Available online at www.sciencedirect.com
SCIENCE DIRECT®

Peptide folding simulations

S Gnanakaran, Hugh Nymeyer, John Portman, Kevin Y Sanbonmatsu and Angel E García*

Developments in the design of small peptides that mimic proteins in complexity, recent advances in nanosecond time-resolved spectroscopy methods to study peptides and the development of modern, highly parallel simulation algorithms have come together to give us a detailed picture of peptide folding dynamics. Two newly implemented simulation techniques, parallel replica dynamics and replica exchange molecular dynamics, can now describe directly from simulations the kinetics and thermodynamics of peptide formation, respectively. Given these developments, the simulation community now has the tools to verify and validate simulation protocols and models (forcefields).

Addresses
Theoretical Biology and Biophysics Group, Theoretical Division, T10 MS K710, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
*e-mail: axg@lanl.gov

Current Opinion in Structural Biology 2003, **13**:168–174

This review comes from a themed issue on
Theory and simulation
Edited by Charles L Brooks III and David A Case

0959-440X/03/$ – see front matter

DOI 10.1016/S0959-440X(03)00040-X

Abbreviations

GB	generalized Born
NOE	nuclear Overhauser effect
PRD	parallel replica dynamics
REMD	replica exchange molecular dynamics
rmsd	root mean square deviation
SA	surface area

Introduction

Understanding the dynamics and mechanism of protein folding continues to be one of the central problems in molecular biology. Peptide folding simulations and experiments characterize the dynamics and molecular mechanisms of the early events of protein folding. Computationally, peptides present a more tractable system than proteins. Experimentally, peptides fold at very fast rates, requiring probing on the nanosecond time resolution. Peptides offer a unique opportunity to bridge the gap between theoretical and experimental understanding of protein folding.

Peptides have many of the features and complexities of proteins. In general, the competition between configurational entropy, hydrogen bond formation, solvation, hydrophobic core formation and ion pair formation determines the folding rate and stability of proteins. This competition plays an essential role throughout the folding process and determines the thermodynamic equilibrium between folded and unfolded states. Modeling this competition is a standing challenge in peptide folding simulations.

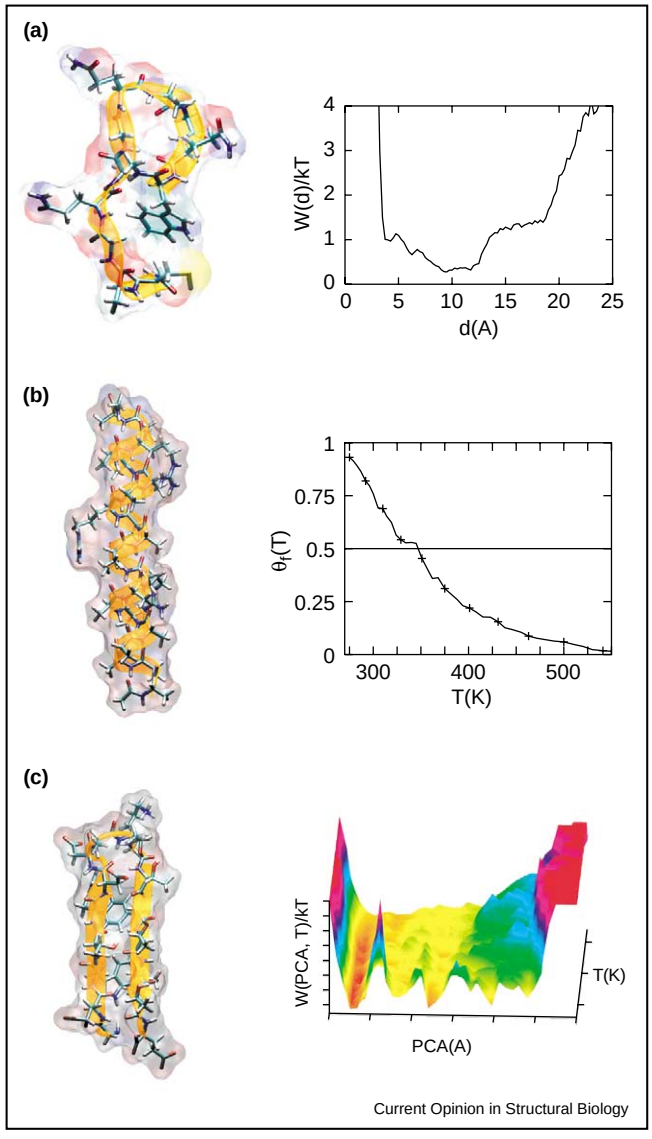
Three major developments have positioned the simulation community to make significant advances toward understanding the mechanism of folding of peptides and small proteins. Firstly, the design of small peptides that mimic proteins in complexity, but are sufficiently small to allow detailed simulation studies [1–4]. Secondly, the development of fast (nanosecond) time-resolved spectroscopy methods to study peptide folding dynamics on the same timescale as computer simulations [5–8,9•,10••,11,12]. Thirdly, the development and implementation of simulation algorithms have helped to overcome the limitations of insufficient sampling [13,14]. Given these conditions, it becomes particularly important to emphasize the validation of simulation techniques and forcefields, and the verification of the simulation results.

