Observation of Noncooperative Folding Thermodynamics in Simulations of 1BBL

Jed W. Pitera,* William C. Swope,* and Farid F. Abraham†
*IBM Almaden Research Center, San Jose, California 95120; and †Chemistry, Materials, and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California 94551

ABSTRACT One of the predictions of the energy landscape theory of protein folding is the possibility of barrierless, ''downhill'' folding under certain conditions. The protein 1BBL has been proposed to fold by such a downhill mechanism, though this is a matter of some dispute. We carried out extensive replica exchange molecular dynamics simulations on 1BBL in explicit solvent to address this controversy and provide a microscopic picture of its folding thermodynamics. Our simulations show two distinct structural transitions in the folding of 1BBL. A low-temperature transition involves a disordering of the protein's tertiary structure without loss of secondary structure. A distinct, higher temperature transition involves the complete loss of secondary structure and dissolution of the hydrophobic core. In contrast, control simulations of the 1BBL homolog E3BD show a single high temperature unfolding transition. Further simulations of 1BBL at high ionic strength show a significant destabilization of helix II but not helix I, suggesting that the apparent folding cooperativity of 1BBL may be highly dependent on experimental conditions. Although our simulations cannot provide definitive evidence of downhill folding in 1BBL, they clearly show evidence of a complex, non-two-state folding process.

INTRODUCTION

Protein folding is often described in the context of two-state models, where there are only two states in which the protein can exist: a native, or folded, state and an unfolded, or denatured, state. Under various thermodynamic conditions, one or the other of these states is more stable. At the melting temperature (T_m), the two states are equally stable and an ensemble of molecules at the melting point is said to be 50% folded in the sense that half the molecules are fully folded and half are fully unfolded. This situation can arise when there is a free energy barrier separating the folded and unfolded manifolds of protein conformations, implying that there is a very low probability of observing partially folded conformations. Fig. 1 shows a free energy landscape projected onto a hypothetical order parameter that indicates the degree of folding. For two-state folding, the figure shows the characteristic barrier and the change in relative stability of the two states with respect to temperature. At any temperature, one observes an equilibrium distribution of values for the order parameter, a distribution that will be bimodal near the melting temperature. The kinetics of two-state folding is generally first order, with relaxation to equilibrium described with a single exponential time constant. Furthermore, because the transition state manifolds are essentially unpopulated, all experimental probes that monitor the concentration of the folded state will indicate a single melting temperature. In the language of energy landscape theory (1), this is referred to as a type 1 scenario.

Energy landscape theory also predicts the possibility of a different picture under conditions that strongly favor the native state where the unfolded state is so destabilized that there is no barrier, known as a type 0 scenario. This is also referred to as barrierless or ''downhill'' folding (Fig. 1, *boxed panel*). The theory makes specific predictions about the thermal dependence of folding and unfolding kinetics in this situation, which is determined by the degree of roughness of the energy landscape and may be multiexponential or even highly nonexponential. In this original development, downhill folding was meant to apply only under strongly folding conditions, and the situation would return to the type 1 scenario under less extreme conditions.

An alternative model (2,3) for downhill folding is that there is a single thermodynamic state for the protein that consists of an ensemble of conformations that gradually loses or gains structure as thermodynamic conditions change. In this ''one-state folding'' view, there is a single minimum in the free energy landscape over a very wide range of thermodynamic conditions (see Fig. 1). This has also been referred to as a second-order transition (4). In this view, the probability distribution for order parameters will show a single peak that shifts gradually as temperature or denaturant concentration changes. In contrast to the two-state situation, at the melting temperature the ensemble of a downhill folding protein consists of members that could each be considered to be partially folded. Therefore, different experimental techniques sensitive to different types of protein structure may show different melting temperatures. This scenario is very exciting, since it may allow the characterization of all of the stages and mechanisms of folding with a series of equilibrium experi-

Submitted October 3, 2007, and accepted for publication February 8, 2008.

Jed W. Pitera, William C. Swope, and Farid F. Abraham contributed equally to this work.

Address reprint requests to Jed W. Pitera, IBM Almaden Research Center, 650 Harry Road, San Jose, CA 95120. Tel.: 408-927-2084; E-mail: pitera@us.ibm.com.

Editor: Angel E. Garcia.

0006-3495/08/06/4837/10 $2.00

doi: 10.1529/biophysj.107.123265

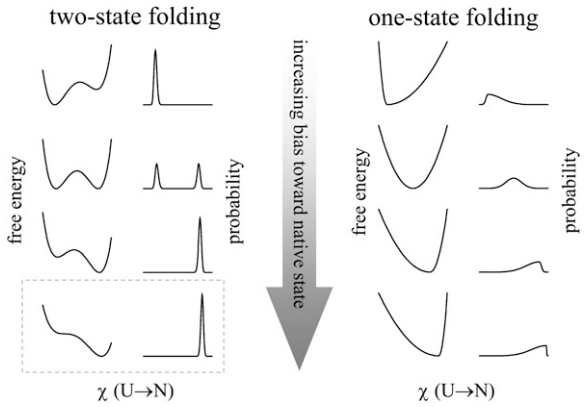

FIGURE 1 Idealized free energy surfaces and probability distributions for one- and two-state folding as a function of a folding order parameter χ at different levels of thermodynamic bias toward the native structure. The original downhill folding scenario proposed by Bryngelson et al. (1) is a limiting case of two-state folding under extreme native bias (*dashed box*), whereas the one-state model corresponds to downhill folding as described by Garcia-Mira et al. (2).

ments performed with different degrees of thermal or chemical denaturation (5,6). In the absence of downhill folding, it is possible to characterize transition states only indirectly, as with ϕ-value analysis (4), where one assumes a linear relation exists between changes in activation and equilibrium free energy differences due to mutations.

In reality, and especially in computer simulations of folding, it may be very difficult to distinguish between these scenarios. Folded proteins possess various types of secondary and tertiary structure that have different degrees of stability. These different types of structure may not all develop at exactly the same temperature, creating a loss of cooperativity and probe-dependent transition temperatures. For small systems even true first-order transitions are rarely sharp, occurring over a broad range of temperatures (7). Multiple, two-state, but temperature-broadened transitions that occur over a narrow temperature range can have many of the signatures ascribed to downhill folders. The difference between downhill and simple exponential kinetic behavior can be very small, if not undetectable, in most observables (8). Similarly, melting curves for downhill folding transitions, although broadened, can show the same sigmoidal shapes as two-state folders (9).

