 2009;17:1453–1464. [PubMed: 19913480] *A large-scale modeling effort of a ribosome-protein complex involving multiple types of simulation resulting in nearly 50ns of simulation of 2.7 million atoms.
22. Yao Y, Harrison CB, Freddolino PL, Schulten K, Mayer ML. Molecular mechanisms of ligand recognition by NR3 subtype glutamate receptors. EMBO Journal 2008;27:2158–2170. [PubMed: 18636091]
23. Ghosh A, Vishveshwara S. A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis. PNAS 2007;104:15711–15716. [PubMed: 17898174]
24. Lee MS, Salsbury FR, Brooks CL. Novel generalized Born methods. J Chem Phys 2002;14:10606–10614.
25. Mongan J, Simmerling C, McCammon JA, Case DA, Onufrit A. Generalized Born with a simple, robust molecular volume correctio. J Chem Theory Comput 2006;3:156–169.
26. Lee MS, Feig M, Salsbury FR, Brooks CL. New analytic approximation to the standard molecular volume definition and its application to generalized born calculations. J Comp Chem 2003;24:1348–1356. [PubMed: 12827676]
27. Thomas LR, Henson A, Reed JC, Salsbury FR, Thorburn A. Direct binding of Fas-associated death domain (FADD) to the tumor necrosis factor-related apoptosis-inducing ligand receptor DR5 is regulated by the death effector domain of FADD. J Biol Chem 2004;31:32780–32785. [PubMed: 15173180]
28. Khalili M, Wales DJ. Computer simulations of peptides from the p53 DNA binding domain. J Chem Theory Comput 2009;5:1380–1392.*A detailed analysis of three implicit solvents and their ability to predict stable structures for peptides. The models disagree although the CHARMM/GBMV model agrees best with experiment.
29. Ganguly D, Chen JH. Atomistic Details of the Disordered States of KID and pKID. Implications in Coupled Binding and Folding. JACS 2009;131:5214–5223.
30. Liu J, Nussinov R. Allosteric effects in the marginally stable von Hippel-Lindau tumor suppressor protein and allostery-based rescue mutant design. PNAS 2008;105:901–906. [PubMed: 18195360]
31. Ma B, Levine AJ. Probing potential binding modes of the p53 tetramer to DNA based on the symmetries encoded in p53 response elements. Nuc Acids Res 2007;35:7733–7747.
32. Salsbury FR. Analysis of errors in Still's equation for macromolecular electrostatic solvation energies. Mol Phys 2006;104:1299–1309.

33. Sigalov G, Fenley A, Onufriev A. Analytical electrostatics for biomolecules: Beyond the generalized Born approximation. J Chem Phys 2006;124:124902. [PubMed: 16599720]
34. Chen JH. Effective Approximation of Molecular Volume Using Atom-Centered Dielectric Functions in Generalized Born Models. J Chem Theory Comput 2010;6:2790–2803.
35. Wang J, Tan CH, Tan YH, Lu Q, Luo R. Poisson-Boltzmann solvents in molecular dynamics simulations. Comm Computer Physics 2008;3:1010–1031.*A recent review of the efforts in implicit solvation.
36. Roe DR, Okur A, Wickstrom L, Hornak V, Simmerling C. Secondary Structure Bias in Generalized Born Solvent Models: Comparison of Conformational Ensembles and Free Energy of Solvent. J Phys Chem B 2007;111:1846–1857. [PubMed: 17256983]
37. Huang ZA, Wong CF. Docking Flexible Peptide to Flexible Protein by Molecular Dynamics Using Two Implicit-Solvent Models: An Evaluation in Protein Kinase and Phosphatase Systems. J Phys Chem B 2009;113:14343–14354. [PubMed: 19845408] *An analysis of protein-peptide docking using two different implicit solvent models. These results should be compared to those of [28], as the conclusions in this study disagree.
38. Yeh IC, Wallqvist A. Structure and Dynamics of End-to-End Loop Formation of the Penta-Peptide Cys-Ala-Gly-Gln-Trp in Implicit Solvents. J Phys Chem B 2009;113:12382–12390. [PubMed: 19685925]
39. Knaggs MH, Salsbury FR, Edgell MH, Salsbury FR. Insights into correlated motions and long-range interactions in CheY derived from molecular dynamics simulations. Biophys J 2007;6:2062–2079. [PubMed: 17172298]
40. Murtagh F. A survey of recent advances in hierarchical clustering algorithms. The Computer Journal 1983;26:354–359.
41. Kelley LA, Gardner SP, Sutcliffe MJ. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Eng 1996;9:1063–1065. [PubMed: 8961360]
42. Madei A. Essential Dynamics of Proteins. Proteins 1993;17:412–423. [PubMed: 8108382]
43. Case DA. Normal-mode analysis of protein dynamics. Curr Opin Struct Bio 1994;4:285–290.
44. Tsai CJ, Del Sol A, Nussinov R. Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. Mol Biosys 2009;5:207–216.**A recent review article on possible mechanism of allostery.
45. Yuan Y, Knaggs MH, Poole LB, Fetrow JS, Salsbury FR. Conformational and Oligomeric Effects on the Cysteine pK(a) of Tryparedoxin Peroxidase. J Biomol Struct Dyn 2010;28:51–70. [PubMed: 20476795] **Pathways are predicted purely based on simulation data, linking two active sites in a homodimer that influence the pKa, a functionally-relevant quantity.
46. Salsbury FR. Effects of Cisplatin Binding to DNA on the Dynamics of the E. coli MutS Dimer. Prot & Pept Lett 2010;17:744–750.
47. Le L, Lee EH, Schulten K, Truong T. Molecular modeling of swine influenza A/H1N1, Spanish H1N1, and avian H5N1 flu N1 neuraminidases bound to Tamiflu and Relenza. PLoS Currents: Influenza 2009;1:RRN1015.**All-atoms simulations combined with molecular models to specifically examine how mutations may lead to drug resistance.
48. Drotschmann K, Topping RR, Clodfelter JE, Salsbury FR. Mutations in the nucleotide-binding domain of MutS homologs uncouple cell death from cell survival. DNA Repair 2004;3:729–742. [PubMed: 15177182]
49. Vasilyeva A, Clodfelter JE, Gorczynski MJ, Gerardi AR, King SB, Salsbury FR, Scarpinato KD. Parameters of Reserpine Analogs That Induce MSH2/MSH6-Dependent Cytotoxic Response. J Nucleic Acids 2010:162018. [PubMed: 20936178]
50. Fetrow JS, Knutson ST, Edgell MH. Mutations in alpha-helical solvent-exposed sites of eglin c have long-range effects: evidence from molecular dynamics simulations. Proteins 2006;63:356–72. [PubMed: 16342264]
51. Scrofani SDB, Chung J, Huntley JJA, Benkovic SJ, Wright PE, Dyson HJ. NMR characterization of the metallo-beta-lactamase from Bacteroides fragilis and its interaction with a tight-binding inhibitor: Role of an active-site loop. Biochem 1999;37:14507–14514. [PubMed: 10545172]

52. Tsai CJ, del Sol A, Nussinov R. Allostery: Absence of a Change in Shape Does Not Imply that Allostery Is Not at Play. J Mol Biol 2008;378:1–11. [PubMed: 18353365]

53. Knight JL, Brooks CL. Validating CHARMM Parameters and Exploring Charge Distribution Rules in Structure-Based Drug Design. J Chem Theory Comput 2006;5:1680–1691. [PubMed: 20046995]