Residual Charge Interactions in Unfolded Staphylococcal Nuclease Can Be Explained by the Gaussian-Chain Model

Huan-Xiang Zhou
Department of Physics, Drexel University, Philadelphia, Pennsylvania 19104, USA

ABSTRACT The discrepancy of the pH dependence of the unfolding free energy for staphylococcal nuclease from what is expected from an idealized model for the unfolded state is accounted for by the recently developed Gaussian-chain model. Residual electrostatic effects in the unfolded state are attributed to nonspecific interactions dominated by charges close along the sequence. The dominance of nonspecific local interactions appears to be supported by some experimental evidence.

INTRODUCTION

The nature of the unfolded state of proteins under physiological conditions has attracted intensive interest in recent years (Bierzynski and Baldwin 1982; Dill and Shortle, 1991; Neri et al., 1992; Logan et al., 1994; Shortle, 1996; Schwalbe et al., 1997; Gillespie and Shortle 1997; Mok et al., 1999; Wrabl and Shortle, 1999; Wong et al., 2000; Yi et al., 2000; Choy and Forman-Kay, 2001). pH dependence of the unfolding free energy, ΔG_{unf}, provides a unique opportunity for gaining insight into the unfolded state. This dependence is governed by (Tanford, 1970)

$$\Delta G_{unf}(\mathrm{pH}) - \Delta G_{unf}(\mathrm{pH}_0) = (k_B T \ln 10) \int_{\mathrm{pH}_0}^{\mathrm{pH}} Q_u \, d\mathrm{pH} - (k_B T \ln 10) \int_{\mathrm{pH}_0}^{\mathrm{pH}} Q_f \, d\mathrm{pH}, \quad (1)$$

where $k_B T$ is the product of the Boltzmann constant and the absolute temperature, and Q_u (Q_f) is the total charge on the protein at a given pH in the unfolded (folded) state. The left-hand side of this equation (ΔG_{unf}) can be obtained by measuring the unfolding free energy under decreasingly denaturing conditions and then extrapolating to physiological conditions, whereas the second term on the right-hand side (Q_f) can be obtained by measuring either the proton release of the folded protein as a function of pH or the pK_a of all its ionizable groups in the folded state. The resulting total charge Q_u on the unfolded protein is a sensitive measure of residual charge–charge interactions.

In the simplest model for predicting Q_u, the unfolded state is thought as a "random coil" that is devoid of any residual interactions. Then,

$$Q_u = -\sum_i 1/(1 + 10^{\mathrm{pK}_{i,0} - \mathrm{pH}}) + N_+, \quad (2)$$

where $pK_{i,0}$ are the pK_a values of model compounds for the ionizable groups, and N_+ is the number of ionizable groups that become charged upon protonation (i.e., Arg, Lys, and the N-terminal). Although Nozaki and Tanford (1967a,b) have shown that Eq. 2 gives good predictions of Q_u measured in 6 M GdnHCl, many recent experimental studies have found that the use of Eq. 2 for Q_u under physiological conditions leads to significant discrepancies between calculated and measured ΔG_{unf} (Oliverberg et al., 1995; Swint-Kruse and Robertson, 1995; Tan et al., 1995; Kuhlman et al., 1999; Whitten and Garcia-Moreno, 2000). The simplest explanation of these discrepancies is that they indicate residual charge–charge interactions in the unfolded state.

That residual charge–charge interactions exist in the unfolded state is not surprising. According to Coulomb's law, two charged residues fully solvated in water have an interaction energy,

$$U_0 = \pm 332/\varepsilon r \ (\mathrm{kcal/mol}), \quad (3)$$

where + (−) is for like (opposite) charges, ε is the dielectric constant of water (=78.5 at room temperature), and r is the distance of the charges (in Å). Though the distance between two residues in the unfolded state is not fixed, Eq. 3 can provide an order-of-magnitude estimate if r is taken to be the mean distance sampled by the residues. Because of the polymer nature of the unfolded protein chain, residues close along the sequence will likely have shorter mean distances (Vijayakumar and Zhou, 2000). At a mean distance of 8 Å, the residual interaction energy between a pair of charges is 0.5 kcal/mol according to Eq. 3.