Inspired by all of these theoretical models, evidence for downhill folding in the protein 1BBL was presented by Garcia-Mira et al. (2). 1BBL is a parallel two-helix bundle that serves as the peripheral subunit binding domain (PSBD) of the E2 in the 2-oxoglutarate dehydrogenase multienzyme complex of *Escherichia coli*. Two fluorescently labeled 40-amino acid variants of the 1BBL protein—a singly labeled version with an N-terminal napthylalanine (Naf-1BBL) and a doubly labeled version that added a dansyl group to the C-terminal lysine (Naf-1BBL-Dan)—were studied and showed a range of broad, but distinct melting transitions ranging from 304 K (increasing fluorescence resonance energy transfer (FRET) efficiency, indicating a decreased end-to-end distance) to 322–325 K (differential scanning calorimetry (DSC), circular dichroism (CD), napthyl quantum yield).

All these experiments were done at pH 7.0, and parallel experiments at pH 3.0 showed no evidence of folding and were used as a representative denatured state for comparison. The broad, overlapping, yet still distinct melting transitions seen for different probes of 1BBL folding indicate that it is not a completely cooperative process. Furthermore, Garcia-Mira et al. used a statistical mechanical model of the folding process to integrate these different experiments and showed that the experimental data could be fit only by a barrierless one-state downhill folding mechanism. In contrast to 1BBL, another member of the PSBD family of proteins, the PSBD from the thermophile *Bacillus stearothermophilus* (E3BD) has previously been shown to be a structured (10), fast folding protein (11) with a two-state melting transition near 328 K (12).

This evidence for downhill folding in 1BBL was challenged by Ferguson et al. (13), who examined the thermal denaturation of three different PSBD proteins: an unlabeled version of 1BBL with an additional four N-terminal residues and one more C-terminal residue (QNND-1BBL), a variant of the E3BD protein from the thermophile *B. stearothermophilus*, and a destabilized homolog from the hyperthermophile *Pyrobaculum aerophilum* (POB). These proteins were studied by near- and far-ultraviolet (UV) CD, UV absorbance, DSC, and NMR chemical shift analysis. These experiments produced a narrow range of T_m values for E3BD (330.0 $\pm$ 0.3 to 330.9 $\pm$ 0.2 K) and broader ranges for POB (343.9 $\pm$ 0.2 to 347 $\pm$ 12 K) and QNND-1BBL (324 $\pm$ 1 to 329.7 $\pm$ 0.3). Each protein was argued to fold with a single cooperative transition. In addition, the authors studied the folding of the doubly labeled 1BBL (Naf-1BBL-Dan) and found that the labels increased the apparent far-UV T_m of the peptide relative to QNND-1BBL (both at pH 7.0). The authors also reported a significant increase in the propensity of doubly labeled 1BBL to aggregate, which prevented them from repeating the previous DSC experiments. The perturbations introduced by labeling were argued to account for the noncooperative folding of Naf-1BBL and Naf-1BBL-Dan, rather than downhill folding.

Further experiments by Naganathan et al. (14) showed that the folding of Naf-1BBL and Naf-1BBL-Dan were similar when studied under nonaggregating conditions (pH 7.0 and low ionic strength) and that the differences between QNND-1BBL and Naf-1BBL could be attributed to the destabilizing interaction between the N-terminal positive charge and helix I in the latter construct. Furthermore, they reanalyzed the DSC data for QNND-1BBL, arguing that it is incompatible with a two-state transition. This report highlighted the significant sensitivity of 1BBL to experimental conditions, including pH, concentration, ionic strength, and denaturant.

The kinetics of PSBD folding were subsequently reported in an extensive study by Ferguson et al. (15), where they

found folding half-times of 5 μs for a QNND-1BBL mutant, ~3.5 μs for POB, and 25 μs for E3BD. The QNND-1BBL variant in these kinetic experiments has His-166 of helix 2 mutated to tryptophan to provide a fluorescence signal of folding, a possibly significant perturbation given its placement in the hydrophobic core and interface between the two helices. It is interesting to note that both the folding and unfolding rates for this QNND-1BBL mutant are quite fast (k_f of 130,000/s $\pm$ 10,000/s and k_u of 6000/s $\pm$ 2000/s at 298 K). These unusual kinetics are ascribed to significant residual helical structure in the denatured state. This study also revisited the evidence for downhill folding in 1BBL, reporting additional NMR data as well as DSC thermograms for QNND-1BBL as a function of ionic strength (21–167 mM). The latter show that QNND-1BBL is destabilized, but not completely denatured by low ionic strength conditions. The 14 ^{13}C shifts show a range of T_m values from 320 $\pm$ 2 to 329 $\pm$ 2 K but can be fit by a single transition with a T_m of 324 $\pm$ 1 K. Ferguson et al. also point out that they have not argued that 1BBL and other small proteins are ''two-state'' folders, rather that they show a ''cooperative'' folding transition. This is arguably a semantic distinction, since finite systems such as proteins cannot undergo true two-state phase transitions with corresponding discontinuities in heat capacity (16). Furthermore, cooperativity is not a sufficient criterion to differentiate one- versus two-state transitions.

The folding of E3BD was further studied by a combination of kinetic studies, ϕ-value experiments, and high temperature unfolding simulations (17). Temperature jump experiments demonstrated that E3BD folds with simple exponential kinetics. The unfolded state of E3BD was shown to have little helical structure based on CD of model peptides of each helix as well as the absence of helical C_α and C_β conformational shifts for the denatured state of the full protein. A set of eight molecular dynamics simulations of the folded structure at 373 K, ranging from 20 to 100 ns in duration, are all described as unfolding in a stepwise fashion, starting with a weakening of tertiary interactions that proceeds through the melting of helix II and finishes with a loss of structure in helix I. The highest C_α-root mean-square deviation (RMSD) from the native state sampled in these simulations was 8 Å, and the denatured state was taken to encompass all conformations with C_α-RMSD above 6 Å.

Interestingly, this simulated denatured state is described as having little repeating helical structure, but the residues of helices I and II show a strong preference for the α-helical region of the (ϕ,ψ) map. As for the transition state for folding, most ϕ-values ranged from 0.3 to 0.9, indicating a diffuse transition state. The diffuse transition state and the absence of structure in the denatured state led the authors to conclude that E3BD folds by a two-state nucleation-condensation mechanism, with simultaneous formation of secondary and tertiary structure. Another simulation study used a Go model to study the folding of 1BBL and compare it with more traditional two-state folders (18). The apparent downhill folding of 1BBL is ascribed to its relatively small number of nonlocal contacts, but this cannot explain the differences seen between 1BBL and its topologically equivalent homologs.

