JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Effects of pH on Proteins: Predictions for Ensemble and Single-Molecule Pulling Experiments

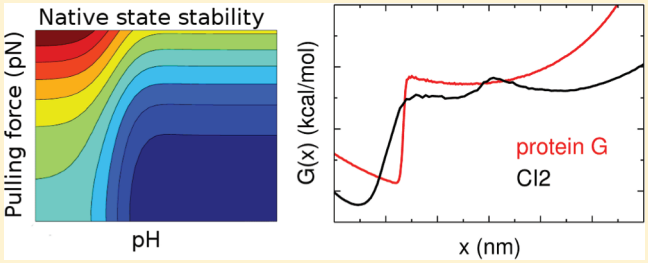
Edward P. O'Brien,[†,‡,§] Bernard R. Brooks,[‡] and D. Thirumalai*[,†]

[†]Biophysics Program, Institute for Physical Science and Technology and Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, United States

[‡]Laboratory of Computational Biology, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

Supporting Information

ABSTRACT: Protein conformations change among distinct thermodynamic states as solution conditions (temperature, denaturants, pH) are altered or when they are subjected to mechanical forces. A quantitative description of the changes in the relative stabilities of the various thermodynamic states is needed to interpret and predict experimental outcomes. We provide a framework based on the Molecular Transfer Model (MTM) to account for pH effects on the properties of globular proteins. The MTM utilizes the partition function of a protein calculated from molecular simulations at one set of solution conditions to predict protein properties at another set of solution conditions. To take pH effects into account, we utilized experimentally measured pK_a values in the native and unfolded states to calculate the free energy of transferring a protein from a reference pH to the pH of interest. We validate our approach by demonstrating that the native-state stability as a function of pH is accurately predicted for chymotrypsin inhibitor 2 (CI2) and protein G. We use the MTM to predict the response of CI2 and protein G subjected to a constant force (f) and varying pH. The phase diagrams of CI2 and protein G as a function of f and pH are dramatically different and reflect the underlying pH-dependent stability changes in the absence of force. The calculated equilibrium free energy profiles as functions of the end-to-end distance of the two proteins show that, at various pH values, CI2 unfolds via an intermediate when subjected to f. The locations of the two transition states move toward the more unstable state as f is changed, which is in accord with the Hammond–Leffler postulate. In sharp contrast, force-induced unfolding of protein G occurs in a single step. Remarkably, the location of the transition state with respect to the folded state is independent of f, which suggests that protein G is mechanically brittle. The MTM provides a natural framework for predicting the outcomes of ensemble and single-molecule experiments for a wide range of solution conditions.



INTRODUCTION

Recent experimental advances, especially in single-molecule techniques,[1,2] have made it possible to obtain detailed mechanistic insights into the folding of proteins over a wide range of external conditions. For example, singe-molecule fluorescence resonance energy transfer (smFRET) experiments have been used to probe the characteristics of the ensemble of unfolded states under folding conditions, providing a glimpse of the nature of the collapse transition in proteins.[3] Developments in single-molecule force spectroscopy, which probe the response of proteins subjected to mechanical force (f), have been used to map the entire folding landscape of proteins described by the end-to-end distance of the molecule.[4−6] These experiments have been remarkably successful in measuring the roughness of the energy landscape,[7] estimating barriers to folding,[8] and characterizing minimum-energy compact structures[9,10] that are sampled during the folding process. In the majority of the smFRET experiments folding or unfolding is initiated using chemical denaturants,[11,12] whereas in pulling experiments external force is applied to select points on the protein to control folding. More recently, the global responses of proteins to f in the presence of osmolytes and denaturants and pH changes have also been reported.[10,13] The wealth of data emerging from these studies demand computational models for which exhaustive simulations at conditions that mimic those used in experiments can be performed.[14]

Besides denaturants,[15−17] protein folding or unfolding can also be initiated by altering the pH.[18] Although a number of experimental studies have reported the pH dependence of protein stability,[19−22] there are very few theoretical approaches which have addressed the thermodynamic aspects of pH-dependent folding. In principle, all-atom molecular dynamics simulations can be used to model pH-dependent effects on protein folding. Such an approach has found success in predicting and interpreting pK_a values of titratable groups within the native and unfolded ensembles.[23−25] However, calculating other pH-dependent protein properties using all-

Received: July 14, 2011
Published: December 8, 2011

   dx.doi.org/10.1021/ja206557y | J. Am. Chem.Soc. 2012, 134, 979–987

atom models is difficult because of inaccuracies in force fields[26] and the inability to adequately sample both the folded and denatured conformational space.

One of the most widely used thermodynamic models for pH effects on proteins was created by Tanford and co-workers, who showed that from knowledge of the pK_a values of titratable groups it is possible to predict the change in protein native-state stability as a function of pH.[17] This thermodynamic model cannot be used to calculate the distribution of pH-dependent properties of proteins because it neglects the ensemble nature of folding. It is limited to making estimates of changes in the stability and the difference in the number of bound protons between the native and unfolded states.

The limitations of the thermodynamic model can be overcome using the Molecular Transfer Model (MTM), which we originally introduced to account for osmolyte effects on proteins.[27] Although the MTM can be used in conjunction with all-atom models for proteins, we used a coarse-grained representation of proteins[27,28] so that the partition function of the system could be precisely computed at a given solution condition. Knowledge of the partition function can be used to compute any thermodynamic property at another solution condition by appropriate reweighting. The MTM utilizes the Tanford model[17,29] to estimate the free energy cost of transferring each microstate (i.e., protein conformation) from one solution condition to another. Thus, the MTM is a post-simulation processing technique, which allows for the rapid prediction of the thermodynamic properties of proteins under a wide range of external conditions by performing simulations at one solution condition.[27,30,31]

Here, we further develop the MTM to model pH effects on protein properties. We validate our approach by demonstrating that the methodology accurately predicts experimentally measured changes in the native-state stability as a function of pH. We further establish the efficacy of the MTM by calculating the pH-dependent response of proteins subjected to an external mechanical force, f. Simulations at constant force and varying pH for chymotrypsin inhibitor 2 (CI2)[32,33] and protein G[34,35] show dramatically different behavior. Both the diagram of states as a function of f and pH and the free energy profiles depend on the protein. Our results are consistent with experimental data from a constant pulling speed force experiment[36] and offer a number of testable predictions.

METHODS

Molecular Transfer Model for pH Effects on Proteins under Tension. The MTM[27] utilizes protein conformations from the C_α side chain model (C_α-SCM) coarse-grained simulations (see below), experimentally measured or theoretically computed amino acid transfer free energies, and weighted histogram equations[37] (WHAM) to predict how changes in the external conditions alter the thermodynamic properties of a protein. The MTM equation for predicting the average of a quantity A at a given pH, temperature, and f value is

$$\langle A(\mathrm{pH}_2, T, f)\rangle = Z(\mathrm{pH}_2, T, f)^{-1} \sum_{k=1}^{R} \sum_{t=1}^{n_k} \frac{A_{k,t} e^{-\beta E_P(k,t,\mathrm{pH}_2 f)}}{\sum_{m=1}^{R} n_m e^{F_m - \beta_m E_m(k,t) - f_m x(k,t)}} \quad (1)$$

where $Z(\mathrm{pH}_2,T,f)$ is the partition function, R is the number of independent simulated trajectories, n_k is the number of protein conformations in the k^{th} simulation, $A_{k,t}$ is the value of property A for the tth conformation, and $\beta = 1/k_B T$, where k_B is the Boltzmann constant and T is the temperature. The potential energy E_P of the tth conformation in the k^{th} simulation at a pH value denoted pH_2 and under external force f is $E_P(k,t,\mathrm{pH}_2 f) = E_P(k,t,\mathrm{pH}_1) + \Delta G_{tr}(k,t,\mathrm{pH}_2) - fx(k,t)$, where $E_P(k,t,\mathrm{pH}_1)$ is the potential energy of the system at pH = pH_1, i.e., the pH conditions at which the simulations are carried out in this study (see below), $\Delta G_{tr}(k,t,\mathrm{pH}_2)$ is the free energy of transferring the tth conformation in the k^{th} simulation from a solution at pH_1 to a solution at pH_2, f is the external pulling force, and x is the end-to-end distance vector of the protein projected onto the direction of the applied force. In the denominator of Eq 1, n_m and F_m are, respectively, the number of conformations and the free energy of the mth simulation. The values of F_m are obtained self-consistently at the simulated solution conditions as described in ref 37.

