

Review

Insights from Coarse-Grained Gō Models for Protein Folding and Dynamics

Ronald D. Hills, Jr.¹ and Charles L. Brooks, III^{1,2,*}

¹ Department of Molecular Biology and Kellogg School of Science and Technology, The Scripps Research Institute, 10550 N. Torrey Pines Rd. TPC6 La Jolla, CA 92037, USA

² Department of Chemistry and Biophysics Program, University of Michigan, 930 N. University Ave, Ann Arbor, MI 48109, USA

* Author to whom correspondence should be addressed; E-Mail: brookscl@umich.edu;
Tel. +1-734-647-6682; Fax: +1-734-647-1604

*Received: 22 January 2009; in revised form: 23 February 2009 / Accepted: 26 February 2009 /
Published: 2 March 2009*

Abstract: Exploring the landscape of large scale conformational changes such as protein folding at atomistic detail poses a considerable computational challenge. Coarse-grained representations of the peptide chain have therefore been developed and over the last decade have proved extremely valuable. These include topology-based Gō models, which constitute a smooth and funnel-like approximation to the folding landscape. We review the many variations of the Gō model that have been employed to yield insight into folding mechanisms. Their success has been interpreted as a consequence of the dominant role of the native topology in folding. The role of local contact density in determining protein dynamics is also discussed and is used to explain the ability of Gō-like models to capture sequence effects in folding and elucidate conformational transitions.

Keywords: Protein folding; Gō models; coarse-graining; energy landscape; conformational transitions.

1. Introduction

Advancing our understanding of biology in the 21st century will involve the structural and dynamical characterization of the countless molecular interactions in the cell. Much can be learned of the structural and dynamical properties of proteins, the workhorses of the cell, from their folding energy landscape. Molecular simulation has emerged as a valuable tool for the detailed exploration of the thermodynamic landscape responsible for guiding the unfolded ensemble through a collection of high energy intermediate structures in order to achieve the proper native structure. We examine recent insights obtained from a simple class of coarse-grained simulation methods for landscape characterization, known as Gō models.

Protein folding is a complex process involving large numbers of degrees of freedom. Therefore, performing traditional molecular dynamics simulations at atomistic detail poses a challenge if one wishes to exhaustively explore the folding landscape of a moderately sized protein with finite computational resources. For this reason coarse-grained representations of the polypeptide chain have been employed in folding simulations. Rather than representing each atom in the protein explicitly groups of atoms can be treated as a single coarse-grained site. In the most common variation, each residue corresponds to a single coarse-grained site placed at the alpha carbon [1-9], although models with multiple sites per residue have been employed [10-14]. Single-site per residue C_α models then make up beads on a string, which are held together with harmonic energy terms in the Hamiltonian for the resulting virtual bond and angle vibrational degrees of freedom. Virtual dihedral angles comprising four successive residues can be subjected to potential functions of varying forms, but all share a similar intent of mimicking backbone chirality and Ramachandran conformational preferences [15]. The underpinning of a Gō model, however, is its treatment of nonbonded interactions.

Early coarse-grained models of folding employed the HP recipe, in which hydrophobic residues interact with a constant favorable contact energy (denoted ϵ) and polar residues either don't interact or are repulsive. HP models enjoyed limited success as they predicted degenerate global energy minima and failed to identify the unique native state [16]. Current coarse-graining developments endeavor to augment such two-flavor models to contain up to 20 flavors, using distinct physicochemical types to represent each amino acid sidechain or multiple coarse-grained sites per residue for larger sidechains [17-20]. Parameterization of such models to ensure consistency with the all-atom thermodynamic ensemble is a challenging problem. Recent empirical implementations rely on dihedral or hydrogen bond restraints to maintain native secondary structure and are thus of limited use for studying protein conformational changes and folding [21-24]. Two ongoing approaches by the Voth [25] and Scheraga [26] groups offer a promising solution to this problem by employing direct parameterization with all-atom simulation data in order to obtain physically accurate peptide equilibria.

In lieu of a coarse-grained model with physically accurate nonbonded interactions, simplistic Gō models have been applied to the study of folding mechanisms with considerable success. An attractive term of the Lennard-Jones variety [12] is employed for pairwise contacts present in a protein's native structure, which consists of a shallow energy basin of depth ϵ that decays to zero for large distances. Residue-residue contacts are typically assigned for nonhydrogen sidechain atoms separated by less than a cutoff distance [12,27], and contacts can also be assigned for backbone hydrogen bonds [2]. All other (non-native) interactions experience a volume-exclusion repulsion. The topology-based approach

of favorably distinguishing between native and non-native interactions is named after the early lattice simulation work of Nobuhiro Gō [28-36], and any potential that is biased toward a target structure can be referred to as Gō-like. In Section 2 we discuss the success of Gō models in describing folding. Section 3 follows with recent insights concerning the role of sequence in folding, and, finally, an outlook is presented for the coarse-graining of conformational transitions.

2. Role of Native Contacts in Folding

Energy landscape theory posits that proteins have evolved a smooth and funnel-like thermodynamic landscape biased toward the native state, allowing them to fold to a unique stable structure on biological timescales [37-39]. A polypeptide does not have sufficient time to explore all possible conformations in search of the native state; therefore native interactions, whether local or nonlocal [40,41], must play a role in guiding the conformational search in the direction of native structure. As topology-centric Gō models let native interactions drive the folding, their large success has been interpreted as a validation of energy landscape theory.

The reduced degrees of freedom in a coarse-grained representation enable one to exhaustively explore the thermodynamic landscape for folding using standard molecular simulation techniques. In particular, high energy regions of conformational space can be characterized such as the transition state or folding intermediate(s) using either brute-force searching at the folding transition temperature or umbrella sampling with carefully chosen biasing restraints [42]. Gō model simulations can thus be readily compared with experimental phi-value analysis of the thermodynamic effects of point mutations. A phi-value is a qualitative measure of the amount of native-like structure a residue contains in the folding transition state [43]. Phi-values can be predicted for each residue in the protein by averaging over all conformations recorded in the simulation trajectory that correspond to the transition state ensemble. The distribution of native-like secondary and tertiary structure present in the transition state can in part describe the relative order of formation of structural elements, the location of the folding nucleus and hence the overall folding mechanism.

Experimental phi-values have been compared in the literature with predictions from simple C_{α} Gō models for a rapidly growing number of proteins [1-4,8,41,42,44-54]. Qualitative consensus is generally obtained between predictions and experiment for individual residue phi-values as well as the major features of the folding mechanism and folding rate [55]. Non-native interactions have been demonstrated to make some important contributions to the folding transition state and mechanism, and variants of the Gō algorithm have been modified to incorporate nonspecific interactions [4,10,13,56-62]. Notwithstanding, the widespread success of native-only Gō models in reproducing folding behavior as observed in experiment and atomistic simulation [63] suggests that the perturbations of non-native interactions on the structures of folding transition states and intermediates are minor in comparison to the dominating influence of native interactions on the funneled energy landscape [64].

