PROTEINS: Structure, Function, and Bioinformatics 65:373–391 (2006)

Criteria for Downhill Protein Folding: Calorimetry, Chevron Plot, Kinetic Relaxation, and Single-Molecule Radius of Gyration in Chain Models With Subdued Degrees of Cooperativity

Michael Knott and Hue Sun Chan*

Department of Biochemistry, and of Medical Genetics and Microbiology, Protein Engineering Network of Centres of Excellence, Faculty of Medicine, University of Toronto, Toronto, Ontario M5S 1A8, Canada

ABSTRACT **Recent investigations of possible downhill folding of small proteins such as BBL have focused on the thermodynamics of non-two-state, "barrierless" folding/denaturation transitions. Downhill folding is noncooperative and thermodynamically "one-state," a phenomenon underpinned by a unimodal conformational distribution over chain properties such as enthalpy, hydrophobic exposure, and conformational dimension. In contrast, corresponding distributions for cooperative two-state folding are bimodal with well-separated population peaks. Using simplified atomic modeling of a three-helix bundle—in a scheme that accounts for hydrophobic interactions and hydrogen bonding—and coarse-grained C_α models of four real proteins with various degrees of cooperativity, we evaluate the effectiveness of several observables at defining the underlying distribution. Bimodal distributions generally lead to sharper transitions, with a higher heat capacity peak at the transition midpoint, compared with unimodal distributions. However, the observation of a sigmoidal transition is not a reliable criterion for two-state behavior, and the heat capacity baselines, used to determine the van't Hoff and calorimetric enthalpies of the transition, can introduce ambiguity. Interestingly we find that, if the distribution of the single-molecule radius of gyration were available, it would permit discrimination between unimodal and bimodal underlying distributions. We investigate kinetic implications of thermodynamic noncooperativity using Langevin dynamics. Despite substantial chevron rollovers, the relaxation of the models considered is essentially single-exponential over an extended range of native stabilities. Consistent with experiments, significant deviations from single-exponential behavior occur only under strongly folding conditions. Proteins 2006;65:373–391.** © 2006 Wiley-Liss, Inc.

Key words: heat capacity; van't Hoff enthalpy; sigmoidal transition; chevron rollover; FRET; single-exponential kinetics

INTRODUCTION

Polymer physics is important to the study of proteins in at least two respects. First, it is a logical starting point for an analysis aimed at providing a physico-chemical basis for the structure and stability of proteins.[1–6] Second, natural proteins are a subset of all possible polypeptides. As polymers, polypeptides can, in principle, possess a broad range of properties shared with other chain molecules. By contrasting the properties of natural proteins with the general behavior of polymers, we can clarify the biological significance of, and the selection pressure for, certain generic protein properties that are remarkable from a polymer perspective. An obvious example is that many natural proteins have an essentially unique native structure, even though homopolymers and polypeptides with random amino acid sequences generally have degenerate ground states. Thus, for these proteins, the biological requirement of functional specificity entails selection of uniquely folding sequences with significantly nonrandom interaction patterns.[7–13]

In-depth comparisons between natural protein behavior and general polymer properties have been made recently with regard to the thermodynamic and kinetic cooperativity[14] of small, single-domain proteins that fold and unfold with apparent two-state kinetics.[15–17] Analyses of this seemingly "all-or-none" switch-like behavior have provided insight into the critical roles of interaction specificity and nonadditivity,[14,18,19] as well as solvation effects,[20–22] in giving rise to the folding barriers[23–26] and the high degrees of cooperativity exhibited by many

Grant sponsors: Premier's Research Excellence Award (Ontario), Protein Engineering Network of Centres of Excellence (PENCE), Canada Foundation for Innovation, and Ontario Innovation Trust; Grant sponsor: Canadian Institutes of Health Research; Grant number: MOP-15323.

*Correspondence to: Hue Sun Chan, Department of Biochemistry, and of Medical Genetics and Microbiology, University of Toronto, 1 King's College Circle, MSB 5363, Toronto, Ontario M5S 1A8, Canada. E-mail: chan@arrhenius.med.toronto.edu

Received 13 March 2006; Revised 19 April 2006; Accepted 26 April 2006

Published online 14 August 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21066

small natural proteins. Meanwhile, experimental study of the noncooperative folding kinetics of a *de novo* designed protein suggests that cooperativity is an evolutionarily selected property,[27] perhaps guarding against aggregation.[28] So, cooperativity, especially in its kinetic aspect, is not simply a corollary of an amino acid sequence's ability to fold to a unique native structure[29]; in other words, it is not a physico-chemical requirement for the folding process.[30] Simulation studies of a continuum simplified atomic model also demonstrate that a protein can be thermodynamically noncooperative, yet have the ability to fold reversibly to an essentially unique native structure.[31,32] The principle is further buttressed by cases where noncooperative, downhill, faster-folding proteins can be engineered experimentally by modifying a wild-type sequence that exhibits more cooperative behavior.[30,33]

In the energy landscape perspective,[2,34] the thermodynamics of folding and unfolding may be viewed as a competition between intraprotein interactions and conformational entropy. As the compact, unique native structure is approached during folding, native-favoring interactions tend to reduce the free energy. At the same time, chain compaction leads to a decrease in conformational freedom, which tends to increase the free energy. If the intraprotein interactions are so specific that only conformations very similar to the native state are strongly favored, the loss of conformational entropy outweighs the favorable interactions during the initial stages of folding, resulting in a free-energy barrier at an intermediate value of some progress variable,[35,36] and a two-state-like (bimodal) conformational distribution. If the native-favoring interactions are less specific, their contribution is significant earlier in the folding process. If it outweighs the loss of conformational entropy, the free-energy profile lacks a barrier, and the conformational distribution is unimodal.

For proteins that are thermodynamically two-state-like, analyses of folding and unfolding trajectories simulated using explicit-chain models suggest strongly that folding/unfolding kinetics are cooperative, with an extensive quasilinear chevron regime, when the folding/unfolding free-energy barrier is sufficiently high.[22,37,38] This is because the high barrier diminishes the population of intermediate conformations that could become kinetic traps and cause chevron rollover.[39] A high barrier tends to imply slower folding than a lower barrier,[36,40] provided that the energy landscapes in both cases are sufficiently smooth. Part, though not all, of this general trend of behavior is also captured by a recent analytical energy landscape theory that does not involve an explicit chain representation.[41] Indeed, in both theory[37] and experiment,[30] cooperative proteins often fold more slowly than their less-cooperative counterparts. Therefore, the question "why is folding so slow?" is as important as (if not more important than[37]) the question "why is folding so fast?" (see e.g., Ref. 42) that is motivated by the so-called Levinthal's paradox.[43,44]

Since chymotrypsin inhibitor 2 (CI2) was first demonstrated in 1991 to be an apparent two-state folder,[15] much of the research on protein folding kinetics has focused on the increasing number of proteins that have been found to be two-state (see e.g., Ref. 45). To match theory to experiment, protein chain models have been systematically evaluated for their ability to reproduce the "smooth" energy landscapes, with no significant kinetic traps, and the remarkable cooperative behavior of two-state proteins.[14,19,35,46]

However, two-state behavior is not a universal property of natural proteins.[14,38] There are at least two possible reasons for significantly non-two-state behavior. First, even if a high degree of cooperativity and landscape smoothness are desirable for certain biological functions, physics and chemistry limit the extent to which these features can be achieved by evolutionary or artificial design of amino acid sequences. In fact, nonnative interactions are often important in protein folding,[47–50] and the folding kinetics of larger proteins, including many investigated in earlier studies, is generally more complex,[23,47,51–53] sometimes involving chevron rollovers that may be viewed[54] as manifestations of glasslike behavior.[2]

Second, there seems to be no reason to think that downhill folding is impossible in principle. Given the diversity of functional requirements, it is reasonable to assume that evolution has exploited not only two-state but also non-two-state behaviors that are to be found among the many possible protein folding energy landscapes,[2,34,55] if they happen to be fit for certain biological functions. A protein that folded in a downhill manner might be able to access different conformations and energy levels, according to the conditions. Indeed, recent research has shown that some natural proteins are intrinsically disordered, and that disordered states[56] can serve biological functions.[57–60] For instance, it has been suggested that the chaperone-like function of the periplasmic protein *hdeA* from *E. coli* is performed by its acid-induced disordered state in the stomach, rather than by its ordered state at neutral pH.[61] Furthermore, dynamic interconversions among energetically similar conformations can entail promiscuity in function (polyreactivity), a trait that is likely to have an important role in evolution.[62] It is therefore very important that we clarify the biophysical implications of possible noncooperative folding behavior.

Kinetic and Thermodynamic Criteria for Downhill Folding

The original argument of Bryngelson et al. for the possibility of downhill folding was based on thermodynamic consideration of free-energy profiles under strongly native conditions. This formulation allows both single- and multiple-exponential kinetic relaxations.[2] However, subsequent treatments have emphasized nonexponential (i.e., multiple- or stretched-exponential) "strange kinetics" as a signature of downhill folding.[63,64] Multiple-exponential

behavior has indeed been shown to be associated with a continuous collapse process in barstar.[65] With the advent of experimental techniques for measuring fast-folding events,[66–68] fast-folding proteins that attain a certain presumed speed limit[69,70] are viewed as prime candidates for downhill folding. This reasoning is inspired by Kramers' theory of reaction barriers in solution,[71] in which the folding speed limit is identified as the pre-exponential (front) factor, estimated experimentally from the kinetic rates of assembly of basic protein structural elements such as α helices[72] and β hairpins,[73] and of forming pairwise interresidue contacts.[74–76] In this framework, folding is downhill at the speed limit because it implies a vanishing exponential barrier in a Kramers-inspired abstraction of the folding process.

In line with this emphasis on nonexponential relaxation as a signature of downhill folding, polypeptide expansion/collapse processes that exhibit an activation enthalpy and single-exponential relaxation have been taken to be two-state.[77] However, although model studies indicate significant deviation from single-exponential behavior[39,78–80] under *strongly* downhill folding conditions, it has also been shown that, under certain downhill folding scenarios, the form of the relaxation can be practically indistinguishable from single-exponential.[81] Therefore, the observation of single-exponential relaxation *per se* does not preclude downhill folding. Moreover, the presence of an enthalpic barrier to folding need not imply a corresponding free-energy barrier, because the gain in solvent entropy upon desolvation of the protein core[21,22] can compensate for the unfavorable enthalpy at the rate-limiting step.

Indeed, connections between folding speed, nonexponential kinetics, and downhill folding are intricate and await elucidation. For instance, ultrafast-folding proteins do not necessarily exhibit nonexponential kinetics. Cases in point include α_3D[67] and a covalently cross-linked dimeric GCN4 coiled-coil.[82] Both fold very fast, with folding time ~3–5 μs, yet their folding kinetics is single-exponential. For GCN4, data from native-state hydrogen exchange argue quite strongly for cooperative rather than downhill folding.[82] More fundamentally, recent experiments, concerning the dependence of folding rates on solvent viscosity, indicate that a Kramers-inspired universal diffusional speed limit may not be a well-conceived concept for protein folding.[83,84] This is because not all of the frictional impediments to folding can be attributed to the solvent; internal friction, which is expected to depend on the protein,[54] can be important in the very fast folding regime.[83,84]

The possibility of downhill folding has also been explored with thermodynamic experiments, most notably on the protein BBL, which is the peripheral subunit-binding domain from *E. coli*'s 2-oxoglutarate dehydrogenase multienzyme complex.[85–87] Indeed, in so far as they are designed to probe the underlying conformational distribution,[2] thermodynamic measurements may be viewed as providing more direct information about downhill folding than kinetic measurements. For BBL, it has been reported that the equilibrium folding/unfolding transitions monitored by different probes (calorimetry and various spectroscopic techniques) do not coincide with one another. As is the case with a more recent observation of probe-dependent kinetics in the downhill folding of a λ_{6-85} mutant,[88] this probe-dependent thermodynamics suggests that different parts of BBL are not folding and unfolding in a cooperative, all-or-none manner, as they would if the process were two-state.[85] According to these reports, the conformational distribution of BBL, with respect to properties monitored by the experimental probes, appears to be unimodal for a wide range of conditions including those of the transition region.

This contrasts with the behavior of certain other proteins, whose folding is downhill (i.e., the conformational distribution is unimodal) under strongly native conditions, but is two-state when the native bias is weakened.[33,63] In certain calorimetrically cooperative proteins that exhibit a bimodal conformational distribution and a high free-energy barrier at the transition midpoint, the denatured ensemble shifts as conditions become more native.[18,89,90] Under conditions strongly favoring the native state, the conformational distribution may change from bimodal to unimodal, and the folding landscape become downhill. In the present article, the term "downhill" is reserved for model proteins that display a unimodal distribution over the entire range of conditions considered, the behavior claimed for BBL.[85]

Modeling as a Way of Classifying Experimental Folding Behavior

Recently, the conclusion of Garcia-Mira et al.[85] that BBL is a downhill folder has been disputed.[91] Apparently, this objection and the subsequent discussions[92,93] hinge on whether, and if so how, various calorimetric and spectroscopic monitoring techniques, and their conventional interpretation in the context of protein folding, permit reliable inferences about whether the conformational distribution is unimodal or bimodal. Seeking to clarify, and ultimately help to resolve, the fundamental issues in this debate, the present computational investigation uses explicit-chain, simplified protein models[32,78,94–97] to delineate the logical relationship between experimental probes and the underlying conformational distribution.

Of course, we cannot use the simulation of simplified models directly to answer the question of whether a specific protein such as BBL folds in a two-state or in a downhill manner. That question must be settled by experiment. However, simulation affords us access to the underlying behavior of a large class of models. This access can be used to classify a model according to its conformational population distribution and the consequent thermodynamic and kinetic properties. Such detailed information is valuable because, currently, direct experimental investigations of the distribution of conformational properties, or of individual kinetic trajectories, are often impractical, although single-molecule

techniques are making rapid advances in closing this gap.[98–107] In the present conceptual framework, a protein investigated experimentally may be fitted into the classification scheme according to its observable properties, and conclusions drawn about its likely underlying conformational distribution and the extent to which its folding process should be described as "downhill" or "two-state." We can also decide which observable characteristics may be relevant, and which irrelevant, to the nature of the folding process. In the event that current experimental criteria are inadequate, the same investigative logic may be used to formulate new criteria.

