THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Modeling the Solvation of Nonpolar Amino Acids in Guanidinium Chloride Solutions

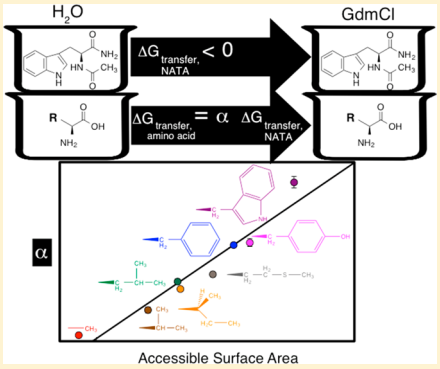
Paul Cohen,[†,§] Ken A. Dill,[‡] and Sheila S. Jaswal*[,†]

[†]Department of Chemistry, Amherst College, P.O. Box 5000, Amherst, Massachusetts 01002, United States

[‡]Laufer Center for Physical & Quantitative Biology and Departments of Chemistry and Physics, Stony Brook University, Stony Brook, New York 11794, United States

Supporting Information

ABSTRACT: It is common to denature proteins by using high temperatures or by adding guanidinium chloride (GdmCl). However, the physical mechanism of denaturation is not well understood. Based on extensive experimental data, we developed a thermodynamic binding-polynomial model for the process of transferring nonpolar amino acids from water into GdmCl solutions, as a function of temperature and GdmCl concentration. To mimic nonpolar amino acids, we utilized the model compound, *N*-acetyl-tryptophanamide (NATA). We find that all nonpolar amino acids behave like NATA, with a scale factor linearly dependent on the surface area. Our model with three thermodynamic parameters fully captures the nonlinear dependencies on both the temperature and GdmCl concentration: binding the first guanidinium ion (Gdm^+) to NATA has favorable entropy and unfavorable enthalpy of desolvation (ΔS = +11.7 cal/mol, ΔH = +3.9 kcal/mol), while cooperativity of binding a second Gdm^+ has a small contribution (K = 0.032 ± 0.003). This model may be useful for a better understanding of protein denaturation by temperature and GdmCl.



INTRODUCTION

The ability of GdmCl to unfold proteins has been exploited for decades to probe the thermodynamics and kinetics of protein folding.[1] The success of the empirically based linear extrapolation model in describing the GdmCl concentration dependence of protein unfolding has enabled the characterization of hundreds of proteins in terms of their stability in the absence of denaturant ($\Delta G_{u,H2O}$) and an empirical *m*-value, describing the dependence of the free energy of unfolding in denaturant ($\Delta G_{u,D}$) on denaturant concentration (c_D):

$$\Delta G_{u,D} = \Delta G_{u,H2O} - mc_D \quad (1)$$

While the *m*-value is observed to correlate with the change in total accessible surface area (ΔASA) and the heat capacity change of unfolding,[2] its physical basis in microscopic solvating structures of water and denaturants is less clear. On the one hand, modern Kirkwood−Buff solution theory is now capable of explaining thermodynamic activity coefficients on the basis of molecular dynamics simulations of potentials of mean force (PMF).[3] On the other hand, our interest here is in finding the best interpretation of the molecular solvating structures for *m*-values from experimental data, where we have no PMFs. Furthermore, $\Delta G_{u,D}$ and in some cases the activation free energy of unfolding ($\Delta G^{\ddagger}_{u,D}$) demonstrate a nonlinear dependence on GdmCl concentration for several proteins.[4−8]

Nozaki and Tanford's observation that GdmCl enhances the solubility of amino acids and their derivatives[9] suggested that the basis for GdmCl action is its ability to decrease the free energy penalty for solvating groups newly exposed upon unfolding. Importantly, the free energy of solvation (ΔG_{solv}) of amino acids displays a nonlinear dependence on GdmCl concentration. Alonso and Dill[10] demonstrated the ability of a model based in solution theory to capture the nonlinear dependence observed for the ΔG_{solv} of leucine on [GdmCl] at 25 °C:

$$\Delta G_{solv} = (g_1 c_D + g_2 c_D^2) \quad (2)$$

The first-order term g_1 accounts for the interaction of leucine with either water or a Gdm^+, and the second-order term g_2 takes into account interactions between water and Gdm^+. Using this model for [GdmCl] dependence, they were able to use leucine as a representative amino acid and predict a cooperative unfolding transition of a simulated protein with increasing GdmCl concentration.[10]

The ability of eq 2 to capture the nonlinear behavior of leucine solvation with physically interpretable parameters and to predict cooperative unfolding behavior on the protein level suggests that application of a solution-theory based model to protein unfolding data could yield parameters that report on the nature of the protein surface exposed upon unfolding. In particular, understanding the relationship between the degree and nature of the exposed surface and solvation of model compounds by GdmCl would enable information about the degree and nature of the exposed surface area in partially

Received: June 26, 2014
Revised: August 19, 2014
Published: August 20, 2014

   dx.doi.org/10.1021/jp506379r | *J. Phys. Chem. B* 2014, 118, 10618−10623

unfolded states to be determined from the observed dependence of $\Delta G_{u,D}$ and $\Delta G^{\ddagger}_{u,D}$ on [GdmCl]. However, the technical requirements of carrying out the densimetric measurements to assess amino acid solubility for each desired concentration limit the range of data available. The original Nozaki and Tanford[9] data set of ΔG_{solv} versus [GdmCl] has only four data points for each amino acid. Using a multiple parameter model to fit the nonlinearity of the limited amino acid solubility data is circumspect.