During the initial preparation of this article, further evidence of downhill folding in 1BBL was reported by Sadqi et al. (19). Extensive measurements of individual proton chemical shifts as a function of temperature for Naf-1BBL at pH 5.3 show a broad range of transition temperatures from 280 to 340 K, even though the midpoint of the average NMR transition (304 K) is in good agreement with CD (308 K). Careful analysis of transitions in chemical shift for each proton shows a complex mechanism of overlapping transitions, with the loss of a large number of intramolecular hydrogen bonds at low temperature (298 K) followed by a loss of tertiary structure, particularly the contacts between the two helices (304 K), and finishing with a loss of backbone structure (309 K) and hydrogen bonds in the interhelical loop (311 K).

One of the key themes running through the experimental investigations of 1BBL folding is a significant sensitivity to experimental conditions. Slight variations in the construct studied, pH, or ionic strength can shift the measured transition temperature by ~10 K. This may be responsible for some of the dispersion in T_m values seen for 1BBL and makes comparisons between different experimental studies difficult. Nonetheless, both the Munoz and Fersht groups have indicated that there is indeed something unusual about the folding of 1BBL, though they still differ in the details.

In general, the strongest test of downhill folding is to actually determine the distribution of conformations along an order parameter relevant for folding at a state point near the folding transition and show that it is unimodal rather than bimodal. One tool for getting such a detailed, microscopic picture of the conformational ensemble sampled by a protein is molecular simulation (20). Building on our experiences using replica-exchange molecular dynamics simulations to elucidate the details of folding in fast folding systems with complex thermodynamics and kinetics, such as the trpzip2 peptide (7,21) and λ-repressor 6–85 (22), we carried out extensive replica-exchange molecular dynamics of the 1BBL protein in explicit solvent over a temperature range from 250 to 625 K. These simulations, complemented by control simulations of QNND-1BBL and E3BD, show clear evidence of noncooperative folding in both 1BBL and QNND-1BBL.

METHODS

Three different protein sequences were simulated in the course of this work: the 40 amino acid peptide from Garcia-Mira et al. (2) (indicated as 1BBL), a longer 45 amino acid sequence of 1BBL with N- and C-terminal extensions studied by Ferguson et al. (13) (indicated as QNND-1BBL), and a 45 amino acid sequence of the 1BBL homolog E3BD. Each protein system was built from the first model of the corresponding NMR structure (Protein Data Bank (23) IDs 1BBL (10), 1W4H, and 1W3D (15), respectively). All ionizable groups, including N- and C-termini, were set to their expected protonation state at pH 7. Unless otherwise specified, each protein was modeled using the

AMBER parm96 (24) force field. The QNND-1BBL sequence was simulated with both the AMBER parm96 (QNND-1BBL) and AMBER parm99SB (25) (QNND-1BBL-ff99SB) force fields. This sequence was also simulated with the parm96 force field and an excess of Na^+ and Cl^- counterions corresponding to an ionic strength of 100 mM (QNND-1BBL-ion). Each protein was solvated in a truncated octahedral box of TIP3P water (26) with a minimum solute-wall distance of 25 Å. In all cases, the systems were neutralized by the addition of chloride counterions as necessary.

After this initial setup, each system was subjected to a protocol of minimization and constant pressure equilibration. The final conformations from these equilibrations were used to start constant volume production runs for each system at all temperatures. All minimizations and molecular dynamics simulations were carried out using the PMEMD module of AMBER8 (27). All production runs used a 9 Å nonbonded cutoff with particle mesh Ewald electrostatics (28), a 2 fs time step with bonds constrained using SHAKE (29), and an Andersen thermostat (30) with a collision period of 10 ps. Further details of the simulated systems, equilibration, and production calculations are included in the Supplementary Material (Data S1).

For 1BBL, we carried out extensive replica exchange molecular dynamics (REMD) (31) with 256 replicas at temperatures exponentially spaced from 250 to 625 K. All replicas were started with the same equilibrated folded conformation. The REMD calculation was run for 32 ns per replica, for a total simulation time of 8 μs. This calculation was carried out on the THUNDER supercomputer at the Lawrence Livermore National Lab, using one four-way Itanium2 node per replica. For the four control calculations, QNND-1BBL, QNND-1BBL-ff99SB, QNND-1BBL-ion, and E3BD, we were unable to do the same extensive REMD simulation as above. Instead, we performed independent NVT molecular dynamics simulations at four or six different temperatures (ranging from 310 K to 450 K) for a minimum of 34 ns each. All control calculations were carried out on an IBM BlueGene/L system, using 64–128 processors per run.

Structural analysis of the protein conformations generated from each simulation was carried out using a combination of tools. Radius of gyration, C_α-RMSDs from reference structures, and geometric parameters (end-to-end distance, interhelical angle) were determined using programs from the AMBER8 package (27). The secondary structure and solvent-accessible surface area (SASA) of each conformation was assigned using the STRIDE (32) package. Interresidue contacts for each conformation were assigned for all residues with side-chain heavy atoms within 5 Å of one another and then filtered to determine the fraction of native contacts and the number of hydrophobic, polar, or ionic contacts. Conformational clustering was carried out at several temperatures over the last 4 ns of REMD data. A k-medoids clustering algorithm was used with a C_α-C_α distance matrix error metric to identify up to 40 clusters at each temperature.

For 1BBL, the SHIFTS 4.1 software package (33) was used to estimate the chemical shifts of H, C_α, and C_β nuclei in each simulation structure. These were averaged and compared against the ^{13}C and ^{1}H chemical shifts reported in Ferguson et al. (15) and Sadqi et al. (19), respectively. We were able to calculate simulated melting curves for 11 of the 14 unique ^{13}C shifts reported by Ferguson et al. and 154 of the 158 ^{1}H shifts reported by Sadqi et al. T_m values were calculated for all simulated melting curves using a nonlinear least squares fitting protocol (see Supplementary Material, Data S1).