THEORY AND MODELS

The Simplified Atomic Model

We first consider a simplified atomic model that explicitly represents all of the backbone N, C_α, and C′ atoms, the O atom attached to the C′, and the H atom attached to the N.[108] This allows hydrogen bonds to be represented realistically as interactions between H and O atoms; it is not an all-atom representation, however, because the H attached to the C_α is neglected, and the side group is represented only by a single virtual C_β atom. This C_β can be polar, hydrophobic, or absent (to represent a glycine residue which facilitates turns). Recently, a slightly modified version of this Ramachandran approach[108] has been applied to the folding of real proteins with some success.[109] Here, we focus on a 54-residue model, containing 162 backbone and 154 side-chain atoms, that is designed to fold to a three-helix bundle native-like conformation.[31,108,110,111]

We have recently studied the thermodynamics of this model in detail, especially with regard to its apparent failure to conform to the calorimetric criterion for two-state cooperativity, an observation that makes the model well suited for exploring possible downhill behavior.[32] The present investigation tackles the folding kinetics of the model to elucidate the relationship between thermodynamic and kinetic aspects of downhill folding. The potential function contains eight terms; as the formalism is the same as before,[32] only two key terms are summarized here.

The hydrophobic interaction

$$V_{\mathrm{hp}} = \varepsilon_{\mathrm{hp}} \sum_{m,n} \left[\left(\frac{\sigma_{\mathrm{hp}}}{l^{m,n}} \right)^{12} - 2 \left(\frac{\sigma_{\mathrm{hp}}}{l^{m,n}} \right)^{6} \right], \quad (1)$$

where $l^{m,n}$ is the distance between atoms m and n, sums over all pairs of virtual C_β that are both hydrophobic. The parameters are $\sigma_{\mathrm{hp}} = 5.0$ Å, $\varepsilon_{\mathrm{hp}} = 0$ if $l^{m,n} > 8.0$ Å, $\varepsilon_{\mathrm{hp}} = 2.2$ if $l^{m,n}/\sigma_{\mathrm{hp}} < 0.89$, and $\varepsilon_{\mathrm{hp}} = 1.0$ otherwise. The hydrogen bonding energy is given by

$$V_{\mathrm{hb}} = \varepsilon_{\mathrm{hb}} \sum_{m,n} \left[5 \left(\frac{\sigma_{\mathrm{hb}}}{l^{m,n}} \right)^{12} - 6 \left(\frac{\sigma_{\mathrm{hb}}}{l^{m,n}} \right)^{10} \right] \cos^2 \alpha \, \cos^2 \beta, \quad (2)$$

where the summation is performed over all pairs of atoms that are made up of one H and one O, $\alpha = 180^\circ - \angle\mathrm{NHO}$, $\beta = 180^\circ - \angle\mathrm{HOC}'$, $\sigma_{\mathrm{hb}} = 2.0$ Å, $\varepsilon_{\mathrm{hb}} = 0$ if $\alpha > 90^\circ$ or $\beta > 90^\circ$, and $\varepsilon_{\mathrm{hb}} = 2.8$ otherwise. No long-range cutoff is imposed on this term. During our simulation, a hydrogen bond is considered to exist between atoms (m, n) if the conditions 1.5 Å $< l^{m,n} <$ 3.0 Å, $\alpha < 45^\circ$ and $\beta < 45^\circ$ are all satisfied. All geometric and energetic parameters in the model, including those just described, are as in our previous study.[32] Unlike those of Gō-like native-centric models[78] (see the next section), the potential function is general and transferable: its definition is not biased toward any particular folded structure.

Coarse-Grained Native-Centric Models

To address the physical origin of variations in cooperativity among proteins, and their likelihood of exhibiting downhill folding, we use C_α native-centric Gō-like modeling[78,112–114] for four small proteins in the Protein Data Bank (PDB), referred to here as BBL, NTL9, HP36, and Psbd41. Specifically, BBL comprises residues 12–48 of the E3-binding domain of the dihydrolipoamide succinyltransferase core from the 2-oxoglutarate dehydrogenase multienzyme complex of *E. coli* (PDB ID: 1BBL).[115] This 37-residue chain is a truncated version of the ~40-residue forms of BBL used in the experiments mentioned earlier.[85–87,91–93] NTL9 comprises residues 1–39 of the N-terminal domain of ribosomal protein L9 (PDB ID: 1CQU).[116] This 39-residue fragment, a shorter form of 56-residue[25] and 51-residue[117] versions of the same protein, has been reported to fold cooperatively.[116] HP36 is a thermostable subdomain made up of residues 41–76 of the chicken villin headpiece (PDB ID: 1VII).[118–120] Psbd41 represents residues 3–43 of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus* (PDB ID: 2PDD).[121] A relative of BBL, it was reported to fold cooperatively,[122] but it has also been argued that it might fold downhill.[85]

These models represent each amino acid by a single position corresponding to the C_α atom in the real protein. The interactions are native-centric: favorable energy terms are assigned on the basis of a native contact set derived from the PDB structure. The potential function takes the same form as those in Refs. 78 and 123 and is similar to that in earlier schemes.[113,114] In the present study, a pair of residues belong to the native contact set if they satisfy two conditions: they are at least four places apart along the chain, and two nonhydrogen atoms, one from each residue, are less than 4.5 Å apart in the PDB structure. This definition, which produces a well-behaved model for certain proteins, is the same as that adopted in two recent studies,[36,124] but is somewhat different from those used previously in other studies. All other parameters here are identical to those of the "without-solvation" model in Ref. 78. This reference also contains a discussion of the biophysical motivation for using native-centric modeling, and of the strengths as well as the limitations of the approach.

Although these without-solvation models have provided useful insights, they tend to be less cooperative than the real proteins they are designed to mimic.[78] A likely reason for this mismatch is that desolvation barriers[20–22,38,125] and many-body effects[38,126–128] are omitted for the sake of computational efficiency. Nonetheless, the results of applying similar models to topology-dependent protein folding[36,114,124] and downhill folding[129,130] lead us to expect that our approach will offer a semiquantitative physical rationalization of variations in cooperativity among different proteins.

Langevin Dynamics

We perform thermodynamic sampling and kinetic simulations of the simplified atomic model and coarse-grained C_α native-centric models using Langevin dynamics[131] in a low-friction regime.[36,131] The variance of the random force is a function of the model temperature, and the (nonrandom) conformational force is the negative gradient of the potential function. The detailed procedure and the parameters adopted are as used in our previous investigations.[32,36,78]

Thermodynamics: Heat Capacity and Radius of Gyration

The heat capacity as a function of temperature is computed from simulation results using the standard relation

$$C_V(T) = \frac{1}{k_B T^2}\left[\langle E^2(T)\rangle - \langle E(T)\rangle^2\right], \quad (3)$$

where $k_B T$ is Boltzmann's constant multiplied by absolute temperature and the angle brackets denote Boltzmann averaging. The total energy E is a sum of potential and kinetic terms.[32,123] Since volume effects are often negligible in protein folding experiments at 1 atm, the simulated energy E and heat capacity C_V may be regarded as equivalent to the experimental enthalpy H and constant-pressure heat capacity C_P. In our discussions, we will simply refer to the simulated C_V as heat capacity.

A peak in the heat capacity is a necessary but not a sufficient condition for calorimetric cooperativity and two-state folding.[18,19] The extent to which folding is cooperative is often judged using the calorimetric criterion, which makes use of the ratio $\kappa_2 = \Delta H_{vH}/\Delta H_{cal}$, where

$$\Delta H_{vH} = 2\sqrt{k_B T_m^2 C_P(T_m)} \quad (4)$$

is the van't Hoff enthalpy at the temperature T_m of the heat capacity peak, and

$$\Delta H_{cal} = \int dT C_P(T) \quad (5)$$

is the calorimetric enthalpy of the transition, which entails an integral of the heat capacity over the transition region. The calorimetric criterion requires $\kappa_2 \approx 1$; if this condition is satisfied, T_m is essentially identical to the midpoint temperature, at which native and denatured states are equally populated.[14,18,19]

Exact satisfaction of $\kappa_2 = 1$ requires an idealized system with two discrete energy values: that is, a system whose density of states comprises two δ-functions. A real protein does not constitute such an idealized system, however. First, interactions between the protein and the solvent imply a temperature-dependent heat capacity even in the absence of conformational change. Second, neither the native nor the denatured state could correspond exactly to a δ-function, and therefore, we expect a contribution to the heat capacity from conformational changes other than the main folding/unfolding transition.[18,19] Consequently, analysis of experimental calorimetric data requires baseline subtractions. The aim of these subtractions is to eliminate effects considered irrelevant to the folding/unfolding transition; the calorimetric criterion is applied to the heat capacity curve after the subtraction. To compare theory and experiment, empirical baseline subtractions must also be applied to simulated heat capacities.[19,32,54]

Another commonly used indicator of cooperative folding is a sigmoidal folding/unfolding transition in some experimental property that is presumed to be a probe of the degree of nativeness. Such properties, which are also used to monitor folding kinetics, include the average radius of gyration[132] and the separation of two parts of the protein.[102] The radius of gyration R_g of a single conformation is defined as

$$R_g^2 = \frac{1}{N}\sum_j |\mathbf{r}_j - \langle \mathbf{r}\rangle|^2, \quad (6)$$

where $\mathbf{r}_j$ is the position vector of the jth atom, N is the total number of atoms, $\langle \mathbf{r}\rangle = \Sigma_j \mathbf{r}_j/N$ is the mean of the position vectors, and $R_g = \sqrt{R_g^2}$. We aim to address the question of how much the changes in R_g, during a transition, can tell us about the underlying conformational distribution.

Folding/Unfolding Kinetics: Cooperativity and Chevron Plots

One of the defining characteristics of apparent two-state folding is a chevron plot[133] whose folding and unfolding arms[15] are linear. As in previous studies by our group,[14,39,78] model chevron plots are constructed here as functions of $\varepsilon/k_B T$, where ε is a parameter for the interaction strength. In the simplified atomic model, the potential takes the form εV, where V is given by Eq. (1) of Ref. 32. In the coarse-grained C_α native-centric models, ε has the same meaning as in Ref. 123. To allow comparison with experimental chevrons at constant T, variations in ε are taken as a rudimentary model for changes in denaturant concentration, and variable-ε simulations at constant T have been used to obtain model chevron plots.[78] Kinetics are determined, in gen-

eral, by the independent variables ε and T. However, for models without elementary desolvation barriers, very similar to those considered here, extensive testing[38] indicates that folding and unfolding rates are essentially controlled only by the ratio ε/T.† This feature allows chevron plots in the present work to be computed at different values of $1/T$, ε being held constant at unity.

One of our goals is to explore whether chevron features can be useful as tests for downhill folding. To construct a chevron plot from simulated trajectories, we need a criterion to decide whether a conformation is folded or unfolded.[38,54] Different choices of criterion may be viewed, roughly, as corresponding to different experimental probes—which may or may not currently be feasible—of the folding/unfolding process. For cooperative protein models with a free-energy barrier (a population minimum) between the compact, native-like and open, disordered conformations, it would be natural either to identify the barrier as a native/denatured demarcation[54] or to exclude part of the barrier (the transition-state region) from the definition of both native and denatured states.[54,78,113] The resulting chevron plot is not expected to be too sensitive to the choice of criterion, within a reasonable range of choices.

However, the choice of criterion is not so straightforward for models that lack a barrier: here, any demarcation would appear arbitrary. In search of consistent criteria that cover such noncooperative behavior as well as cooperative folding, we have adopted here, as a first step, rather stringent folding/unfolding definitions for the coarse-grained model, while noting that the precise relationship between our theoretical criteria and experimental measurements remains to be better elucidated. We classify a conformation as "native" when all the native contacts are in place (the proportion of native contacts $Q = 1$), and as "nonnative" when the number of native contacts q is less than a certain threshold number q_u. For the simplified atomic model, we consider folding to have occurred when at least 34 native-state hydrogen bonds are present (indicating helical structure), and, importantly, none of the pairs of hydrophobic groups that are in contact in the native state is separated by >14.5 Å (indicating that the protein is packed into a bundle). The model is considered to have unfolded when, of the pairs of hydrophobic groups that are in contact in the native state, at least one is separated by at least 29 Å, because this feature is sufficient to indicate that the bundle structure has been broken.

For an apparent two-state folder, the folding and unfolding arms of the experimental chevron plot intersect at the thermodynamic transition midpoint.[15] This feature is reproduced in chain models that are sufficiently cooperative.[22,38] However, in noncooperative models, the intersection point of the arms can depend strongly on the somewhat arbitrary folding/unfolding criterion (see e.g., Fig. 15 in Ref. 38, and discussion later). From a purely theoretical standpoint, it is almost always possible to choose a criterion—if one is not concerned about its physical implications and experimental feasibility—so that the chevron arms cross at the thermodynamic midpoint. In view of this consideration, the coincidence of a chevron-defined and a thermodynamic midpoint is best regarded as a necessary but not as a sufficient condition for two-state folding. The more important features for our attention are the curvature of the arms and the relation between mean and median transition times.

Exponential and Nonexponential Relaxation

Real two-state proteins exhibit single-exponential relaxation[15] following a sudden denaturing change in external conditions (e.g., a jump in temperature or in denaturant concentration). For two-state proteins, this probably reflects the existence of an ensemble of bottleneck, or transition-state, conformations through which all folding routes must pass. However, essentially single-exponential relaxation can also accompany non-two-state folding.[81]

To characterize the extent to which model folding trajectories conform to single-exponential relaxation, we use simulated values of p_{unf}, the proportion of the chain population that has not yet undergone a single folding transition. By definition, this quantity will decay to zero because given enough time every chain would have been folded at least once, even though some of them would unfold subsequently. This quantity is not to be confused with the proportion of the chain population that is unfolded; the latter would take account of "back-transitions" and would generally decay to a finite equilibrium value. We fit p_{unf} to a function $f(t)$, where $f = 1$ for $t \le t_0$ and $f = e^{-k(t-t_0)}$ for $t \ge t_0$. Here, k is the folding rate; as has been discussed,[78] the lower bound t_0 accounts for the average time that the unfolded ensemble may take to pre-equilibrate in adjusting to the new temperature or solvent conditions. It follows that the interval t_0 can depend on the folding conditions.