To enable more ready determination of solubility through simple absorbance measurements, the Clarke group employed a tryptophan derivative NATA as a model for nonpolar amino acids (Figure 1).[11] Their comprehensive data set includes

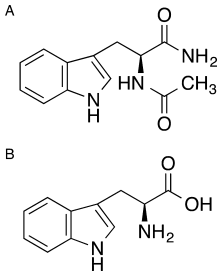


Figure 1. Model compound structures. (A) *N*-Acetyl-tryptophanamide. (B) Tryptophan.

ΔG_{solv} data for NATA ($\Delta G_{solv,NATA}$) at 15 GdmCl concentrations from 0.4 to 6 M, and at four temperatures, 15, 25, 35, and 45 °C. NATA demonstrates the canonical nonlinear dependence of ΔG_{solv} on [GdmCl] at all temperatures and, thus, serves as an experimentally accessible substitute for nonpolar amino acids.

Here we apply a binding polynomial formalism based in solution theory and exploit Clarke's comprehensive NATA data set to provide the foundation for extracting physically relevant parameters from protein data in GdmCl. We demonstrate that this model captures the [GdmCl] dependence of ΔG_{solv} for nonpolar amino acids ($\Delta G_{solv,aa}$) using [GdmCl]-dependent $\Delta G_{solv,NATA}$-derived parameters as constants with just a single scale factor as a fit parameter. The scale factors show a remarkable correlation with the accessible surface area (ASA) of the amino acid residues, validating NATA as a model compound for nonpolar amino acid solvation. In addition, by incorporating enthalpic and entropic components for the GdmCl interaction to the model we can capture the complete temperature and GdmCl dependence of NATA solvation using only three parameters. We discuss the implications of our results for understanding the interaction of Gdm^+ with a nonpolar surface and the potential application to protein solvation and unfolding data in GdmCl.

METHODS

Amino Acid Solvation Data. The combined amino acid data used for analysis can be found in Supplementary Table 1. $\Delta G_{solv,aa}$ data (cal/mol) as a function of [GdmCl] was obtained for the nonpolar amino acids alanine, valine, leucine, methionine, phenylalanine, tyrosine, and tryptophan from Nozaki and Tanford[9] and for isoleucine from Sarker and Bolen.[12]

***N*-Acetyl Tryptophanamide Solvation Data.** Data for $\Delta G_{solv,NATA}$ (cal/mol) as a function of [GdmCl] and temperature were obtained from Parker et al.[11] and can be found in Supplementary Table 2.

Solution Theory Formalism for GdmCl Concentration Dependence of ΔG_{solv}. The following is an explanation of how the free energy of transfer of an amino acid solute between two media that differ in temperature and concentration of denaturant (cosolute) can be computed. Consider the thermodynamic cycle in Figure 2 for transferring a solute **s**

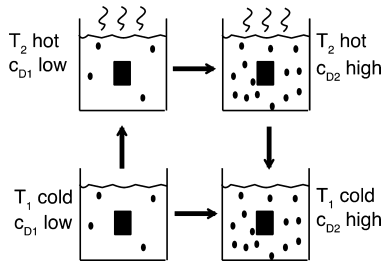


Figure 2. Thermodynamic cycle for solute transfer.

from a water solution at temperature $T = T_1$ and denaturant concentration $c_D = c_{D1}$ to a second water solution at temperature $T = T_2$ and denaturant $c_D = c_{D2}$.

Now, we can compute the free energy of this transfer by using the semigrand ensemble of thermodynamics, $\tilde{G}(T, p, \mu_s, \mu_d)$, where T is temperature, p is pressure, μ_s is the solute chemical potential, and μ_d is the denaturant chemical potential. Or, more simply $\tilde{G} = \tilde{G}(T, P, c_s, c_D)$ where c_s is the solute concentration. Now, we can express $\tilde{G}$ in terms of its partition function $\tilde{Q}$:[13]

$$\tilde{G} = -RT \ln \tilde{Q} \tag{3}$$

Standard solution theory suggests the following form[10,14,15]

$$\tilde{Q} = ac_s[1 + K_1c_D + K_2{c_D}^2] \tag{4}$$

Note that when $c_d = 0$, eq 4 reduces to the standard expression for solutes in pure water:

$$\tilde{Q} = ac_s \Rightarrow \tilde{G} = \tilde{G}_0 = \mu_0 + RT \ln c_s \text{ (where } \mu_0 = -RT \ln Q) \tag{5}$$

For the remainder of the article, we will drop the subscript $_D$, so that c will represent c_D.