Electrostatic effects in the unfolded state have been treated in several models (Stigter et al. 1991; Elcock, 1999; Zhou, 2002a; Kundrotas and Karshikoff, 2002) and their importance on protein stability has now been recognized (Elcock, 1999; Pace et al., 2000; Zhou, 2002a). In our

Submitted January 31, 2002 and accepted for publication July 25 2002.

Address reprint requests to author's present address, Huan-Xiang Zhou, Institute of Molecular Biophysics and Department of Physics, Florida State University, Tallahassee, FL 32306. Tel.: 850-644-7052; Fax: 850-644-0098; E-mail: zhou@sb.fsu.edu.

0006-3495/02/12/2981/06 $2.00

Gaussian-chain model (Zhou, 2002a), the interaction energy between two ionizable groups at a distance r is given by the Debye–Hückel theory,

$$U = \pm 332 \exp(-\kappa r)/\varepsilon r, \quad (4)$$

where $\kappa = (8\pi I e^2/\varepsilon k_B T)^{1/2}$ ($= I^{1/2}/3.04$ Å^{-1} at room temperature) and I is the ionic strength. The distance between two groups is assumed to have a Gaussian distribution,

$$p(r) = 4\pi r^2 (3/2\pi d^2)^{3/2} \exp(-3r^2/2d^2), \quad (5a)$$

where d is the root-mean-square distance. This mean distance depends on l, the number of peptide bonds separating the two residues, and is given by

$$d = bl^{1/2} + s, \quad (5b)$$

where the effective bond length b was set to 7.5 Å and the shift distance s (to account for the fact that the distance of interest is between two sidechains) was set to 5 Å. The mean interaction energy then has the magnitude,

$$W_{ij} = 332(6/\pi)^{1/2}[1 - \pi^{1/2} x \exp(x^2)\mathrm{erfc}(x)]/\varepsilon d, \quad (6)$$

where $x = \kappa d/6^{1/2}$ and, erfc(x) is the complementary error function.

The residual interactions shift the pK_a of the ionizable groups in the unfolded state. As a result, the total charge Q_u on the protein is significantly different from what is predicted by Eq. 2. The Gaussian-chain model has been found to give accurate predictions for the pH dependence of ΔG_{unf} for barnase, chymotrypsin inhibitor 2, Ovomucoid third domain, and ribonucleases A and T1 (Zhou, 2002a) and the N-terminal domain of protein L9 (Zhou, 2002b) over wide pH ranges.

In this paper, we use the Gaussian-chain model to predict the total charge on staphylococcal nuclease (SNase) in the unfolded state. We show that the pH dependence of ΔG_{unf} calculated using Q_u thus predicted agrees well with the experimental results of Whitten and Garcia-Moreno (2000). This indicates that the residual electrostatic effects in the unfolded state of SNase can be attributed to nonspecific interactions dominated by charges close along the sequence.

METHODS

Prediction of Q_u

In the Gaussian-chain model (Zhou, 2002a), it is assumed that, in the absence of residual charge–charge interactions, the pK_a of the ionizable groups take "standard" model-compound values p$K_{i,0}$. The distribution of the protonation states x_i of the ionizable groups is governed by the Hamiltonian (Zhou and Vijayakumar, 1997)

$$\mathcal{H} = (k_B T \ln 10) \sum_i (\mathrm{pH} - \mathrm{pK}_{i,0}) x_i + \frac{1}{2} \sum_{i \neq j} W_{ij}(x_i - x_{i0})(x_j - x_{j0}), \quad (7)$$

where $x_i = 0$ (1) when group i is unprotonated (protonated), and x_{i0} is the protonation state when the group is charge neutral (1 for Asp, Glu, Tyr, and the C-terminal and 0 for Arg, Lys, His, and the N-terminal). The average protonation of group i at a given pH is

$$\bar{x}_i = \frac{\sum_{\{x_i\}} x_i \exp(-\mathcal{H}/k_B T)}{\sum_{\{x_i\}} \exp(-\mathcal{H}/k_B T)}, \quad (8)$$

which was evaluated by a Monte Carlo simulation. The total charge on the protein is

$$Q_u = \sum_i (\bar{x}_i - x_{i0}). \quad (9)$$

Calculation of ΔG_{unf}

The value ΔG_{unf} was calculated from Eq. 1 by numerically integrating Q_u over pH. The total charge Q_f on the folded protein, required in Eq. 1, was measured by Whitten and Garcia-Moreno (2000) potentiometrically.