For a set of M trajectories simulated under the same conditions and yielding folding times t_i, $i = 1,2, \ldots, M$, the fitting parameters t_0 and k in $f(t)$ are determined by numerical minimization of

$$\chi^2 = \sum_{i=1}^{M} \left(\frac{f(t_i) - p_{unf}(t_i)}{\sigma(t_i)} \right)^2 . \quad (7)$$

To estimate $\sigma(t_i)$, we consider a theoretical system comprising M exemplars, each of which has a probability $f(t_i)$ of not having folded by time t_i. The distribution of the number of never-folded exemplars can be modeled by an M-trial binomial distribution with single-trial probability $f(t_i)$. The standard deviation of the distribution is $\sqrt{Mf(t_i)[1 - f(t_i)]}$, and we divide this by M to esti-

†It should be noted, however, that for certain models with pairwise desolvation barriers and more complex thermodynamics for the intrachain interactions,[21,22] rates simulated for the same ε/T can be significantly different for different values of T.

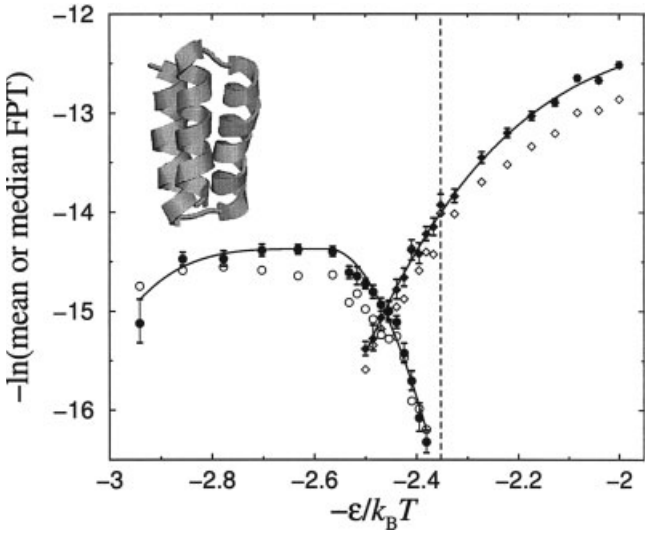

Fig. 1. Chevron plot for the simplified atomic model, showing the negative natural logarithms of the mean first passage times (MFPTs) (filled symbols) of folding (circles) and unfolding (diamonds) as functions of the interaction strength parameter ε/k_BT; also the median FPTs divided by ln 2 (open symbols). $\varepsilon = 1$ is used in the simulations. Curves fitted to the mean FPT values are a guide for the eye. Error bars are standard deviations of the mean. The vertical dashed line marks $\varepsilon/k_BT = 2.35$, corresponding to the heat capacity peak of the thermodynamic folding/unfolding transition (see Fig. 3). The ribbon diagram shows a typical folded conformation at $\varepsilon/k_BT = 2.86$.

mate $\sigma(t_i)$, the margin of error on the simulated p_{unf}.[‡] Accordingly, we apply this approximation and set $\sigma(t_i) = \sqrt{f(t_i)[1-f(t_i)]/M}$ in Eq. (7). To ensure that σ remains nonzero for t_i around t_0 and below, a lower bound is introduced such that it cannot be smaller than the value derived for M trials and single-trial probability $(M-1)/M$, i.e., $(M\sigma)^2 = \max\{1 - 1/M,\ Mf(t_i)[1 - f(t_i)]\}$. We use χ^2/M, calculated from the minimized χ^2, to quantify the deviation from linearity.

RESULTS AND DISCUSSION

Chevron Rollover and Relaxation Behavior in a Noncooperative Folding Scenario: The Simplified Atomic Model

We begin with kinetic analyses of the simplified atomic model; Figure 1 shows a simulated chevron plot for this three-helix bundle. The plot required extensive simulation, because the system contains 316 atoms, which is considerably larger than the Gō-like continuum models for which chevron plots have been simulated.[22,38,78] To our knowledge, chevron plots have not previously been obtained for continuum, nonnative-centric models. One reason for studying the chevron behavior of this model is that it helps us determine the generality of the trends observed in continuum Gō-like constructs: for those proteins which are two-state-like around the transition midpoint, previous model simulations have suggested that folding under conditions of high native stability is associated with severe chevron rollover and nonexponential relaxation.[39,54] More importantly, thermodynamic analyses of the simplified atomic model have shown that its folding/unfolding transition is noncooperative and has downhill characteristics.[32] It is worthwhile, therefore, to use this model to explore the kinetic consequences of downhill folding.

In Figure 1, as in other chevron plots in this study, more negative $-\varepsilon/k_BT$ values correspond to conditions under which the native state is more favored thermodynamically. The mean and median folding times as functions of $-\varepsilon/k_BT$ are derived from simulations which were first run at a high model temperature ($T = 1$) for a random number ($\geq 10^5$) of time steps, then quenched to the desired T. For unfolding, simulations were initialized in the native conformation at low T, and then the temperature was increased. Each mean first passage time in Figure 1 represents the average of 50–90 trajectories, except in one slow-folding case where the average is of 37 trajectories.

The folding arm of the chevron plot in Figure 1 shows a rollover, with a maximum folding rate at $-\varepsilon/k_BT \approx -2.63$; a less severe rollover is also evident for the unfolding arm at $-\varepsilon/k_BT > -2.3$. Nonetheless, there is a quasilinear regime[37] between $-\varepsilon/k_BT \approx -2.45$ and -2.38. The near coincidences of the mean first passage times (FPTs) with the median FPTs divided by ln 2, for this range of $-\varepsilon/k_BT$ values, indicate that kinetic relaxation is essentially single-exponential[54,134] under the quasilinear regime. In contrast, the discrepancies elsewhere in the plot between the mean FPTs and the median FPTs divided by ln 2 show that relaxation deviates more from single-exponential under stronger or weaker folding conditions. We include error bars on the mean FPTs to emphasize that the discrepancies are significant.

Figure 2 focuses on folding relaxation, showing in more detail the crossover from single-exponential to non-single-exponential behavior.[39,78,135–137] Here, p_{unf} represents the distribution of folding times at six different ε/k_BT values: when $\ln p_{\mathrm{unf}} = -3$, for example, the probability of the model having folded once is about 0.95. In Figure 2, the upper panel contains folding time distributions for ε/k_BT values under the quasilinear regime, whereas the lower panel covers more strongly folding conditions.

At $\varepsilon/k_BT = 2.38$ and 2.44, which are, respectively, near to and slightly more favorable than the ε/k_BT value at the heat capacity peak (see Fig. 3), the decay of the unfolded state appears to be single-exponential. We know this because the inclined dashed lines in the upper panel of Figure 2 are a good fit to the unfolding curve. Using the procedure specified by Eq. (7), the fitted folding rates are, respectively, $k = 9.6 \times 10^{-8}$ and 4.3×10^{-7}, with $t_0/10^6 = 1.46$ and 1.41; the minimized values of χ^2/M, quantifying the deviation from linearity, are 0.28 and 0.35.

This essentially single-exponential behavior implies that, after allowing a short time, t_0, for pre-equilibration, the folding probability of an individual protein is

[‡]Fitting procedures used in several other recent investigations of folding time distributions (e.g., Refs. 78 and 80) effectively adopt a constant σ.[139]

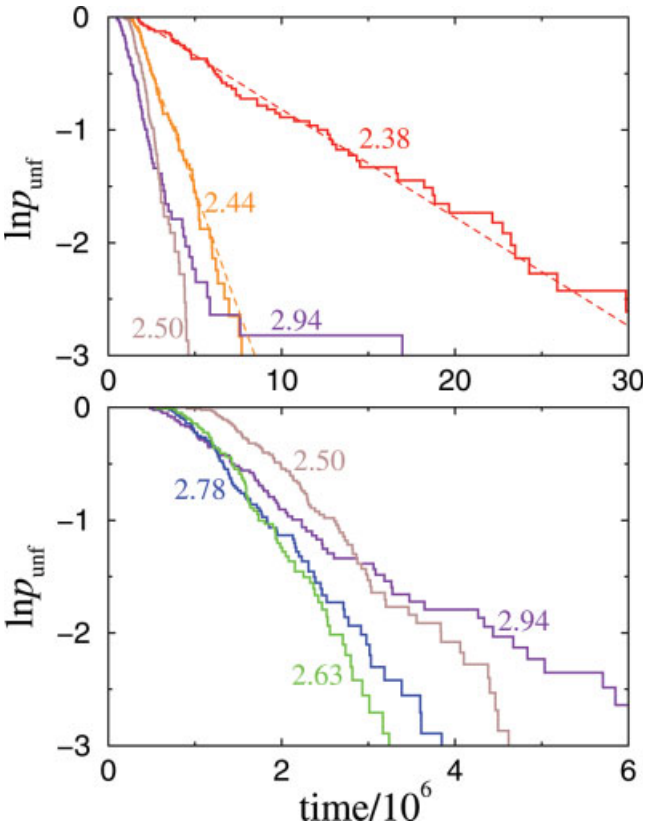

Fig. 2. Folding time distributions for the simplified atomic model under different folding conditions, as characterized by ε/k_BT values indicated in the plots. Data in this figure are from the same simulations as the folding arm of the chevron plot in Figure 1. Starting from fully unfolded ensembles, each curve shows the logarithm of the probability p_{unf} that a protein chain remains unfolded, as a function of time (number of simulation time steps). Different scales are used for the horizontal axes of the two panels, for a clearer presentation. Inclined dashed lines in the upper panel represent single-exponential fits [see Eq. (7)] to data for $\varepsilon/k_BT = 2.38$ and 2.44.

independent of time under the quasilinear chevron regime.[§] Under this regime, folding rate increases with more favorable folding conditions, while the decay remains exponential. However, the p_{unf} plot curves downward under stronger folding conditions (e.g., at $\varepsilon/k_BT = 2.63$ in the lower panel of Figure 2, for which $\chi^2/M = 1.04$, substantially larger than the values quoted earlier for $\varepsilon/k_BT = 2.38$ and 2.44). This suggests that, perhaps, because of the more downhill shape of the folding landscape under these conditions, it may no longer be meaningful to distinguish the "pre-equilibration" process of adjusting to the initial quench from the folding process itself. It would be interesting to further analyze this prediction, and its experimental implications, in future studies.

[§]The nonnegligible pre-equilibration times observed in folding simulations (Eq. (4) of Ref. 78) is a kinetic feature that can, in principle, be corroborated by single-molecule experiments. This behavior is also relevant[138] to the accuracy of folding time estimations obtained from certain methodologies in large-scale distributive computing simulations.

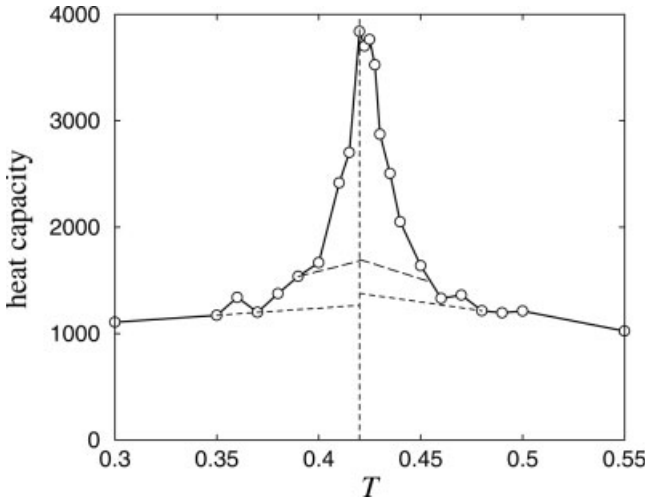

Fig. 3. Heat capacity of the simplified atomic model as a function of temperature T (in model units). Circles represent the simulation results; lines joining them are a guide for the eye. The vertical dashed line corresponds to the temperature $T_m = 0.42$ of the heat capacity peak; inclined dashed lines are the baselines used in the calculation of the van't Hoff to calorimetric enthalpy ratio $\kappa_2^{(s)}$.

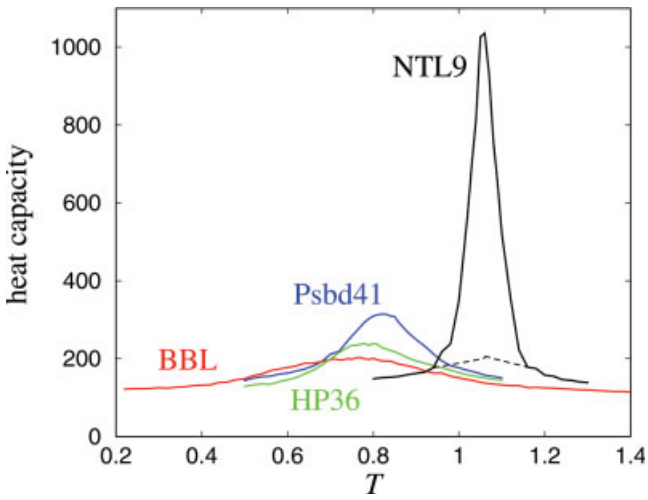

Fig. 4. Heat capacities of the coarse-grained C_α native-centric models of fragments NTL9 (black), BBL (red), HP36 (green), and Psbd41 (blue). Inclined dashed lines indicate baselines used in the calculation of the van't Hoff to calorimetric enthalpy ratio $\kappa_2^{(s)}$ for the NTL9 model.

When folding conditions become even stronger ($-\varepsilon/k_BT < -2.63$), the folding arm "rolls down"; its gradient becomes positive. Early in the folding process, the rate appears to follow the same trend as at weaker folding conditions (i.e., it increases as folding conditions become stronger). Subsequently, it slows down ($\varepsilon/k_BT = 2.94$ curve in the upper panel of Fig. 2). This is probably caused by kinetic trapping: strongly folding conditions in a nonnative-centric model favor intrachain contacts, largely irrespective of their consistency with the native fold.[39] Because of this slowdown, which gives rise to non-single-exponential relaxation, the time taken to reach $p_{unf} \approx 0.05$ (~95% have undergone folding once), for example, is about the same at $\varepsilon/k_BT = 2.94$ as at

2.44 (compare crossing of these two curves in the upper panel of Fig. 2). As has been pointed out, the theoretical possibility of a chevron rolldown may not be realizable under common experimental conditions, as the required level of native stability may be higher than is physically achievable even at zero denaturant concentration.[54] It would be interesting to explore whether stabilizing agents such as osmolytes[140] can offer access to a wider range of chevron behavior.