More interesting here is when $c_s = 1$ (standard state of the solute). Then we have

$$\Delta\tilde{G}_{solv} = -RT \ln[1 + K_1c + K_2c^2] \tag{6}$$

K_1 is related to the strength of the interaction between Gdm^+ and the solute, K_2 incorporates the excess of interaction beyond two Gdm^+ ions interacting independently with the solute, which is essentially the cooperativity between the first and second binding events of Gdm^+ to the solute.

Using NATA-Derived Parameters to fit ΔG_{solv} for nonpolar amino acids. Equation 6 is a ΔG_{solv} per unit surface area of the solute. In order to account for the different surface areas of the different nonpolar amino acids, we introduce an amino acid specific surface-area scaling factor (σ_{aa}). Now, the variables K_1 and K_2 are replaced with the constants $K_{1,NATA}$ and $K_{2,NATA}$ obtained from fitting $\Delta G_{solv,NATA}$ at 25 °C:

$$\Delta G_{solv,aa} = -\sigma_{aa}RT \ln(1 + K_{1,NATA}c + K_{2,NATA}c^2) \tag{7}$$

dx.doi.org/10.1021/jp506379r | J. Phys. Chem. B 2014, 118, 10618−10623

Predicting the Relationship between Scale Factor and the Residue ASA. ASA in Å^2 for amino acid residues in the context of Gly-X-Gly was obtained from Miller et al.[16] The relationship between σ_{aa} and the residue ASA for the amino acids is fit to a linear regression:

$$\sigma_{aa} = y - m\mathrm{ASA} \tag{8}$$

Incorporating Temperature Dependence of $\Delta G_{solv,NATA}$. K_1 and K_2 may have temperature dependences. These can be found by plotting ln K versus $1/T$ and using the following expression in terms of enthalpy and entropy:

$$\ln K = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \tag{9}$$

Assuming that the primary temperature dependence is in K_1, substituting eq 9 into eq 6 yields a model for how the solvation free energy of NATA depends on T and c:

$$\Delta G_{solv,NATA} = -RT\ln(1 + be^{-a/T}c + K_2c^2) \tag{10}$$

where $a = \Delta H/R$ and $b = e^{\Delta S/R}$.

Software. Mathematica version 8.0 was used for all analysis and linear and nonlinear regression.

RESULTS

Solution-Theory Binding Polynomial Formalism Fits NATA Solubility Data in GdmCl. Alonso and Dill had shown that the two-parameter solution-theory based model captures the nonlinear dependence of the ΔG_{solv} for leucine on [GdmCl] at 25 °C, using the available $\Delta G_{solv,Leu}$ measurements at four concentrations.[10] We wanted to test the related binding polynomial model on the larger data set of the nonpolar model compound, NATA, to more rigorously validate the model's ability to capture the nonlinearity of nonpolar solvation as a function of [GdmCl]. Fitting the 25 °C $\Delta G_{solv,\ NATA}$ data at 15 GdmCl concentrations with eq 6 leads to an excellent fit (solid gray line in Figure 3A, R = 0.999) and K_1 = 0.56 ± 0.02 and K_2 = 0.035 ± 0.006. We conclude that this simple model accurately describes NATA solvation.

Nonpolar Amino Acid Solvation Scales NATA Solvation Behavior. We next determined whether NATA solvation is representative of nonpolar amino acid solvation in general. To test this, we fit each of the data sets for the dependence of $\Delta G_{solv,aa}$ on [GdmCl] at 25 °C for alanine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine, and tryptophan with eq 7 using the NATA-derived $K_{1,NATA}$ and $K_{2,NATA}$ parameters as constants and σ_{aa} as the only variable. This NATA-based model fits the nonpolar amino acid data very well (solid lines in Figure 3A), with R values above 0.996 for all except valine (0.98) and alanine (0.75, which appears to have a linear dependence on [GdmCl]) (Table 1). Thus, NATA is a valid model of the exposed nonpolar surface in proteins.

Factor That Scales NATA Parameters for Nonpolar Amino Acid Fits Is Correlated with Residue Accessible Surface Area. What is the relationship between the scale factor that is required to predict $\Delta G_{solv,aa}$, given the $\Delta G_{solv,NATA}$? The σ_{aa} values resulting from fitting the nonpolar amino acid solvation data to eq 8 range from 0.05 for alanine, the smallest amino acid of the set, to 1.24 for tryptophan, the largest (Table 1).

The σ_{aa} values increase with residue ASA over the eight amino acids (Figure 3B). Linear regression yielded a remarkably good correlation (R = 0.96), leading to the following equation:

$$\sigma_{aa} = (-1.0 \pm 0.2) + (0.0079 \pm 0009)\ \mathrm{ASA\ (Å^2)} \tag{11}$$

Thus, the degree to which NATA parameters need to be scaled to fit the dependence on [GdmCl] of $\Delta G_{solv,aa}$ for nonpolar amino acids reports on the extent of residue ASA being solvated.