In what follows, we briefly describe newly implemented, highly parallel simulation techniques that allow the sampling of the kinetics and thermodynamics of peptides directly from molecular simulations. We describe how these techniques have been used to study peptide systems. We divide the systems according to the timescales that characterize the formation of basic structures. The fastest events are described by loop closing (10 ns timescale), α-helix formation (200 ns timescale), β-hairpin folding (1–10 μs timescale) and mini-protein folding (1–10 μs timescale). Examples of systems that form on these timescales and that have been studied in detail are shown in Figure 1. Given the extent of the simulations being conducted, we now have a good opportunity to validate the simulation techniques and forcefields by doing careful comparisons of the simulation and experimental results. We briefly describe instances in which these comparisons have been made.

New simulation methods

The central problems in molecular dynamics simulations are adequate sampling, accurate forcefield parameters and adequate timescales. These problems are also present in folding simulations of peptides, but are more tractable than for large protein systems and give us an opportunity to explore the well-known problem of sampling in molecular

Figure 1



Examples of the structures and free energy or thermodynamic profiles involved in the three elementary steps of secondary structure formation, which occur on different timescales. **(a)** Formation of an end-to-end contact (loop closing), which forms on a 10 ns timescale [8]. The curve on the right-hand side shows the free energy profile as a function of the end-to-end distance [44•]. This energy profile does not deviate much from a quadratic shape, expected for a random chain. **(b)** α-Helix formation, with a 200 ns timescale [5]. The right-hand side shows a profile of the fraction of helical amino acids as a function of temperature [45••]. **(c)** β-Hairpin formation, with a characteristic folding timescale of 1–10 μs [6]. The right-hand side shows the rugged free energy landscape as a function of temperature and a principal component that best describes the system fluctuations [34•,65].

dynamics. The first priority of computer simulations should be to combine detailed experimental and theoretical studies of the structure, thermodynamic equilibrium and folding kinetics for a large set of peptide systems, and then validate the theoretical methods and verify the accuracy of the results. By doing this, we will be able to detect deficiencies in forcefields, solvation models and water models. In cases in which there is good correlation between theory and experiment, we can then start the synergistic interpretation of experimental data, whereby theoretical models provide an atomic picture of the events that are consistent with the observation.

New sampling techniques applied to protein simulations have helped to overcome the sampling of the equilibrium and kinetics of peptide folding. These methods make use of embarrassingly parallel schemes to enhance sampling, thus making efficient use of multiprocessor, low-cost cluster machines. The four most widely used of these techniques are umbrella sampling, multicanonical sampling, replica exchange molecular dynamics (REMD) and parallel replica dynamics (PRD). Brief descriptions and applications of these methods are described next.

The simplest parallel sampling method is to run many uncoupled copies of the same system with different initial conditions. This has been used effectively by Ferrara *et al.* [15••] to study the relaxation behavior, transition rates and equilibrium properties of α peptides and β peptides. The same method was used to study the folding/unfolding of two *de novo* β-sheet-forming peptides [16–18] and a helical peptide [19]. The massive parallelism inherent in this method has been useful in projects such as Folding@Home [20], which uses the excess compute cycles of weakly coupled private computers. This simple parallel simulation method is most successful for systems with implicit solvent, the use of which increases the slowest relaxation rates by factors of 100–1000 [15••]. With explicit solvent, most single simulations are short compared to the system relaxation time and are strongly influenced by the initial conditions, with one exception described below [21•].

A more sophisticated method invented to increase the range of simulated rates is the PRD method [22]. In this method, independent simulations are started from the same conformational basin. When one of these simulations exits a basin, all the other simulations are restarted from the new basin position. In the ideal case that barrier crossing is fast and waiting times are exponential, this method yields a linear increase in the rates of conformational transitions [14]. This method has been applied to study the folding of the C-terminal β hairpin of protein G (GB1) [23••]; however, this method is suspect when applied to proteins because of the difficulty of identifying when a barrier-crossing event has occurred [24•].

One of the oldest methods to enhance the calculation of static properties is umbrella sampling. In the umbrella sampling method, one or more separate simulations are carried out with modified potential functions. These

separate simulations are then combined and the resulting phase space distribution is corrected to determine what it would have been if the sampling had been done with the original potential. In the context of protein folding, this method has been most extensively used by Brooks and co-workers ([25,26•]; for recent reviews, see [27,28]).