The chevron behavior of this simplified atomic model may be compared to that of lattice[37,39,54,127] and Gō-like continuum[38,78] models. A feature common to all model chevron plots determined so far is that relaxation is essentially single-exponential for the quasilinear part of the folding arm, while kinetic trapping renders folding-arm rollover more severe, and introduces non-single-exponential relaxation[37,39,54,127] as folding conditions get stronger. Chevron plots for certain lattice models with hydrophobic-like interactions lack a quasilinear regime, because of trapping effects; their folding rate decreases with increasing native stability even around the transition midpoint [see Fig. 5(B) of Ref. 14 for a 3-letter 27mer model or Fig. 15 of Ref. 38 for a 20-letter 15mer side-chain model]. Thus, kinetic trapping appears less severe in the simplified atomic model than that in some lattice models. Apparently, kinetic trapping is inhibited by the specificity provided by the hydrogen bonding term, allowing a quasilinear chevron regime to emerge. Alternatively, the difference might be partially due to the reduced number of degrees of freedom in the lattice model. Here, however, the most important lesson of Figure 2 is that downhill folding can be consistent with a narrow—but not extended—regime of quasilinear chevron behavior and single-exponential relaxation.[93]

Calorimetric Cooperativity Criteria and Baseline Subtractions

To highlight downhill characteristics of the simplified atomic model, we now briefly revisit its thermodynamics.[32] Figure 3 shows its heat capacity as a function of temperature. Computed values are shown by circles in the plot: each is derived from at least 3×10^5 conformations sampled from 1.5×10^8 simulation time steps for $0.38 \leq T \leq 0.47$, or from at least 1×10^5 conformations sampled from 5×10^7 simulation time steps elsewhere. This is equivalent, with more extensive data, to the upper panel of Figure 3 in Ref. 32.

Our previous investigation of the simplified atomic model indicates that it is a thermodynamically downhill folder.[32] Its energy distribution has a single peak, which shifts with changing temperature. The distribution is never bimodal, but is most spread out in the transition region (Fig. 9 of Ref. 32), giving rise to a heat capacity peak. However, such a peak by itself does not imply that the transition is calorimetrically two-state.[18,19] In principle, a more quantitative criterion is that the van't Hoff to calorimetric enthalpy ratio $\Delta H_{\mathrm{vH}}/\Delta H_{\mathrm{cal}} \approx 1$.

However, this criterion requires caution because, practically, the derivation of ΔH_{vH} and ΔH_{cal} from experimental measurements calls for baseline subtractions.[14,18,141] Empirical procedures for setting calorimetric baselines (which effectively define the native state ensemble[19,54]) can involve considerable uncertainties; this can have a significant impact on the conclusions drawn about the underlying conformational distribution and the cooperativity of the transition.[19,87,92] Do the tails observed in simulated heat capacity curves simply result from a model's failure to mimic the real protein?[19] Or do they correspond (whatever other faults the model may suffer from) to tails that are present in real systems, and that are lost in the experimental baseline subtraction? Is it possible that information indicating a transition to be noncooperative could be destroyed by an experimental baseline subtraction, leading us to conclude wrongly that the transition is cooperative?

TABLE I. Geometric and Thermodynamic Properties of the Coarse-Grained C_α Native-Centric Models

	BBL	HP36	Psbd41	NTL9
Number of residues	37	36	41	39
Number of native contacts	39	50	56	93
Native contacts per residue	2.11	2.78	2.73	4.77
T_{m}	0.77	0.78	0.82	1.06
$\Delta H_{\mathrm{vH}}/\Delta H_{\mathrm{cal}}$	0.51	0.68	0.71	0.94

These considerations are illustrated by the two sets of baselines in Figure 3. The baselines are constructed empirically as approximate tangents extending from the low- and high-T tails of the heat capacity to the transition region. Both sets of baselines are intuitively reasonable, yet the van't Hoff to calorimetric enthalpy ratios resulting from the two sets are significantly different: the lower set gives[19] $\Delta H_{\mathrm{vH}}/\Delta H_{\mathrm{cal}} = \kappa_2^{(\mathrm{s})} = 0.56$, whereas the upper set gives $\kappa_2^{(\mathrm{s})} = 0.80$. Nonetheless, the calorimetric criterion—though not always definitive—can be discriminating to a degree: higher baselines than the upper set in Figure 3 would not have appeared reasonable. The fact that even the upper baselines fail to satisfy the criterion $\Delta H_{\mathrm{vH}}/\Delta H_{\mathrm{cal}} \approx 1$ serves to emphasize the downhill nature of the transition and the underlying unimodal energy distribution.[32]

Thermodynamic simulations of our coarse-grained Gō-like models of BBL, NTL9, HP36, and Psbd41 allow us to explore the influence of native topology on thermodynamic cooperativity. Figure 4 shows the heat capacities: each value is derived from at least 2×10^5 conformations sampled from 1×10^8 simulation time steps. It is clear from the very different heat capacity profiles that the sharpness of the folding/unfolding transition varies considerably among the four models, even though their chain lengths are all very similar (Table I). The heat capacities possess peaks of varying height: NTL9 has a sharp peak at a model transition temperature $T_{\mathrm{m}} = 1.06$, while the other three have less prominent peaks.

Interestingly, the degree of thermodynamic cooperativity is related to the topology of the native state, at least in this coarse-grained modeling scheme. As shown in Table I, the numbers of native contacts are quite differ-

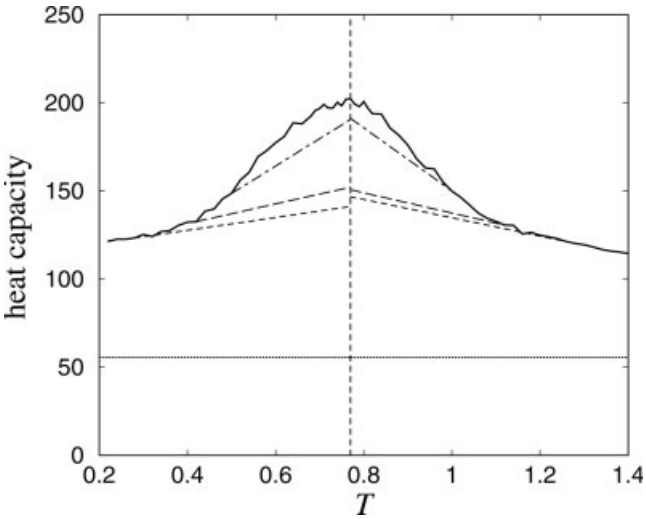

Fig. 5. Enlargement of the model BBL heat capacity curve from Figure 4, showing the three pairs of baselines used to calculate the van't Hoff to calorimetric enthalpy ratio $\kappa_2^{(s)}$ for this coarse-grained C_α native-centric model. The vertical dashed line marks the temperature of the heat capacity peak. The horizontal dotted line indicates the kinetic contribution $3Nk_B/2$ to the heat capacity[32] ($N = 37$, $k_B = 1$ in the present units).

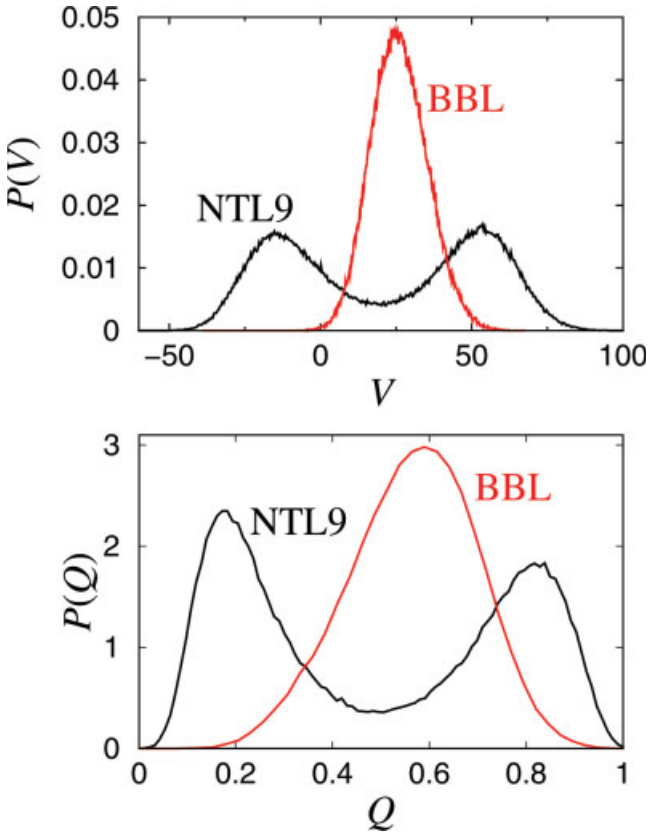

Fig. 6. Probability distributions P_i as a function of potential energy V (upper panel), and as a function of fractional number of native contacts Q (lower panel), for the coarse-grained C_α native-centric models of NTL9 (black) and BBL (red), computed at their respective transition temperatures T_m.

ent in the different fragments. For instance, each residue in BBL makes an average of 2.11 native contacts, compared to 4.77 in NTL9. Within this sample, there is a clear positive correlation between the contacts-per-residue aspect of native topology and the apparent degree of calorimetric cooperativity. The values of $\Delta H_{vH}/\Delta H_{cal}$ listed in Table I are obtained using baseline subtractions[19]; the results are consistent with a very similar trend first noted by Zuo et al.,[129,130] who use κ_2 values[19] without baseline subtractions for a different, and larger, set of coarse-grained models that includes BBL.

The different behaviors of the coarse-grained models for NTL9 and BBL may be used to contrast two-state-like (bimodal) and downhill folding scenarios. As emphasized earlier, the models are not expected to be quantitatively accurate for every property of the proteins they seek to mimic. In particular, it is almost certain that the cooperativity (and the $\Delta H_{vH}/\Delta H_{cal}$ ratio) is lower than that of the real protein,[21,22,78] owing to the neglect of many-body interactions and detailed desolvation effects. Notwithstanding these limitations, this coarse-grained topology-based modeling (Table I) successfully captures part of the physics that underlies the difference in experimental cooperativity. Take NTL9 and BBL as examples: despite the short length of the NTL9 fragment, experimental evidence that it folds in an apparent two-state manner is persuasive, especially since its chevron plot has an extended linear regime.[116] On the other hand, although the question of whether BBL is a downhill folder remains a controversial one,[85,91–93] there is experimental evidence for non-two-state behavior. This includes an apparent mismatch between m-values determined by thermodynamic methods and that deduced from a chevron plot that spans a rather narrow stability range (page 441 of Ref. 93). Because of the small size of BBL, such non-two-state features are unlikely a manifestation of discrete intermediate states separated from both the native and denatured ensembles by free-energy barriers, as has been proposed for other proteins.[142] In view of this consideration and our simulation results, the experimental non-two-state properties may be seen as evidence for a very low free-energy barrier or barrierless downhill folding for BBL.

Consistent with the trend discussed earlier, Table I shows that the van't Hoff to calorimetric enthalpy ratio is much higher for NTL9 than for BBL. The NTL9 model's prominent heat capacity peak allows a more straightforward choice of baselines. An intuitively reasonable set is indicated in Figure 4; these baselines lead to a value of $\kappa_2^{(s)} = 0.94 \approx 1$. In contrast, the choice of a "reasonable" pair of baselines for the BBL model is problematic. Figure 5 considers three pairs: the lowest pair leads to $\kappa_2^{(s)} = 0.51$ (Table I), the middle pair to $\kappa_2^{(s)} = 0.58$, and the top pair to $\kappa_2^{(s)} = 1.01$. The top pair is clearly unreasonable. It does not correspond to a set of baselines that would have been chosen for experimental data. Our purpose of showing these unreasonable baselines is to emphasize that the BBL model can appear cooperative only if an unreasonable set of baselines is chosen. Even from the

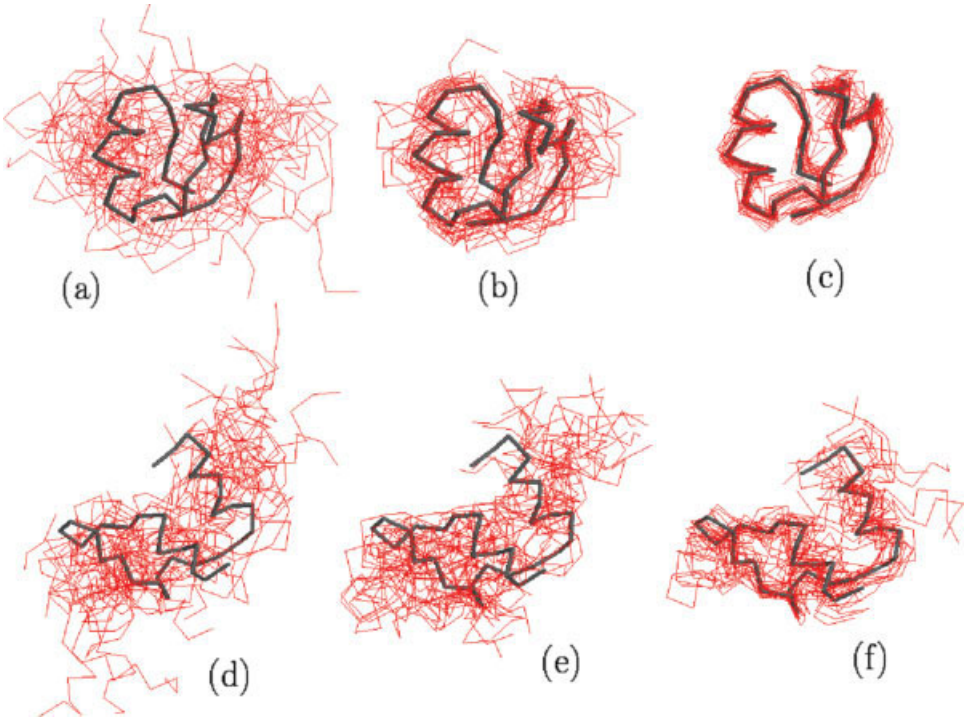

Fig. 7. Typical conformations of the coarse-grained C_α native-centric models of NTL9 **(a)–(c)** and BBL **(d)–(f)**. PDB native structures are depicted in thick black lines; 20 conformations for each set of conditions are drawn in thinner red lines. The conformations are sampled at various Q values; conditions are as in Figure 6. Positions of these Q values in the probability distributions $P(Q)$ can be readily ascertained from the lower panel of Figure 6. Numbers of native contacts and Q values are as follows. (a)–(c) NTL9: (a) at the low-Q peak of $P(Q)$, with 17 native contacts, $Q = 0.183$; (b) around the $P(Q)$ trough, with 45–49 native contacts, $0.484 \leq Q \leq 0.527$; (c) at the high-Q peak of $P(Q)$, with 77–78 native contacts, $0.828 \leq Q \leq 0.839$. (d)–(f) BBL: (d) at the low-Q tail of the $P(Q)$ distribution, with 8–10 contacts, $0.205 \leq Q \leq 0.256$; (e) at the single $P(Q)$ peak, with 23 contacts, $Q = 0.590$; and (f) at the high-Q tail of the $P(Q)$ distribution, with 33–37 contacts, $0.846 \leq Q \leq 0.949$.

heat capacity alone, we reach a robust conclusion that this coarse-grained BBL model is not two-state.