Temperature Dependence of Nonpolar Solvation. Having established the validity of both NATA as a good model for nonpolar amino acids, and the relation in eq 6 to determine the physical parameters K_1 and K_2 describing the interaction between GdmCl and nonpolar surface, we investigated the underlying thermodynamics of nonpolar solvation in GdmCl. We analyzed the comprehensive $\Delta G_{solv,NATA}$ data set (Supplementary Table 2, which demonstrates increasingly favorable solvation by GdmCl as the temperature is increased from 15 to 45 °C[11]) to extract the

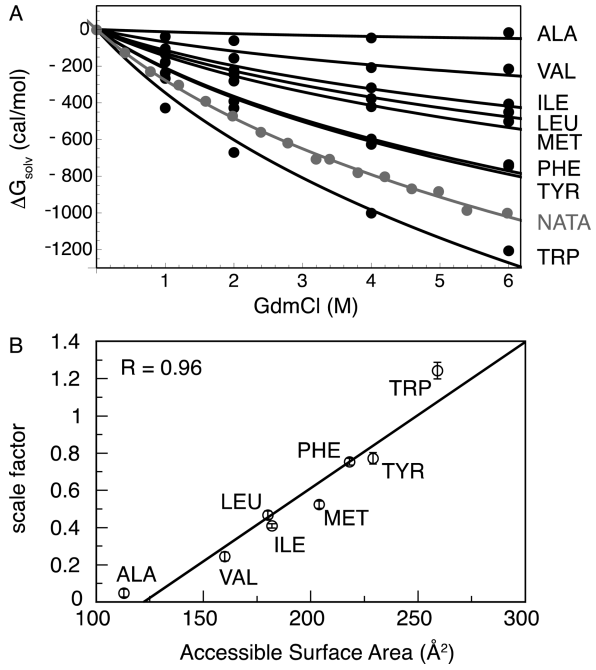


Figure 3. Nonpolar solvation in GdmCl captured by solution-theory parameters for NATA. (A) ΔG_{solv} as a function of [GdmCl] at 25 °C for NATA (gray symbols) compared to nonpolar amino acids (black symbols) as listed. Gray line is a fit to NATA data using eq 6 $\Delta G_{solv} = -RT\ \ln[1 + K_1c + K_2c^2]$ Black lines are fits to nonpolar amino acid data using eq 7, $\Delta G_{solv,aa} = -\sigma_{aa}RT\ \ln(1 + K_{1,NATA}c + K_{2,NATA}c^2)$, where $K_{1,NATA}$ = 0.56 and $K_{2,NATA}$ =0.035. (B) σ_{aa} demonstrates a linear correlation with residue ASA.

Table 1. Results of Fitting Nonpolar $\Delta G_{solv,aa}$ Dependence on [GdmCl] with NATA Parameters

amino acid	σ_{aa}	R value
alanine	0.05 ± 0.02	0.758
isoleucine	0.41 ± 0.01	0.998
leucine	0.47 ± 0.02	0.996
methionine	0.52 ± 0.02	0.997
phenylalanine	0.75 ± 0.03	0.999
tryptophan	1.24 ± 0.04	0.997
tyrosine	0.77 ± 0.03	0.997
valine	0.24 ± 0.02	0.981

dx.doi.org/10.1021/jp506379r | J. Phys. Chem. B 2014, 118, 10618−10623

enthalpic and entropic contributions to nonpolar solvation by GdmCl. Fitting the [GdmCl]-dependent $\Delta G_{solv,NATA}$ data at each temperature independently with eq 6 leads to excellent fits (Table 2, $R > 0.999$ in all cases).

Table 2. Temperature Dependence of NATA K_1 and K_2 Parameters

temp, K	K_1	K_2
288	0.46 ± 0.02	0.043 ± 0.004
298	0.56 ± 0.02	0.035 ± 0.006
308	0.60 ± 0.02	0.051 ± 0.006
318	0.80 ± 0.03	0.035 ± 0.009

While $K_{1,NATA}$ increases with temperature from 0.46 at 15 °C to 0.80 at 45 °C, $K_{2,NATA}$ does not demonstrate a clear trend with temperature, remaining within 20% of an average of 0.041. Plotting the data using the van't Hoff form (Figure 4A) highlights the temperature dependent increase in $K_{1,NATA}$ and the temperature insensitivity of $K_{2,NATA}$.