A less obvious method of umbrella sampling is to use a biasing potential that is solely a function of the potential energy. Determining this bias self-consistently, so that all potential energies are equally sampled, allows the system to do a random walk in potential energy space and easily surmount large enthalpic barriers. This is the multicanonical method created by Berg and Neuhaus [29]. This method was applied to peptides by Hansmann and Okamoto [30], and recently by Alves and Hansmann [31], and by Kamiya *et al.* [32]. Although widely used, the determination of the biasing function is difficult, especially for systems with explicit solvent. The multicanonical method and its variants are limited by the need to self-consistently determine a biasing function. The need to determine these factors can be eliminated through a clever trick invented independently on several occasions (for a review, see [33]). This method is commonly referred to as simulated tempering or replica exchange. The parallel version of this algorithm was adapted for use with molecular dynamics as REMD [13]. REMD has many advantages. It is particularly easy to implement and requires no expensive fitting to be done; it produces information over a range of temperatures and is easily adapted for use with implicit or explicit solvent. Relaxation times with this method are typically decreased by factors of 20 or more for peptides at room temperature [34•].

Other sampling techniques are being developed and applied to peptide systems [35–39].

We now go on to describe simulations and open issues concerning model systems that represent the basic elements of structure formed during peptide and small protein folding. Simulations of these disordered systems are used to determine the timescales of loop closing (10 ns timescale), and formation of α-helical peptides (200 ns timescale), β hairpins and mini-proteins (1–10 μs timescale).

Loop-closing kinetics

A fundamental timescale in peptide dynamics is the time it takes to form an intermolecular contact. Lapidus *et al.* [40] measured the viscosity and temperature dependence of cysteine-tryptophan quenching in $C(AGQ)_kW$ peptides (k = 1–6; denoted CW_k) to isolate the diffusion-limited and reaction-limited quenching rates. They found that their measurements could be well fit by a model of a stiff polymer with a few adjustable parameters. The diffusion-limited quenching time is found to be on the order of 10 ns, increasing monotonically with peptide length. Two studies of $(SG)_k$ (k = 1–4) peptides show some discrepancy in the length dependence of the quenching rate using different probes. Although the measured quenching rate is also on the order of 10 ns, Bieri *et al.* [41] found the measured rate increases monotonically with peptide length, whereas Hudgins *et al.* [42] found a marked turnover in the rate. This short timescale for diffusion-limited quenching is accessible to molecular dynamics simulations. Yeh and Hummer [21•] found that the measured quenching rates were consistent with the simulations [8,43].

Diffusion-limited quenching times are typically found to be 10–100 times slower than times calculated from simple polymer models, which are often parameterized as a reduced effective diffusion coefficient. Determining the extent to which the small effective diffusion coefficient reflects the timescale of the end-to-end dynamics is complicated by the approximations required to compute mean contact times. Lapidus *et al.* [40] emphasized that a single reaction coordinate model with an adjustable effective diffusion coefficient captures both the timescale of the end-to-end distance relaxation and the mean contact time of Langevin simulations of CW_k peptides. Alternatively, all-atom simulations of CW_3 were analyzed using bounds on the mean contact times for a general Gaussian chain [44•]. The main source of the slow dynamics is probably the nonlocal interactions along the chain or the coupling to solvation kinetics [40,44•].

α-Helix formation

Peptides consisting of a single α helix present a unique opportunity to compare theory with experiment — their small size is amenable to extensive sampling by simulation and their fast folding time is feasible to simulate [19,31,45••,46,47••,48–50].

Recent REMD simulations of the helix-coil transition have revealed the importance of backbone shielding by sidechains during α-helix formation [45••]. García and Sanbonmatsu simulated the thermodynamics of two 21 amino acid peptides: Ala-21 and the Fs peptide [where Fs stands for 'folded short' peptide with sequence $Ac\text{-}A_5(AAARA)_3A$-methyl amide]. In the Fs peptide simulation, the arginine sidechain was observed to desolvate the backbone carbonyls while the helical content simultaneously increased. The correlation between backbone shielding and higher helical stability was suggested previously by Vila *et al.* [46].

In these simulations, sampling was sufficient to achieve similar melting curves and helical content profiles for both folding and unfolding simulations, as well as for simulations with distinct initial conditions. Sampling quality for α helices has also been addressed recently by Smith *et al.* [47••], who revealed that different sampling measures have different convergence times.

Whereas the rmsd and intermolecular interaction energy stabilize quickly (~3 ns), the number of hydrogen bonds and number of clusters require significantly more time to equilibrate (30–50 ns).