To estimate $\Delta G_{tr}(k,t,\mathrm{pH}_2)$, we use a model developed by Tanford and co-workers[17] in which the free energy of transferring a titratable group p in conformation l of the protein from pH_1 to pH_2 is

$$\delta g_{p,l} = -k_B T \times \ln\left[\frac{10^{\mathrm{pH}_2} + \Theta_N(l) \times 10^{\mathrm{p}K_{N,p}} + \Theta_D(l) \times 10^{\mathrm{p}K_{D,p}}}{10^{\mathrm{pH}_1} + \Theta_N(l) \times 10^{\mathrm{p}K_{N,p}} + \Theta_D(l) \times 10^{\mathrm{p}K_{D,p}}}\right] \quad (2)$$

where $\Theta_N(l)$ and $\Theta_D(l)$ are Heaviside step functions that identify a conformation l as being either native or denatured. $\Theta_N(l)$ ($\Theta_D(l)$) equals 1 if conformation l is native (denatured) and 0 otherwise. $\mathrm{p}K_{N,p}$ and $\mathrm{p}K_{D,p}$ are the pK_a values for group p in the native and denatured states, respectively. We use $\mathrm{p}K_{N,p}$ values that have been determined experimentally[34,38] (see Table I, Supporting Information). Details on classifying conformations as native and denatured are given below. Finally, $\Delta G_{tr}(K,t,\mathrm{pH}_2) = \sum_{p=1}^{N_p} \delta g_{p,l}$ which is the sum of the $\delta g_{p,l}$ values that are calculated using Eq 2.

Coarse-Grained Models for Proteins. We model the 65-residue-long protein CI2 and 56-residue protein G using the C_α-SCM[27,28] in which each amino acid is represented as two interaction sites, one of which is located at the α-carbon position of the backbone. For all amino acids except glycine, the other interaction site is located at the center-of-mass of the side chain. We use a Go model version[39] of the C_α-SCM. Thus, side chains that are in contact or backbone groups that form hydrogen bonds in the crystal structure have attractive nonbonded Lennard-Jones interactions, while all other nonbonded interactions are repulsive. Sequence-dependent effects are modeled using nonbonded interaction parameters based on the Miyazawa–Jernigan statistical potential.[40] The excluded volume of an amino acid side chain is proportional to its experimentally measured partial molar volume in solution.

The potential energy E_P of a C_α-SCM conformation is $E_P = E_A + E_{HB} + E_{NB}^N + E_{NB}^{NN}$, which is the sum of potential energy terms corresponding to angles (E_A), hydrogen bonds (E_{HB}), and native (E_{NB}^N) and non-native (E_{NB}^{NN}) nonbonded interactions. We use the Shake algorithm[41] to hold the bond lengths fixed in the simulations; hence, there is no energy term corresponding to this constraint. The functional forms of the terms in the C_α-SCM force field are

$$E_A = \sum_{i=1}^{N} K_A(\theta_i - \theta_{i,0})^2 + \sum_{i=1}^{N_D} \sum_{j=1}^{3} K_{D_j} (1 + \cos(n_j \varphi_i - \delta_{ij})) + \sum_{i=1}^{N_{Ch}} K_{Ch}(\Psi_i - \Psi_{i,0})^2 \quad (3)$$

$$E_{HB} + E_{NB}^N = \sum_{i=1}^{N_{HB}} \varepsilon_{HB}\left[\left(\frac{r_{HB,i}^0}{r_i}\right)^{12} - 2\left(\frac{r_{HB,i}^0}{r_i}\right)^{6}\right] + \sum_{i=1}^{N_N} \varepsilon_i^N\left[\left(\frac{r_{\min,i}}{r_i}\right)^{12} - 2\left(\frac{r_{\min,i}}{r_i}\right)^{12}\right] \quad (4)$$

dx.doi.org/10.1021/ja206557y | J. Am. Chem.Soc. 2012, 134, 979–987

$$E_{NB}^{NN} = \sum_{i=1}^{N_{NN}} \varepsilon_i^{NN}\left[\left(\frac{r_{\min,i}}{r_i}\right)^{12}\right] \quad (5)$$

On the right-hand side (RHS) of Eq 3, and from left to right, the summations correspond, respectively, to bond angle, dihedral angle, and improper dihedral angle energy terms. The improper dihedral term is used to model chirality about the C_α interaction site. On the RHS of Eq 4, we model hydrogen bonds found in the crystal structure as a Lennard-Jones potential (first term) with a well depth set to ε_{HB} and the positions of the minima $r^0_{HB,i}$ set by the interaction site distance in the crystal structure. The second summation term in Eq 4 accounts for native interactions between sites and is treated using the Lennard-Jones potential whose well depth is set using the Miyazawa–Jernigan statistical potential and location of the minima $r_{\min,i}$ corresponds to the distance in the crystal structure. Non-native interactions (Eq 5) between sites are treated as short-ranged and repulsive, with $r_{\min,i}$ being proportional to the experimentally measured partial molar volume of the amino acid type. The force field parameters used for CI2 and protein G are given in Table II in the Supporting Information, and additional details on the model can be found in our previous study.[27] We use the crystal structures with PDB codes 2CI2[42] and 1GB1[43] for CI2 and protein G, respectively.

Simulation Details. We use Hamiltonian replica exchange (HREX)[44,45] in the canonical ensemble to obtain equilibrium simulations of CI2 and protein G at constant force (f) applied in the positive x direction to the C-terminal C_α interaction site of each protein. In the simulations the N-terminal C_α interaction site was fixed at the origin. In the HREX simulation, independent trajectories (replicas) are simulated at different temperatures and at different f values using Langevin dynamics[46] with a damping coefficient of 0.8 ps^{-1} and an integration time step of 6 fs. The nonbonded interactions were truncated at 20 Å with a switch function applied starting at 18 Å. We used CHARMM (version c33b2) to simulate the time evolution of the replicas.[47] Every 5000–7000 integration time steps, the coordinates of the proteins were saved for each replica and then exchanged, either between neighboring temperatures or between neighboring external tensions (i.e., Hamiltonians) according to exchange criteria that preserve detailed balance.[44] It total, 90 000 exchanges, alternating between temperature and force values, were attempted. The first 10 000 exchanges were discarded to allow for equilibration.