For most observables including the chemical shifts, data were linearly averaged over the last 10 ns of simulation at each temperature. The end-to-end distance was both linearly averaged as well as averaged as the inverse sixth power ($\langle r^{-6}\rangle^{-1/6}$) for more direct comparison with the FRET experiment of Garcia-Mira et al. (2). In addition to simple averages, the temperature-dependent distributions of observables were calculated by creating histograms of each property at each temperature to produce a potential of mean force (PMF), again using the last 10 ns of simulation. These PMFs, expressed in units of k_BT, were examined for evidence of discontinuities or first-order phase transitions.

The QNND-1BBL, QNND-1BBL-ff99SB, QNND-1BBL-ion, and E3BD control simulations were analyzed in a similar fashion to the 1BBL REMD data. The only difference was that all results were averaged over the last 5 ns of each control trajectory rather than the last 10 ns, and the chemical shift and PMF analyses were not performed.

To assess the convergence of the REMD calculations, we also looked at the time-dependent behavior of each observable as a function of temperature. This analysis (described in the Supplementary Material, Data S1) indicates that high-temperature transitions (>370 K) are well converged but that convergence is incomplete at lower temperatures.

RESULTS

Fig. 2 shows the temperature-dependent behavior of 1BBL observables related to the overall geometry of the protein structure, including the interhelical angle between the axes of helices I and II, the radius of gyration, the fraction of native contacts, and the fractional helicity for each helix. At low temperatures, the two helices of 1BBL are nearly parallel, with an interhelical angle near 20°. Near 300 K, the two helices begin sampling a range of relative orientations, and they are disordered relative to one another (average angle 90°) at 350 K. In contrast, the radius of gyration gradually increases with temperature, particularly above 350 K. Both the linearly averaged and $\langle r^{-6}\rangle^{-1/6}$ averaged end-to-end distances (not shown) show a weak decrease with temperature from 250 to 350 K, corresponding to the low-temperature increase of FRET efficiency seen by Garcia-Mira et al. (2). At low temperature, the parallel helices of 1BBL keep the N- and C-terminus well separated, whereas the loss of interhelical order with a slight increase in temperature allows the two helices to sample antiparallel orientations that bring the two termini closer together.

The loss of interhelical order is also reflected in a loss of native side-chain contacts (Fig. 2, *middle panel*), which shows a very broad transition starting below 300 K and finishing around 400 K. In contrast, the number of hydrophobic contacts is stable until 370 K, where it decreases until reaching a high temperature baseline value at 425 K. The SASA of hydrophobic residues shows a somewhat broader transition with a similar midpoint (data not shown). A possible driving force for the low-temperature loss of interhelical order is the formation of nonnative salt bridges between helices I and II. Zero at low temperature, the number of salt bridges begins to increase at around 300 K and there is an average of approximately one salt bridge between the two helices at all temperatures between 350 and 500 K.

The loss of native side-chain interactions parallels the broad loss of native-like secondary structure between 275 and 425 K. The α-helices of 1BBL start to disorder at 300 K (Fig. 2), with helix II melting around 350 K and helix I undergoing a broader melting transition centered at 390 K. Above 350 K, we observed the formation of a small fraction (10%–15%) of ordered β-sheet structure. Above 425 K, 1BBL is denatured in our simulations and there is little or no ordered secondary structure. If we assign secondary structure based only on (ϕ,ψ) analysis, the disordered chain is 8% helical, 8% parallel β, 16% antiparallel β, and 16% polyproline-II helix.

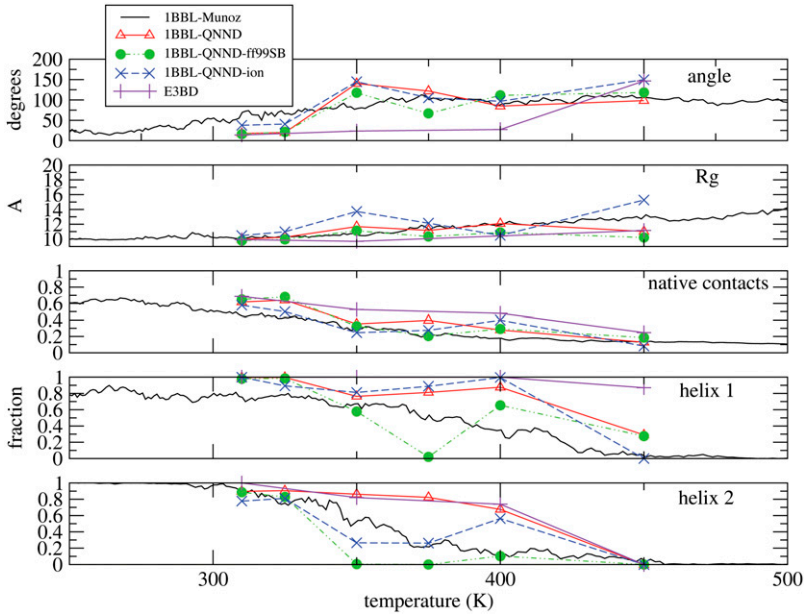

FIGURE 2 Comparison of various simulation observables versus temperature for the 1BBL REMD (*solid line*), QNND-1BBL (*red line* with *triangles*), QNND-1BBL-ff99SB (*green dotted line* with *filled circles*), QNND-1BBL-ion (*blue dashed line* with *Xs*), and E3BD (*solid purple line* with *plus signs*) simulations. From top to bottom, the panels show the interhelical angle, the radius of gyration, the fraction of native side chain-side chain van der Waals contacts, and the fractional α-helicity of helices I and II.

Other observables that report on local structure are the ^{13}C C_α and C_β chemical shifts reported by Ferguson et al. (15), which are primarily sensitive to adjacent (ϕ,ψ) and χ angles (34). A better indicator of tertiary structure is the large set of 111 simulated ^1H chemical shifts. The simulated T_m values for all these observables are shown in Fig. 3 as a scatter plot by residue. The scatter plot indicates that low T_m values are associated mostly with residues between the two helical regions. This plot also shows a cluster between 320 and 370 K associated with residues in helix II, as well as a larger cluster between 350 and 420 K associated with residues in helix I. Attempts were made to perform an atom-by-atom comparison of the experimental ^1H chemical shift melting curves (19) with those from the simulation. We found limited correlation between the simulated melting temperatures and experimental values. However, this is probably not a surprising result when an overall melting transition is broad in temperature due to the fact that it is made up of many close and overlapping transitions of different structural elements. In these cases, due to inevitable limitations in force fields, chemical shift models, sampling, and equilibration, one is not likely to see strong and direct correspondence between simulation and experiment, though they agree in the qualitative view that different structural elements of 1BBL disorder at different temperatures.