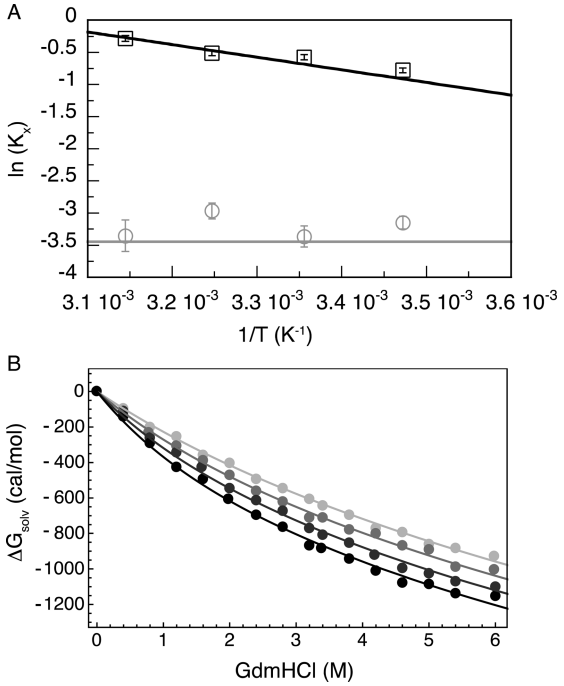


Figure 4. Temperature dependence of NATA solvation. (A) Dependence of $K_{1,NATA}$ and $K_{1,NATA}$ on temperature from individual fits at each temperature. (B) $\Delta G_{solv,\ NATA}$ as a function of [GdmCl] and temperature (black 15 °C, dark gray 25 °C, gray 35 °C, and light gray 45 °C) fit with eq 10, $\Delta G_{solv,\ NATA} = -RT \ln(1 + be^{-a/T} c + K_2 c^2)$, where $a = 1960 \pm 70$, $b = 360 \pm 80$, and $K_2 = 0.032 \pm 0.003$.

With the understanding of the effect of temperature on $K_{1,\ NATA}$ and $K_{2,\ NATA}$, we tested whether a single model could describe NATA solvation as a function of both temperature and [GdmCl]. Using eq 10, we performed a global analysis of the data at all temperatures and GdmCl concentrations simultaneously. Using three parameters to fit all 60 measurements leads to an excellent fit (Figure 4B, $R = 0.9996$). The apparent enthalpy and entropy underlying the process described by K_1 obtained from the a and b parameters of the global fit are 3.89 kcal/mol and 11.7 cal/(mol K), respectively.

DISCUSSION

The results of our analysis demonstrate that using a simple solution-theory based model to fit solvation data for the convenient experimental model compound NATA accurately captures the relevant features of nonpolar amino acid solvation by GdmCl. The nonlinear dependencies of $\Delta G_{solv,aa}$ on [GdmCl] observed for the amino acids valine, leucine, methionine, phenylalanine, tyrosine, tryptophan, and isoleucine are each well fit with just a single variable, σ_{aa}, when the NATA-derived parameters are used as constants in the model. Furthermore, those scale factors strongly correlate with residue ASA, demonstrating that NATA solvation captures the essence of solvation of nonpolar amino acid residues. This supports our long-term strategy of exploiting the broad range of NATA solvation data as a function of GdmCl concentration and temperature to understand nonpolar solvation in proteins through insights from the physically based solution theory parameters.

Application of Solution Theory to Model NATA Solvation in Guanidinium Chloride. Solution theory accurately models the nonlinear [GdmCl] dependence of nonpolar amino acid $\Delta G_{solv,aa}$ via eq 6. While the limited solvation data for each amino acid (measurements at four GdmCl concentrations)[9] are easily fit with eq 6 (data not shown), the application of the higher order model is better warranted by the comprehensive data for the model compound NATA (15 measurements and a wider concentration range). K_1 and K_2 reflect the GdmCl impact on ΔG_{solv} in terms of M^{-1} and M^{-2} respectively. Based on the parameters derived for NATA, nonpolar amino acid solvation by GdmCl at 25 °C arises from independent interactions between Gdm^+ and the amino acid surface represented by the first-order term $K_{1,NATA}$ (0.56), and slight cooperativity between the first and second binding events of Gdm^+ to the amino acid surface represented by the second-order term $K_{2,NATA}$ (0.035).

Our best model indicates that $K_{1,NATA}$ is strongly temperature-dependent, while $K_{2,NATA}$ is temperature-independent. This suggests that the primary interaction between a single Gdm^+ and NATA molecule (represented by $K_{1,NATA}$) involves a change in enthalpy, but the cooperativity of binding an additional Gdm^+ ion (represented by $K_{2,NATA}$) does not. Analysis of the temperature dependence of NATA solvation in GdmCl via eq 10 reveals a favorable entropy (+11.7 cal/(mol K)) and unfavorable enthalpy (+3.89 kcal/mol) associated with the interaction of a single Gdm^+ ion and NATA molecule. The favorable entropy is consistent with an increase in solvent disorder when Gdm^+ replaces water in binding to the hydrophobic NATA. On the other hand, when NATA replaces water in binding to Gdm^+, favorable hydrogen bonds are lost, leading to an unfavorable change in enthalpy. The lack of an enthalpy contribution to K_2 suggests that entropy governs the cooperativity of binding a second Gdm^+ ion to NATA.[17] Based on this foundation, additional future experiments and analysis can eventually deconvolute the specific energetic contributions of each process.