Calculations were done at the same conditions of the experiments: T = 293 K, I = 100 mM, and pH from 3.5 to 9. SNase undergoes acid denaturation with a mid-pH of pH_{mid} = 3.71. Near this pH, the measured total charge under native conditions may be viewed as the equilibrium average of the folded and unfolded states,

$$Q(\mathrm{measured}) = \frac{Q_f + Q_u \exp(-\Delta G_{unf}/k_B T)}{1 + \exp(-\Delta G_{unf}/k_B T)}. \quad (10)$$

The total charge Q_f on the folded protein alone was obtained from this equation using experimental ΔG_{unf} of Whitten and Garcia-Moreno and calculated Q_u. Only a small correction near pH_{mid} was introduced.

The model compound pK_a were taken to be those used by Whitten and Garcia-Moreno to fit the proton titration curve of SNase in 6 M GdnHCl. These are: Asp, 3.9; Glu, 4.4; His, 6.6; Tyr, 10.0; Lys, 10.4; Arg, 12.0; N-terminal, 7.5; and C-terminal, 3.5. They differ very little from standard values used in our previous study (Zhou, 2002a). Calculated ΔG_{unf} changed little when these standard model compound pK_a were used.

Calculation of p$K_{i,u}$

It is convenient to describe the proton titration of the unfolded protein by pKa of the ionizable groups. These are denoted as p$K_{i,u}$ and were obtained by fitting the pH dependence of the average protonation x_i to the Hill equation,

$$\log \frac{\bar{x}_i}{1 - \bar{x}_i} = n_i(\mathrm{pK}_{i,u} - \mathrm{pH}). \quad (11)$$

Note that, regardless of the value of the Hill coefficient n_i, p$K_{i,u}$ equals the pH at which the group is 50% protonated.

Residual charge interactions in unfolded state

The electrostatic interaction energy of a charged group with the rest of the protein can be assessed by the effect of a charge-neutralization mutation on the free energy of the protein. In the unfolded state, the Gaussian-chain model predicts the interaction energy of the mth charge by (Zhou, 2002a)

$$\exp(\Delta G_u^{int}/k_B T) = \left\langle \exp\left[(x_m - x_{m0}) \sum_{i \neq m} W_{mi}(x_i - x_{i0})/k_B T \right] \right\rangle. \quad (12)$$

The average is over a Boltzmann distribution of the protonation states (cf. Eqs. 7 and 8).