β-Hairpin folding

β-Hairpin systems are an equally important building block of proteins as α helices and provide an opportunity to verify forcefields that might be favorable for α-helical systems. The GB1 peptide is probably the most widely studied β-hairpin peptide [15••,23••,34•,50–52,53••, 54–56]. The structure, dynamics [1] and kinetics of folding [6] of this peptide have been studied experimentally. At least 40% of the peptide was shown to adopt β-hairpin structure and the peptide folds/unfolds in 6 μs. REMD simulations by García and Sanbonmatsu [34•] have explored the free energy landscape of GB1 over a wide range of temperatures. In this calculation, the replicas were simulated in explicit solvent for 3.5 ns/replica, using the Parm94 (AMBER94) forcefield. They found that the β hairpin is the most stable state (40% of the configurations at 300K), in addition to a significant population of α helix (15% at 300K). Using the energy landscape theory [57], the folding rate for this peptide was estimated to be 1–7 μs. In the energy landscape theory, it is assumed that folding dynamics can be properly described as a diffusion process of an ensemble of protein configurations over a low-dimensional free-energy surface, which may be constructed using different order parameters. García and Sanbonmatsu used the free energy profile shown in Figure 1c, at 300K, and estimated the diffusion coefficient in the peptide configuration space from constant temperature molecular dynamics simulations. In this calculation, only the order of magnitude (1 μs) is credible.

Zhou *et al.* [52] conducted a similar study, but using the OPLS forcefield and Ewald summations to treat the electrostatics. They also found the β hairpin to be the most stable state, but the α-helical state is not significantly populated. We believe that these short (3.5 ns/replica) REMD simulations [34•,52] are not sufficient to reproduce equilibrium distributions for this hairpin.

Zagrovic *et al.* [23••] simulated the GB1 hairpin with the PRD approach, using the OPLS forcefield and Still's generalized Born/surface area (GB/SA) implicit solvent model [58] at 300K. The simulation was carried out for a total of ~38 μs 'folding time' and resulted in eight trajectories reaching the folded state (0.3%). Many trajectories sampled a semihelical intermediate. The upper bound for the folding time, directly calculated from the simulations, was estimated to be 4.7 μs, in agreement with the measured time of 0.9–6 μs. Jang *et al.* [50] studied the same peptide, using the CHARMM19 forcefield and a GB solvation model [59] in relatively short (15 ns) molecular dynamics simulations. They found that β hairpins are formed quickly, but that the α-helical conformations have a much lower energy (11 kcal/mol) than the β hairpin.

Interestingly, Zhou and Berne [53••] conducted another REMD study of the GB1 peptide, but this time using the GB/SA implicit solvent model of Ghosh *et al.* [60]. This calculation should have revealed very similar equilibrium behavior as the calculations by Zagrovic *et al.* [23••]. Nevertheless, Zhou and Berne found that GB1 does not form a β hairpin, but rather a structure with exposed hydrophobic sidechains and buried charged groups [53••]. They conclude that the GB/SA implicit solvent models need to improve the treatment of electrostatic interactions and properly screen ion pair formation. We assume that surface area (SA) implicit solvent treatment models have the same problems [61], which are sometimes solved by neutralizing charged groups [15••]. These results are not conclusive because most implementations of the GB and GB/SA models are different from Still's [58], and therefore are difficult to compare [59,60]. For example, a study of the free energy surfaces of the mini-protein betanova obtained with explicit solvent, GB and GB/SA approximations by Bursulaya and Brooks [26•] shows that the three models are similar.

Mini-proteins

Mini-proteins are short peptides that have either more than one hairpin, or a mixture of β and α structures. In some instances, there is a small hydrophobic core. Examples of these are the BBA5 mini-protein [10••], betanova [1] and the WW peptide [7]. A combined theoretical and experimental study of the folding kinetics of BBA5 was conducted by Snow *et al.* [10••]. Multiple molecular dynamics simulations totaling 700 μs were performed to directly determine the folding rate of this mini-protein. The measured and calculated folding rates (1–7 μs) were in very good agreement. Calculations on other mini-proteins have been recently performed by Jang *et al.* [50], Bursulaya and Brooks [26•], and Ferrara and Caflisch [16].

Comparison of simulation and experimental results

In the literature reporting computer simulations, one often finds a statement such as 'the simulation results are in agreement with the experimental results'. This kind of statement has sometimes been used loosely or without much consideration. Even a qualitative or an incidental agreement is taken to strengthen and support the validity of the methodology or calculations. In reality, however, such an agreement can be a result of lack of sampling, an important concern for peptide simulations.