3. Sequence Dependence of Folding

Over the last decade the folding community has emphasized the native topology as the main determinant of folding behavior [39,65]. The topology-centric paradigm has resulted from the rise of energy landscape theory, the prevalence and success of Gō modeling studies, and the notion of contact

order. Defined as the average sequence separation between contacting residues in the native state, contact order is inversely correlated with experimental folding rates, allowing kinetic predictions to be made based on the local or nonlocal nature of complex backbone topologies. Small helical proteins are known to fold faster than large β -sheet proteins due to the more local nature of their contacts [66,67]. Nonetheless, it is a protein's sequence that determines the particular topology it chooses to adopt. Single residue mutations can significantly alter the kinetics of folding and the stability of the native state. Amyloid disease-causing mutations lead to partially unfolded forms that self-associate to form pathogenic oligomers and aggregates [68,69]. The precise role of sequence in folding remains to be understood.

Methods have been employed for extending simple G \ddot{o} models to examine sequence effects in folding. Rather than weighting native contacts equally by assigning the same interaction energy ϵ to all contacts, energetic heterogeneity can be introduced by scaling residue-residue interactions according to physically motivated criteria. Native interactions have been optimized in Src-SH3, CI2, protein G and S6 using thermodynamic perturbation data from mutation experiments [64,70]. Clementi and coworkers optimized native and non-native contact energies in Src-SH3 using an iterative procedure to maximize the energy gap between the folded and unfolded states [57]. Dokholyan and coworkers scaled contact energies based on atomistic simulations with the CHARMM force field that identified a small set of critical residues in SOD1 folding [71]. Karanicolas and Brooks [2] employed a general set of residue-residue interactions using the tabulated potentials of Miyazawa and Jernigan (MJ) derived from statistical analysis of contacts in the Protein Data Bank [72]. Native contact energies have also been weighted by crystallographic B-factors [73] and hydrogen exchange protection factors [74] with success. One last method that can be considered to constitute a 'flavored' G \ddot{o} model is the all-atom G \ddot{o} approach [75-85]. The protein is represented at atomic resolution, and nonbonded atom-atom pairs within a cutoff distance are assigned a favorable contact potential. As larger sidechains are more likely to be assigned a higher number of atom-atom contacts, they are apt to interact more favorably than smaller sidechains.

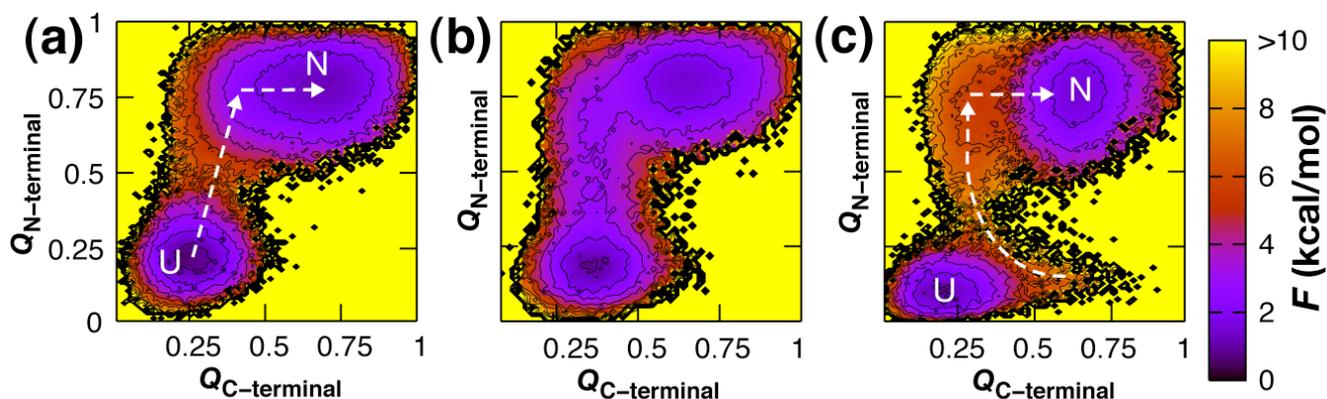
Energetic heterogeneity adds ruggedness to the folding landscape and, more importantly, allows for the investigation of sequence effects in, for example, the folding of proteins with identical topology. The MJ-flavored model of Karanicolas and Brooks was used to investigate symmetry breaking in protein L and protein G, two proteins that share a common topology consisting of N- and C-terminal β -hairpins and a central α -helix. The balance between enthalpic interactions and chain entropy was observed to lead to N-terminally nucleated folding in protein L and C-terminal nucleation in protein G, consistent with experimental findings [2]. The same model was recently employed to compare the folding mechanisms for three members of the common flavodoxin fold [42,86,87]. CheY, NtrC and Spo0F each consist of five $\beta\alpha$ -repeats arranged into a central β -sheet and surrounding α -helices. Consistent with experimental findings for CheY [88,89], all three proteins were observed to obey an N-terminal nucleation mechanism in simulations due to abundant contacts in the N-terminus. A larger number of contacts in the C-terminal subdomain of Spo0F were observed to lead to some differences in its folding mechanism, however.

While Clementi *et al.* had previously observed N-terminal nucleation in simulations of CheY with a simple G \ddot{o} model [1], a more detailed analysis was undertaken by Hills *et al.* to compare the role of sequence in CheY homologs [42,86,87]. The free energy landscape was characterized as a function of

the formation of the various secondary and tertiary structural elements within the 120-residue proteins in order to map out the relative order of structure formation. Though more rigorous computational methods exist for determining the folding transition state ensemble [90,91], using structural coordinates to monitor folding progress in simulations has proven valuable in the elucidation of complex folding mechanisms in large proteins when multiple reaction coordinates are needed to explain folding behavior [49,92,93]. Transition path sampling methods were recently employed to analyze the reaction coordinate for the 20-residue Trp-cage mini-protein [94]. Two structural coordinates were needed to best describe its parallel folding pathways of helix and tertiary contact formation. The fraction of native contacts formed, denoted Q , is a progress variable of considerable utility [92,95-97]. Folding in CheY and homologs was shown to be well described by native contact formation in the N- and C-terminal subdomains [42,86,87].

While CheY and NtrC contain an alanine-rich C-terminal cavity that is dynamic in the unphosphorylated state and important for function, in Spo0F this region is filled in with bulkier residues such as isoleucine and exhibits rigidity, having lost the C-terminal conformational allostery of its evolutionary homologs [98-101]. The enhanced stability and van der Waals contacts in this region were shown to lead to competitive frustration in the folding landscape (Figure 1). Whereas in CheY and NtrC the N-terminal subdomain formed easily and rapidly nucleated C-terminal folding, in Spo0F C-terminal contacts were seen to form ahead of and temporarily preclude N-terminal folding.