For CI2, five temperature windows (300, 317, 330, 345, and 380 K) and eight f values (0.00, 0.35, 3.47, 8.68, 9.03, 9.38, 9.73, 10.42, and 13.89 pN) were used for a total of forty replicas. For protein G, four temperature windows (310, 320, 330, and 370 K) and ten f values (0.00, 0.35, 1.60, 2.85, 4.10, 5.35, 6.60, 7.85, 9.10, 10.42, and 13.89 pN) were used for a total of forty replicas. Swap acceptance ratios of between 10% and 40% were achieved in the HREX runs. The equilibrium properties of the proteins at temperatures and constant pulling forces other than those explicitly simulated were calculated using Eqs 1 and 6.

Analysis. A conformation is native if the root-mean-squared distance (RMSD) of the C_α interaction sites is within 5 Å, for protein G, or 11 Å, for CI2, of the corresponding C_α atoms in the crystal structure; otherwise it is classified as denatured. These RMSD thresholds were determined as the upper limit on the integral of the RMSD probability densities at the melting temperature (i.e., the maximum in the heat capacity trace) that yielded a value of 0.5. This method is illustrated in Figure S1 (Supporting Information). This means we assumed that at the melting temperature CI2 is a two-state system. Structurally, CI2's larger threshold arises from the disordered random coil regions in the native-state ensemble (see Figure 1A).

A key step in applying the MTM is in the choice of the reference pH ("pH_1" in the equations above) in the postsimulation analysis. To choose the reference pH, we first calculated the native stability of these two proteins at 300 K. For CI2 we identified a pH value for which the calculated and experimentally measured stabilities are similar.[38] Using this criterion, we obtained a value of 3.5 for the reference pH. We then

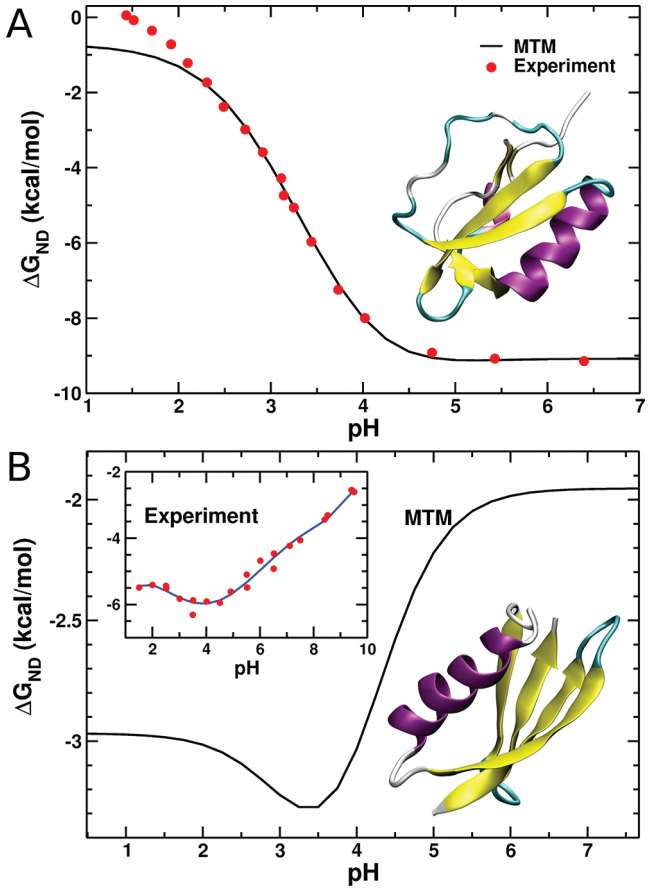


Figure 1. Stability of the native state, relative to the denatured-state ensemble ($\Delta G_{ND} = -k_B T \ln(P_N/P_D)$, where P_N and P_D are the probabilities of being in the native and denatured ensembles, respectively) as a function of pH. Panel A is for CI2, and panel B is for protein G. Native-state structures are shown of CI2 and protein G in a secondary structure representation based on crystal structures with PDB accession codes of 2CI2 and 1GB1, respectively. Experimental data (red circles) in (A) are from ref 38. Because experimental data for wild-type protein G are unavailable, we show in the inset in (B) experimental data (red circles) for a triple mutant protein G (T2Q, N8D, N37D).[35] The blue line is a fifth-order polynomial fit to the data and is used to guide the eye. For CI2, the temperature in the simulations was 302 K and in the experiment it was 298 K. For protein G, the simulation temperature was 317 K and in the experiment it was 298 K.

determined the temperature at which the calculated stability exactly equaled the experimental value of −6.0 kcal/mol at pH 3.5. This occurs at a simulation temperature of 302 K. In this way we set the overall free energy of this system to match the experiment at a single pH value. This procedure provided a reference solution condition (T = 302 K, pH 3.5) from which predictions at all other pH values are made by reweighting of the partition function using the MTM procedure in Eqs 1 and 6. Thus, despite the fact that there are no hydrogens in the coarse-grained simulation model, the thermodynamic effects of differential proton binding to N and D can still be accounted for within the MTM theory, as demonstrated by the successful comparisons between simulations and experiments. Experimental data of pK_a values and stability versus pH for wild-type protein G do not exist. Therefore, we chose a simulation temperature (317 K) that resulted in a native stability typical for such small proteins (−3.0 kcal/mol) and set pH_1 to 2.3. The trends and conclusions presented below are insensitive to the choice of reference pH, especially when such stability matching is carried out.

dx.doi.org/10.1021/ja206557y | J. Am. Chem.Soc. 2012, 134, 979–987

Two-dimensional native-state stability phase diagrams (e.g., $\Delta G_{ND}(f,\mathrm{pH})$, etc.) are computed by rewriting Eq 1 in terms of the probability of being folded as a function of f and pH using[48]

$$P_N(f,\mathrm{pH}) = Z(\mathrm{pH}_2,T,f)^{-1}\sum_{k=1}^{R}\sum_{t=1}^{n_k}\frac{\Theta_N(k,t)e^{-\beta E_P(k,t,\mathrm{pH}_2,f)}}{\sum_{m=1}^{R} n_m e^{F_m-\beta_m E_m(k,t)-f_m x(k,t)}} \quad (6)$$

where $\Delta G_{ND}(f,\mathrm{pH}) = -k_BT\ln(P_N(f,\mathrm{pH})/(1-P_N(f,\mathrm{pH})))$. All terms in Eq 6 are the same as in Eq 1, except we use the Heaviside step function $\Theta_N(k,t)$, which is 1 if conformation (k,t) is native and 0 otherwise. We calculate f_m using $P_N(f_m,\mathrm{pH}) = 0.5$.

RESULTS

Molecular Transfer Model for pH Effects on Proteins. The theory of the MTM hinges on the observation that if the partition function $Z(\mathrm{A})$ $(=\sum_j e^{-\beta E(j,\mathrm{A})})$ is known at some solution condition A, and if the free energy cost $\Delta G_{tr}(\mathrm{A}\rightarrow\mathrm{B})$ of transferring each protein conformation from A to an arbitrary solution condition B is known, then the partition function in B is $Z(\mathrm{B}) = \sum_j e^{-\beta E(j,\mathrm{A})-\beta\Delta G_{tr}(j,\mathrm{A}\rightarrow\mathrm{B})}$. In other words, the potential energy of the jth conformation in B is the sum of its potential energy in A $(E(j,\mathrm{A}))$ and the reversible work of transferring conformation j from A to B. In the current study, A and B differ in pH. In practice, the precision of the MTM, which is a mean-field-like approximation to the exact partition function, is limited only by the accuracy of the protein model Hamiltonian (i.e., the force field), the errors in the ΔG_{tr} model, and the extent of sampling in A.