One of the issues that arise when comparing simulation results with experimental measurements is the relevant temperature at which the comparison has to be made. Daura *et al.* [62••] have done a detailed conformational

study on a β-hexapeptide in methanol and address such a concern. Peter *et al.* [63] have shown that a more accurate comparison with measured nuclear Overhauser effects (NOEs) and J-couplings can be made by explicitly calculating relaxation rates from molecular dynamics simulations and relating them to NOESY (NOE spectroscopy) and ROESY (rotating frame NOESY) intensities directly. Especially for small peptide systems, the timescales of both internal and overall rotational motions are accessible to molecular dynamics simulations, and should enable the direct calculation of relaxation rates. Generally, the interproton distances inferred from NOE intensities are directly compared with the average distances calculated from the simulations. Such a treatment neglects the effects due to internal dynamics and spin diffusion. Feenstra *et al.* [64••] calculated the NMR cross-relaxation rates (NOE intensities) in three different ways for a nine-residue peptide from the protein HPr. Interestingly, it was found that, regardless of the methodology chosen, the correlations between experimental and theoretical intensities were identical. It was suggested that the inadequacies of the current forcefields and parameters limited the agreement of the results from the detailed calculation with the experimental measurements. The most important result from this study was that, regardless of the methodology chosen, the end result was only meaningful when the ensemble contained all relevant conformations. When comparisons are made to NMR experimental observations, the extent of sampling is more important than the details of the atomic motion [64••].

Conclusions

The most difficult question to answer is how good are the forcefields we use. Forcefields are not physical laws, but rather parameterizations of the system energy as a function of its atomic configuration. These parameterizations are accordingly subject to modifications and calibration with experimental data. As better sampling methods emerge, the accuracy of all forcefields must be revisited. However, in light of the results described above, the performance of existing forcefields has exceeded our expectations. Considering that current forcefields were predominately calibrated with folded proteins, the balance between multiple conformations in folding simulations and the observation of folding/unfolding transitions at temperatures within 10–20% of experimentally observed values is reason for optimism. Small changes in the forcefield will most probably result in better agreement with experiments. Although close agreement between simulation and experiment is important, we emphasize that care should be taken before stating that 'simulation results are in agreement with experiment'. Areas of disagreement between simulation and experiment are equally important, and are essential to improving simulation methods. Conversely, hasty claims of agreement with experiment may stunt the growth of the simulation community in a manner akin to 'sweeping the problem under the rug'.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. de Alba E, Antoro J, Rico MA, Jimenez MA: ***De novo* design of a monomeric three-stranded anti-parallel beta-sheet**. *Protein Sci* 1999, **8**:854-865.

2. Ramirez-Alvarado M, Blanco FJ, Serrano L: **Elongation of the BH8 beta-hairpin peptide: electrostatic interactions in beta-hairpin formation and stability**. *Protein Sci* 2001, **10**:1381-1392.

3. Pastor MT, de la Paz ML, Lacroix E, Serrano L, Perez-Paya E: **Combinatorial approaches: a new tool to search for highly structured beta-hairpin peptides**. *Proc Natl Acad Sci USA* 2002, **99**:614-619.

4. Neidigh JW, Fesinmeyer RM, Andersen NH: **Designing a 20-residue protein**. *Nat Struct Biol* 2002, **9**:425-430.

5. Williams S, Causgrove TP, Gilmanshin R, Fang KS, Callender RH, Woodruff WH, Dyer RB: **Fast events in protein folding: helix melting and formation in a small peptide**. *Biochemistry* 1996, **35**:691-697.

6. Munoz V, Thompson PA, Hofrichter JA, Eaton WA: **Folding dynamics and mechanism of beta-hairpin formation**. *Nature* 1997, **390**:196-199.

7. Crane JC, Koepf EK, Kelly JW, Gruebele M: **Mapping the transition state of the WW domain beta-sheet**. *J Mol Biol* 2000, **298**:283-292.

8. Lapidus LJ, Eaton WA, Hofrichter J: **Measuring the rate of intramolecular contact formation in polypeptides**. *Proc Natl Acad Sci USA* 2000, **97**:7220-7225.

9. • Sporlein S, Carstens H, Satzger H, Renner C, Behrendt R, Moroder L, Tavan P, Zinth W, Wachtveitl J: **Ultrafast spectroscopy reveals subnanosecond peptide conformation dynamics and validates molecular dynamics simulation**. *Proc Natl Acad Sci USA* 2002, **99**:7998-8002.

A careful comparison between molecular dynamics simulations of DMSO-solvated linear and cyclic APB peptides, and femtosecond time-resolved spectroscopy attempted to validate the simulations. Absorbance spectra were compared for several 1 ns simulations in a spectral region that possessed no ground state absorption. The simulations were in good agreement with the experiments, which occurred on identical timescales, leading the authors to associate ballistic *cis/trans* isomerization motions with the observed spectral features.

10. •• Snow CD, Nguyen N, Pande VS, Gruebele M: **Absolute comparison of simulated and experimental protein-folding dynamics**. *Nature* 2002, **42**:102-106.

The authors carried out a detailed comparison of the measured and calculated folding rates for the mini-protein BBA5. The calculation consists of thousands of 10–50 ns (700 μs total simulation time) simulations using a GB/SA implicit solvent model. There is a very good agreement between measured and calculated folding rates. This study might represent the first example of a truly synergistic collaboration between theoretical and experimental work.