Figure 1. N-terminally nucleated folding landscapes for structural homologs CheY (a), NtrC (b) and Spo0F (c). The equilibrium free energy was computed as a function of the fraction of native contacts formed within the N- ($Q_{N\text{-terminal}}$) and C-terminal ($Q_{C\text{-terminal}}$) subdomains and is shown at the transition temperature at which the folded and unfolded states are equally populated [42,87].



Simulation results are in accord with experimental analysis of the three homologs; the kinetics of folding were found to be slowest in Spo0F, suggesting an interesting relationship between function and folding mechanism in the CheY family [87,88]. Comparison of the folding mechanisms for the CheY family, T4 lysozyme and interleukin-1 β suggests that subdomain competition is a general property of multimodule proteins [42,47,102]. It is interesting to note that the contact energies assigned by the flavored G \ddot{o} model were uniformly distributed throughout the three CheY-like proteins. The local contact density, defined as the number of native contacts per residue, could alone explain folding

behavior, implying that an unflavored G δ model could also have possibly captured the main sequence effects in the three homologs. This notion is supported by a recent comparison study of results obtained from flavored and unflavored G δ models [103]. For β -sheet proteins in which tertiary contacts are well distributed throughout the protein, introducing energetic heterogeneity was observed to negligibly impact the folding mechanism; for α -helical bundles held together by only a handful of tertiary interactions, interaction heterogeneity was observed to alter folding.

4. Conclusions and Outlook

The role of contact density in dictating protein folding and dynamics is increasingly being recognized. Contact density has been used to explain B-factors [104] and downhill folding [54]. A comprehensive survey of experimental and simulation data for homodimers revealed that high interface hydrophobicity and a large ratio of interfacial to monomeric contacts are indicative of an obligate two-state association mechanism, in which dimerization precedes monomer folding [50,51]. Two-state dimers commonly have an intertwined or pseudo domain-swapped interface. Cooperativity in hydrophobic packing was shown to be sufficient for nucleating amyloid aggregation [105]. Residue fluctuations and the location of folding nuclei have been explained by treating proteins as small-world networks consisting of a regular lattice with a fraction of the edges between vertices replaced with new random connections [106-112]. Clusters of the branched aliphatic sidechains isoleucine, leucine and valine have been identified as particularly important in stability and folding [87,88,113,114]. In order to predict hydrogen exchange protection factors, a combinatorial algorithm was used to generate an ensemble of unfolded conformational states and the thermodynamic probability of each state was estimated using considerations of chain entropy and solvent exposed surface area [115]. A survey of large two-subdomain proteins revealed that native contact density can distinguish stable, nucleating subdomains from dynamic subdomains involved in function [86]. Correlations between contact density and folding/dynamics help to explain the success of coarse-grained models in biomolecular simulation. Nowhere is this success more surprising than in the case of analytical Ising models of folding that consider chain entropy and native contact enthalpy while neglecting the three-dimensional spatial representation of the peptide chain [116-120].

The role of native contacts in protein dynamics is best appreciated in the context of elastic network normal mode analysis (EN-NMA). To calculate vibrational modes in large systems proteins can be represented as an elastic network of C α -beads in which residues within a cutoff distance experience harmonic interactions. The lowest frequency normal modes of vibration are then easily computed for the elastic network model. Often, a single or small subset of low frequency vibrations has been shown to reproduce the large amplitude structural rearrangements associated with biological function [19,121-123]. EN-NMA has successfully reproduced the functional dynamics of systems as diverse as adenylate kinase, myosin, RNA polymerase, the ribosome, GroEL-GroES and viral capsids [124,125]. The success of EN-NMA in capturing functional motions has been interpreted as a consequence of nature's exploitation of shape, topology and packing in the evolution of protein conformational dynamics [124,126-128].

A disadvantage of EN-NMA is that 'closed' to 'open' conformational transitions cannot be observed as they would involve the breaking of harmonic bonds in the network. Alternative methods

exist for computing the minimum free energy path between two conformational states based on interpolation between the two reference elastic networks for each state. Optimization algorithms can be employed to find between the two global minima either a single minimum energy pathway [129] or an ensemble of low energy paths reflecting the ruggedness of protein landscapes [130]. The double-well approach has recently been employed successfully with Gō models, in which native interaction potentials were included for residue contacts present in either of two reference conformations [131-133]. The so-called dual, or switching, Gō model was used to study conformational transitions in myosin, Arc repressor and F1-ATPase.

The pervasive role of contact density can explain the robustness of the many varieties of Gō models that have been used to study folding. The results obtained with a Gō model are sensitive to the model parameters and various energy functional forms used [5,6,12,134-136]. A common means to safeguard against this is to compare folding results for multiple proteins obtained using the same model in light of available experimental data. Simulation results can also be influenced by the experimental structure from which the Gō model is derived, particularly the method used in its solution [137,138]. Note that NtrC exhibits a lower folding barrier than CheY and Spo0F in Figure 1, a result of having fewer native contacts overall and therefore a smoother landscape. The sparser native contacts in NtrC can be attributed to its being based on an NMR solution structure as opposed to the crystal structures for CheY and Spo0F [87]. It is also not uncommon for Gō models to be employed with native dihedral and/or angle restraints in addition to the typical nonbonded native interactions in order to make folding more cooperative [1,139]. Introducing a small desolvation barrier in the nonbonded native interaction potential has been employed as an alternate means of ensuring folding cooperativity [2,12,61,140].

Sulkowska and Cieplak recently compared a total of 62 different variations of the Gō potential in reproducing the experimental mechanical unfolding data of 28 proteins [12]. Several of the common functional forms employed for the native contact potential along with both uniform and heterogeneous energy scales were able to accurately reproduce the maximum unfolding force in stretching simulations ($R^2 \approx 0.84$). Most of the other models investigated came in a close second tier with $R^2 > 0.70$ for the correlation between predicted and experimental unfolding forces. Whitford *et al.* have compared results obtained from C_α and all-atom Gō models of three proteins [85]. While the C_α models neglected sidechain packing considerations, the correct global folding mechanisms were still obtained. Such studies suggest that the many variations of Gō models that have been employed to date are more alike than they are dissimilar.