We use the Aune–Tanford pH model,[49] which is among one of the most widely used theories to account for pH effects on protein stability,[17] to compute the free energy, $\Delta G_{tr}(\mathrm{pH}_1\rightarrow\mathrm{pH}_2)$, of transferring a protein conformation from pH_1 to pH_2. The change in the experimentally measured native (N) state stability with respect to the denatured (D) state stability (ΔG_{ND}) due to a change in pH is fit using

$$\Delta G_{tr}(l,\mathrm{pH}_1\rightarrow\mathrm{pH}_2) = -k_BT\sum_{k=1}^{N_t}\ln\left[\frac{10^{\mathrm{pH}_2}+10^{\mathrm{p}K_{k,l}}}{10^{\mathrm{pH}_1}+10^{\mathrm{p}K_{k,l}}}\right] \quad (7)$$

where the summation is over the N_t titratable groups and $\mathrm{p}K_{k,l}$ is the $\mathrm{p}K_a$ value of the k^{th} titratable group in the lth protein conformation. It can be shown that Eq 7 is a mean-field result obtained by integrating over all possible protonation states of a protein with N_t independent titratable groups in the native and denatured states.[50] The success of Eq 7 in modeling experimental ΔG_{ND} versus pH not only offers insight into the mechanism of pH denaturation, but also provides a means to estimate the free energy cost of transferring individual protein conformations from one solution pH to another.

To implement the MTM for pH effects, we use Hamiltonian replica exchange simulations[44] of the C_α side chain coarse-grained model[28] of protein G and CI2 to calculate the partition function $Z(\mathrm{A})$. We use experimental $\mathrm{p}K_a$ values (Table I, Supporting Information) to estimate the free energy cost of transferring the conformations (Eq 7) sampled in the simulations that we classify as belonging to either the native or denatured ensembles on the basis of suitable order parameters. To optimize the accuracy of the calculated partition function from the simulation, we use the weighted histogram equations (see the Methods for details).[37]

The Molecular Transfer Model Accurately Models pH Denaturation. We first calculated the thermodynamic properties of CI2 and protein G as a function of pH at $f = 0$. The MTM prediction of the dependence of $\Delta G_{ND}(\mathrm{pH})$ on pH for CI2 is in excellent agreement with experiment (Figure 1A).[38] Just as in the experiment, we find that the stability of CI2 decreases monotonically in a sigmoidal fashion as the pH decreases. Although there are small differences between the experimental and simulation $\Delta G_{ND}(\mathrm{pH})$ data at pH less than 2, the overall agreement with experiment shows that the MTM accurately models pH effects on the thermodynamics of folding and unfolding.

We also calculated ΔG_{ND} as a function of pH (Figure 1B) for wild-type protein G for which experimental data are not available. However, pH-dependent $\Delta G_{ND}(\mathrm{pH})$ for a triple mutant (T2Q, N8D, N37D) of protein G[35] has been measured. Although mutations can alter the native-state stability, it is likely that the response of ΔG_{ND} to pH for the wild-type and the mutant will be qualitatively similar. With this caveat, we note that the overall shape of the calculated ΔG_{ND} as a function of pH for wild-type protein G is similar to the experimental data from the triple mutant (Figure 1B). They both exhibit nonmonotonic trends with minima located in the pH range of 3–4. In addition, the differences in stability at two different pH values, $\Delta G_{ND}(\mathrm{pH}\ 7) - \Delta G_{ND}(\mathrm{pH}\ 3.4)$ are similar for the wild type and the triple mutant, with values of 1.3 and 1.7 kcal/mol, respectively. This suggests that the three mutations to titratable groups in the wild-type protein do not drastically alter the characteristics of the thermodynamic response of protein G to pH changes. The nonmonotonic dependence of ΔG_{ND} on pH observed for protein G contrasts with the monotonic dependence observed for CI2 (Figure 1A).

Force–pH and Force–Temperature Phase Diagrams. The versatility of the MTM is illustrated by probing the response of protein G and CI2 when, as is done in constant force single-molecule pulling experiments, a tensile force f is applied to their N and C termini at various pH values. Such constant force pulling simulations are at equilibrium. We calculated the phase diagram for both CI2 and protein G as a function of pH and f (Figure 2). On the basis of the destabilization of the native state of CI2 at acidic pH at $f = 0$ (Figure 1A), we expect that the midpoint force required to unfold CI2 should decrease as the pH decreases. This expectation is borne out in the pH range from 2 to 9 (Figure 2A). For CI2, decreasing pH facilitates force unfolding by stabilizing the denatured state. As a result, the force required to unfold CI2 decreases as the pH decreases (Figure 2A).

The f–pH phase diagram for protein G (Figure 2B) differs qualitatively from that for CI2. In contrast to CI2, for protein G, increasing the pH above 3.4 destabilizes the native-state ensemble, which implies that smaller forces are needed to unfold protein G (Figure 2B). These results show that the mechanical responses of proteins are strongly pH-dependent and can have opposite trends, which reflects the underlying stability of the proteins.

We also calculated for CI2 the f–T phase diagram at two pH values (Figure 3). The locus of points separating the folded, partially folded (see below), and unfolded structures is reminiscent of previously calculated f–T phase diagrams for simpler lattice and off-lattice models.[51,52] Not surprisingly, the region of stability increases as the temperature decreases

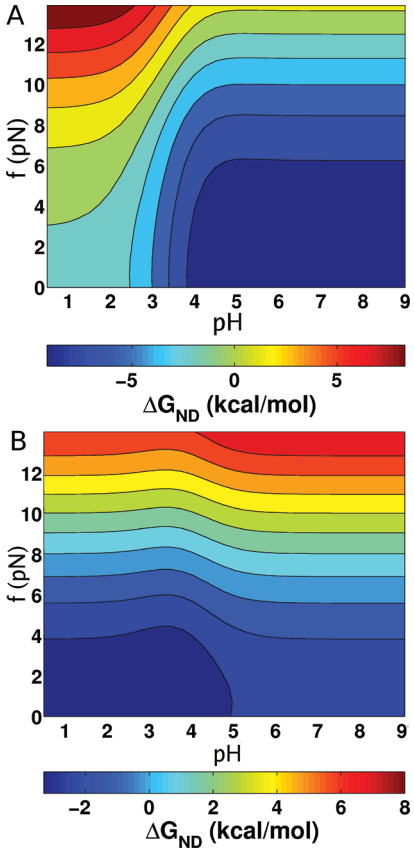


Figure 2. Force–pH phase diagram. (A) The f–pH diagram displays $\Delta G_{ND}(f,\mathrm{pH})$ for CI2 at a simulation temperature of 302 K. The solid lines correspond to lines of isostability. The scale for $\Delta G_{ND}(f,\mathrm{pH})$ is given at the bottom of the panel. (B) Same as (A), except it is for protein G at a simulation temperature of 317 K.

(compare the extent of blue regions in part A and B of Figure 3). The force required to destabilize CI2's native state increases as the temperature is lowered. Furthermore, at T = 280 K and pH 3.5, only at f > 12 pN (Figure 3B) is the native state unstable, whereas at pH 1.0, this occurs for f > 8 pN.