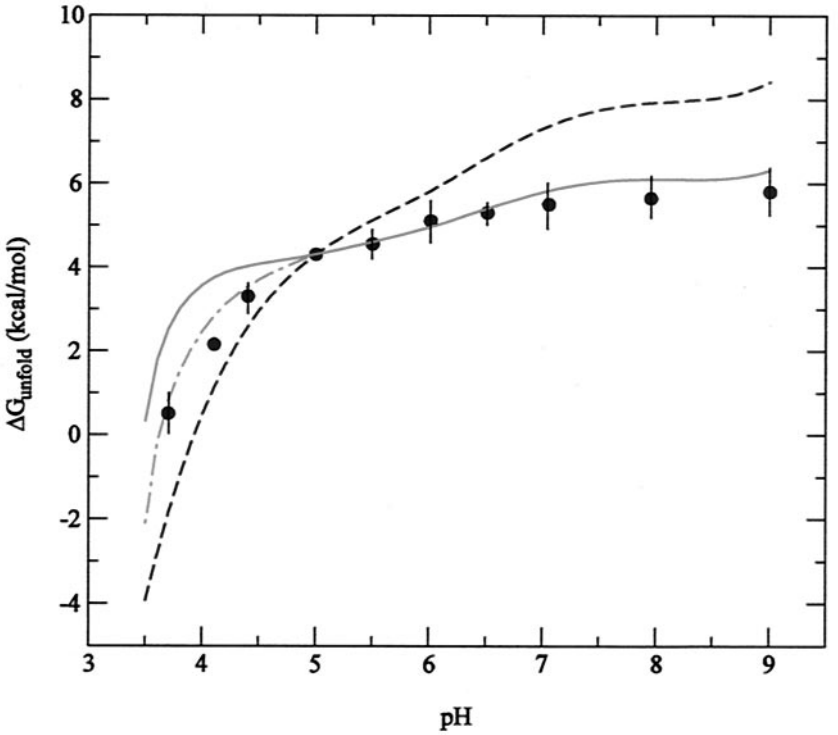

FIGURE 1 Calculated and experimental pH dependence of ΔG_{unf} for SNase at $T = 293$ K and $I = 100$ mM. *Circles with error bars*, experiment; *solid curve*, Gaussian-chain model with standard b of 7.5 Å; *dot-dashed curve*, Gaussian-chain model with b increased to 10 Å; *dashed curve*, idealized model.

RESULTS AND DISCUSSION

Calculated pH dependence of ΔG_{unf}

The unfolding free energy calculated with the Gaussian-chain model is displayed in Fig. 1. From pH 5 to 9, it agrees with the experimental results of Whitten and Garcia-Moreno (2000) to within experimental error. As noted by Whitten and Garcia-Moreno, the idealized model, with Q_u predicted by Eq. 2, significantly overestimates the increase in stability on going from pH 5 to 9. Experimentally, ΔG_{unf} increases by 1.5 kcal/mol in this pH range. The Gaussian-chain model predicts an increase of 2 kcal/mol, whereas the idealized model predicts an increase of 4.1 kcal/mol.

At lower pH, the Gaussian-chain model slightly underestimates the decrease in ΔG_{unf}. This perhaps indicates a limitation of the Gaussian-chain model with a fixed effective bond length. At such pH, the acidic groups will be neutralized by protonation and the protein chain has only positive charges. It is thus reasonable to expect that the chain will expand somewhat (leading to a larger b) (Pace et al., 1990; Dill and Shortle, 1991). This expansion will reduce the strengths of charge–charge interactions. The resulting ΔG_{unf} will move toward what is predicted by the idealized model and be in better agreement with experiment. Indeed, agreement with experiment for pH < 5 can be achieved by $b = 10$ Å. Deterioration of the Gaussian-chain model (with a fixed effective bond length) at extreme pH has been noted previously (Zhou, 2002a).

Predicted pK_a in the unfolded state

SNase has a total of 61 ionizable groups: 8 Asp, 12 Glu, 5 Arg, 23 Lys, 4 His, 7 Tyr, and the N- and C-terminals (see Table 1). The pK_a shifts of these groups from model-compound values $pK_{i,0}$ predicted by the Gaussian-chain model (with $b = 7.5$ Å) are listed in Table 1. Several points are worth noting. 1) The pK_a are down shifted for Asp and Glu and up shifted for Lys and Arg, reflecting the favorable interactions between the acidic and basic groups. 2) The average shifts are 0.31, 0.37, 0.36, and 0.15, respectively, for Asp, Glu, Arg, and Lys. The shifts for Lys are smaller because the favorable interactions with Asp and Glu are tampered by unfavorable interactions among the 23 Lys and with Arg. 3) The four His residues have down shifted pK_a (averaging 0.21), and more details are given below. 4) The seven Tyr have barely perturbed pK_a.

The pH dependence of the folding stability between pH 5 and 9 is mainly responsible by the deprotonation of the four histidines and the N-terminal. In the unfolded state, the pK_a of His-8, His-46, His-121, and His-124 are 6.29, 6.30, 6.54, and 6.44, respectively, as predicted by the Gaussian-chain model with a model compound pK_a of 6.6 at $T = 293$ K and $I = 100$ mM. The larger $pK_{i,u}$ shifts of His-8 and His-46 can be simply explained by the presence of other charged residues in nearby positions along the sequence. Each of these residues is bounded by Lys on both sides, with the potential repulsion between a protonated His and the positive charges on Lys serving to depress the pK_a. In contrast, His-121 and

TABLE 1 Predicted pK_a shifts, Hill coefficients, and residual interaction energies and experimental results for $\Delta\Delta G_{unf}$ due to charge nutralization