Insights into Nonpolar Amino Acid Solvation by Guanidinium Chloride. The NATA parameters accurately fit the limited solubility data available for nonpolar amino acids as a function of GdmCl concentration with a single additional amino acid specific parameter: ΔG_{solv}(aa, T, GdmCl) = $\sigma_{aa}\Delta G_{solv}$ (NATA, T, GdmCl). Therefore, the $K_{1,NATA}$ and $K_{2,NATA}$ parameters are able to capture the dominant physical

dx.doi.org/10.1021/jp506379r | J. Phys. Chem. B 2014, 118, 10618−10623

basis underlying nonpolar amino acid solvation, despite the fact that NATA is a neutral compound and the amino acids are zwitterionic. σ_{aa} reports on the extent of the surface being solvated in the amino acid compared to NATA and ranges from 0.05 for alanine to 1.24 for tryptophan.

The enhanced solubility observed for the nonpolar amino acids and NATA in the presence of GdmCl appears to require a minimum exposure of surface area. The correlation line describing the dependence of σ_{aa} on ASA, whether in terms of residue (eq 11, Figure 3B), free amino acid, or nonpolar side chain (data not shown), is systematically offset from the origin. A scale factor of 0 corresponds to 170 Å^2 of free amino acid ASA, 122 Å^2 of residue ASA, or 80 Å^2 of nonpolar side chain ASA. Alanine, at 140 Å^2 free amino acid ASA, 113 Å^2 residue ASA, and 67 Å^2 nonpolar side chain ASA, falls below that threshold. In fact, the ΔG_{solv} values for alanine and the smaller amino acid glycine are very slight and are less favorable at 6 M GdmCl than at 2 M.[9] Furthermore, the NATA-derived K_1 and K_2 does not adequately fit ΔG_{solv} for glycine (data not shown), and the fit for alanine (Figure 3B) leads to a very small σ_{ala} (0.049) and a relatively poor R value (0.758) compared to the larger amino acids (>0.981). This indicates that alanine and glycine are too small for Gdm^+ to exert its increasingly favorable effect on amino acid solubility with increasing concentration: a minimum side chain nonpolar surface of 80 Å^2 appears to be required.

σ_{aa} correlates most strongly with the total residue ASA (R = 0.96, Figure 3B) followed by free amino acid ASA (R = 0.92, data not shown) and side chain nonpolar ASA (R = 0.88, data not shown). While the intercepts vary as mentioned above depending on surface area type used, the values for the slopes obtained from the correlation are the same within error. We have chosen to focus on the better correlation with residue ASA (which includes the polar backbone), given that our ultimate goal is to derive information from the scale factor on the exposure of buried amino acid residues during protein unfolding in GdmCl.

As both NATA and the amino acids possess polar groups in addition to their nonpolar side chains (Figure 1), a strong correlation with total residue area is to be expected. By counting C and H atoms as nonpolar and N and O atoms as polar, NATA and tryptophan are both 85% nonpolar by atom count. Using residue ASA values for tryptophan from Miller et al.[16] this corresponds to ~70% nonpolar surface. Since ~60% of the difference in ASA between native and unfolded proteins is a nonpolar surface,[18] scaling the parameters derived from modeling the NATA solvation data should provide reasonable fits to protein unfolding data and a qualitative estimate of the degree of surface exposure.

Application to Protein Unfolding. Several methods have been very successful at providing insights from the denaturant dependence of $\Delta G_{u,D}$ into the amount of surface exposed upon unfolding. The m-value from the linear extrapolation model was also shown to correlate more strongly with the ΔASA between the native and unfolded rather than the difference in the nonpolar surface area alone.[2] Importantly, the ratio of the $m^{\ddagger}$-value for the dependence of the $\Delta G^{\ddagger}_{u,D}$ for kinetic unfolding to the transition state, to the m-value for equilibrium unfolding ($m^{\ddagger}/m_U$, known as the beta-Tanford value) has been used to obtain an estimate of the degree of unfolding in the transient transition state.[19] Although the linear extrapolation model is empirical, application of other models based in solvent-transfer theory,[10] denaturant binding,[1,20,21] and local bulk domain analysis[22] have yielded near-linear predictions of $\Delta G_{u,D}$ vs c_D. However, linearity is most robustly predicted for unfolding behavior in urea,[18,22] while [GdmCl] dependence is nonlinear in a number of experimental studies of equilibrium[4−6] and kinetic unfolding.[7,8] Furthermore, due to the narrow range of GdmCl concentrations over which unfolding measurements are experimentally accessible, nonlinearity is likely to be present in many more protein systems than have been reported.

The importance and necessity of accounting for the nonlinearity of nonpolar amino acid solubility in GdmCl has previously been recognized, specifically by researchers in the Clarke group.[11,23,24] Using, first, nonpolar amino acids[23,24] and then NATA solubility measurements[11] in GdmCl, they implemented an analysis transforming GdmCl concentration into a denaturant activity. This corrected for the nonpolar solubility effect and enabled them to fit the nonlinearity in their unfolding and folding kinetics data for the N-terminal domain of phosphoglycerate kinase from *Bacillus stearothermophilus* and domain 1 of the T-cell adhesion protein CD2. While their empirical model worked well for those examples, the physically based parameters derived from their NATA solvation data, using our model, should have more general application to other protein systems.