Force Midpoints. From the phase diagrams the pH-dependent midpoint unfolding force, f_m, can be determined using the criterion $\Delta G_{ND}(f_m,\mathrm{pH}) = 0$. Similarly, T-dependent f_m can be computed using $\Delta G_{ND}(f_m,T) = 0$. At high (>5) and low (<2) pH values, we find f_m is largely unchanged for both proteins (Figure 4). For CI2, the interplay between native-state

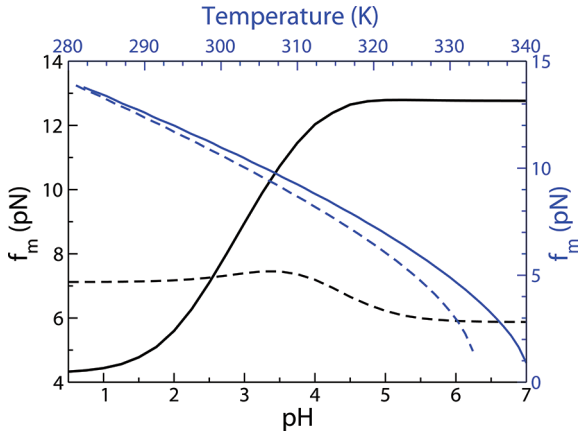


Figure 4. Force midpoint at various temperatures and pH. The temperature scale is on top in blue, and the corresponding scale for f_m is on the right. Solid lines are for CI2, and the dotted lines are for protein G, with blue corresponding to temperature and black to pH changes. Unless otherwise stated, the solution conditions for CI2 are 302 K and pH 3.5 and for protein G the conditions are 317 K and pH 2.3.

stabilization with increasing pH and the counteracting f-induced destabilization results in a population of partially structured conformations at $f < f_m$ and pH > 4 (see the blue regions in Figure 2A). At intermediate pH values (2 < pH < 5), f_m for CI2 is an increasing function of pH (Figure 4), which is a reflection of the enhanced stability of the native state at f = 0 (Figure 1A). In contrast, at intermediate values of pH, f_m for protein G exhibits nonmonotonic behavior with a maximum at pH 3.4, which coincides with the pH at which the native-state stability is largest when f = 0 (Figure 1B).

Although the dependence of f_m on pH differs greatly for the two proteins, the T-dependent f_m results exhibit qualitatively similar behavior (blue curves in Figure 4). Increasing temperature affects all interactions, whereas changing pH

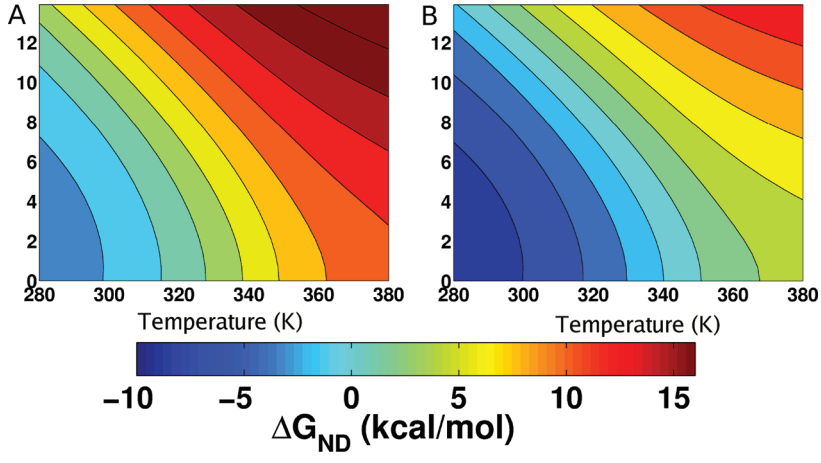


Figure 3. Force–temperature phase diagram for CI2. (A) Contours of this phase diagram at pH 1.0 are lines of isostability in $\Delta G_{ND}(f,T)$. Blue regions correspond to a thermodynamically stable native state, while red regions correspond to the unfolded state. (B) The f–T diagram is for pH 3.5. Enhanced stability at higher pH is reflected in large f–T regions in which the native state is stable.

dx.doi.org/10.1021/ja206557y | J. Am. Chem.Soc. 2012, 134, 979–987

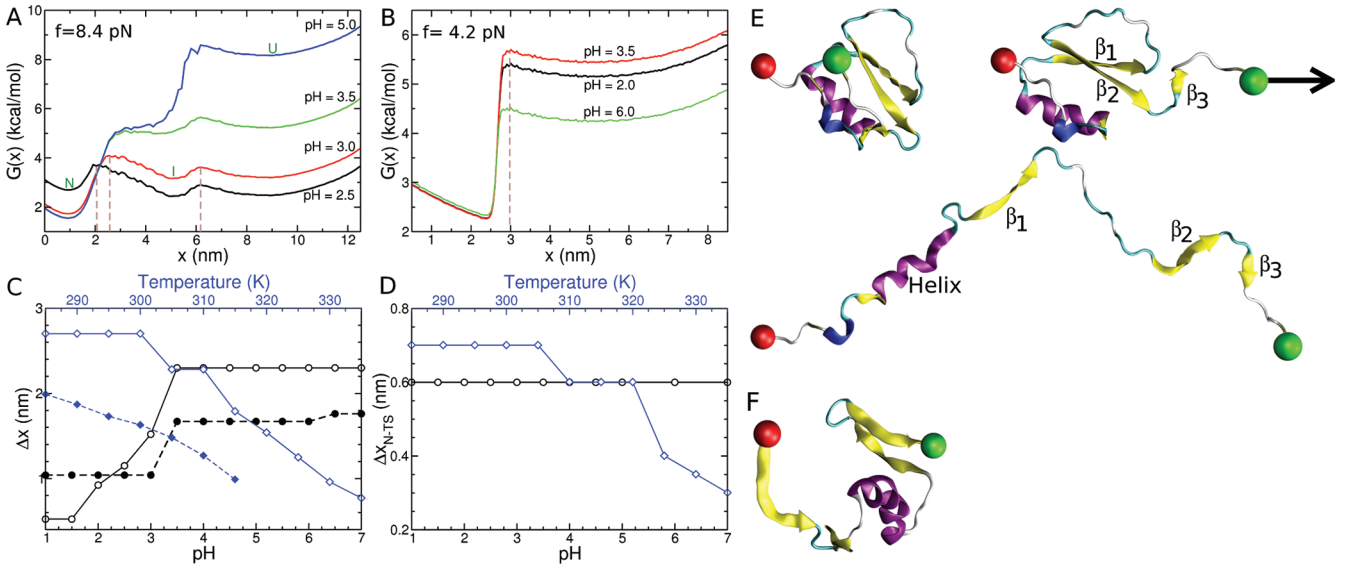


Figure 5. pH and temperature effects on the mechanical response of CI2 and protein G at constant tension forces of 8.4 and 4.2 pN, respectively. The free energy profile $G(x) = -k_BT \ln(P(x))$, where $P(x)$ is the probability of finding a given x value as a function of the end-to-end distance of the protein, projected onto the pulling vector, for (A) CI2 and (B) protein G at different pH values as labeled. The temperature is 302 and 317 K in (A) and (B), respectively. Brown dashed lines indicate transition-state locations at pH 2.5 and 3.0 in (A) and pH 3.5 in (B). The location of the native, intermediate, and fully unfolded basins of attraction in $G(x)$ are marked by the labels N, I, and U, respectively. (C) For CI2, the distances, Δx, between the native state and first transition state (Δx_{N-TS}) and intermediate state and second transition state (Δx_{I-TS}) are shown as a function of pH (lower axis) and temperature (upper axis). The black symbols correspond to pH, and blue symbols are for temperature. In both cases, solid lines show Δx_{N-TS} and dashed lines correspond to Δx_{I-TS}. (D) Same as (C) but for protein G. No intermediate basin of attraction exists for protein G, so only the distance between the native and transition states (Δx_{N-TS}) is reported. (E) Sample conformations (top to bottom) from the native, intermediate, and unfolded states of CI2 during simulations at 300 K and f = 8.68 pN. β-strands 1−3 are labeled, and the direction in which the constant tension is applied to the C-terminus (green sphere) is indicated by the black arrow. The N-terminal residue, fixed in space during the simulation, is shown as a red sphere. (F) Simulation structure of protein G with x = 3.09 nm from the replica at T = 320 K and f = 4.1 pN in the replica exchange simulations.