Current non-Gō coarse-grained models depend on some form of secondary structure restraints to ensure native state stability [21-24]. Improving the physical accuracy of these non-Gō interactions will be a necessary step in removing biasing potentials from coarse-grained models so that large scale conformational transitions can be more readily studied. Native structure restraints aside, empirical coarse-graining techniques can often be considered Gō-like for their very nature of neglecting detailed atomistic interactions whilst favoring the essential interactions of the system [141-144]. In lieu of a physically accurate and transferable non-Gō coarse-grained model for proteins, Gō-like models will continue to find application in the study of protein folding and dynamics. The mechanical unfolding of titin, ubiquitin, bacteriorhodopsin and knotted transcarbamylase has recently been simulated using Gō potentials [76,145-149]. A Gō modeling survey of 7,510 structures in the Protein Data Bank was used to further delineate the topological determinants of mechanical stability [150]. A 297-residue outer

membrane protease has been studied using a hybrid atomistic/coarse-grained approach in which the active site was represented with an atomistic force field and the protein scaffold was described using a Gō model [151-153]. Symmetrized Gō potentials have been employed for domain swapping and amyloid aggregation in which intermonomer residue contacts are treated identical to intramonomer native contacts [13,14,154]. Gō model simulations have been performed to explore the influence of macromolecular crowding agents [155,156] and confinement [62] on folding. Electrostatic interactions were incorporated between a Gō model of a transcription factor and coarse-grained DNA, which caused the protein to visit partially unfolded forms and suggested a fly casting mechanism of target recognition [157]. Lastly, in light of the connections with contact density, the link between function and folding mechanism is becoming a topic of considerable investigation [86,158-164].

Acknowledgements

R.D.H. thanks the Burroughs Wellcome Fund La Jolla Interfaces in Science Training Program for financial support, the Center for Theoretical Biological Physics (ctbp.ucsd.edu) for providing a stimulating intellectual environment and D.A. Case for providing laboratory space. C.L.B. acknowledges support from NIH grant GM48807. Present address for Dr. Hills: Center for Biophysical Modeling and Simulation, University of Utah, 315 S 1400 E HEB2020, Salt Lake City, Utah 84112.

References and Notes

1. Clementi, C.; Nymeyer, H.; Onuchic, J.N. Topological and energetic factors: What determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. *J. Mol. Biol.* **2000**, *298*, 937-953.
2. Karanicolas, J.; Brooks, C.L., III. The origins of asymmetry in the folding transition states of protein L and protein G. *Protein Sci.* **2002**, *11*, 2351-2361.
3. Koga, N.; Takada, S. Roles of native topology and chain-length scaling in protein folding: A simulation study with a Go-like model. *J. Mol. Biol.* **2001**, *313*, 171-180.
4. Paci, E.; Vendruscolo, M.; Karplus, M. Validity of Go models: Comparison with a solvent-shielded empirical energy decomposition. *Biophys. J.* **2002**, *83*, 3032-3038.
5. Prieto, L.; de Sancho, D.; Rey, A. Thermodynamics of Go-type models for protein folding. *J. Chem. Phys.* **2005**, *123*, 154903.
6. Prieto, L.; Rey, A. Influence of the chain stiffness on the thermodynamics of a Go-type model for protein folding. *J. Chem. Phys.* **2007**, *126*, 166103.
7. Rhee, Y.M.; Pande, V.S. On the role of chemical detail in simulating protein folding kinetics. *Chem. Phys.* **2006**, *323*, 66-77.
8. Shea, J.E.; Onuchic, J.N.; Brooks, C.L., III. Exploring the origins of topological frustration: Design of a minimally frustrated model of fragment B of protein A. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12512-12517.
9. Dokholyan, N.V.; Buldyrev, S.V.; Stanley, H.E.; Shakhnovich, E.I. Discrete molecular dynamics studies of the folding of a protein-like model. *Fold. Des.* **1998**, *3*, 577-587.

10. Cheung, M.S.; Finke, J.M.; Callahan, B.; Onuchic, J.N. Exploring the interplay between topology and secondary structural formation in the protein folding problem. *J. Phys. Chem. B* **2003**, *107*, 11193-11200.
11. Lam, A.R.; Borreguero, J.M.; Ding, F.; Dokholyan, N.V.; Buldyrev, S.V.; Stanley, H.E.; Shakhnovich, E. Parallel folding pathways in the SH3 domain protein. *J. Mol. Biol.* **2007**, *373*, 1348-1360.
12. Sulkowska, J.I.; Cieplak, M. Selection of optimal variants of Go-like models of proteins through studies of stretching. *Biophys. J.* **2008**, *95*, 3174-3191.
13. Ding, F.; Dokholyan, N.V.; Buldyrev, S.V.; Stanley, H.E.; Shakhnovich, E.I. Molecular dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. *J. Mol. Biol.* **2002**, *324*, 851-857.
14. Barton, S.; Jacak, R.; Khare, S.D.; Ding, F.; Dokholyan, N.V. The length dependence of the PolyQ-mediated protein aggregation. *J. Biol. Chem.* **2007**, *282*, 25487-25492.
15. Kwiecinska, J.I.; Cieplak, M. Chirality and protein folding. *J. Phys.-Condes. Matter* **2005**, *17*, S1565-S1580.
16. Yue, K.; Fiebig, K.M.; Thomas, P.D.; Chan, H.S.; Shakhnovich, E.I.; Dill, K.A. A test of lattice protein folding algorithms. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 325-329.
17. Han, W.; Wan, C.K.; Wu, Y.D. Toward a coarse-grained protein model coupled with a coarse-grained solvent model: Solvation free energies of amino acid side chains. *J. Chem. Theory Comput.* **2008**, *4*, 1891-1901.
18. Han, W.; Wu, Y.D. Coarse-grained protein model coupled with a coarse-grained water model: Molecular dynamics study of polyalanine-based peptides. *J. Chem. Theory Comput.* **2007**, *3*, 2146-2161.
19. Sherwood, P.; Brooks, B.R.; Sansom, M.S.P. Multiscale methods for macromolecular simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 630-640.
20. Tozzini, V. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144-150.
21. Bond, P.J.; Sansom, M.S.P. Insertion and assembly of membrane proteins via simulation. *J. Am. Chem. Soc.* **2006**, *128*, 2697-2704.
22. Monticelli, L.; Kandasamy, S.K.; Periole, X.; Larson, R.G.; Tieleman, D.P.; Marrink, S.J. The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819-834.
23. Shih, A.Y.; Arkhipov, A.; Freddolino, P.L.; Schulten, K. Coarse grained protein-lipid model with application to lipoprotein particles. *J. Phys. Chem. B* **2006**, *110*, 3674-3684.
24. Yap, E.H.; Fawzi, N.L.; Head-Gordon, T. A coarse-grained alpha-carbon protein model with anisotropic hydrogen-bonding. *Proteins* **2008**, *70*, 626-638.
25. Thorpe, I.F.; Zhou, J.; Voth, G.A. Peptide folding using multiscale coarse-grained models. *J. Phys. Chem. B* **2008**, *112*, 13079-13090.
26. Makowski, M.; Sobolewski, E.; Czaplewski, C.; Oldziej, S.; Liwo, A.; Scheraga, H.A. Simple physics-based analytical formulas for the potentials of mean force for the interaction of amino acid side chains in water. IV. Pairs of different hydrophobic side chains. *J. Phys. Chem. B* **2008**, *112*, 11385-11395.