Conformational Distributions: Nativeness Measure and Structural Diversity

Figure 6 confirms the calorimetric inference that the distribution of energy (essentially equivalent to enthalpy) is bimodal for the NTL9 model but unimodal for the BBL model, under transition conditions. The same applies to the distribution of the nativeness measure Q (fractional number of native contacts). The BBL model is clearly non-two-state, as its distributions have only one maximum. In contrast, the distributions of the NTL9 model have two maxima, so we characterize this model as "two-state-like." But the data also show that the NTL9 model is not highly cooperative: intermediate values of V and Q have substantial probabilities. In this respect, the NTL9 model is less cooperative than similar Gō-like models for larger proteins, which have much lower intermediate-Q probabilities.[22,78] As has been emphasized,[19] these results imply that $\kappa_2^{(s)} \approx 1$ cannot rule out a substantial chain population with intermediate enthalpies. Furthermore, as a general feature of this class of Gō-like models without desolvation barriers (see above), the low-V and low-Q maxima in the NTL9 distributions do not correspond to the $Q = 1$ PDB structure. Indeed, the NTL9 native peak in Figure 6 at $Q \approx 0.85$ is rather broad: it represents a conformational ensemble with considerable structural diversity (compare discussion in Ref. 22). Although $\kappa_2^{(s)} \approx 1$ is an informative observation, the criterion allows a wider range of behavior than the highly cooperative two-state folding to which it is intended to correspond.

Figure 7 further illustrates differences between the conformational distributions of two-state-like and downhill folders under transition conditions. At low Q values, both NTL9 and BBL models adopt a diverse set of mainly open conformations, as expected [Fig. 7(a,d)]. However, the two models are obviously different at intermediate Q. For the NTL9 model, intermediate values of Q are minimally populated (see Fig. 6), and conformational diversity is significantly more limited than at low Q [compare Fig. 7(a,b)]. For a two-state-like protein chain model, intermediate-Q conformations are closely related to conformations in the transition-state ensemble.[36,113] Figure 7(b) shows that intermediate-Q conformations of the NTL9 model cluster loosely around the PDB structure. On the other hand, intermediate Q values of the downhill-folding BBL model are maximally populated (see Fig. 6), while conformational diversity

is high and comparable to that at low Q [compare Fig. 7(d,e)]. The difference between the two-state-like and the downhill folder is also conspicuous at high Q values. Here, conformations of the NTL9 model cluster tightly around the PDB structure [Fig. 7(c)]; the downhill-folding BBL model shows more conformational diversity [Fig. 7(f)].

Distribution of Radius of Gyration and Intraprotein Distances as Signatures for Downhill Folding Behavior

As Figure 7 suggests, the influence on chain dimension—or conformational size—of changing folding conditions may be used to discriminate between two-state-like and downhill behavior. The chain dimension of proteins is commonly characterized by the radius of gyration R_g (see above), accessible experimentally by means of small angle X-ray scattering (SAXS).[143–146] Figure 8 shows how the average R_g of our four coarse-grained models varies with folding conditions (upper panel). To various degrees, the mean R_g of each model undergoes a sigmoidal transition from low $\langle R_g \rangle$ values at low T (folding conditions) to high $\langle R_g \rangle$ values at high T.

For these models, all of which have very similar chain lengths, a higher degree of cooperativity is associated with a sharper transition in $\langle R_g \rangle$ around T_m. The slopes of the $\langle R_g \rangle$ curves at T_m, and therefore the suggested degrees of cooperativity, follow the decreasing order NTL9 > Psbd41 > HP36 > BBL, which is identical to the ordering determined from calorimetric cooperativity (Table I and Fig. 4).

A sharp transition does not, however, necessarily imply an underlying bimodal distribution,[147] let alone conformity to the more stringent calorimetric criterion[19] or highly cooperative two-state folding. For this reason, it would be desirable to use the *distribution* of R_g or of some intraprotein distance as an alternative test for two-state behavior. A protein would qualify as two-state if the distribution showed a sharp, temperature-independent peak (corresponding to the native conformation) well below T_m, and if this peak became less prominent under less native conditions, and finally disappeared, without changing its position. In contrast, the peak would move continuously into the denatured region in a downhill folding scenario.

The lower panel of Figure 8 shows the distributions of R_g for the fragments at their respective transition temperatures T_m. Varied behavior is observed: The NTL9 model has a sharp peak at low R_g, with a broader, nearly flat-topped distribution (which does not quite achieve a second peak) from $R_g \approx 9.5$ Å to $R_g > 13$ Å. This behavior is attributable to a well-populated native-like low-R_g ensemble at T_m: the overall R_g distribution for the NTL9 model is a superposition of the native-like peak and the broad distribution of the unfolded conformations. The BBL model exhibits a single, wide, nearly symmetric round-topped distribution with its peak located at intermediate R_g. Such a distribution is consistent with the unimodal distributions of V and Q (see Fig. 6). The Psbd41 and HP36 models are intermediate between NTL9 and BBL. Unlike the BBL model, their distributions are significantly asymmetric. Whereas the Psbd41 model exhibits a shoulder at $R_g \approx 11.5$ Å ($P(R_g)$ concave downward), the less cooperative HP36 model has a thick "tail" but no shoulder ($P(R_g)$ concave upward at $R_g \approx 11.5$ Å). Since the population peak is not at $Q = 1$ for this class of coarse-grained models (as discussed earlier—compare Fig. 6), in no case does the peak of the R_g distribution correspond to the PDB structure.

Figure 9 presents simulation results for the dependence of the R_g distribution on folding conditions. The top panel indicates that, for the most cooperative of our models (NTL9), the distribution changes from a sharp low-R_g peak at low T to a broad, flat distribution at high T. In contrast, the downhill-folding BBL model (middle panel) does not show a sharp peak even well below T_m, and there is no dramatic change in behavior as T increases: the peak simply broadens, shortens, and shifts. The simplified atomic model is intermediate between the NTL9 and BBL models, as might be expected, given the calorimetric results (Figs. 3–5 and Table I).

Figures 8 and 9 suggest that qualitative features of R_g distributions may distinguish downhill folders with a shifting unimodal distribution, like the BBL model, from two-state-like folders such as the NTL9 model (see Fig. 6). Experimental data on the R_g distribution, if available, would help to uncover possible downhill behavior. However, Figure 9 also indicates that high-precision quantitative measurement may be needed to distinguish between a downhill folder such as the simplified atomic model, which is more cooperative than the BBL model (compare Fig. 6 of this article and Fig. 9 of Ref. 32), and a two-state-like but only *weakly cooperative* protein such as the coarse-grained NTL9 model (compare top and bottom panels of Fig. 9): their qualitative features are comparable. Direct measurement of R_g distributions is difficult with current SAXS techniques. Physically motivated distributions (such as that of Flory and Fisk[1,148]), that are consistent with data from SAXS and other structural techniques, may be proposed in some cases (see e.g., Ref. 144). These fitted distributions provide important insights, but remain hypothetical, and several distributions can fit a set of experimental data equally well. In view of these limitations, it is desirable to develop new experimental techniques that permit essentially assumption-free determination of R_g distributions.

These considerations also suggest that we might use other measures of intraprotein distances to discriminate between two-state and downhill folding. For instance, spatial separations between two parts of the chain can be determined using time-resolved fluorescence resonance energy transfer (FRET).[149–151] Single-molecule FRET[101–103,107] may also provide distributions of intraprotein separations. In principle, such data could yield information similar to that provided by the R_g distributions in Figures 8 and 9. As tests for downhill folding, these methods would avoid baseline subtraction prob-

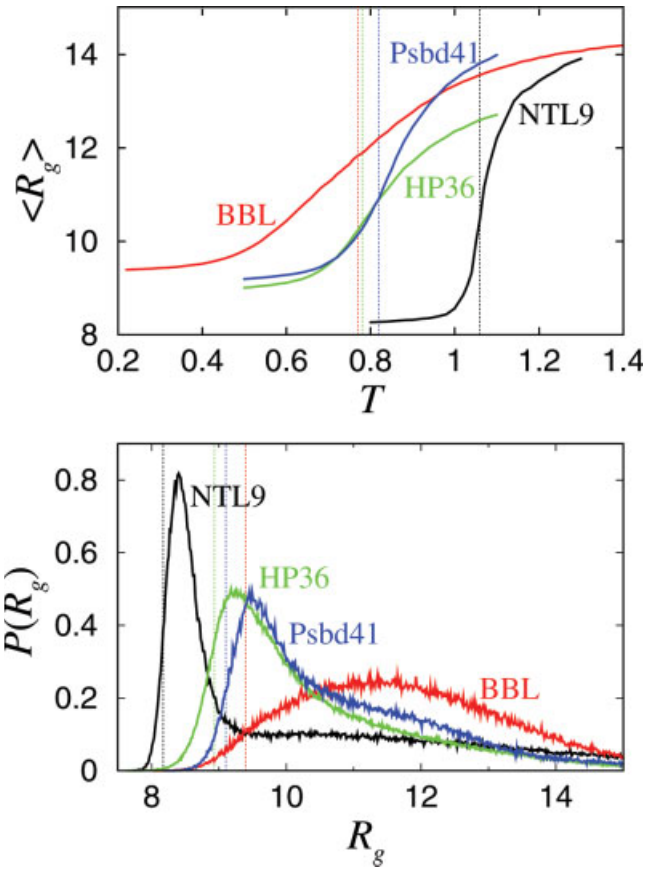

Fig. 8. Radius of gyration as a probe for folding/unfolding cooperativity. Upper panel: average radius of gyration $\langle R_g \rangle$ of the coarse-grained C_α native-centric models of NTL9 (black), BBL (red), HP36 (green) and Psbd41 (blue), as functions of temperature T. Vertical dotted lines indicate temperatures T_m of the heat capacity peaks (compare Fig. 4). Lower panel: population (probability) distribution $P(R_g)$ of radius of gyration, at the temperature T_m of the peak heat capacity of each of the models. Radii of gyration of their PDB native structures are indicated by vertical dotted lines (same color coding). All R_g values in this article are given in units of Å.

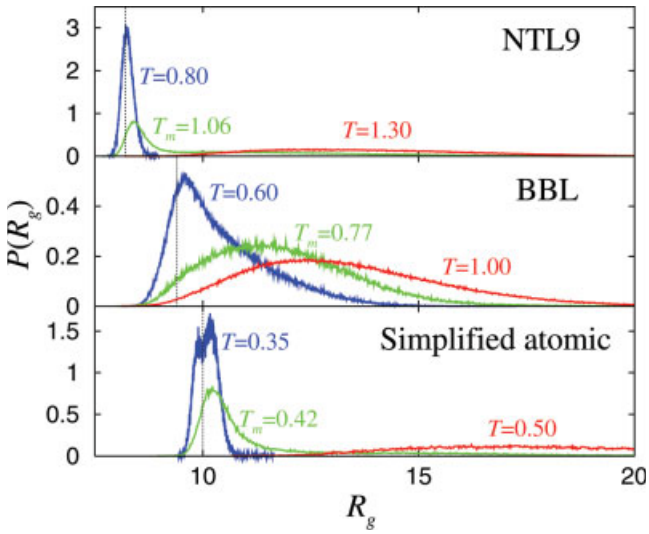

Fig. 9. Distribution $P(R_g)$ of radius of gyration at various temperatures T, for the coarse-grained C_α native-centric models of NTL9 (top panel) and BBL (middle panel), and the simplified atomic model (bottom panel). Radii of gyration of the native states are marked by vertical dotted lines.

lems, and might also mitigate uncertainties associated with the interpretation of sigmoidal unfolding curves.

As the distinction between two-state and downhill folding scenarios is ultimately a question of conformational distribution, experimental probes that offer a more thorough characterization of a conformation are better equipped to provide more definitive answers. In as much as single-molecule properties are concerned, structural properties such as R_g and intraprotein distances can be more reliably determined than the energy or free energy of a conformation. Moreover, as suggested by the study on BBL, "interaction" probes such as FRET that detect intraprotein distances appear to be more sensitive to possible noncooperativity of the folding process (by showing more discrepancies with other measurements such as calorimetry) than probes that measure the local environment of a given residue or probes such as circular dichroism that measure overall structural features (Figs. 2 and 3 of Ref. 85).