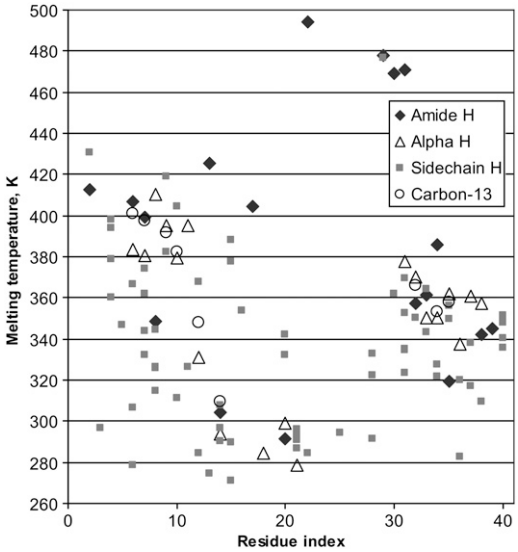

FIGURE 3 Scatter plot of transition temperatures by residue for simulated ^1H and ^{13}C chemical shifts from the 1BBL REMD data.

The potentials of mean force for several order parameters as a function of temperature are shown in Fig. 4, contoured in units of 0.5 k_BT. The fraction of native contacts and interhelical angle show broad, barrierless transitions at a relatively low temperature. In contrast, the fraction of native secondary structure and number of hydrophobic contacts show a more cooperative transition around 400 K. Of these, only the number of hydrophobic contacts shows anything resembling two-state behavior, with some small barriers at intermediate values near the transition temperature. These plots also serve to summarize the two classes of transitions we see in our 1BBL data. There is a group of low-temperature transitions ($<$325 K) that corresponds to a loss of tertiary structure (interhelical angle, native contacts, a subset of ^1H shifts) and a separate group of higher-temperature transitions (350–400 K) that corresponds to a loss of hydrophobic interactions and the melting of helices I and II (number of hydrophobic contacts, native secondary structure). This is schematized in Fig. 5, which shows the centroids of the five largest clusters at several temperatures that span these two transitions.

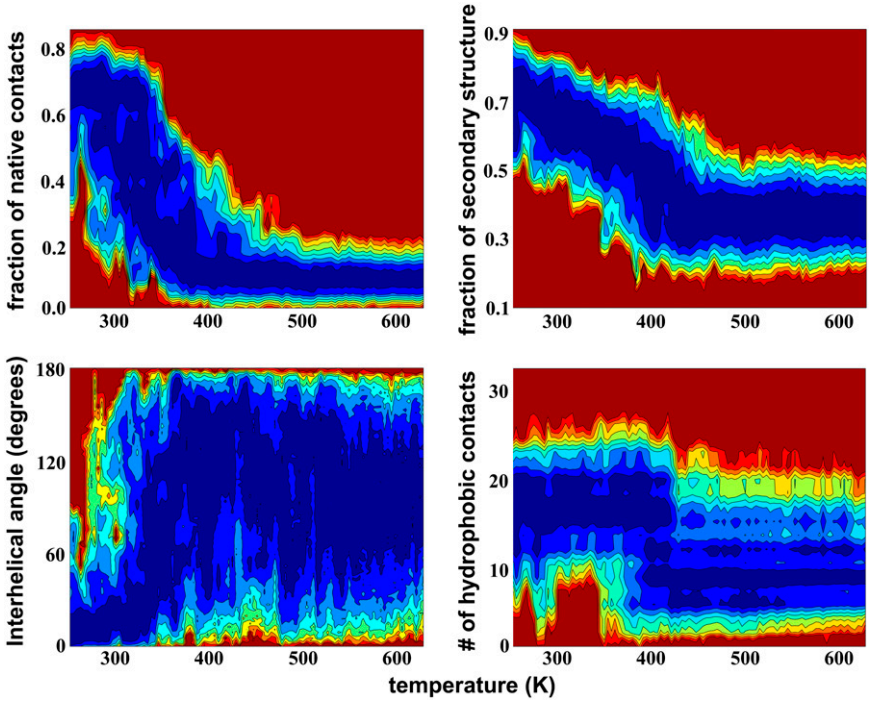

FIGURE 4 Contour plots of the temperature-dependent potentials of mean force in different simulation observables determined from the 1BBL REMD data. Contours are at intervals of 0.5 k_BT and range from dark blue (<0.5 k_BT) to red (>5 k_BT).

A possible physical origin of the noncooperative behavior of 1BBL is suggested by the intraprotein interaction energies seen in the REMD data (Fig. 6). Protein conformations from the last 10 ns of REMD at each temperature were analyzed in terms of their total, bonded (bonds, angles, dihedrals, and one to four interactions), van der Waals, and electrostatic interactions. The van der Waals term includes only the interactions between protein atoms, omitting solute-solvent effects. The electrostatic term, however, includes both intraprotein Coulomb interactions and the effect of solvent via a continuum generalized Born model (35) with a solvent dielectric of 80. The total and bonded energies both increase with temperature, as expected. The nonbonded terms, however, both show anomalous behavior at low temperatures. In particular, the electrostatic energy of 1BBL is actually higher between 250 and 325 K (where the protein is largely folded) than at

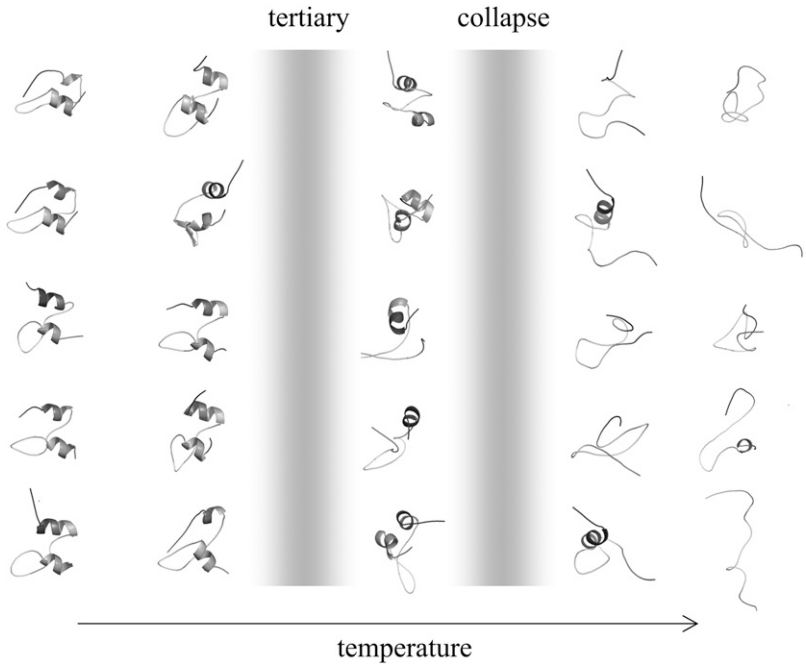

FIGURE 5 Structures representing the five largest clusters populated by 1BBL as a function of temperature during the REMD simulation. The locations of the broad transition involving the loss of tertiary structure and interhelical order (''tertiary'') and the sharper transition involving the loss of hydrophobic contacts (''collapse'') are indicated by gray bars.