27. da Silveira, C.H.; Pires, D.E.V.; Minardi, R.C.; Ribeiro, C.; Veloso, C.J.M.; Lopes, J.C.D.; Meira, W.; Neshich, G.; Ramos, C.H.I.; Habesch, R.; Santoro, M.M. Protein cutoff scanning: A comparative analysis of cutoff dependent and cutoff free methods for prospecting contacts in proteins. *Proteins* **2009**, *74*, 727-743.
28. Taketomi, H.; Ueda, Y.; Go, N. Studies on protein folding, unfolding and fluctuations by computer simulation. 1. Effect of specific amino acid sequence represented by specific inter-unit interactions. *Int. J. Pept. Protein Res.* **1975**, *7*, 445-459.
29. Go, N.; Scheraga, H.A. On the use of classical statistical mechanics in the treatment of polymer chain conformations. *Macromolecules* **1976**, *9*, 535-542.
30. Ueda, Y.; Go, N. Theory of large-amplitude conformational fluctuations in native globular proteins: Independent fluctuating site model. *Int. J. Pept. Protein Res.* **1976**, *8*, 551-558.
31. Go, N.; Taketomi, H. Respective roles of short-range and long-range interactions in protein folding. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 559-563.
32. Ueda, Y.; Taketomi, H.; Go, N. Studies on protein folding, unfolding and fluctuations by computer simulation. 2. A three-dimensional lattice model of lysozyme. *Biopolymers* **1978**, *17*, 1531-1548.
33. Go, N.; Taketomi, H. Studies on protein folding, unfolding and fluctuations by computer simulation. 3. Effect of short-range interactions. *Int. J. Pept. Protein Res.* **1979**, *13*, 235-252.
34. Go, N.; Taketomi, H. Studies on protein folding, unfolding and fluctuations by computer simulation. 4. Hydrophobic interactions. *Int. J. Pept. Protein Res.* **1979**, *13*, 447-461.
35. Go, N. Theoretical studies of protein folding. *Annu. Rev. Biophys. Bioeng.* **1983**, *12*, 183-210.
36. Taketomi, H.; Kano, F.; Go, N. The effect of amino acid substitution on protein folding and protein unfolding transition studied by computer simulation. *Biopolymers* **1988**, *27*, 527-559.
37. Brooks, C.L., III; Onuchic, J.N.; Wales, D.J. Statistical thermodynamics: Taking a walk on a landscape. *Science* **2001**, *293*, 612-613.
38. Dill, K.A.; Ozkan, S.B.; Shell, M.S.; Weikl, T.R. The protein folding problem. *Ann. Rev. Biophys.* **2008**, *37*, 289-316.
39. Onuchic, J.N.; Wolynes, P.G. Theory of protein folding. *Curr. Opin. Struct. Biol.* **2004**, *14*, 70-75.
40. Ozkan, S.B.; Wu, G.A.; Chodera, J.D.; Dill, K.A. Protein folding by zipping and assembly. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 11987-11992.
41. Prieto, L.; Rey, A. Influence of the native topology on the folding barrier for small proteins. *J. Chem. Phys.* **2007**, *127*, 175101.
42. Hills, R.D., Jr.; Brooks, C.L., III. Subdomain competition, cooperativity, and topological frustration in the folding of CheY. *J. Mol. Biol.* **2008**, *382*, 485-495.
43. Fersht, A.R.; Sato, S. Phi-Value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7976-7981.
44. Chavez, L.L.; Gosavi, S.; Jennings, P.A.; Onuchic, J.N. Multiple routes lead to the native state in the energy landscape of the beta-trefoil family. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10254-10258.
45. Dokholyan, N.V.; Buldyrev, S.V.; Stanley, H.E.; Shakhnovich, E.I. Identifying the protein folding nucleus using molecular dynamics. *J. Mol. Biol.* **2000**, *296*, 1183-1188.

46. Ferreiro, D.U.; Cho, S.S.; Komives, E.A.; Wolynes, P.G. The energy landscape of modular repeat proteins: Topology determines folding mechanism in the ankyrin family. *J. Mol. Biol.* **2005**, *354*, 679-692.
47. Gosavi, S.; Chavez, L.L.; Jennings, P.A.; Onuchic, J.N. Topological frustration and the folding of interleukin-1 beta. *J. Mol. Biol.* **2006**, *357*, 986-996.
48. Hubner, I.A.; Oliveberg, M.; Shakhnovich, E.I. Simulation, experiment, and evolution: Understanding nucleation in protein S6 folding. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8354-8359.
49. Karanicolas, J.; Brooks, C.L., III. Improved Go-like models demonstrate the robustness of protein folding mechanisms towards non-native interactions. *J. Mol. Biol.* **2003**, *334*, 309-325.
50. Levy, Y.; Cho, S.S.; Onuchic, J.N.; Wolynes, P.G. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *J. Mol. Biol.* **2005**, *346*, 1121-1145.
51. Levy, Y.; Wolynes, P.G.; Onuchic, J.N. Protein topology determines binding mechanism. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 511-516.
52. Roy, M.; Chavez, L.L.; Finke, J.M.; Heidary, D.K.; Onuchic, J.N.; Jennings, P.A. The native energy landscape for interleukin-1 beta. Modulation of the population ensemble through native-state topology. *J. Mol. Biol.* **2005**, *348*, 335-347.
53. Settanni, G.; Hoang, T.X.; Micheletti, C.; Maritan, A. Folding pathways of prion and doppel. *Biophys. J.* **2002**, *83*, 3533-3541.
54. Zuo, G.H.; Wang, J.; Wang, W. Folding with downhill behavior and low cooperativity of proteins. *Proteins* **2006**, *63*, 165-173.
55. Kouza, M.; Li, M.S.; O'Brien, E.P.; Hu, C.K.; Thirumalai, D. Effect of finite size on cooperativity and rates of protein folding. *J. Phys. Chem. A* **2006**, *110*, 671-676.
56. Clementi, C.; Plotkin, S.S. The effects of nonnative interactions on protein folding rates: Theory and simulation. *Protein Sci.* **2004**, *13*, 1750-1766.
57. Das, P.; Matysiak, S.; Clementi, C. Balancing energy and entropy: A minimalist model for the characterization of protein folding landscapes. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10141-10146.
58. Gu, Z.; Rao, M.K.; Forsyth, W.R.; Finke, J.M.; Matthews, C.R. Structural analysis of kinetic folding intermediates for a TIM barrel protein, indole-3-glycerol phosphate synthase, by hydrogen exchange mass spectrometry and Go model simulation. *J. Mol. Biol.* **2007**, *374*, 528-546.
59. Paci, E.; Vendruscolo, M.; Karplus, M. Native and non-native interactions along protein folding and unfolding pathways. *Proteins* **2002**, *47*, 379-392.
60. Zarrine-Afsart, A.; Wallin, S.; Neculai, A.M.; Neudecker, P.; Howell, P.L.; Davidson, A.R.; Chan, H.S. Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 9999-10004.
61. Zhang, Z.; Chan, H.S. Native topology of the designed protein Top7 is not conducive to cooperative folding. *Biophys. J.* **2009**, *96*, L25-L27.
62. Griffin, M.A.; Friedel, M.; Shea, J.E. Effects of frustration, confinement, and surface interactions on the dimerization of an off-lattice beta-barrel protein. *J. Chem. Phys.* **2005**, *123*, 174707.