Kinetic Characteristics of Weakly Cooperative and Noncooperative Downhill Folding: Single-Molecule Trajectories and Chevron Plots

We now explore the kinetic properties of the coarse-grained models. Figure 10 compares representative folding trajectories for the two-state-like NTL9 model and the downhill-folding BBL model. Folding is considered to have occurred when the chain first reaches $Q = 1$. At the transition temperature (upper panels), the example trajectory for NTL9 undergoes a rough transition after $\approx 4 \times 10^5$ simulation time steps from an unfolded regime ($Q < 0.5$ most of the time) to a native-like regime ($Q > 0.5$ most of the time), although it does not show the sharp hopping behavior of models that are even more cooperative (e.g., Fig. 2 of Ref. 38). In contrast, the BBL model, at its transition temperature, fluctuates over a wide range centered on $Q \approx 0.5$. The features of this trajectory are consistent with those determined by Zuo et al.[130] A similar contrast is seen under stronger folding conditions (lower panels). Once $Q > 0.5$ is achieved, the trajectory of NTL9 leads to folding shortly after, whereas the BBL model spends more time at intermediate values of Q before it eventually folds. In real proteins, such differences should be detectable by single-molecule measurements.[103,105,106]

Figure 11 shows chevron plots for all four coarse-grained model fragments. The simulation procedure is similar to that used for the simplified atomic model. Folding data are derived from simulations which were first run at a high temperature ($T = 2$) for a random number ($\geq 10^5$) of time steps, then quenched to the desired T; unfolding data are derived from simulations which were initialized in the PDB structure at low T, and then raised to the desired T. Mean first passage times are averaged from 200 trajectories, with the excep-

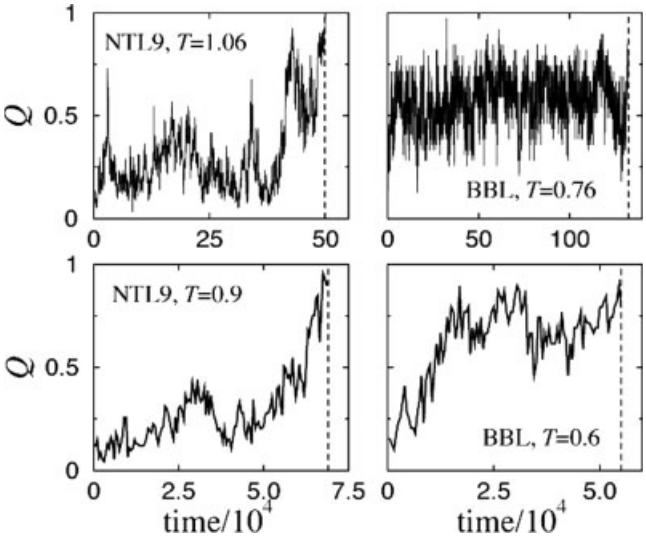

Fig. 10. Representative folding trajectories for coarse-grained C_α native-centric models of NTL9 and BBL, showing how the fractional number of native contacts, Q, evolves with time. The plots show trajectories recorded at temperatures around the transition midpoints, corresponding to the positions of the heat capacity peaks (upper panels), as well as under strongly folding conditions, i.e., at substantially lower temperatures (lower panels). Trajectories are constructed from Q values recorded every 500 simulation time steps; their folding times are marked by vertical dashed lines.

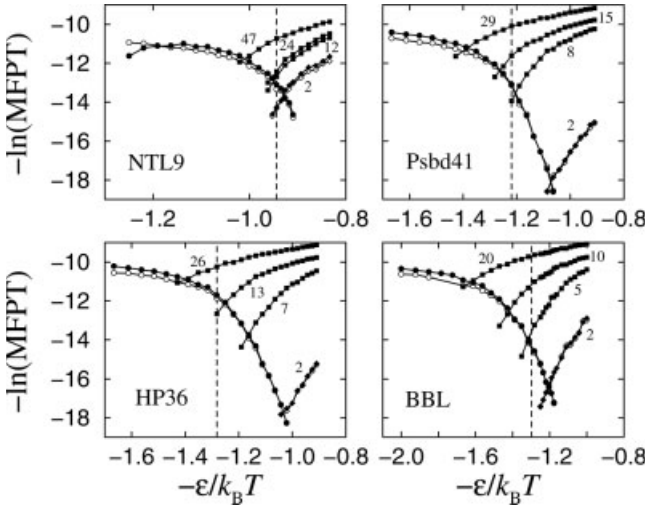

Fig. 11. Chevron plots for the coarse-grained C_α native-centric models of NTL9, BBL, HP36, and Psbd41, showing negative logarithms of the mean first passage times of folding (filled circles) and unfolding ($q_u = 2$, filled diamonds), as functions of the interaction strength parameter ε/k_BT. Negative logarithms of median FPTs of folding (open circles) and unfolding (open diamonds), divided by ln 2, are also plotted for comparison (compare Fig. 1). Filled squares indicate mean FPTs of unfolding using alternative definitions of the unfolded state, with q_u values given. Lines joining datapoints of simulation results are a guide for the eye. The vertical dashed lines mark the heat capacity peaks.

tion of a few slow-folding and slow-unfolding datapoints for the BBL, HP36, and Psbd41 models. The filled diamonds in Figure 11 constitute unfolding chevron arms simulated using a "fully unfolded" $q_u = 2$ definition of unfolding. Included for comparison are the unfolding arms indicated by the filled squares, which use alternative definitions (q_u values marked on the figure).

The folding arms exhibit significant rollovers for all four models, including the two-state-like NTL9 model. Chevron rollovers are observed in Gō-like models[78] because mild kinetic trapping, due to steric constraints and transient formation of partially nonnative topologies, is possible in this construct, even though the pairwise interactions in these models are native-centric (see explicit examples in Ref. 39). Here, the maximum folding rates of the four models under strongly folding conditions are all very similar ($-\ln$ (MFPT) ≈ -10 to -11). The coarse-grained Gō-like modeling scheme is likely to be able to capture the general trend, or even the rank ordering, of cooperativity among different proteins.[130] However, as was noted in a prior study of CI2,[78] its essentially pairwise additive interactions tend to underestimate the cooperativity of the proteins it seeks to model, and is therefore insufficient for quantitative prediction of the degree of cooperativity of real proteins. This consideration accounts for the chevron rollover in Figure 11 for the NTL9 model, although the chevron plot for the real 39-residue NTL9 fragment has no rollover.[116] This modeling limitation aside, an important message from Figure 11 is that broad-brush features of chevron plots, including their rollover behavior and maximum folding rates, cannot by themselves be used to distinguish between a weakly cooperative folder like the present NTL9 model (or even certain more cooperative models of CI2[38,78]) and a downhill folder such as the BBL model.

To test the robustness of our conclusions, we determined folding times for the NTL9 model using an alternative folding criterion, $Q \geq 0.817$ (76 native contacts, corresponding to the high-Q peak in the lower panel of Fig. 6). Naturally, recorded folding times were shorter than with the criterion $Q = 1$, but the general features of the chevron arm were the same. It is noteworthy that one subtle feature distinguishing the NTL9 model from the other three, which is probably a consequence of its more cooperative properties, is the form of the dependence of its unfolding chevron arm on the definition of unfolding. For the Psbd41, HP36, and BBL models, the unfolding arms resulting from different choices of q_u are all well separated, and no "natural" intersection point between the folding and unfolding arms reveals itself. On the other hand, while the unfolding arm of the NTL9 model shifts substantially between $q_u = 2$ and $q_u = 12$, and between $q_u = 24$ and $q_u = 47$, it shifts little between $q_u = 12$ and $q_u = 24$. Moreover, the intersection point between the folding and unfolding arms for $q_u = 12$ and 24 is very close to the heat capacity peak. This behavior results from the peak in the conformational distribution at $Q \approx 0.2$ (which corresponds to a denatured free-energy minimum[78]). The $q_u = 12$ and 24 conditions are equivalent to setting the folded/unfolded boundary at $Q \approx 0.13 - 0.25$, the most probable part of the denatured state (see Fig. 6). In contrast, downhill folders have no such denatured peak. As a result, the probability of

reaching a q_u-defined denatured state decreases monotonically with decreasing q_u. It remains to be explored whether experimental probes corresponding to different unfolding definitions can be developed to exploit this subtle distinction between weakly cooperative and downhill folding.

Folding Time Distributions Under Weakly and Strongly Folding Conditions: Downhill Folding can be Consistent With an Essentially Single-Exponential Relaxation

We have also investigated the extent to which kinetic relaxation behavior can be used to discriminate between two-state-like and downhill folding. The MFPT and (median FPT)/ln 2 nearly coincide, over extended $-\varepsilon/k_BT$ ranges, for all four coarse-grained models, which suggests that the relaxation is essentially single-exponential under these regimes.[54,134] Figure 12 covers a larger range of ε/k_BT values, and compares in more detail the folding relaxation behavior of the NTL9 model (two-state-like but only weakly cooperative) and the downhill-folding BBL model. Figure 12 shows the natural logarithms of the probabilities p_{unf} of a chain never having folded, as a function of simulation time. The behavior is broadly similar to that of the simplified atomic model (see Fig. 2). But, perhaps, the salient features are more effectively illustrated by these coarse-grained models because they are simpler, more trajectories can be simulated, and stronger folding conditions are computationally accessible.

Figure 12 shows that, under transition and weakly folding conditions (small ε/k_BT), folding relaxation is essentially single-exponential both for the NTL9 model ($\varepsilon/k_BT = 1.00$ and 1.11) and for the BBL model ($\varepsilon/k_BT = 1.43$): the time dependence of ln p_{unf} is essentially linear. However, under folding conditions much stronger than those considered in the chevron plots of Figure 11, relaxation becomes non-single-exponential ($\varepsilon/k_BT = 1.25$ and 3.33 for the NTL9 and BBL models, respectively). Under these conditions, the ln p_{unf} plots are nonlinear, and folding time distributions show two distinct regions for both models. Some members of the ensemble fold fast, so that the early sections of the ln p_{unf} curves seem to continue the trend of increasing folding speed with larger ε/k_BT. However, as ε/k_BT becomes larger, an increasing proportion of the members of the ensemble take a very long time to fold, and the folding rate slows dramatically and remains slow. This is probably a consequence of increased kinetic trapping under strongly folding conditions.[38,39,78]

The crossover from essentially single-exponential to non-single-exponential relaxation is quantified in the two insets by the variation of χ^2/M with respect to $-\varepsilon/k_BT$. By construction [Eq. (7)], $\chi^2/M \approx 0$ corresponds to single-exponential relaxation, while a large value for χ^2/M implies large deviation from single-exponential behavior. Thus, χ^2/M serves a similar role to the FPT-moment measures introduced by Leite et al. in a 27mer lattice model study[80] (see also related analytical consideration in Ref. 152). The insets in Figure 12 show that, for both models, χ^2/M remains $\sim$0 for a range of $-\varepsilon/k_BT$ values around the transition point (vertical dashed lines), and that χ^2/M deviates greatly from zero only under strongly folding conditions. The similarity between the NTL9 and BBL models suggests that relaxation behavior alone cannot distinguish between two-state-like and downhill folding. Indeed, the lower panel of Figure 12 shows that, under weak folding conditions, downhill folding can exhibit essentially single-exponential relaxation, as is observed for cooperative folding.

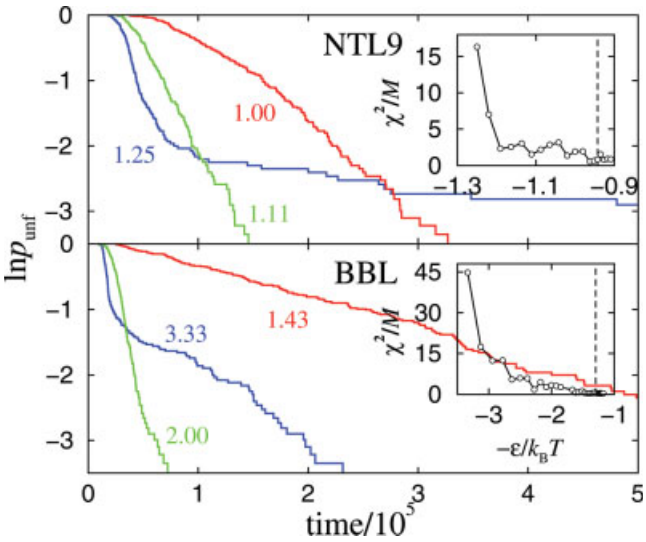

Fig. 12. Folding time distributions for the coarse-grained C_α native-centric models of NTL9 (upper panel) and BBL (lower panel), at various interaction strengths parametrized by ε/k_BT, as in Figure 2. Insets: χ^2/M [see Eq. (7)] as a function of ε/k_BT, quantifying the extent to which the decay of the unfolded state deviates from single-exponential relaxation during the folding process. The vertical dashed line in each inset marks the ε/k_BT value corresponding to the heat capacity peak.

CONCLUDING REMARKS

The main focus of the present investigation has been on the type of possible downhill folding that entails a unimodal conformational distribution and noncooperative behavior even in the thermodynamic folding/unfolding transition region or under weakly folding conditions. To address the question of how such behavior might be experimentally detected, we have studied a simplified atomic model together with four coarse-grained models of protein fragments with various degrees of cooperativity.

The modeling effort does not aim to determine whether a particular real protein is a downhill folder. That is a question to be settled ultimately by experiments. Nonetheless, the general trend exhibited by our coarse-grained analysis suggests that real BBL may be less cooperative than the other three protein fragments modeled here, especially the conspicuously more cooper-

ative NTL9 fragment. This is suggested, for instance, by individual folding trajectories, by the heat capacity, and by the distribution of the radius of gyration. The degree of cooperativity appears to be strongly influenced by the topology of the native state: the higher the average number of contacts per residue, the higher the degree of cooperativity. If we had to choose, on the basis of these results, which of the four fragments would be most likely to show downhill behavior, BBL would be the best candidate.