We propose adapting the solution theory-based model in eq 9 to take into account the inherent nonlinearity in amino acid solvation as a function of [GdmCl] by scaling the NATA-derived parameters to also fit [GdmCl]-dependent protein unfolding data:

$$\Delta G_{u,D} = \Delta G_{u,H2O} - \sigma RT \ln(1 + K_{1,NATA}c + K_{2,NATA}c^2) \quad (12)$$

$$\Delta G^{\ddagger}_{u,D} = \Delta G^{\ddagger}_{u,D} - \sigma RT \ln(1 + K_{1,NATA}c + K_{2,NATA}c^2) \quad (13)$$

This model provides a physically based method to more accurately fit protein data varying in degree of nonlinearity and to qualitatively estimate the degree of surface exposure upon unfolding from the scale factor. As most protein unfolding transitions induced in the presence of denaturant expose more than 122 Å^2 of total surface, this model could be employed to fit [GdmCl] dependent data for processes involving surface exposure by proteins to obtain an estimate of the ASA involved from eq 6. This treatment could be particularly helpful for gaining structural insight into partially unfolded intermediates and transition states from GdmCl-induced unfolding, as it provides a direct estimate of surface exposure and does not require knowledge of the equilibrium m-value for global unfolding.

Auton and Bolen have used amino acid solvation data in urea to construct a very accurate predictive model for the energetic effects of urea on protein unfolding.[25,26] In contrast to their observations with urea, GdmCl unfolding cannot be fit assuming a primarily backbone effect: the solvation of amino acids such as glycine and alanine in GdmCl is much smaller than that for the other amino acids.[9] Additional solubility measurements on polar and charged amino acids and model compounds in GdmCl will enable future work to refine our solution-theory based model for analysis of GdmCl-induced unfolding data and to develop a similarly predictive quantitative model for the energetic effects of GdmCl on protein unfolding.

Conclusions. Solvation of nonpolar amino acids and protein unfolding exhibit a nonlinear dependence on [GdmCl]. We have demonstrated that a solution theory-

based model with three parameters captures the nonlinearity observed for the solvation of the model compound NATA as a function of [GdmCl] and temperature. Applying the NATA-based parameters to the primarily nonpolar amino acids with solubility data available in GdmCl enables a single variable, σ_{aa}, which correlates with ASA, to fit their nonlinear behavior. Nonpolar area accounts for 70% of NATA's ASA and 60% of the surface exposed upon protein unfolding. Therefore, analysis of equilibrium and kinetics unfolding data for proteins in GdmCl using eqs 12 and 13 respectively, to incorporate the NATA-based parameters, will reflect the underlying physical interactions between Gdm^+ and the accessible surface, account for nonlinearity, and with eq 11 provide a direct estimate of surface area exposure during unfolding.

■ ASSOCIATED CONTENT

Supporting Information

Supplementary Tables 1 and 2 that include nonpolar amino acid residue ASA and solvation data and NATA solvation data, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sjaswal@amherst.edu. Telephone: 413-542-5153. Fax: 413-542-2735.

Present Address

[§]Paul Cohen, Vanderbilt University Medical Center, 1211 Medical Center Drive, Nashville, TN 37232, USA.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge Nese Kurt Yilmaz for careful reading of the manuscript and editorial help. S.S.J. acknowledges support from Amherst College Start-up Funds. P.C. acknowledges support from the Amherst College White Prize and Amherst College HHMI Summer Research Funding.

■ REFERENCES

(1) Pace, C. N.; Hermans, J. The Stability of Globular Protein. *Crit. Rev. Biochem. Mol. Biol.* **1975**, *3*, 1−43.

(2) Myers, J. K.; Nick Pace, C.; Martin Scholtz, J. Denaturant M-Values and Heat Capacity Changes: Relation to Changes in Accessible Surface Areas of Protein Unfolding. *Protein Sci.* **1995**, *4*, 2138−2148.

(3) Pierce, V.; Kang, M.; Aburi, M.; Weerasinghe, S.; Smith, P. E. Recent Applications of Kirkwood-Buff Theory to Biological Systems. *Cell Biochem. Biophys.* **2008**, *50*, 1−22.

(4) Yi, Q.; Baker, D. Direct Evidence for a Two-State Protein Unfolding Transition From Hydrogen-Deuterium Exchange, Mass Spectrometry, and NMR. *Protein Sci.* **1996**, *5*, 1060−1066.

(5) Matouschek, A.; Matthews, J. M.; Johnson, C. M.; Fersht, A. R. Extrapolation to Water of Kinetic and Equilibrium Data for the Unfolding of Barnase in Urea Solutions. *Protein Eng.* **1994**, *7*, 1089−1095.