63. Zhou, Y.Q.; Karplus, M. Folding thermodynamics of a model three-helix-bundle protein. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 14429-14432.
64. Matysiak, S.; Clementi, C. Optimal combination of theory and experiment for the characterization of the protein folding landscape of S6: How far can a minimalist model go? *J. Mol. Biol.* **2004**, *343*, 235-248.
65. Du, R.; Pande, V.S.; Grosberg, A.Y.; Tanaka, T.; Shakhnovich, E. On the role of conformational geometry in protein folding. *J. Chem. Phys.* **1999**, *111*, 10375-10380.
66. Ivankov, D.N.; Garbuzynskiy, S.O.; Alm, E.; Plaxco, K.W.; Baker, D.; Finkelstein, A.V. Contact order revisited: Influence of protein size on the folding rate. *Protein Sci.* **2003**, *12*, 2057-2062.
67. Plaxco, K.W.; Simons, K.T.; Baker, D. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **1998**, *277*, 985-994.
68. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197-208.
69. Thirumalai, D.; Klimov, D.K.; Dima, R.I. Emerging ideas on the molecular basis of protein and peptide aggregation. *Curr. Opin. Struct. Biol.* **2003**, *13*, 146-159.
70. Sutto, L.; Tiana, G.; Broglia, R.A. Sequence of events in folding mechanism: Beyond the Go model. *Protein Sci.* **2006**, *15*, 1638-1652.
71. Khare, S.D.; Ding, F.; Dokholyan, N.V. Folding of Cu, Zn superoxide dismutase and familial amyotrophic lateral sclerosis. *J. Mol. Biol.* **2003**, *334*, 515-525.
72. Miyazawa, S.; Jernigan, R.L. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.* **1996**, *256*, 623-644.
73. Rao, M.K.; Chapman, T.R.; Finke, J.M. Crystallographic B-factors highlight energetic frustration in aldolase folding. *J. Phys. Chem. B* **2008**, *112*, 10417-10431.
74. Dixon, R.D.S.; Chen, Y.W.; Feng, D.; Khare, S.D.; Prutzman, K.C.; Schaller, M.D.; Campbell, S.L.; Dokholyan, N.V. New insights into FAK signaling and localization based on detection of a FAT domain folding intermediate. *Structure* **2004**, *12*, 2161-2171.
75. Clementi, C.; Garcia, A.E.; Onuchic, J.N. Interplay among tertiary contacts, secondary structure formation and side-chain packing in the protein folding mechanism: All-atom representation study of protein L. *J. Mol. Biol.* **2003**, *326*, 933-954.
76. Kleiner, A.; Shakhnovich, E. The mechanical unfolding of ubiquitin through all-atom Monte Carlo simulation with a Go-type potential. *Biophys. J.* **2007**, *92*, 2054-2061.
77. Linhananta, A.; Boer, J.; MacKay, I. The equilibrium properties and folding kinetics of an all-atom Go model of the Trp-cage. *J. Chem. Phys.* **2005**, *122*, 114901.
78. Linhananta, A.; Zhou, Y.Q. The role of sidechain packing and native contact interactions in folding: Discontinuous molecular dynamics folding simulations of an all-atom Go model of fragment B of Staphylococcal protein A. *J. Chem. Phys.* **2002**, *117*, 8983-8995.
79. Meinke, J.H.; Hansmann, U.H.E. Protein simulations combining an all-atom force field with a Go term. *J. Phys.-Condes. Matter* **2007**, *19*, 285215.
80. Shimada, J.; Kussell, E.L.; Shakhnovich, E.I. The folding thermodynamics and kinetics of crambin using an all-atom Monte Carlo simulation. *J. Mol. Biol.* **2001**, *308*, 79-95.

81. Shimada, J.; Shakhnovich, E.I. The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11175-11180.
82. Zhou, Y.Q.; Linhananta, A. Thermodynamics of an all-atom off-lattice model of the fragment B of Staphylococcal protein A: Implication for the origin of the cooperativity of protein folding. *J. Phys. Chem. B* **2002**, *106*, 1481-1485.
83. Luo, Z.L.; Ding, J.D.; Zhou, Y.Q. Temperature-dependent folding pathways of pin1 WW domain: An all-atom molecular dynamics simulation of a Go model. *Biophys. J.* **2007**, *93*, 2152-2161.
84. Luo, Z.L.; Ding, J.D.; Zhou, Y.Q. Folding mechanisms of individual beta-hairpins in a Go model of Pin1 WW domain by all-atom molecular dynamics simulations. *J. Chem. Phys.* **2008**, *128*, 225103.
85. Whitford, P.C.; Noel, J.K.; Gosavi, S.; Schug, A.; Sanbonmatsu, K.Y.; Onuchic, J.N. An all-atom structure-based potential for proteins: Bridging minimal models with all-atom empirical forcefields. *Proteins*. In press, **2009**, doi:10.1002/prot.22253.
86. Hills, R.D., Jr.; Brooks, C.L., III. Coevolution of function and the folding landscape: Correlation with density of native contacts. *Biophys. J.* **2008**, *95*, L57-L59.
87. Hills, R.D., Jr.; Kathuria, S.V.; Wallace, L.A.; Matthews, C.R.; Brooks, C.L., III. Functional diversity and shifting cores of stability modulate the folding mechanisms of flavodoxin fold proteins. *J. Mol. Biol.* **2009**, Pending publication.
88. Kathuria, S.V.; Day, I.J.; Wallace, L.A.; Matthews, C.R. Kinetic traps in the folding of beta alpha-repeat proteins: CheY initially misfolds before accessing the native conformation. *J. Mol. Biol.* **2008**, *382*, 467-484.
89. LopezHernandez, E.; Serrano, L. Structure of the transition state for folding of the 129 aa protein CheY resembles that of a smaller protein, CI-2. *Fold. Des.* **1996**, *1*, 43-55.
90. Du, R.; Pande, V.S.; Grosberg, A.Y.; Tanaka, T.; Shakhnovich, E.S. On the transition coordinate for protein folding. *J. Chem. Phys.* **1998**, *108*, 334-350.
91. Snow, C.D.; Rhee, Y.M.; Pande, V.S. Kinetic definition of protein folding transition state ensembles and reaction coordinates. *Biophys. J.* **2006**, *91*, 14-24.
92. Cho, S.S.; Levy, Y.; Wolynes, P.G. P versus Q: Structural reaction coordinates capture protein folding on smooth landscapes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 586-591.
93. Piana, S.; Laio, A. Advillin folding takes place on a hypersurface of small dimensionality. *Phys. Rev. Lett.* **2008**, *101*, 208101.
94. Juraszek, J.; Bolhuis, P.G. Rate constant and reaction coordinate of Trp-cage folding in explicit water. *Biophys. J.* **2008**, *95*, 4246-4257.
95. Shakhnovich, E. Protein folding thermodynamics and dynamics: Where physics, chemistry, and biology meet. *Chem. Rev.* **2006**, *106*, 1559-1588.
96. Shakhnovich, E.; Farztdinov, G.; Gutin, A.M.; Karplus, M. Protein folding bottlenecks: A lattice Monte-Carlo simulation. *Phys. Rev. Lett.* **1991**, *67*, 1665-1668.
97. Faccioli, P. Characterization of protein folding by dominant reaction pathways. *J. Phys. Chem. B* **2008**, *112*, 13756-13764.