Group	$pK_{i,u}-pK_{i,0}$	n_i	ΔG_u^{int}	$\Delta\Delta G_{unf}$*
N-terminal	−0.14	0.98		
Lys5	−0.03	0.89	+0.33	+0.3
Lys6	−0.02	0.87	+0.34	+0.3
His8	−0.31	1.00	+0.01	−0.4
Lys9	+0.13	0.88	+0.11	**−1.4**
Glu10	−0.50	0.97	−0.46	−1.3
Lys16	+0.14	0.89	+0.04	−0.2
Asp19	−0.30	0.93	−0.20	−0.1
Asp21	−0.31	0.94	−0.20	**+0.7**
Lys24	+0.16	0.88	+0.01	−0.2
Tyr27	−0.13	0.84		
Lys28	+0.15	0.87	+0.14	−0.7
Arg35	+0.25	0.96	+0.01	**−1.4**
Asp40	−0.35	0.93	−0.21	**−0.2**
Glu43	−0.39	0.91	−0.29	+0.3
Lys45	+0.16	0.88	+0.12	+0.3
His46	−0.30	0.96	+0.01	−0.5
Lys48	+0.07	0.82	+0.29	+0.1
Lys49	+0.09	0.81	+0.26	−0.3
Glu52	−0.49	0.90	−0.44	−0.1
Lys53	+0.34	0.81	0	−0.4
Tyr54	+0.01	0.82		
Glu57	−0.32	0.92	−0.26	−0.2
Lys63	+0.10	0.82	+0.21	**−0.5**
Lys64	+0.11	0.83	+0.21	+0.1
Glu67	−0.40	0.93	−0.34	**−1.0**
Lys70	+0.17	0.83	+0.13	−0.1
Lys71	+0.21	0.83	+0.01	−0.4
Glu73	−0.34	0.88	−0.22	−1.4
Glu75	−0.27	0.86	−0.13	**−2.2**
Asp77	−0.45	0.89	−0.27	−3.1
Lys78	+0.30	0.87	−0.01	−0.6
Arg81	+0.45	0.89	+0.01	−1.1
Asp83	−0.52	0.88	−0.45	**−3.8**
Lys84	+0.29	0.82	+0.04	+0.2
Tyr85	−0.08	0.84		
Arg87	+0.46	0.92	+0.12	**−0.9**
Tyr91	+0.06	0.86		
Tyr93	+0.12	0.84		
Asp95	−0.33	0.91	−0.25	−3.3
Lys97	+0.25	0.87	−0.03	−0.1
Glu101	−0.28	0.94	−0.24	−1.9
Arg105	+0.31	0.92	+0.01	**−1.4**
Lys110	+0.12	0.87	+0.14	−1.3
Tyr113	−0.02	0.84		
Tyr115	−0.06	0.84		
Lys116	+0.22	0.83	+0.12	+0.7
His121	−0.06	0.98	+0.01	−3.1
Glu122	−0.51	0.93	−0.31	**−0.4**
His124	−0.16	0.96	+0.05	+0.4
Arg126	+0.32	0.90	+0.18	**−1.7**
Lys127	+0.08	0.90	+0.15	+0.2
Glu129	−0.43	0.92	−0.34	−2.4
Lys133	+0.14	0.85	+0.12	−1.4
Lys134	+0.19	0.84	+0.07	+0.1
Glu135	−0.49	0.91	−0.45	**−0.7**
Lys136	+0.25	0.88	−0.06	−0.9
Asp142	−0.11	0.88	+0.17	−0.3
Asp143	+0.02	0.86	+0.20	−0.1
Glu146	−0.07	0.92	+0.15	−0.2
C-terminal	−0.11	0.94		

*Taken from Meeker et al. (1996) for charge mutations to Ala. Results for residues forming salt bridges are in bold.

H124 border a positive charge on one side and a negative charge (Glu-122) on the other, leading to cancellation of the potential repulsive and attractive interactions with the nearby charges. The small downward shifts of $pK_{i,u}$ for His-121 and His-124 can then be attributed to more distant charges (28 Lys/Arg versus 20 Glu/Asp).