only affects titratable groups. Temperature effects are global, and pH effects are more localized. As a consequence, $\Delta G(f,\text{pH})$ and $\Delta G(f,T)$ are different, which leads to the predictions in Figure 4.

pH-Dependent Free Energy Profiles. Single-molecule force experiments are routinely used to obtain the f-dependent free energy profiles $G(x)$, where x is the end-to-end distance of the protein projected onto the pulling direction.[53,54] The free energy profiles in principle can yield both the barrier height to unfolding and the location of the transition state assuming that x is an appropriate reaction coordinate. It is also possible to extract the intrinsic ruggedness of the folding landscape at $f = 0$ using the f-dependent kinetics of unfolding at different temperatures.[55,56] This would require doing explicit kinetic simulations or experimentally measuring the unfolding rates. We first calculated $G(x)$ at several pH values for CI2 at f = 8.4 pN and protein G at f = 4.2 pN (Figure 5A,B).

For CI2, we observe three basins of attraction in $G(x)$ over a range of pH values (Figure 5A), which suggests that CI2 undergoes a two-stage force-induced unfolding transition in which a partially folded state is populated between the fully folded and fully unfolded basins. To obtain structural insights into the nature of the intermediate, we calculated the fraction of native contacts for various structural elements within the native topology of CI2 (see the structure in Figure 1A) as a function of x. The analysis indicates (Figure S2, Supporting Information) that the transition from the native to the intermediate basin (located between 2 and 6 nm in $G(x)$) corresponds to the unfolding of β-strand 3 (residues 75 and 76 in PDB 2CI2), resulting in the loss of tertiary interactions with β-strand 2 (residues 65−71) and the α-helix (residues 32−42). The transition to the unfolded basin (located at x > 7 nm) corresponds to the unfolding of the rest of the structural elements in the protein (i.e., β-strands 1 and 2 and interaction of these strands with the α-helix). Sample structures of the native, intermediate, and unfolded conformations from the simulations of CI2 under these conditions are consistent with this analysis (Figure 5E).

The pH-dependent free energy profiles for protein G under tension have only two basins of attraction (Figure 5B), which implies that force-induced unfolding occurs in a single step. The free energy barrier to unfolding increases from 2.1 kcal/mol at pH 6.0 to 3.4 kcal/mol at pH 3.5. Because the curvatures near the native basin and barrier top are roughly independent of pH (Figure 5B), it follows from the Kramers theory that transition rates between the folded and unfolded states are determined entirely by the barrier height. Thus, the calculated $G(x)$ profiles in conjunction with the Kramers theory predict that the unfolding rate $k_U(f)$ increases by a factor of 10 as the pH increases from 6.0 to 3.5. The predictions for free energy profiles and the inferred changes in unfolding rates are amenable to experimental tests.

To ascertain the generality of our conclusions, we show in Figure 6 the free energy profiles over a wide range of forces. Just as in Figure 5, we find that for CI2 there is a force-induced folding intermediate, suggesting that the two transition states (TSs) persist at all relevant f values. Remarkably, the invariance of the location of the TS is preserved at all forces. Both these figures show in a rather dramatic manner that the mechanical responses of CI2 and protein G are very different, which

dx.doi.org/10.1021/ja206557y | J. Am. Chem.Soc. 2012, 134, 979−987

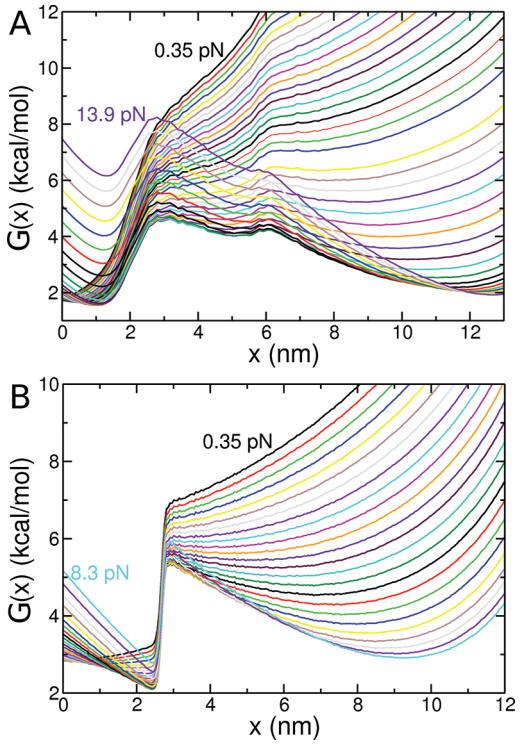


Figure 6. Equilibrium free energy profiles $G(x)$ at various constant tension forces at pH 3.5 for (A) CI2 at 302 K and (B) protein G at 317 K. As indicated in the panels, the force values range from 0.35 to 13 pN for CI2 and from 0.35 to 8.3 pN for protein G. Successive profiles differ by approximately 0.35 pN of applied tension. Comparisons between (A) and (B) reveal vividly the dramatic differences in the compliance between these two proteins, thus underscoring the importance of the native structure.

reflects the underlying variations in the native topology[52] (see below for additional discussions).

pH and Temperature-Dependent Movements in the Transition-State Location Suggest Hammond–Leffler Behavior. According to the Hammond–Leffler postulate,[57] the TS should resemble the least stable species in the reaction. Although originally proposed for reactions of small organic molecules, Hyeon and Thirumalai[58] showed that the Hammond postulate is also applicable to force unfolding of biomolecules regardless of the nature of the reaction coordinate. For proteins under tension, this implies that the location, x_{TS}, of the TS, should either be independent of f or move toward the native state when f increases.