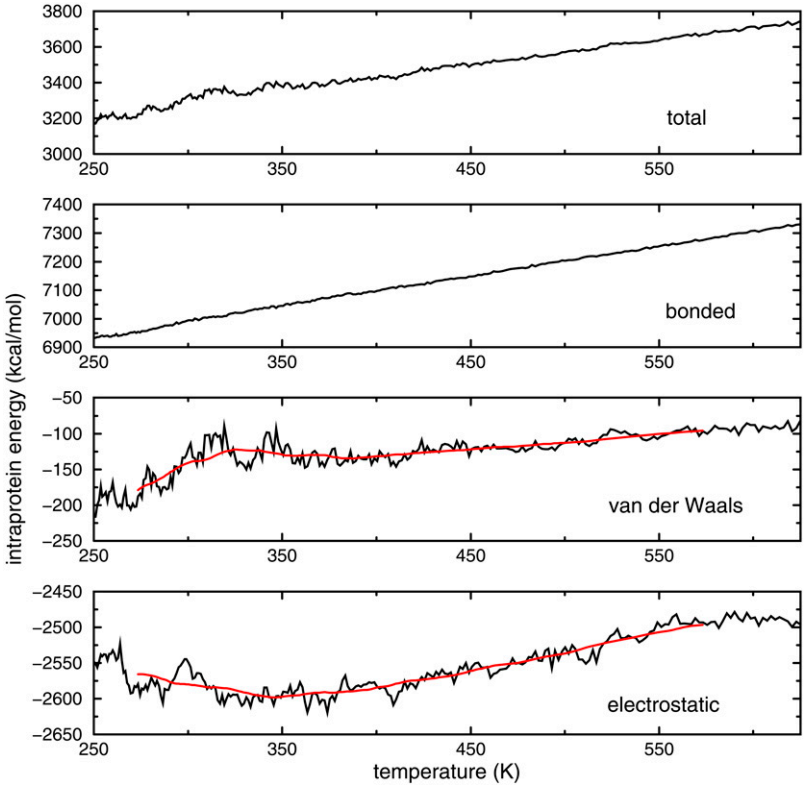

FIGURE 6 Components of the average intramolecular interaction energy for 1BBL from the REMD simulations. The van der Waals energy includes only interactions between protein atoms, whereas the electrostatic term includes both intraprotein Coulomb interactions and a generalized Born approximation for the protein-solvent interaction. A 25 K running average (*red line*) is shown for these two terms.

350 K (where the interhelical order has been lost). This unfavorable electrostatic interaction is compensated for by a significantly lower van der Waals energy. In effect, the folded state of 1BBL appears to be energetically frustrated, with unfavorable electrostatic interactions that are necessary to permit optimal van der Waals packing. Any alteration of the strength of one of these interactions (e.g., by modifying the ionic strength) might significantly alter the folding thermodynamics of 1BBL.

To verify our observation of two sets of transition temperatures in the 1BBL REMD data and test some of these ideas, we ran control simulations of two related sequences, QNND-1BBL and E3BD, and analyzed them in a similar fashion. The QNND-1BBL sequence was also simulated with a different force field (QNND-1BBL-ff99SB) and at an elevated (~100 mM NaCl) ionic strength (QNND-1BBL-ion). The five data sets are compared in Fig. 2. For all three simulations of the QNND-1BBL sequence, it is clear that the interhelical angle and radius of gyration undergo a change at low temperature (between 325 and 350 K), whereas the helices melt at a higher temperature (above 375 K), similar in behavior to the two sets of transitions seen for 1BBL. In contrast, the interhelical angle in E3BD does not change until temperatures above 400 K, where the helices also begin to melt. There is clearly a decoupling of tertiary order from secondary structure in 1BBL and QNND-1BBL that is absent in E3BD. This decoupling is somewhat reduced in the QNND-1BBL-ff99SB simulations, where helices I and II both show signs of melting at lower temperatures (350–400 K for I, 325–350 K for II). However, there is still residual helical content in helix I at temperatures where helix II is completely melted and all interhelical order is lost. A different scenario occurs in the QNND-1BBL-ion simulations at high ionic strength. These calculations are largely similar to the QNND-1BBL data except that the T_m of helix II is significantly depressed (to between 325 and 350 K), whereas helix I is relatively unaffected (T_m still >375 K). Further examination of the QNND-1BBL-ion simulations reveals that helix II has a high density of charged residues (6 out of 10) and appears to be significantly stabilized by salt bridges between arginine 156 and glutamines at positions 161 and 164. When this electrostatic interaction is disrupted by the increase in ionic strength, helix II is significantly destabilized.

These control simulations can also be compared against the simulations of E3BD unfolding at 373 K reported by Ferguson et al. (17). Although those prior calculations unfolded to structures no more than 8 Å C_α-RMSD from the native state, our replica-exchange simulations of 1BBL are completely unfolded by their criterion (>6 Å C_α-RMSD) at that temperature, ranging from 6.4 Å to 16.5 Å C_α-RMSD. Our 375 K simulation of QNND-1BBL also shows complete unfolding, with a C_α-RMSD range from 6.6 to 11.3 Å over the last 5 ns. Similar to the Ferguson et al. simulations, our 400 K simulation of E3BD is 51% unfolded by the same measure, sampling C_α-RMSD values from 4.7 to 7.0 Å over

the last 5 ns. Visual inspection of our 400 K E3BD simulation shows it to still be largely folded, with extensive native-like secondary structure and tertiary contacts.