In the folded state, Alexandrescu et al. (1988) determined the pK_a of the four histidines at $T = 298$ K and $I = 300$ mM. Their results are 6.8, 5.8, 5.5, 6.1, respectively, with His-8 showing a small upward shift whereas the other three His residues show substantial downward shifts. That the pK_a of the last three histidines in the unfolded state are also calculated to be downward shifted explains the much more moderate increase, in agreement with experiment, in the predicted ΔG_{unf} between pH 5 and 9 relative to the idealized model.

Charge–charge interactions in the unfolded state

The interaction energies of the 52 charged residues at pH 7 with the rest of the protein are listed in Table 1. Echoing the observations on the shifts in $pK_{i,u}$ noted above, the ΔG_u^{int} results have the following salient features. 1) Overall the interactions are favorable for Asp and Glu and unfavorable for Lys and Arg, due to the excess net positive charge on the protein. 2) The average interaction energies are −0.16, −0.27, 0.07, and 0.12 kcal/mol, respectively, for Asp, Glu, Arg, and Lys. 3) The four histidines have negligible interaction energies (averaging just 0.02 kcal/mol) with the rest of the proteins, reflecting the fact that, at pH 7, all the His residues are mostly charge-neutral via deprotonation.

Meeker et al. (1996) have measured the effects of neutralizing all the 52 charged residues at pH 7 by mutations to Ala. Their results for $\Delta\Delta G_{unf}$ are also listed in Table 1. In a search to find factors that are important in determining the variability of these $\Delta\Delta G_{unf}$ results, Meeker et al. investigated potential correlations with a large number of parameters describing residue environment. It was found that the only parameters correlated with $\Delta\Delta G_{unf}$ were ones measuring the local packing density (i.e., number of C_α atoms within a sphere of 10-Å radius, fraction of sidechain burial, and temperature factor) in the folded state. Even for these parameters, the correlations were only moderate, with correlation coefficients ranging from 0.35 to 0.55. We actually found a modest correlation between $\Delta\Delta G_{unf}$ and ΔG_u^{int}. With 13 residues that form salt bridges excluded, the correlation coefficient is 0.47 (inclusion of these 13 residues reduced the correlation coefficient to 0.39).

Connections to other experimental studies

Although details about charge–charge interactions in the unfolded state can be calculated with the Gaussian-chain model, these are difficult to assess experimentally. There are two experimental studies that shed light on charge–charge interactions in the unfolded state. Flanagan et al.

(1992) measured the chemical shift dispersion in the NMR spectrum of Δ137–149, an SNase fragment serving as a model for the unfolded state. They found little chemical shift dispersion for the four histidines in Δ137–149, indicating "it is unlikely that the four histidines at positions 8, 46, 121, and 124 are involved in residual stable structure." Given this finding, the basic assumption of the Gaussian-chain model that residual electrostatic effects are dominated by local charges may not be unreasonable for the four histidines of SNase.

Sinclair and Shortle (1999) analyzed extensive CD and NMR data on Δ131Δ (another model for the unfolded state of SNase) and a large number of mutants. Their basic conclusion was that the overall topology of the unfolded state ensemble is "determined by many coupled local interactions rather than a few highly specific long-range interactions." This conclusion is heartening, because the Gaussian-chain model is premised on dominance of local interactions. Directly relevant to the present study are the nine charge mutations investigated by Sinclair and Shortle: E52V, K71A, E75A, K84A, R87A, R105A, K110A, H121A, and K133A. Sinclair and Shortle found that these charge neutralizations did not affect the structural ordering of the SNase fragment, suggesting that these charges are not involved in specific interactions in the unfolded state. Three of the charges, E75, R87, and R105, form salt bridges in the folded state. In light of the Sinclair–Shortle work, these charges are unlikely to participate in native-like structures.

Shortle and Akerman (2001) recently designed an experiment to study the persistence of long-range structural ordering of Δ131Δ under various concentrations of urea. The protein fragment was confined in the pores of polyacrylamide gels. The main finding of this work was that long-range ordering, as evidenced by residual dipolar coupling, persists even in 8 M urea. However, conformation sampling in a confined environment will likely be different from that in bulk solution, because "more expanded conformations will be disfavored relative to compact ones" (Shortle and Akerman, 2001). Our recent theoretical calculations indicate that the excluded-volume effect of confinement can significantly shift the equilibrium from the unfolded state toward the folded state (Zhou and Dill, 2001). Indeed, Klimov et al. (2002) have explicitly suggested that "the presence of significant native interactions in Δ131Δ even at strongly denaturing conditions is caused by the conformational restrictions introduced by the surrounding gel."