(6) Johnson, C. M.; Fersht, A. R. Protein Stability as a Function of Denaturant Concentration: the Thermal Stability of Barnase in the Presence of Urea. *Biochemistry* **1995**, *34*, 6795−6804.

(7) Otzen, D. E.; Kristensen, O.; Proctor, M.; Oliveberg, M. Structural Changes in the Transition State of Protein Folding: Alternative Interpretations of Curved Chevron Plots. *Biochemistry* **1999**, *38*, 6499−6511.

(8) Jaswal, S. S.; Truhlar, S. M. E.; Dill, K. A.; Agard, D. A. Comprehensive Analysis of Protein Folding Activation Thermodynamics Reveals a Universal Behavior Violated by Kinetically Stable Proteases. *J. Mol. Biol.* **2005**, *347*, 355−366.

(9) Nozaki, Y.; Tanford, C. The Solubility of Amino Acids, Diglycine, and Triglycine in Aqueous Guanidine Hydrochloride Solutions. *J. Biol. Chem.* **1970**, *245*, 1648−1652.

(10) Alonso, D. O.; Dill, K. A. Solvent Denaturation and Stabilization of Globular Proteins. *Biochemistry* **1991**, *30*, 5974−5985.

(11) Parker, M. J.; Lorch, M.; Sessions, R. B.; Clarke, A. R. Thermodynamic Properties of Transient Intermediates and Transition States in the Folding of Two Contrasting Protein Structures. *Biochemistry* **1998**, *37*, 2538−2545.

(12) Sarker, S.; Bolen, D. *Temperature Dependence of Solubility of Amino Acids and Model Peptide in Water and Different Concentrations of Guanidine Hydrochloride*; Department of Chemistry and Biochemistry in the Graduate School of Southern Illinois University at Carbondale.

(13) Dill, K. A.; Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Biology*, 1st ed.; Taylor & Francis: New York, 2003.

(14) Schellman, J. A. Selective Binding and Solvent Denaturation. *Biopolymers* **1987**, *26*, 549−559.

(15) Schellman, J. A. The Thermodynamics of Solvent Exchange. *Biopolymers* **1994**, *34*, 1015−1026.

(16) Miller, S.; Janin, J.; Lesk, A. M.; Chothia, C. Interior and Surface of Monomeric Proteins. *J. Mol. Biol.* **1987**, *196*, 641−656.

(17) Miklavc, A. Temperature-Nearly-Independent Binding Constant in Several Biochemical Systems. *Biochem. Pharmacol.* **1996**, *51*, 723−729.

(18) Courtenay, E. S.; Capp, M. W.; Record, M. T. Thermodynamics of Interactions of Urea and Guanidinium Salts with Protein Surface: Relationship Between Solute Effects on Protein Processes and Changes in Water-Accessible Surface Area. *Protein Sci.* **2001**, *10*, 2485−2497.

(19) Fersht, A. R. *Structure and Mechanism in Protein Science*; W. H. Freeman and Company: 1999.

(20) Tanford, C. Protein Denaturation. C. Theoretical Models for the Mechanism of Denaturation. *Adv. Protein Chem.* **1970**, *24*, 1−95.

(21) Record, M. T.; Zhang, W.; Anderson, C. F. Analysis of Effects of Salts and Uncharged Solutes on Protein and Nucleic Acid Equilibria and Processes: a Practical Guide to Recognizing and Interpreting Polyelectrolyte Effects, Hofmeister Effects, and Osmotic Effects of Salts. *Adv. Protein Chem.* **1998**, *51*, 281−353.

(22) Courtenay, E. S.; Capp, M. W.; Saecker, R. M.; Record, M. T. Thermodynamic Analysis of Interactions Between Denaturants and Protein Surface Exposed on Unfolding: Interpretation of Urea and Guanidinium Chloride M-Values and Their Correlation with Changes in Accessible Surface Area (ASA) Using Preferential Interaction Coefficients and the Local-Bulk Domain Model. *Proteins* **2000**, *4* (Suppl), 72−85.

(23) Staniforth, R. A.; Burston, S. G.; Smith, C. J.; Jackson, G. S.; Badcoe, I. G.; Atkinson, T.; Holbrook, J. J.; Clarke, A. R. The Energetics and Cooperativity of Protein Folding: a Simple Experimental Analysis Based Upon the Solvation of Internal Residues. *Biochemistry* **1993**, *32*, 3842−3851.

(24) Parker, M. J.; Spencer, J.; Clarke, A. R. An Integrated Kinetic Analysis of Intermediates and Transition States in Protein Folding Reactions. *J. Mol. Biol.* **1995**, *253*, 771−786.

(25) Auton, M.; Holthauzen, L. M. F.; Bolen, D. W. Anatomy of Energetic Changes Accompanying Urea-Induced Protein Denaturation. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15317−15322.

(26) Auton, M.; Bolen, D. W. Application of the Transfer Model to Understand How Naturally Occurring Osmolytes Affect Protein Stability. *Meth. Enzymol.* **2007**, *428*, 397−418.