DISCUSSION

REMD simulations of 1BBL and conventional molecular dynamics simulations of the related QNND-1BBL sequence under varying conditions all show evidence of noncooperative folding thermodynamics in this putative downhill folder. Although our calculations cannot provide definitive evidence of a one-state folding transition, we observe broad, weakly cooperative unfolding transitions in most order parameters. Different observables show different transition temperatures, with a group of transitions around 325 K in observables related to the tertiary order of the protein. A second set of transitions between 360 and 400 K represent the melting of helices I and II and the loss of nonspecific hydrophobic contacts. Direct calculation of ^{1}H and ^{13}C chemical shifts from the simulation data supports this view with a clearly multimodal distribution of T_m values. In a rough sense, our simulations suggest a hierarchical mechanism for folding in 1BBL, where hydrophobic collapse and helix formation can occur at temperatures where the native structure is not stable. In this state, the two helices are free to sample a number of different relative orientations and packing arrangements. The ensemble is dominated only by structures with native-like tertiary structure at very low temperatures.

Though extensive, our simulations are not long enough to be fully converged at lower temperatures. Our analysis indicates that some of the high temperature transitions are stable, but there is still a significant downward shift in the T_m of low temperature transitions with increasing simulation time. Since the simulations were started using folded conformations, we expect anomalously high apparent T_m values until equilibration is achieved. As a result, we do not expect that any of our reported T_m values are quantitatively correct. Given the stability of the high temperature transitions, however, we do not expect the qualitative features of 1BBL's folding thermodynamics to change with further simulation.

The ionic strength sensitivity of 1BBL can be combined with the microscopic details seen in our simulations to provide a possible explanation for its noncooperative folding. The two helices of 1BBL have a number of charged residues at pH 7.0, yet there are no native salt bridges between them. The low-temperature loss of tertiary structure is associated with the formation of an average of one nonnative salt bridge between the two helices. The folded state has, on average, a higher intramolecular electrostatic energy than the ensemble of partially unfolded structures. This is compensated for by more favorable van der Waals interactions in the folded state. The folded state of 1BBL appears to be energetically frustrated, with tightly packed tertiary structure formed at the expense of unfavorable electrostatic interactions. ϕ-value analysis of E3BD folding also suggests the presence of unfavorable electrostatic interactions in the native and transition states (36).

Alterations to the pH or ionic strength significantly alter these electrostatic interactions and shift the conformational ensemble. This hypothesis is supported by the observations of Ferguson et al. (15) that the folded state of QNND-1BBL is destabilized at low ionic strength. The specific nature of this destabilization is evident from our simulations of QNND-1BBL in 100 mM NaCl, where we observe a significant decrease in the T_m of helix II but little change in the behavior of helix I. Helix I may, therefore, constitute the residual helical structure in the unfolded state of 1BBL, whereas variations in buffer conditions may tune the melting of helix II to be either coincident or distinct from other structural transitions.

One clear issue with our simulations is the significant shift we observe in simulated T_m values relative to experiment. Many of the transitions we observe for 1BBL, QNND-1BBL, and E3BD occur at temperatures from 350 to 400 K, whereas the experimental melting temperatures for these proteins range from 280 to 340 K. This is a common problem, and many other simulations of protein folding thermodynamics have shown significantly elevated melting temperatures (37,38). Part of the issue may be equilibration, as noted above. The other explanation may lie with the specific force field and simulation conditions used. The high T_m values we see for helices I and II in 1BBL may in part be due to the AMBER parm96 force field used (24), though some researchers have reported this force field to preferentially stabilize β- rather than α-structure (39). The high pressure at these elevated simulation temperatures is expected to have little effect on α-helical stability (40). Further control simulations with the AMBER parm99SB force field (25) showed a significant decrease in T_m values for both helices, though the melting of helix I still occurs at temperatures well above the loss of interhelical order.

CONCLUSIONS

With the simulation data reported in this work and the preponderance of experimental data now available (2,13–15,17,19,36), there is significant evidence for noncooperative folding thermodynamics in 1BBL. Although our simulations cannot definitively label 1BBL as a one-state downhill folder, we find broad, barrierless transitions and a range of T_m values for different observables supporting that possibility. Our simulations also provide detailed insight into the structural changes occurring in 1BBL as a function of temperature. For both 1BBL and the related QNND-1BBL sequence, these transitions appear to form two classes: a low temperature loss of native tertiary structure that preserves nonspecific hydrophobic interactions and a high-temperature loss of helical secondary structure and disruption of all hydrophobic contacts. This noncooperativity of structural transitions is reflected in the range of transition temperatures seen for experimental (19) and simulated (Fig. 3) ^{1}H and ^{13}C chemical shifts.

Also, these results are at least somewhat robust to changes in the protein force field, as two different AMBER force field variants both show multiple melting temperatures for different structural transitions. Further simulations at experimental ionic strengths revealed that the stability of helix II is highly sensitive to the counterion environment, whereas helix I is relatively unaffected. In contrast with experiment, however, our simulations at high ionic strength show a destabilization of 1BBL rather than the experimentally observed stabilization. Regardless, the differing ionic strength sensitivity of the two helices may account for some of the variability in experimental studies of 1BBL, since different ionic strengths might cause the helical T_m values to coincide (as in cooperative folding) or separate (as in a downhill scenario).

In contrast to 1BBL and QNND-1BBL, simulations of the related E3BD protein show a single set of high temperature transitions, in agreement with prior simulation and experimental data (17) and supporting the idea that the 1BBL protein folds differently from its homologs. Finally, there is evidence from our simulations that the native state of 1BBL is energetically frustrated, a possible source of its unusual folding behavior and an interesting issue for the use of frustration-free models (41) in simulation studies of folding.

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

J.W.P. and W.C.S. thank Prof. David Case for helpful discussions regarding NMR chemical shift estimation.