Using extensive simulations, we have evaluated the effectiveness of several common experimental criteria for distinguishing two-state folding from noncooperative downhill folding. Our analysis suggests that ensemble-averaged properties such as heat capacity and average radius of gyration can, in some instances, supply useful information about the relative degrees of cooperativity of different proteins. However, they cannot by themselves definitively determine whether a protein is a noncooperative downhill folder. For example, empirical baseline subtractions are necessary for the interpretation of experimental calorimetric data, but the baseline subtraction procedure may inadvertently discard important conformational information.[19,87]

A chevron plot with linear folding and unfolding arms under all experimentally accessible conditions, providing m-values consistent with those obtained from thermodynamic measurements, would be a definitive signature of highly cooperative two-state folding.[14,127] However, a chevron plot with a rollover can be consistent with both weakly cooperative (bimodal conformational distribution at transition midpoint) and noncooperative (unimodal conformational distribution) behavior. Simulations here and elsewhere indicate that both weakly cooperative two-state-like proteins and noncooperative downhill-folding proteins can show chevron rollovers. Thus, chevron rollover cannot be regarded as a signature of downhill folding. Similarly, kinetic relaxation in the transition region and under weakly folding conditions can be essentially single-exponential, and have similar speeds, for both a weakly cooperative two-state-like protein and a downhill-folding protein. It follows that the observation of essentially single-exponential kinetics around the folding/unfolding transition cannot be used to rule out an underlying unimodal distribution. However, experimental probes corresponding to different definitions of the transition might allow the degree of cooperativity to be quantified using the probe-dependence of the chevron arms.

Noncooperative downhill behavior is defined by its underlying conformational distribution. Therefore, questions about the existence of such behavior in natural proteins might be addressed by direct observations of trajectories of the relevant conformational distributions using single-molecule techniques under a wide range of folding conditions. The present simulations suggest that, if available, distributions of single-molecule radius of gyration, or distributions of other measures of single-molecule chain dimension that may be accessible by FRET or other techniques, would allow clearer discrimination between noncooperative and two-state behavior. We are confident that further experimental advances in this direction would lead to definitive answers about downhill folding behavior in proteins.

ACKNOWLEDGMENTS

We thank Osman Bilsel, Neil Ferguson, Feng Gai, Martin Gruebele, Steve Hagen, Victor Muñoz, Jed Pitera, Jose Sanchez-Ruiz, Tobin Sosnick, and Jin Wang for helpful discussions about the possibility of downhill folding, Dan Raleigh for information regarding the folding cooperativity of small protein fragments, Ben Schuler for discussion about single-molecule techniques, and Seb Doniach and Hiroshi Kihara for insights into small angle X-ray scattering. We are grateful to Wei Wang for sharing his team's computational work on downhill folding at Nanjing University before publication. H.S.C. holds a Canada Research Chair in Proteomics, Bioinformatics, and Functional Genomics.

NOTE ADDED IN PROOF

Our 41-residue native-centric model for Psbd41, for which experimental folding data are available,[122] is based on the 43-residue 2PDD structure[121] (which is only two residues longer). The more recently determined 55-residue structure of the same protein (PDB ID: 1W3D) is significantly different from the 2PDD structure in the loop region.[153] If the same 41-residue part of the 1W3D structure were used in our modeling, the number of native contacts would be increased by $\sim$10, resulting in an enhancement in cooperativity, e.g., the peak heat capacity value in Figure 4 for Psbd41 would increase to $\sim$400 (Zhirong Liu and H.S.C., unpublished results). But the ordering of the degrees of cooperativity of our four coarse-grained models remains unchanged. We are grateful to Tim Sharpe and Alan Fersht for bringing the 1W3D structure to our attention.

REFERENCES

1. Chan HS, Dill KA. Polymer principles in protein structure and stability. Annu Rev Biophys Biophys Chem 1991;20:447–490.
2. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins 1995;21:167–195.
3. Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS. Principles of protein folding—a perspective from simple exact models. Protein Sci 1995;4:561–602.
4. Thirumalai D, Woodson SA. Kinetics of folding of proteins and RNA. Acc Chem Res 1996;29:433–439.
5. Dobson CM, Šali A, Karplus M. Protein folding: a perspective from theory and experiment. Angew Chem Int Ed Engl 1998;37:868–893.
6. Mirny L, Shakhnovich E. Protein folding theory: from lattice to all-atom models. Annu Rev Biophys Biomol Struct 2001;30: 361–396.
7. Chan HS, Dill KA. Comparing folding codes for proteins and polymers. Proteins 1996;24:335–344.
8. Govindarajan S, Goldstein RA. Why are some protein structures so common? Proc Natl Acad Sci USA 1996;93:3341–3345.
9. Li H, Helling R, Tang C, Wingreen N. Emergence of preferred structures in a simple model of protein folding. Science 1996; 273:666–669.

10. Chan HS. Folding alphabets. Nat Struct Biol 1999;6:994–996.
11. Irbäck A, Sandelin E. On hydrophobicity correlations in protein chains. Biophys J 2000;79:2252–2258.
12. Cui Y, Wong WH, Bornberg-Bauer E, Chan HS. Recombinatoric exploration of novel folded structures: a heteropolymer-based model of protein evolutionary landscapes. Proc Natl Acad Sci USA 2002;99:809–814.
13. Xia Y, Levitt M. Simulating protein evolution in sequence and structure space. Curr Opin Struct Biol 2004;14:202–207.
14. Chan HS, Shimizu S, Kaya H. Cooperativity principles in protein folding. Methods Enzymol 2004;380:350–379.
15. Jackson SE, Fersht AR. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. Biochemistry 1991;30:10428–10435.
16. Jackson SE. How do small single-domain proteins fold? Fold Des 1998;3:R81–R91.
17. Baker D. A surprising simplicity to protein folding. Nature 2000;405:39–42.
18. Chan HS. Modeling protein density of states: additive hydrophobic effects are insufficient for calorimetric two-state cooperativity. Proteins 2000;40:543–571.
19. Kaya H, Chan HS. Polymer principles of protein calorimetric two-state cooperativity. Proteins 2000;40:637–661. Erratum: 2001;43:523.
20. Cheung MS, García AE, Onuchic JN. 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.
21. Liu Z, Chan HS. Desolvation is a likely origin of robust enthalpic barriers to protein folding. J Mol Biol 2005;349:872–889.
22. Liu Z, Chan HS. Solvation and desolvation effects in protein folding: native flexibility, kinetic cooperativity, and enthalpic barriers under isostability conditions. Phys Biol 2005;2:S75–S85.
23. Matthews CR. Pathways of protein folding. Annu Rev Biochem 1993;62:653–683.
24. Scalley ML, Baker D. Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability. Proc Natl Acad Sci USA 1997;94:10636–10640.
25. Kuhlman B, Luisi DL, Evans PA, Raleigh DP. Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9. J Mol Biol 1998;284:1661–1670.
26. Bilsel O, Matthews CR. Barriers in protein folding reactions. Adv Protein Chem 2000;53:153–207.
27. Scalley-Kim M, Baker D. Characterization of the folding energy landscapes of computer generated proteins suggests high folding free energy barriers and cooperativity may be consequences of natural selection. J Mol Biol 2004;338:573–583.
28. Dobson CM. Protein misfolding, evolution and disease. Trends Biochem Sci 1999;24:329–332.
29. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. Design of a novel globular protein fold with atomic-level accuracy. Science 2003;302:1364–1368.
30. Gruebele M. Downhill protein folding: evolution meets physics. C R Biol 2005;328:701–712.
31. Favrin G, Irbäck A, Samuelsson B, Wallin S. Two-state folding over a weak free-energy barrier. Biophys J 2003;85:1457–1465.
32. Knott M, Chan HS. Exploring the effects of hydrogen bonding and hydrophobic interactions on the foldability and cooperativity of helical proteins using a simplified atomic model. Chem Phys 2004;307:187–199.
33. Yang WY, Gruebele M. Folding λ-repressor at its speed limit. Biophys J 2004;87:596–608.
34. Dill KA, Chan HS. From Levinthal to pathways to funnels. Nat Struct Biol 1997;4:10–19.
35. Wallin S, Chan HS. A critical assessment of the topomer search model of protein folding using a continuum explicit-chain model with extensive conformational sampling. Protein Sci 2005;14:1643–1660.
36. Wallin S, Chan HS. Conformational entropic barriers in topology-dependent protein folding: perspectives from a simple native-centric polymer model. J Phys Condens Matter 2006;18: S307–S328.
37. Kaya H, Chan HS. Simple two-state protein folding kinetics requires near-Levinthal thermodynamic cooperativity. Proteins 2003;52:510–523.
38. Kaya H, Liu Z, Chan HS. Chevron behavior and isostable enthalpic barriers in protein folding: successes and limitations of simple Gō-like modeling. Biophys J 2005;89:520–535.
39. Kaya H, Chan HS. Origins of chevron rollovers in non-two-state protein folding kinetics. Phys Rev Lett 2003;90:258104.
40. Naganathan AN, Sanchez-Ruiz JM, Muñoz V. Direct measurement of barrier heights in protein folding. J Am Chem Soc 2005;127:17970–17971.
41. Wang J, Lee C, Stell G. The cooperative nature of hydrophobic forces and protein folding kinetics. Chem Phys 2005;316:53–60.
42. Schonbrun J, Dill KA. Fast protein folding kinetics. Proc Natl Acad Sci USA 2003;100:12678–12682.
43. Levinthal C. Are there pathways for protein folding? J Chim Phys 1968;65:44–45.
44. Levinthal C. How to fold graciously. In: Debrunner P, Tsibris JCM, Münck E, editors. Mössbauer spectroscopy in biological systems. Urbana, Illinois: University of Illinois Press; 1969. pp 22–24. Proceedings of a meeting held at Allerton House, Monticello, Illinois, 1969.
45. Maxwell KL, Wildes D, Zarrine-Afsar A, de los Rios MA, Brown AG, Friel CT, Hedberg L, Horng JC, Bona D, Miller EJ, Vallee-Belisle A, Main ERG, Bemporad F, Qiu LL, Teilum K, Vu ND, Edwards AM, Ruczinski I, Poulsen FM, Kragelund BB, Michnick SW, Chiti F, Bai Y, Hagen SJ, Serrano L, Oliveberg M, Raleigh DP, Wittung-Stafshede P, Radford SE, Jackson SE, Sosnick TR, Marqusee S, Davidson AR, Plaxco KW. Protein folding: defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins. Protein Sci 2005;14:602–616.
46. Gillespie B, Plaxco KW. Using protein folding rates to test protein folding theories. Annu Rev Biochem 2004;73:837–859.
47. Shiraki K, Nishikawa K, Goto Y. Trifluoroethanol-induced stabilization of the α-helical structure of β-lactoglobulin: implication for non-hierarchical protein folding. J Mol Biol 1995;245:180–194.
48. Nishimura C, Wright PE, Dyson HJ. Role of the B helix in early folding events in apomyoglobin: evidence from site-directed mutagenesis for native-like long range interactions. J Mol Biol 2003; 334:293–307.
49. Clementi C, Plotkin SS. The effects of nonnative interactions on protein folding rates: theory and simulation. Protein Sci 2004;13:1750–1766.
50. Di Nardo AA, Korzhnev DM, Stogios PJ, Zarrine-Afsar A, Kay LE, Davidson AR. Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. Proc Natl Acad Sci USA 2004;101:7954–7959.
51. Kim PS, Baldwin RL. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. Annu Rev Biochem 1982;51:459–489.
52. Kim PS, Baldwin RL. Intermediates in the folding reactions of small proteins. Annu Rev Biochem 1990;59:631–660.
53. Jennings PA, Finn BE, Jones BE, Matthews CR. A reexamination of the folding mechanism of dihydrofolate reductase from *Escherichia coli*: verification and refinement of a four-channel model. Biochemistry 1993;32:3783–3789.
54. Kaya H, Chan HS. Towards a consistent modeling of protein thermodynamic and kinetic cooperativity: how applicable is the transition state picture to folding and unfolding? J Mol Biol 2002;315:899–909.
55. Chan HS, Dill KA. Protein folding in the landscape perspective: chevron plots and non-Arrhenius kinetics. Proteins 1998;30:2–33.
56. Choy W-Y, Forman-Kay JD. Calculation of ensembles of structures representing the unfolded state of an SH3 domain. J Mol Biol 2001;308:1011–1032.
57. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. Intrinsically disordered protein. J Mol Graph Model 2001;19:26–59.
58. Tompa P. Intrinsically unstructured proteins. Trends Biochem Sci 2002;27:527–533.
59. Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R. Extended disordered proteins: targeting function with less scaffold. Trends Biochem Sci 2003;28:81–85.

60. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 2005;6:197–208.
61. Hong W, Jiao W, Hu J, Zhang J, Liu C, Fu X, Shen D, Xia B, Chang Z. Periplasmic protein HdeA exhibits chaperone-like activity exclusively within stomach pH range by transforming into disordered conformation. J Biol Chem 2005;280:27029–27034.
62. James LC, Tawfik DS. Conformational diversity and protein evolution—a 60-year-old hypothesis revisited. Trends Biochem Sci 2003;28:361–368.
63. Sabelko J, Ervin J, Gruebele M. Observation of strange kinetics in protein folding. Proc Natl Acad Sci USA 1999;96:6031–6036.
64. Eaton WA. Searching for "downhill scenarios" in protein folding. Proc Natl Acad Sci USA 1999;96:5897–5899.
65. Sinha KK, Udgaonkar JB. Dependence of the size of the initially collapsed form during the refolding of barstar on denaturant concentration: evidence for a continuous transition. J Mol Biol 2005;353:704–718.
66. Eaton WA, Muñoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, Hofrichter J. Fast kinetics and mechanisms in protein folding. Annu Rev Biophys Biomol Struct 2000;29:327–359.
67. Zhu Y, Alonso DOV, Maki K, Huang C-Y, Lahr SJ, Daggett V, Roder H, DeGrado WF, Gai F. Ultrafast folding of α_3D: a *de novo* designed three-helix bundle protein. Proc Natl Acad Sci USA 2003;100:15486–15491.
68. Zhu Y, Fu X, Wang T, Tamura A, Takada S, Savan JG, Gai F. Guiding the search for a protein's maximum rate of folding. Chem Phys 2004;307:99–109.
69. Yang WY, Gruebele M. Folding at the speed limit. Nature 2003; 423:193–197.
70. Kubelka J, Hofrichter J, Eaton WA. The protein folding 'speed limit'. Curr Opin Struct Biol 2004;14:76–88.
71. Kramers HA. Brownian motion in a field of force and the diffusion model of chemical reactions. Physica 1940;7:284–304.
72. Thompson PA, Muñoz V, Jas GS, Henry ER, Eaton WA, Hofrichter J. The helix-coil kinetics of a heteropeptide. J Phys Chem B 2000;104:378–389.
73. Xu Y, Oyola R, Gai F. Infrared study of the stability and folding kinetics of a 15-residue β-hairpin. J Am Chem Soc 2003; 125:15388–15394.
74. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J. Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: testing a 1-dimensional diffusion model for peptide dynamics. J Phys Chem B 2002;106:11628–11640.
75. Krieger F, Fierz B, Bieri O, Drewello M, Kiefhaber T. Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. J Mol Biol 2003;332:265–274.
76. Krieger F, Fierz B, Axthelm F, Joder K, Meyer D, Kiefhaber T. Intrachain diffusion in a protein loop fragment from carp parvalbumin. Chem Phys 2004;307:209–215.
77. Hagen SJ, Eaton WA. Two-state expansion and collapse of a polypeptide. J Mol Biol 2000;301:1019–1027.
78. Kaya H, Chan HS. 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 Corrigendum: 2004;337:1069–1070.
79. Wang J. The complex kinetics of protein folding in wide temperature ranges. Biophys J 2004;87:2164–2171.
80. Leite VBP, Onuchic JN, Stell G, Wang J. Probing the kinetics of single molecule protein folding. Biophys J 2004;87:3633–3641.
81. Hagen SJ. Exponential decay kinetics in "downhill" protein folding. Proteins 2003;50:1–4.
82. Meisner WK, Sosnick TR. Barrier-limited, microsecond folding of a stable protein measured with hydrogen exchange: implications for downhill folding. Proc Natl Acad Sci USA 2004;101: 15639–15644.
83. Qiu L, Hagen SJ. A limiting speed for protein folding at low solvent viscosity. J Am Chem Soc 2004;126:3398–3399.
84. Hagen SJ, Qiu L, Pabit SA. Diffusional limits to the speed of protein folding: fact or friction? J Phys Condens Matter 2005; 17:S1503–S1514.
85. Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Muñoz V. Experimental identification of downhill protein folding. Science 2002;298:2191–2195.
86. Oliva FY, Muñoz V. A simple thermodynamic test to discriminate between two-state and downhill folding. J Am Chem Soc 2004;126:8596–8597.
87. Muñoz V, Sanchez-Ruiz JM. Exploring protein-folding ensembles: a variable-barrier model for the analysis of equilibrium unfolding experiments. Proc Natl Acad Sci USA 2004;101: 17646–17651.
88. Ma H, Gruebele M. Kinetics are probe-dependent during downhill folding of an engineered λ_{6-85} protein. Proc Natl Acad Sci USA 2005;102:2283–2287.
89. Englander SW, Mayne L, Bai Y, Sosnick TR. Hydrogen exchange: the modern legacy of Linderstrϕm-Lang. Protein Sci 1997;6:1101–1109.
90. Mello CC, Barrick D. An experimentally determined protein folding energy landscape. Proc Natl Acad Sci USA 2004;101: 14102–14107.
91. Ferguson N, Schartau PJ, Sharpe TD, Sato S, Fersht AR. One-state downhill *versus* conventional protein folding. J Mol Biol 2004;344:295–301.
92. Naganathan AN, Perez-Jimenez R, Sanchez-Ruiz JM, Muñoz V. Robustness of downhill folding: guidelines for the analysis of equilibrium folding experiments on small proteins. Biochemistry 2005;44:7435–7449.
93. Ferguson N, Sharpe TD, Schartau PJ, Sato S, Allen MD, Johnson CM, Rutherford TJ, Fersht AR. Ultrafast barrier-limited folding in the peripheral subunit-binding domain family. J Mol Biol 2005;353:427–446.
94. Thirumalai D, Klimov DK. Deciphering the timescales and mechanisms of protein folding using minimal off-lattice models. Curr Opin Struct Biol 1999;9:197–207.
95. Chan HS, Kaya H, Shimizu S. Computational methods for protein folding: scaling a hierarchy of complexities. In: Jiang T, Xu Y, Zhang MQ, editors. Current topics in computational biology. Cambridge, MA: MIT Press; 2002. pp 403–447.
96. Karanicolas J, Brooks CL. The importance of explicit chain representation in protein folding models: an examination of Ising-like models. Proteins 2003;53:740–747.
97. Head-Gordon T, Brown S. Minimalist models for protein folding and design. Curr Opin Struct Biol 2003;13:160–167.
98. Lu H, Xun L, Xie X. Single-molecule enzymatic dynamics. Science 1998;282:1877–1882.
99. Zhuang X, Bartley LE, Babcock HP, Russell R, Ha T, Herschlag D, Chu S. A single-molecule study of RNA catalysis and folding. Science 2000;288:2048–2051.
100. Yang H, Xie X. Probing single-molecule dynamics photon by photon. J Chem Phys 2002;117:10965–10979.
101. Schuler B, Lipman EA, Eaton WA. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. Nature 2002;419:743–747.
102. Lipman EA, Schuler B, Bakajin O, Eaton WA. Single-molecule measurement of protein folding kinetics. Science 2003;301: 1233–1235.
103. Rhoades E, Gussakovsky E, Haran G. Watching proteins fold one molecule at a time. Proc Natl Acad Sci USA 2003;100: 3197–3202.
104. Rhoades E, Cohen M, Schuler B, Haran G. Two-state folding observed in individual protein molecules. J Am Chem Soc 2004; 126:14686–14687.
105. Xie Z, Srividya N, Sosnick TR, Pan T, Scherer NF. Single-molecule studies highlight conformational heterogeneity in the early folding steps of a large ribozyme. Proc Natl Acad Sci USA 2004;101:534–539.
106. Cecconi C, Shank EA, Bustamante C, Marqusee S. Direct observation of the three-state folding of a single protein molecule. Science 2005;309:2057–2060.
107. Schuler B. Single-molecule fluorescence spectroscopy of protein folding. Chemphyschem 2005;6:1206–1220.
108. Irbäck A, Sjunnesson F, Wallin S. Three-helix-bundle protein in a Ramachandran model. Proc Natl Acad Sci USA 2000;97: 13614–13618.
109. Zhang M, Chen C, He Y, Xiao Y. Improvement on a simplified model for protein folding simulation. Phys Rev E 2005;72:051919.
110. Takada S, Luthey-Schulten Z, Wolynes PG. Folding dynamics with nonadditive forces: a simulation study of a designed helical protein and a random heteropolymer. J Chem Phys 1999;110: 11616–11629.

111. Irbäck A, Sjunnesson F, Wallin S. Hydrogen bonds, hydrophobicity forces and the character of the collapse transition. J Biol Phys 2001;27:169–179.
112. Micheletti C, Banavar JR, Maritan A, Seno F. Protein structures and optimal folding from a geometrical variational principle. Phys Rev Lett 1999;82:3372–3375.
113. Clementi C, Nymeyer H, Onuchic JN. 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.
114. Koga N, Takada S. Roles of native topology and chain-length scaling in protein folding: a simulation study with a Gō-like model. J Mol Biol 2001;313:171–180.
115. Robien MA, Clore GM, Omichinski JG, Perham RN, Appella E, Sakaguchi K, Gronenborn AM. 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 1992;31:3463–3471.
116. Horng J-C, Moroz V, Raleigh DP. Rapid cooperative two-state folding of a miniature α-β protein and design of a thermostable variant. J Mol Biol 2003;326:1261–1270.
117. Luisi DL, Kuhlman B, Sideras K, Evans PA, Raleigh DP. Effects of varying the local propensity to form secondary structure on the stability and folding kinetics of a rapidly folding mixed α/β protein: characterization of a truncation mutant of the N-terminal domain of the ribosomal protein L9. J Mol Biol 1999;289:167–174.
118. McKnight CJ, Doering DS, Matsudaira PT, Kim PS. A thermostable 35-residue subdomain within villin headpiece. J Mol Biol 1996;260:126–134.
119. Wang M, Tang Y, Sato SS, Vugmeyster L, McKnight CJ, Raleigh DP. Dynamic NMR line-shape analysis demonstrates that the villin headpiece subdomain folds on the microsecond time scale. J Am Chem Soc 2003;125:6032–6033.
120. Brewer SH, Vu DM, Tang Y, Li Y, Franzen S, Raleigh DP, Dyer RB. Effect of modulating unfolded state structure on the folding kinetics of the villin headpiece subdomain. Proc Natl Acad Sci USA 2005;102:16662–16667.
121. Kalia YN, Brocklehurst SM, Hipps DS, Appella E, Sakaguchi K, Perham RN. The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. J Mol Biol 1993;230:323–341.
122. Spector S, Young P, Raleigh DP. Nativelike structure and stability in a truncation mutant of a protein minidomain: the peripheral subunit-binding domain. Biochemistry 1999;38:4128–4136.
123. Knott M, Kaya H, Chan HS. Energetics of protein thermodynamic cooperativity: contributions of local and nonlocal interactions. Polymer 2004;45:623–632.
124. Chavez LL, Onuchic JN, Clementi C. Quantifying the roughness on the free energy landscape: entropic bottlenecks and protein folding rates. J Am Chem Soc 2004;126:8426–8432.
125. Sheinerman FB, Brooks CL. Calculations on folding of segment B1 of streptococcal protein G. J Mol Biol 1998;278:439–456.
126. Jewett AI, Pande VS, Plaxco KW. Cooperativity, smooth energy landscapes and the origins of topology-dependent protein folding rates. J Mol Biol 2003;326:247–253.
127. Kaya H, Chan HS. Contact order dependent protein folding rates: kinetic consequences of a cooperative interplay between favorable nonlocal interactions and local conformations preferences. Proteins 2003;52:524–533.
128. Ejtehadi MR, Avall SP, Plotkin SS. Three-body interactions improve the prediction of rate and mechanism in protein folding models. Proc Natl Acad Sci USA 2004;101:15088–15093.
129. Zuo G-H, Zhang J, Wang J, Wang W. Folding behaviour for proteins BBL and E3BD with Gō-like models. Chin Phys Lett 2005;22:1809–1812.
130. Zuo G, Wang J, Wang W. Folding with downhill behavior and low cooperativity of proteins. Proteins 2006;63:165–173.
131. Veitshans T, Klimov D, Thirumalai D. Protein folding kinetics: timescales, pathways and energy landscapes in terms of sequence-dependent properties. Fold Des 1997;2:1–22.
132. Plaxco KW, Millett IS, Segel DJ, Doniach S, Baker D. Chain collapse can occur concomitantly with the rate-limiting step in protein folding. Nat Struct Biol 1999;6:554–556.
133. Matthews CR. Effect of point mutations on the folding of globular proteins. Methods Enzymol 1987;154:498–511.
134. Gutin A, Sali A, Abkevich V, Karplus M, Shakhnovich EI. Temperature dependence of the folding rate in a simple protein model: search for a "glass" transition. J Chem Phys 1998;108: 6466–6483.
135. Abkevich VI, Gutin AM, Shakhnovich EI. Free energy landscape for protein folding kinetics: intermediates, traps, and multiple pathways in theory and lattice model simulations. J Chem Phys 1994;101:6052–6062.
136. Zhou Y, Zhang C, Stell G, Wang J. Temperature dependence of the distribution of the first passage time: results from discontinuous molecular dynamics simulations of an all-atom model of the second β-hairpin fragment of protein G. J Am Chem Soc 2003;125:6300–6305.
137. Kaya H, Chan HS. Explicit-chain model of native-state hydrogen exchange: implications for event ordering and cooperativity in protein folding. Proteins 2005;58:31–44.
138. Marianayagam NJ, Fawzi NL, Head-Gordon T. Protein folding by distributed computing and the denatured state ensemble. Proc Natl Acad Sci USA 2005;102:16684–16689.
139. Snow CD, Nguyen N, Pande VS, Gruebele M. Absolute comparison of simulated and experimental protein-folding dynamics. Nature 2002;420:102–106.
140. Auton M, Bolen DW. Predicting the energetics of osmolyte-induced protein folding/unfolding. Proc Natl Acad Sci USA 2005;102:15065–15068.
141. Zhou Y, Hall CK, Karplus M. The calorimetric criterion for a two-state process revisited. Protein Sci 1999;8:1064–1074.
142. Sanchez IE, Kiefhaber T. Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. J Mol Biol 2003;325:367–376.
143. Sosnick TR, Trewhella J. Denatured states of ribonuclease A have compact dimensions and residual secondary structure. Biochemistry 1992;31:8329–8335.
144. Choy W-Y, Mulder FAA, Crowhurst KA, Muhandiram DR, Millett IS, Doniach S, Forman-Kay JD, Kay LE. Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. J Mol Biol 2002;316:101–112.
145. Millet IS, Doniach S, Plaxco KW. Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. Adv Protein Chem 2002;62:241–262.
146. Yamada Y, Yajima T, Fujiwara K, Arai M, Ito K, Shimizu A, Kihara H, Kuwajima K, Amemiya Y, Ikeguchi M. Helical and expanded conformation of equine β-lactoglobulin in the cold-denatured state. J Mol Biol 2005;350:338–348.
147. Chan HS, Bromberg S, Dill KA. Models of cooperativity in protein folding. Philos Trans R Soc Lond B 1995;348:61–70.
148. Flory PJ, Fisk S. Effect of volume exclusion on the dimensions of polymer chains. J Chem Phys 1966;44:2243–2248.
149. Magg C, Schmid FX. Rapid collapse precedes the fast two-state folding of the cold shock protein. J Mol Biol 2004;335:1309–1323.
150. Sridevi K, Lakshmikanth GS, Krishnamoorthy G, Udgaonkar JB. Increasing stability reduces conformational heterogeneity in a protein folding intermediate ensemble. J Mol Biol 2004; 337: 699–711.
151. Ratner V, Amir D, Kahana E, Haas E. Fast collapse but slow formation of secondary structure elements in the refolding transition of *E. coli* adenylate kinase. J Mol Biol 2005;352: 683–699.
152. Lee CL, Stell G, Wang J. First-passage time distribution and non-Markovian diffusion dynamics of protein folding. J Chem Phys 2003;118:959–968.
153. Allen MD, Broadhurst RW, Solomon RG, Perham RN. Interaction of the E2 and E3 components of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*: use of a truncated protein domain in NMR spectroscopy. FEBS J 2005;272:259–268.