98. Sola, M.; Lopez-Hernandez, E.; Cronet, P.; Lacroix, E.; Serrano, L.; Coll, M.; Parraga, A. Towards understanding a molecular switch mechanism: Thermodynamic and crystallographic studies of the signal transduction protein CheY. *J. Mol. Biol.* **2000**, *303*, 213-225.
99. Stock, A.M.; Guhaniyogi, J. A new perspective on response regulator activation. *J. Bacteriol.* **2006**, *188*, 7328-7330.
100. Varughese, K.I.; Tsigelny, I.; Zhao, H.Y. The crystal structure of beryll fluoride Spo0F in complex with the phosphotransferase Spo0B represents a phosphotransfer pretransition state. *J. Bacteriol.* **2006**, *188*, 4970-4977.
101. Volkman, B.F.; Lipson, D.; Wemmer, D.E.; Kern, D. Two-state allosteric behavior in a single-domain signaling protein. *Science* **2001**, *291*, 2429-2433.
102. Cellitti, J.; Bernstein, R.; Marqusee, S. Exploring subdomain cooperativity in T4 lysozyme II: Uncovering the C-terminal subdomain as a hidden intermediate in the kinetic folding pathway. *Protein Sci.* **2007**, *16*, 852-862.
103. Cho, S.S.; Levy, Y.; Wolynes, P.G. Quantitative criteria for native energetic heterogeneity influences in the prediction of protein folding kinetics. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 434-439.
104. Halle, B. Flexibility and packing in proteins. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1274-1279.
105. Hills, R.D., Jr.; Brooks, C.L., III. Hydrophobic cooperativity as a mechanism for amyloid nucleation. *J. Mol. Biol.* **2007**, *368*, 894-901.
106. Aftabuddin, M.; Kundu, S. Hydrophobic, hydrophilic, and charged amino acid networks within protein. *Biophys. J.* **2007**, *93*, 225-231.
107. Atilgan, A.R.; Akan, P.; Baysal, C. Small-world communication of residues and significance for protein dynamics. *Biophys. J.* **2004**, *86*, 85-91.
108. Bode, C.; Kovacs, I.A.; Szalay, M.S.; Palotai, R.; Korcsmaros, T.; Csermely, P. Network analysis of protein dynamics. *FEBS Lett.* **2007**, *581*, 2776-2782.
109. Dokholyan, N.V.; Li, L.; Ding, F.; Shakhnovich, E.I. Topological determinants of protein folding. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8637-8641.
110. Greene, L.H.; Higman, V.A. Uncovering network systems within protein structures. *J. Mol. Biol.* **2003**, *334*, 781-791.
111. Higman, V.A.; Greene, L.H. Elucidation of conserved long-range interaction networks in proteins and their significance in determining protein topology. *Physica A* **2006**, *368*, 595-606.
112. Vendruscolo, M.; Dokholyan, N.V.; Paci, E.; Karplus, M. Small-world view of the amino acids that play a key role in protein folding. *Phys. Rev. E* **2002**, *65*, 061910-1-4.
113. Kathuria, S.V.; Matthews, C.R. Clusters of isoleucine, leucine and valine side chains define cores of stability in globular proteins: Sequence determinants of structure, stability and folding. **2009**, Pending publication.
114. Wu, Y.; Vadrevu, R.; Kathuria, S.; Yang, X.Y.; Matthews, C.R. A tightly packed hydrophobic cluster directs the formation of an off-pathway sub-millisecond folding intermediate in the alpha subunit of tryptophan synthase, a TIM barrel protein. *J. Mol. Biol.* **2007**, *366*, 1624-1638.
115. Hilser, V.J.; Freire, E. Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J. Mol. Biol.* **1996**, *262*, 756-772.

116. Alm, E.; Baker, D. Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11305-11310.
117. Galzitskaya, O.V.; Finkelstein, A.V. A theoretical search for folding/unfolding nuclei in three-dimensional protein structures. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11299-11304.
118. Karanicolas, J.; Brooks, C.L., III. The importance of explicit chain representation in protein folding models: An examination of Ising-like models. *Proteins* **2003**, *53*, 740-747.
119. Munoz, V.; Eaton, W.A. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11311-11316.
120. Takada, S. Go-ing for the prediction of protein folding mechanisms. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11698-11700.
121. Doruker, P.; Jernigan, R.L.; Bahar, I. Dynamics of large proteins through hierarchical levels of coarse-grained structures. *J. Comput. Chem.* **2002**, *23*, 119-127.
122. Tama, F.; Sanejouand, Y.H. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **2001**, *14*, 1-6.
123. Petrone, P.; Pande, V.S. Can conformational change be described by only a few normal modes? *Biophys. J.* **2006**, *90*, 1583-1593.
124. Tama, F.; Brooks, C.L., III. Symmetry, form, and shape: Guiding principles for robustness in macromolecular machines. *Annu. Rev. Biophys. Biomolec. Struct.* **2006**, *35*, 115-133.
125. Yang, L.; Song, G.; Jernigan, R.L. How well can we understand large-scale protein motions using normal modes of elastic network models? *Biophys. J.* **2007**, *93*, 920-929.
126. Bagci, Z.; Jernigan, R.L.; Bahar, I. Residue packing in proteins: Uniform distribution on a coarse-grained scale. *J. Chem. Phys.* **2002**, *116*, 2269-2276.
127. Keskin, O.; Jernigan, R.L.; Bahar, I. Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophys. J.* **2000**, *78*, 2093-2106.
128. Lu, M.Y.; Ma, J.P. The role of shape in determining molecular motions. *Biophys. J.* **2005**, *89*, 2395-2401.
129. Maragakis, P.; Karplus, M. Large amplitude conformational change in proteins explored with a plastic network model: Adenylate kinase. *J. Mol. Biol.* **2005**, *352*, 807-822.
130. Chu, J.W.; Voth, G.A. Coarse-grained free energy functions for studying protein conformational changes: A double-well network model. *Biophys. J.* **2007**, *93*, 3860-3871.
131. Best, R.B.; Chen, Y.G.; Hummer, G. Slow protein conformational dynamics from multiple experimental structures: The helix/sheet transition of arc repressor. *Structure* **2005**, *13*, 1755-1763.
132. Koga, N.; Takada, S. Folding-based molecular simulations reveal mechanisms of the rotary motor F-1-ATPase. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5367-5372.
133. Takagi, F.; Kikuchi, M. Structural change and nucleotide dissociation of myosin motor domain: Dual Go model simulation. *Biophys. J.* **2007**, *93*, 3820-3827.
134. Cieplak, M.; Hoang, T.X. Universality classes in folding times of proteins. *Biophys. J.* **2003**, *84*, 475-488.
135. Kaya, H.; Chan, H.S. Solvation effects and driving forces for protein thermodynamic and kinetic cooperativity: How adequate is native-centric topological modeling? *J. Mol. Biol.* **2003**, *326*, 911-931.