The $G(x)$ profiles for CI2 at f = 8.4 pN (Figure 5A) show that there are two transition states, one between the native and intermediate states, whose distance is Δx_{N-TS} with respect to the location of the native state, and the other between the intermediate and fully unfolded ensemble, whose distance is Δx_{I-TS}. Figure 5C shows that Δx_{N-TS} and Δx_{I-TS} are independent of pH when the pH exceeds 3.5. As the pH increases, resulting in enhanced stability of both N with respect to I and I with respect to U (Figure 5A), Δx_{N-TS} and Δx_{I-TS} increase with a dramatic jump at pH 3.0. These results imply that the locations of the two transition states move closer to the less stable species, which is in accord with the Hammond–Leffler postulate. As a corollary, we expect and find (Figure 5C) that upon an increase in temperature Δx_{N-TS} and Δx_{I-TS} should decrease as both the folded and intermediate states are destabilized relative to the unfolded state. Although these observations do not establish the adequacy of the one-dimensional reaction coordinate to describe f-induced unfolding of CI2, they support the generality of the Hammond–Leffler postulate for interpreting force spectroscopy results.[55]

The abrupt change in Δx_{N-TS} and Δx_{I-TS} at pH 3 ranges from about 1.7 to 2.0 nm (Figure 5C). Similarly, the change in Δx_{N-TS} is 2.3 nm as the temperature is increased from 300 to 340 K. Such large changes in TS locations are not typically observed in constant loading rate atomic force microscopy (AFM) experiments. For example, the maximum value of x_{N-TS} observed in filamin is ~0.7 nm.[59] The observed changes for CI2 are similar to the values obtained in the transition from the intermediate to the unfolded state in RNase H using laser optical tweezer experiments.[4] The large variations in Δx_{N-TS} and Δx_{I-TS} for CI2 show that besides experimental conditions the native-state topology must also play a critical role in response to f.

In sharp contrast to CI2, the TS changes in the unfolding of protein G are very different. The TS location Δx_{N-TS} is independent of pH (Figure 5B,), which implies that protein G behaves as a brittle material when subjected to f at all pH values. As the temperature increases, Δx_{N-TS} decreases (Figure 5D) in two steps, one at 305 K and the other at 320 K. In comparison to CI2, the values of Δx_{N-TS} are roughly in the range observed for several proteins using AFM experiments. The value of Δx_{N-TS} changes by 0.4 nm as the temperature increases from 280 to 340 K. The decrease in Δx_{N-TS}, reflecting the movement of the TS closer to the native state, as temperature increases is consistent with the Hammond–Leffler postulate.

■ DISCUSSION

We have introduced a way to account for pH effects on proteins within the framework of the Molecular Transfer Model. Our formulation overcomes the key limitations of the Tanford thermodynamic model,[49] which is restricted to predicting only changes in protein stability due to changes in pH. Besides accomplishing this goal, the MTM also offers a molecular interpretation of folding and unfolding over a broad range of external conditions, including the response to f and pH. In principle, the MTM can be combined with all-atom simulations to calculate pH effects on proteins. As a matter of practice, however, currently such simulations undersample the partition function of proteins and therefore do not yield statistically significant results for the self-assembly of proteins.

Applications of the f-dependent response to pH of proteins using the MTM have revealed a number of surprising predictions. In particular, we found that f_m had a nonlinear dependence on pH; f_m increases at acidic pH for protein G, whereas it decreases for CI2. These results correlate with the pH dependence of ΔG_{ND} at f = 0, a conclusion also reached from single-molecule force experiments.[36] We have also shown that the movement of the transition-state location follows Hammond–Leffler behavior at all forces and solution conditions examined here. Large or discontinuous changes in the transition-state location inferred from the free energy profiles provide structural evidence for plasticity or brittleness of forced unfolding of CI2 and protein G. We note that the sequence of events during an unfolding event cannot be directly calculated from the Hamiltonian replica exchange simulations, which are restricted to obtaining the measurable equilibrium free energy profiles. Brownian dynamics or all all-atom molecular dynamics simulations have to be performed to

dx.doi.org/10.1021/ja206557y | J. Am. Chem.Soc. 2012, 134, 979–987

obtain the force-induced unfolding kinetics. Nevertheless, using our previous study,[52] which established a link between the structure of the native state and potential unfolding pathways, we can suggest a plausible structural origin of the brittleness of protein G. Our earlier work[52] showed that upon application of force unfolding a shearing-type motion occurs between β-sheets that are arranged in an antiparallel manner. Using this result, we surmise that protein G unfolds by shearing (or sliding) of the strands in the β-sheets (most likely the C-terminal strands) with respect to each other. Hence, the transition is cooperative involving an f-independent transition state (Figures 5 and 6). In contrast, unfolding is gradual in the plastic protein CI2 in which the TSs move in response to f (Figures 5 and 6). Explicit kinetic simulations are needed to enumerate the force-induced unfolding pathways and to further confirm the drastically different responses to f predicted for these proteins.

Our results can be compared at a qualitative level to single-molecule constant pulling speed experiments on ubiquitin[36] in which pH effects were studied. Those experiments found[36] that the unfolding force of ubiquitin is constant over a range of pH values (6−10) and decreases at acidic pH. These findings are qualitatively consistent with our results for CI2 (Figure 4A). A major prediction from our results is that such pH-dependent trends depend critically on the specific protein under study. For example, f_m for protein G shows the opposite of the trend observed in CI2; f_m increases slightly at more acidic pH values (Figure 4B).

In single-molecule pulling experiments, with x as the only experimentally accessible coordinate, identification of the TS location with the ensemble of TS structures in the multi-dimensional landscape is a challenging problem. It is possible that at $f > f_m$ the pulling coordinate is a good reaction coordinate because at large forces the molecule is likely to be aligned along the f direction, thus forcing it to unfold along the coordinate conjugate to f. Recently, we showed that the suitability of x as a reaction coordinate is determined by the interplay between compaction (determined by protein stability) and tension (dependent on x_{TS} and the barrier to unfolding).[60] A test of the adequacy of x as a reaction coordinate is captured by the experimentally measurable molecular tensegrity parameter, $s = f_c/f_m$, where the unfolding critical force, f_c, equals $\Delta G^{\ddagger}/x_{TS}$. Here $\Delta G^{\ddagger}$ is the height of the free energy barrier. For CI2, with its two transition states, the values of $s_1(N\rightarrow TS)$ and $s_2(I\rightarrow TS_2)$ are 0.019 and 0.005 at pH 3.0. The theory of Morrison et al.[60] predicts that at this pH x is a good reaction coordinate for both the transitions because it is likely that the ensemble of conformations starting from Δx_{N-TS} (Δx_{I-TS}) would reach I and N (I and U) with equal probability ($p_{fold} \approx 0.5$). For protein G at pH 6.0, $s = 0.058$, which also lies in the range for which x is likely to be a good reaction coordinate as assessed by the theory outlined in ref 60.

A number of assumptions underlie our application of the MTM, including the temperature independence of pH transfer free energies ($\Delta G_{tr}(k,t,pH_2)$) and the use of the independent site model of titration. In addition, there are other assumptions[17,50] that are inherent to the Aune−Tanford model that should be kept in mind in specific applications of our theory. However, the excellent agreement between experiments and simulations demonstrated here (Figure 1) and in previous applications of MTM[27,31] suggests that this assumption is reasonable for the proteins and solution conditions studied here.

To predict pH effects on proteins, we utilized experimentally measured pK_a values of titratable side chains of protein G and CI2.[34,38] In the absence of such experimental data, pK_a values calculated using quantum chemical methods can be utilized. Alternatively, the MTM could potentially be utilized to solve the inverse problem of predicting pK_a values either from simulation structures alone or from known changes of protein properties as a function of pH.[61] The MTM could also be used to test different functional forms that go beyond the two-state mean-field assumption of Eq 7.[50]

The Molecular Transfer Model is a significant advance in our ability to model in a natural way the effects of osmolytes and pH on the folding of proteins. By combining thermodynamic models and physiochemical data, the MTM incorporates the effects of osmolytes and pH into simulations in a physically transparent and theoretically rigorous manner. Consequently, reliable simulations can be performed to predict measurable quantities, which enables a direct comparison to experiments.[27,30,31,62]

■ ASSOCIATED CONTENT

Supporting Information

Additional data analysis, figures, and force field parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

thirum@umd.edu

Present Address

[§]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.