Work by F.F.A. was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

REFERENCES

1. Bryngelson, J. D., J. N. Onuchic, N. D. Socci, and P. G. Wolynes. 1995. Funnels, pathways, and the energy landscape of protein-folding: a synthesis. *Proteins.* 21:167–195.
2. Garcia-Mira, M. M., M. Sadqi, N. Fischer, J. M. Sanchez-Ruiz, and V. Munoz. 2002. Experimental identification of downhill protein folding. *Science.* 298:2191–2195.
3. Munoz, V. 2002. Thermodynamics and kinetics of downhill protein folding investigated with a simple statistical mechanical model. *Int. J. Quantum Chem.* 90:1522–1528.
4. Fersht, A. R. 1999. Structure and Mechanism in Protein Science. W. H. Freeman, New York.
5. Eaton, W. A. 1999. Searching for ''downhill scenarios'' in protein folding. *Proc. Natl. Acad. Sci. USA.* 96:5897–5899.
6. Kubelka, J., J. Hofrichter, and W. A. Eaton. 2004. The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* 14:76–88.
7. Yang, W. Y., J. W. Pitera, W. C. Swope, and M. Gruebele. 2004. Heterogeneous folding of the trpzip hairpin: full atom simulation and experiment. *J. Mol. Biol.* 336:241–251.
8. Hagen, S. J. 2003. Exponential decay kinetics in ''downhill'' protein folding. *Proteins.* 50:1–4.
9. Oliva, F. Y., and V. Munoz. 2004. A simple thermodynamic test to discriminate between two-state and downhill folding. *J. Am. Chem. Soc.* 126:8596–8597.
10. Robien, M. A., G. M. Clore, J. G. Omichinski, R. N. Perham, E. Appella, K. Sakaguchi, and A. M. Gronenborn. 1992. Three-dimensional solution structure of the E3-binding domain of the dihydrolipoamide succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of Escherichia coli. *Biochemistry.* 31:3463–3471.
11. Spector, S., and D. P. Raleigh. 1999. Submillisecond folding of the peripheral subunit-binding domain. *J. Mol. Biol.* 293:763–768.
12. Spector, S., B. Kuhlman, R. Fairman, E. Wong, J. A. Boice, and D. P. Raleigh. 1998. Cooperative folding of a protein mini domain: the peripheral subunit-binding domain of the pyruvate dehydrogenase multienzyme complex. *J. Mol. Biol.* 276:479–489.
13. Ferguson, N., P. J. Schartau, T. D. Sharpe, S. Sato, and A. R. Fersht. 2004. One-state downhill versus conventional protein folding. *J. Mol. Biol.* 344:295–301.
14. Naganathan, A. N., R. Perez-Jimenez, J. M. Sanchez-Ruiz, and V. Munoz. 2005. Robustness of downhill folding: guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry.* 44:7435–7449.
15. Ferguson, N., T. D. Sharpe, P. J. Schartau, S. Sato, M. D. Allen, C. M. Johnson, T. J. Rutherford, and A. R. Fersht. 2005. Ultra-fast barrier-limited folding in the peripheral subunit-binding domain family. *J. Mol. Biol.* 353:427–446.
16. Barber, M. N. 1983. Finite size scaling. *In* Phase Transitions and Critical Phenomena. C. Domb and J. L. Lebowitz, editors. Academic Press, New York. 145.
17. Ferguson, N., R. Day, C. M. Johnson, M. D. Allen, V. Daggett, and A. R. Fersht. 2005. Simulation and experiment at high temperatures: ultrafast folding of a thermophilic protein by nucleation-condensation. *J. Mol. Biol.* 347:855–870.
18. Zuo, G., J. Wang, and W. Wang. 2006. Folding with downhill behavior and low cooperativity of proteins. *Proteins.* 63:165–173.
19. Sadqi, M., D. Fushman, and V. Munoz. 2006. Atom-by-atom analysis of global downhill protein folding. *Nature.* 442:317–321.
20. van Gunsteren, W. F., D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hunenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. van der Vegt, and H. B. Yu 2006. Biomolecular modeling: goals, problems, perspectives. *Angew. Chem. Int. Ed.* 45:4064–4092.
21. Pitera, J. W., I. Haque, and W. C. Swope. 2006. Absence of reptation in the high-temperature folding of the trpzip2 beta-hairpin peptide. *J. Chem. Phys.* 124:141102.
22. Larios, E., J. W. Pitera, W. C. Swope, and M. Gruebele. 2006. Correlation of early orientational ordering of engineered lambda(6–85) structure with kinetics and thermodynamics. *Chem. Phys.* 323:45–53.
23. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.
24. Cornell, W. D., J. W. Caldwell, and P. A. Kollman. 1997. Calculation of the phi-psi maps for alanyl and glycyl dipeptides with different additive and non-additive molecular mechanical models. *J. Chim. Phys.* 94:1417–1435.
25. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. 2006. Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins.* 65:712–725.
26. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.
27. Case, D. A., T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. 2005. The AMBER biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.
28. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald—an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
29. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of Cartesian equations of motion of a system with

constraints—molecular dynamics of N-alkanes. *J. Comput. Phys.* 23: 327–341.

30. Andersen, H. C. 1980. Molecular dynamics simulations at constant pressure and-or temperature. *J. Chem. Phys.* 72:2384–2393.
31. Sugita, Y., and Y. Okamoto. 1999. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* 314:141–151.
32. Frishman, D., and P. Argos. 1995. Knowledge-based protein secondary structure assignment. *Proteins Struct. Funct. Genet.* 23:566–579.
33. Xu, X. P., and D. A. Case. 2001. Automated prediction of N-15, C-13(alpha), C-13(beta) and C-13′ chemical shifts in proteins using a density functional database. *J. Biomol. NMR.* 21:321–333.
34. Xu, X. P., and D. A. Case. 2002. Probing multiple effects on N-15, C-13 alpha, C-13 beta, and C-13′ chemical shifts in peptides using density functional theory. *Biopolymers.* 65:408–423.
35. Tsui, V., and D. A. Case. 2000. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers.* 56:275–291.
36. Ferguson, N., T. D. Sharpe, C. M. Johnson, and A. R. Fersht. 2006. The transition state for folding of a peripheral subunit-binding domain contains robust and ionic-strength dependent characteristics. *J. Mol. Biol.* 356:1237–1247.
37. Pitera, J. W., and W. Swope. 2003. Understanding folding and design: replica-exchange simulations of ''Trp-cage'' miniproteins. *Proc. Natl. Acad. Sci. USA.* 100:7587–7592.
38. Zhou, R. H., B. J. Berne, and R. Germain. 2001. The free energy landscape for beta hairpin folding in explicit water. *Proc. Natl. Acad. Sci. USA.* 98:14931–14936.
39. Sorin, E. J., and V. S. Pande. 2005. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys. J.* 88:2472–2493.
40. Paschek, D., S. Gnanakaran, and A. E. Garcia. 2005. Simulations of the pressure and temperature unfolding of an alpha-helical peptide. *Proc. Natl. Acad. Sci. USA.* 102:6765–6770.
41. Go, N. 1983. Theoretical studies of protein folding. *Annu. Rev. Biophys. Bioeng.* 12:183–210.