11. Qiu LL, Pabit SA, Roitberg AE, Hagen SJ: **Smaller and faster: the 20-residue Trp-cage protein folds in 4 micros**. *J Am Chem Soc* 2002, **124**:12952-12953.

12. Werner JH, Dyer RB, Fesinmeyer RM, Andersen NH: **Dynamics of the primary processes of protein folding: helix nucleation**. *J Phys Chem B* 2002, **106**:487-494.

13. Sugita Y, Okamoto Y: **Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape**. *Chem Phys Lett* 2000, **329**:261-270.

14. Shirts MR, Pande VS: **Mathematical analysis of coupled parallel simulations**. *Phys Rev Lett* 2001, **86**:4983-4987.

15. •• Ferrara P, Apostolakis J, Caflisch A: **Thermodynamics and kinetics of folding of two model peptides investigated by molecular dynamics simulations**. *J Phys Chem B* 2000, **104**:5000-5010.

This manuscript describes multiple folding and unfolding simulations of a helix-forming 15-mer [$(AAQAA)_3$] and a β-hairpin-forming 12-mer [Val_5-^{D}Pro-Gly-Val_5] using an implicit solvent model. The forcefield is good enough to fold both β and α peptides. Folding and unfolding rates show the non-Arrhenius temperature dependence seen in simple lattice models. Folding is more cooperative for the β peptide than the α peptide, and is slower for the β peptide than the α peptide.

16. Ferrara P, Caflisch A: **Folding simulations of a three-stranded antiparallel beta-sheet peptide**. *Proc Natl Acad Sci USA* 2000, **20**:10780-10785.

17. Ferrara P, Caflisch A: **Native topology or specific interactions: what is more important for protein folding?** *J Mol Biol* 2001, **306**:837-850.

18. Cavalli A, Ferrara P, Caflisch A: **Weak temperature dependence of the free energy surface and folding pathways of structured peptides**. *Proteins* 2002, **47**:305-314.

19. Hiltpold A, Ferrara P, Gsponer J, Caflisch A: **Free energy surface of the helical peptide Y(MEARA)$_6$**. *J Phys Chem B* 2000, **104**:10080-10086.

20. Shirts M, Pande VS: **Computing: screen savers of the world unite!** *Science* 2000, **290**:1903-1904.

21. • Yeh IC, Hummer G: **Peptide loop-closure kinetics from a microsecond molecular dynamics simulation in explicit solvent**. *J Am Chem Soc* 2002, **124**:6563-6568.
Extensive all-atom, explicit solvent simulations (microsecond sampling) were carried out on CW_1 and CW_2 peptides using Amber and Charm forcefields. Whereas the potentials of mean force from the two potentials were markedly different, the end-to-end contact times were found to be quite similar. The authors found that the measured quenching rates were consistent with the simulations [40].

22. Voter AF: **Parallel replica method for dynamics of infrequent events**. *Phys Rev B-Cond Matt* 1998, **57**:13985-13988.

23. •• Zagrovic B, Sorin EJ, Pande VS: **β-hairpin folding simulations in atomistic detail using an implicit solvent model**. *J Mol Biol* 2002, **313**:151-169.
In this work, the folding of a peptide has been simulated for the longest time that has ever been reported for a peptide of this size. This simulation has allowed the authors to directly observe folding events that occur on the microsecond timescale.

24. • Fersht AR: **On the simulation of protein folding by short time scale molecular dynamics and distributed computing**. *Proc Natl Acad Sci USA* 2002, **99**:14122-14125.
In this paper, Fersht criticizes the use of multiple short simulations to study folding kinetics. He argues that protein folding is not an elementary kinetic step, because it involves a series of early conformational steps that lead to lag phases at the beginning of the kinetics. The presence of these lag phases can bias short simulations toward selecting minor pathways that have fewer or faster lag steps, and so the major folding pathways could be missed.

25. Boczko EM, Brooks CL: **First-principles calculation of the folding free-energy of a three-helix bundle protein**. *Science* 1995, **269**:393-396.

26. • Bursulaya BD, Brooks CL: **Comparative study of the folding free energy landscape of a three-stranded beta-sheet protein with explicit and implicit solvent models**. *J Phys Chem B* 2000, **104**:12378-12383.
An excellent illustration of the use of umbrella sampling to sample the free energy surface of a peptide, betanova. The free energy surfaces obtained with explicit and implicit solvent models are very similar.

27. Brooks CL: **Protein and peptide folding explored with molecular simulations**. *Acc Chem Res* 2002, **35**:447-454.

28. Shea JE, Brooks CL: **From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding**. *Annu Rev Phys Chem* 2001, **52**:499-535.

29. Berg BA, Neuhaus T: **Multicanonical algorithms for 1st order phase-transitions**. *Phys Lett B* 1991, **267**:249-253.