136. Kaya, H.; Liu, Z.R.; Chan, H.S. Chevron Behavior and isostable enthalpic barriers in protein folding: Successes and limitations of simple Go-like modeling. *Biophys. J.* **2005**, *89*, 520-535.
137. Prieto, L.; Rey, A. Simulations of the protein folding process using topology-based models depend on the experimental structure. *J. Chem. Phys.* **2008**, *129*, 115101.
138. Rey-Stolle, M.F.; Enciso, M.; Rey, A. Topology-based models and NMR structures in protein folding simulations. *J. Comp. Chem.* In press. **2009**, doi:10.1002/jcc.21149.
139. Settanni, G.; Cattaneo, A.; Maritan, A. Role of native-state topology in the stabilization of intracellular antibodies. *Biophys. J.* **2001**, *81*, 2935-2945.
140. Cheung, M.S.; Garcia, A.E.; Onuchic, J.N. Protein folding mediated by solvation: Water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 685-690.
141. Bellesia, G.; Shea, J.E. Self-assembly of beta-sheet forming peptides into chiral fibrillar aggregates. *J. Chem. Phys.* **2007**, *126*, 245104.
142. Fawzi, N.L.; Kohlstedt, K.L.; Okabe, Y.; Head-Gordon, T. Protofibril assemblies of the arctic, dutch, and flemish mutants of the Alzheimer's A beta(1-40) peptide. *Biophys. J.* **2008**, *94*, 2007-2016.
143. Nguyen, H.D.; Hall, C.K. Spontaneous fibril formation by polyalanines; Discontinuous molecular dynamics simulations. *J. Am. Chem. Soc.* **2006**, *128*, 1890-1901.
144. Nguyen, H.D.; Reddy, V.S.; Brooks, C.L., III. Deciphering the kinetic mechanism of spontaneous self-assembly of icosahedral capsids. *Nano Lett.* **2007**, *7*, 338-344.
145. Cieplak, M.; Filipek, S.; Janovjak, H.; Krzysko, K.A. Pulling single bacteriorhodopsin out of a membrane: Comparison of simulation and experiment. *Biochim. Biophys. Acta-Biomembr.* **2006**, *1758*, 537-544.
146. Cieplak, M.; Hoang, T.X.; Robbins, M.O. Folding and stretching in a Go-like model of titin. *Proteins* **2002**, *49*, 114-124.
147. Sulkowska, J.I.; Sulkowski, P.; Szymczak, P.; Cieplak, M. Stabilizing effect of knots on proteins. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19714-19719.
148. Sulkowska, J.I.; Kloczkowski, A.; Sen, T.Z.; Cieplak, M.; Jernigan, R.L. Predicting the order in which contacts are broken during single molecule protein stretching experiments. *Proteins* **2008**, *71*, 45-60.
149. Li, M.S.; Kouza, M.; Hu, C.K. Refolding upon force quench and pathways of mechanical and thermal unfolding of ubiquitin. *Biophys. J.* **2007**, *92*, 547-561.
150. Sulkowska, J.I.; Cieplak, M. Mechanical stretching of proteins: A theoretical survey of the Protein Data Bank. *J. Phys.-Condes. Matter* **2007**, *19*, 283201.
151. Neri, M.; Anselmi, C.; Carnevale, V.; Vargiu, A.V.; Carloni, P. Molecular dynamics simulations of outer-membrane protease T from E-coli based on a hybrid coarse-grained/atomistic potential. *J. Phys.-Condes. Matter* **2006**, *18*, S347-S355.
152. Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. Coarse-grained model of proteins incorporating atomistic detail of the active site. *Phys. Rev. Lett.* **2005**, *95*, 218102.
153. Neri, M.; Baaden, M.; Carnevale, V.; Anselmi, C.; Maritan, A.; Carloni, P. Microseconds dynamics simulations of the outer-membrane protease T. *Biophys. J.* **2008**, *94*, 71-78.

154. Yang, S.C.; Levine, H.; Onuchic, J.N. Protein oligomerization through domain swapping: Role of inter-molecular interactions and protein concentration. *J. Mol. Biol.* **2005**, *352*, 202-211.
155. Homouz, D.; Perham, M.; Samiotakis, A.; Cheung, M.S.; Wittung-Stafshede, P. Crowded, cell-like environment induces shape changes in aspherical protein. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11754-11759.
156. Pincus, D.L.; Thirumalai, D. Crowding effects on the mechanical stability and unfolding pathways of ubiquitin. *J. Phys. Chem. B* **2009**, *113*, 359-368.
157. Levy, Y.; Onuchic, J.N.; Wolynes, P.G. Fly-casting in protein-DNA binding: Frustration between protein folding and electrostatics facilitates target recognition. *J. Am. Chem. Soc.* **2007**, *129*, 738-739.
158. Ferreiro, D.U.; Hegler, J.A.; Komives, E.A.; Wolynes, P.G. Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19819-19824.
159. Gosavi, S.; Whitford, P.C.; Jennings, P.A.; Onuchic, J.N. Extracting function from a beta-trefoil folding motif. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10384-10389.
160. Jager, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendle, M.; Bowman, M.E.; Noel, J.P.; Gruebele, M.; Kelly, J.W. Structure-function-folding relationship in a WW domain. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10648-10653.
161. Miyashita, O.; Onuchic, J.N.; Wolynes, P.G. Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12570-12575.
162. Miyashita, O.; Wolynes, P.G.; Onuchic, J.N. Simple energy landscape model for the kinetics of functional transitions in proteins. *J. Phys. Chem. B* **2005**, *109*, 1959-1969.
163. Whitford, P.C.; Miyashita, O.; Levy, Y.; Onuchic, J.N. Conformational transitions of adenylate kinase: Switching by cracking. *J. Mol. Biol.* **2007**, *366*, 1661-1671.
164. Karanicolas, J.; Brooks, C.L., III. Integrating folding kinetics and protein function: Biphasic kinetics and dual binding specificity in a WW domain. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3432-3437.