■ ACKNOWLEDGMENTS

We are grateful to Changbong Hyeon and Greg Morrison for useful discussions. E.P.O. thanks Professor D. Wayne Bolen for suggesting the use of the Henderson−Hasselbach equation to estimate group transfer free energies upon a change in pH. This work was supported in part by grants from the National Science Foundation (NSF; 09-14033) and National Institutes of Health (NIH; GM089685) to D.T. and an NIH Graduate Partnerships Program (GPP) Biophysics Fellowship and NSF postdoctoral fellowship to E.P.O. In this study we utilized the high-performance computational capabilities of the Biowulf Linux cluster at the NIH, Bethesda, MD (http://biowulf.nih.gov).

■ REFERENCES

(1) Schuler, B.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16−26.

(2) Borgia, A.; Williams, P. M.; Clarke, J. *Annu. Rev. Biochem.* **2008**, *77*, 101−125.

(3) Camacho, C. J.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6369−6372.

(4) Cecconi, C.; Shank, E. A.; Bustamante, C.; Marqusee, S. *Science* **2005**, *309*, 2057−2060.

(5) Bertz, M.; Wilmanns, M.; Rief, M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13307−13310.

(6) Fernandez, J. M.; Li, H. B. *Science* **2004**, *303*, 1674−1678.

(7) Hyeon, C. B.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10249−10253.

(8) Hummer, G.; Szabo, A. *Acc. Chem. Res.* **2005**, *38*, 504−513.

(9) Camacho, C. J.; Thirumalai, D. *Phys. Rev. Lett.* **1993**, *71*, 2505−2508.

(10) Garcia-Manyes, S.; Dougan, L.; Fernandez, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 10540−10545.
(11) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528−1533.
(12) Rhoades, E.; Cohen, M.; Schuler, B.; Haran, G. *J. Am. Chem. Soc.* **2004**, *126*, 14686−14687.
(13) Cao, Y.; Li, H. *J. Mol. Biol.* **2008**, *375*, 316−324.
(14) Thirumalai, D.; O'Brien, E. P.; Morrison, G.; Hyeon, C. *Annu. Rev. Biophys.* **2010**, *39*, 159−183.
(15) Robinson, D. R.; Jencks, W. P. *J. Am. Chem. Soc.* **1965**, *87*, 2462−2470.
(16) Bolen, D. W.; Santoro, M. M. *Biochemistry* **1988**, *27*, 8063−8068.
(17) Tanford, C. *Adv. Protein Chem.* **1970**, *24*, 1−95.
(18) Tanford, C.; Roxby, R. *Biochemistry* **1972**, *11*, 2192−2198.
(19) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1993**, *231*, 459−474.
(20) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415−436.
(21) Onufriev, A.; Case, D. A.; Ullmann, G. M. *Biochemistry* **2001**, *40*, 3413−3419.
(22) Whitten, S. T.; Wool, J. O.; Razeghifard, R.; Garcia-Moreno, B.; Hilser, V. J. *J. Mol. Biol.* **2001**, *309*, 1165−1175.
(23) Khandogin, J.; Brooks, C. L. *Biochemistry* **2006**, *45*, 9363−9373.
(24) Fitch, C. A.; Karp, D. A.; Lee, K. K.; Stites, W. E.; Lattman, E. E.; Garcia-Moreno, B. *Biophys. J.* **2002**, *82*, 3289−3304.
(25) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 738−752.
(26) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. *Biophys. J.* **2011**, *100*, L47−L49.
(27) O'Brien, E. P.; Ziv, G.; Haran, G.; Brooks, B. R.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13403−13408.
(28) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2544−2549.
(29) Auton, M.; Bolen, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15065−15068.
(30) O'Brien, E. P.; Brooks, B. R.; Thirumalai, D. *Biochemistry* **2009**, *48*, 3743−3754.
(31) Liu, Z.; Reddy, G.; O'Brien, E. P.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 7787−7792.
(32) Jackson, S. E.; Fersht, A. R. *Biochemistry* **1991**, *30*, 10428−10435.
(33) Jackson, S. E.; Moracci, M.; Elmasry, N.; Johnson, C. M.; Fersht, A. R. *Biochemistry* **1993**, *32*, 11259−11269.
(34) Khare, D.; Alexander, P.; Antosiewicz, J.; Bryan, P.; Gilson, M.; Orban, J. *Biochemistry* **1997**, *36*, 3580−3589.
(35) Lindman, S.; Linse, S.; Mulder, F. A. A.; Andre, I. *Biophys. J.* **2007**, *92*, 257−266.
(36) Chyan, C. L.; Lin, F. C.; Peng, H. B.; Yuan, J. M.; Chang, C. H.; Lin, S. H.; Yang, G. L. *Biophys. J.* **2004**, *87*, 3995−4006.
(37) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(38) Tan, Y. J.; Oliveberg, M.; Davis, B.; Fersht, A. R. *J. Mol. Biol.* **1995**, *254*, 980−992.
(39) Go, N. *Annu. Rev. Biophys. Bioeng.* **1983**, *12*, 183−210.
(40) Miyazawa, S.; Jernigan, R. L. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 49−68.
(41) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(42) McPhalen, C. A.; James, M. N. *Biochemistry* **1987**, *26*, 261−269.
(43) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657−661.
(44) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6051.
(45) Kouza, M.; Hu, C. K.; Li, M. S. *J. Chem. Phys.* **2008**, *128*, 045103.
(46) Veitshans, T.; Klimov, D.; Thirumalai, D. *Folding Des.* **1997**, *2*, 1−22.
(47) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(48) Klimov, D. K.; Thirumalai, D. *J. Phys. Chem. B* **2001**, *105*, 6648−6654.
(49) Aune, K. C.; Tanford, C. *Biochemistry* **1969**, *8*, 4579−4585.
(50) Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 9556−9561.
(51) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6166−6170.
(52) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7254−7259.
(53) Gebhardt, J. C. M.; Bornschlogla, T.; Rief, M. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2013−2018.
(54) Greenleaf, W. J.; Frieda, K. L.; Foster, D. A. N.; Woodside, M. T.; Block, S. M. *Science* **2008**, *319*, 630−633.
(55) Hyeon, C.; Thirumalai, D. *J. Phys.: Condens. Matter* **2007**, *19*, 113101.
(56) Nevo, R.; Brumfeld, V.; Kapon, R.; Hinterdorfer, P.; Reich, Z. *EMBO Rep.* **2003**, *6*, 482−486.
(57) Leffler, J. E. *Science* **1952**, *117*, 340−341.
(58) Hyeon, C. B.; Thirumalai, D. *Biophys. J.* **2006**, *90*, 3410−3427.
(59) Schlierf, M.; Rief, M. *J. Mol. Biol.* **2005**, *354*, 497−503.
(60) Morrison, G.; Hyeon, C.; Hinczewski, M.; Thirumalai, D. *Phys. Rev. Lett.* **2011**, *106*, 138102.
(61) Elcock, A. H. *J. Mol. Biol.* **1999**, *294*, 1051−1062.
(62) O'Brien, E. P.; Morrison, G.; Brooks, B. R.; Thirumalai, D. *J. Chem. Phys.* **2009**, *130*, 124903.