30. Hansmann UHE, Okamoto Y: **Prediction of peptide conformation by multicanonical algorithm: new approach to the multiple-minima problem**. *J Comput Chem* 1993, **14**:1333-1338.

31. Alves NA, Hansmann UHE: **Helix formation and folding in an artificial peptide**. *J Chem Phys* 2002, **117**:2337-2343.

32. Kamiya N, Higo J, Nakamura H: **Conformational transition states of a beta-hairpin peptide between the ordered and disordered conformations in explicit water**. *Protein Sci* 2002, **11**:2297-2307.

33. Mitsutake A, Sugita Y, Okamoto Y: **Generalized-ensemble algorithms for molecular simulations of biopolymers**. *Biopolymers* 2001, **60**:96-123.

34. • Garcia AE, Sanbonmatsu KY: **Exploring the energy landscape of a beta hairpin in explicit solvent**. *Proteins* 2001, **42**:345-354.
An REMD simulation of the GB1 peptide in explicit solvent shows that the peptide samples multiple free energy minima with different secondary structure.

35. Zaloj V, Elber R: **Parallel computations of molecular dynamics trajectories using the stochastic path approach**. *Comp Phys Comm* 2000, **128**:118-127.

36. Eastman P, Gronbech-Jensen N, Doniach S: **Simulation of protein folding by reaction path annealing**. *J Chem Phys* 2001, **114**:3823-3841.

37. Hansmann UHE: **Generalized ensemble techniques and protein folding simulations**. *Comp Phys Comm* 2002, **147**:604-607.

38. Hansmann UHE, Wille LT: **Global optimization by energy landscape paving**. *Phys Rev Lett* 2002, **88**:068105.

39. Krivov SV, Chekmarev SF, Karplus M: **Potential energy surfaces and conformational transitions in biomolecules: a successive confinement approach applied to a solvated tetrapeptide**. *Phys Rev Lett* 2002, **88**:8101-8104.

40. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J: **Effects of chain stiffness on the dynamics of loop formation in polypeptides. Testing a 1-dimensional diffusion model of peptide dynamics**. *J Phys Chem B* 2002, **106**:11628-11640.

41. Bieri O, Wirz J, Hellrung B, Schutkowski M, Drewello M, Kiefhaber T: **The speed limit for protein folding measured by triplet-triplet energy transfer**. *Proc Natl Acad Sci USA* 1999, **96**:9597-9601.

42. Hudgins RR, Huang F, Gramlich G, Nau WM: **A fluorescence-based method for direct measurement of submicrosecond intramolecular contact formation in biopolymers: an exploratory study with polypeptides**. *J Am Chem Soc* 2002, **124**:556-564.

43. Lapidus LJ, Eaton WA, Hofrichter J: **Dynamics of intramolecular contact formation in polypeptides: distance dependence of quenching rates in a room-temperature glass**. *Phys Rev Lett* 2001, **87**:258101.

44. • Portman JJ: **Non-Gaussian dynamics from a simulation of a short peptide: loop closure rates and effective diffusion coefficients**. *J Chem Phys* 2003, **118**:2381-2391.
All-atom simulations of CW_3 were analyzed using bounds on the mean contact times for a general Gaussian chain. Although correlation functions of the end-to-end vector were found to be non-Gaussian, the simulated diffusion-limited contact time agrees with the Gaussian bounds if the diffusion coefficient is reduced to match the simulated end-to-end distance relaxation.

45. •• Garcia AE, Sanbonmatsu KY: **α-Helical stabilization by side chain shielding of backbone hydrogen bonds**. *Proc Natl Acad Sci USA* 2002, **99**:2782-2787.
REMD simulations in explicit solvent examined the mechanism of the helix-coil transition and addressed the issues of adequate sampling and forcefield accuracy. It was found that Parm94 exaggerates helix formation, whereas Parm96, designed to address this deficiency, produces predominately β-hairpin conformations for the Ala-21 system. A simple modification of the Parm94 forcefield was shown to produce σ values with the correct order of magnitude (0.004) and reasonable values for the α-helix melting temperature.

46. Vila JA, Ripoll DR, Scheraga HA: **Physical reasons for the unusual alpha-helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides**. *Proc Natl Acad Sci USA* 2002, **97**:13075-13079.

47. •• Smith LJ, Daura X, van Gunsteren WF: **Assessing equilibrium and convergence in biomolecular simulations**. *Proteins* 2002, **48**:487-496.

The authors attempt to establish criteria for assessing the quality of sampling achieved during a simulation. The intermolecular energy, rmsd, number of clusters as determined from a cluster analysis, number of intermolecular hydrogen bonds and number of torsion angle transitions were used as measures of equilibration. These quantities were compared for various molecular dynamics simulations of α helices and β hairpins at several different temperatures. This study revealed that different measures have different convergence times. Whereas the rmsd and intermolecular interaction energy stabilize quickly ($\sim$3 ns), the number of hydrogen bonds and number of clusters require significantly more time to equilibrate (30–50 ns).

48. Favrin G, Irback A, Wallin S: **Folding of a small helical protein using hydrogen bonds and hydrophobicity forces**. *Proteins* 2002, **47**:99-105.

49. Peng Y, Hansmann UHE: **Solvation model dependency of helix-coil transition in polyalanine**. *Biophys J* 2002, **82**:3269-3276.

50. Jang S, Shin S, Pak Y: **Molecular dynamics study of peptides in implicit water: *ab initio* folding of beta-hairpin, beta-sheet, and beta-beta-alpha-motif**. *J Am Chem Soc* 2002, **124**:4976-4979.

51. Ma B, Nussinov R: **Molecular dynamics simulations of a beta-hairpin fragment of protein G: balance between side-chain and backbone forces**. *J Mol Biol* 2000, **296**:1091-1104.

52. Zhou RH, Berne BJ, Germain R: **The free energy landscape for beta hairpin folding in explicit water**. *Proc Natl Acad Sci USA* 2001, **98**:14931-14936.

53. •• Zhou RH, Berne BJ: **Can a continuum solvent model reproduce the free energy landscape of a beta-hairpin folding in water?** *Proc Natl Acad Sci USA* 2002, **99**:12777-12782.

REMD calculations of the free energy surface of the GB1 β hairpin in explicit and implicit (GB/SA) solvent models are in great disagreement. The authors argue that GB/SA and implicit solvent models need to improve on the energetics of ion pair formation.

54. Wu HW, Wang SM, Brooks BR: **Direct observation of the folding and unfolding of a beta-hairpin in explicit water through computer simulation**. *J Am Chem Soc* 2002, **124**:5282-5283.

55. Roccatano D, Colombo G, Fioroni M, Mark AE: **Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: a molecular dynamics study**. *Proc Natl Acad Sci USA* 2002, **99**:12179-12184.

56. Lee J, Shin S: **Two-dimensional correlation analysis of peptide unfolding: molecular dynamics simulations of beta hairpins**. *J Phys Chem B* 2002, **106**:8796-8802.

57. Plotkin SS, Onuchic JN: **Understanding protein folding with energy landscape theory: part I: basic concepts**. *Q Rev Biophys* 2002, **35**:111-167.

58. Still WC, Tempczyk A, Hawley RC, Hendrickson T: **Semianalytical treatment of solvation for molecular mechanics and dynamics**. *J Am Chem Soc* 1990, **112**:6127-6129.

59. Bashford D, Case DA: **Generalized Born models of macromolecular solvation effects**. *Annu Rev Phys Chem* 2000, **51**:129-152.

60. Ghosh A, Rapp CS, Friesner RA: **Generalized Born model based on a surface integral formulation**. *J Phys Chem* 1998, **102**:10983-10990.

61. Shen MY, Freed KF: **Long time dynamics of met-enkephalin: comparison of explicit and implicit solvent models**. *Biophys J* 2002, **82**:1791-1808.

62. •• Daura X, Gademann K, Schafer H, Jaun B, Seebach D, van Gunsteren WF: **The beta-peptide hairpin in solution: conformational study of a beta-hexapeptide in methanol by NMR spectroscopy and MD simulation**. *J Am Chem Soc* 2001, **123**:2393-2404.

A detailed conformational study on a β-hexapeptide in methanol at two different temperatures, 298K and 340K. Reversible folding to the hairpin conformation was observed at both temperatures. It was found that the ensemble of structures from the simulation at 340K reproduced the room temperature NMR-derived data, such as NOEs, better than those obtained from the simulation at 298K. It was suggested that the disagreement in the temperature could be a result of improper parameterization.

63. Peter C, Daura X, van Gunsteren WF: **Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations**. *J Biomol NMR* 2001, **20**:297-310.

64. •• Feenstra KA, Peter C, Scheek RM, van Gunsteren WF, Mark AE: **A comparison of methods for calculating NMR cross-relaxation rates (NOESY and ROESY intensities) in small peptides**. *J Biomol NMR* 2002, **23**:181-194.

The authors calculated NMR cross-relaxation rates (NOE intensities) in three different ways. The first approach considers internal dynamics by incorporating the fluctuations in interproton distance, time correlation function from the motion of interproton vectors and coupling to overall tumbling. The second approach corresponds to calculating $<1/r^6>$ for clustered conformations and, subsequently, spectral densities were obtained from a Lorentz function using these averages. In the third approach, the average distances are assumed to be proportional to intensities. It was found that, regardless of the methodology chosen, the correlations between experimental and theoretical intensities were identical.

65. Garcia AE: **Large-amplitude non-linear motions in proteins**. *Phys Rev Lett* 1992, **68**:2696-2700.