Evaluation of the Gaussian-chain model

The Gaussian-chain model has now been found to give accurate predictions for the pH dependence of folding stability for barnase, chymotrypsin inhibitor 2, Ovomucoid third domain, ribonucleases A and T1 (Zhou, 2002a), the N-terminal domain of protein L9 (Zhou, 2002b), and, here, for SNase over wide pH ranges. In judging the Gaussian-chain model, its simplicity is worth noting.

The model emphasizes the interactions between charged residues close along the sequence. For unfolded SNase, there appears to be some experimental evidence for the dominance of local interactions. The fact that no specific tertiary interactions (of the salt-bridge type) are included may have contributed to the success of the Gaussian-chain model, because such interactions sometimes could lead to excessive pK_a shifts in the unfolded state (Elcock 1999; Zhou 2002a). In contrast, the unfolded state of some proteins have been characterized as being "compact." (Neri et al., 1992; Mok et al., 1999; Choy et al., 2001). The Gaussian-chain model cannot adequately account for possible long-range interactions implicated in these cases. The model of Elcock (1999) accounts for native-like long-range interactions and is probably more realistic under these circumstances.

In summary, we have shown that the discrepancy of the pH dependence of ΔG_{unf} for SNase from the idealized model as observed by Whitten and Garcia-Moreno (2000) can be accurately accounted for by the Gaussian-chain model. The residual electrostatic effects in the unfolded state have been attributed to nonspecific interactions dominated by charges close along the sequence. We expect that the Gaussian-chain model will continue to help characterizing the unfolded state.

This work was supported in part by National Institutes of Health grant GM58187.

REFERENCES

Alexandrescu, A. T., D. A. Mills, E. L. Ulrich, M. Chinami, and J. L. Markley. 1988. NMR assignments of the four histidines of staphylococcal nuclease in native and denatured states. *Biochemistry.* 27: 2158–2165.

Bierzynski, A., and R. L. Baldwin. 1982. Local secondary structure in ribonuclease A denatured by guanidine. HCl near 1 degree C. *J. Mol. Biol.* 162:173–186.

Choy, W.-Y., and J. D. Forman-Kay. 2001. Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J. Mol. Biol.* 308:1011–1032.

Dill, K. A., and D. Shortle. 1991. Denatured states of proteins. *Annu. Rev. Biochem.* 60:795–825.

Elcock, A. H. 1999. Realistic modeling of the denatured states of proteins allows accurate calculations of the pH dependence of protein stability. *J. Mol. Biol.* 294:1051–1062.

Flanagan, J. M., M. Kataoka, D. Shortle, and D. M. Engelman. 1992. Truncated staphylococcal nuclease is compact but disordered. *Proc. Natl. Acad. Sci. U.S.A.* 89:748–752.

Gillespie, J. R., and D. Shortle. 1997. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* 268:170–184.

Klimov, D. K., D. Newfield, and D. Thirumalai. 2002. Simulations of β-hairpin formation in spherical pores using distributed computing. *Proc. Natl. Acad. Sci. U.S.A.* 99:8019–8024.

Kuhlman, B., D. L. Luisi, P. Young, and D. P. Raleigh. 1999. pKa values and the pH dependent stability of the N-terminal domain of L9 as probes

of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions. *Biochemistry.* 38:4896–4903.

Kundrotas, P. J., and A. Karshikoff. 2002. Model for calculation of electrostatic interactions in unfolded proteins. *Phys. Rev. E.* 65:011901.

Logan, T. M., Y. Theriault, and S. W. Fesik. 1994. Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. *J. Mol. Biol.* 236:637–648.

Meeker, A. K., B. Garcia-Moreno, E., and D. Shortle. 1996. Contributions of the ionizable amino acids to the stability of staphylococcal nuclease. *Biochemistry.* 35:6443–6449.

Mok, Y.-K., C. M. Kay, L. E. Kay, and J. Forman-Kay. 1999. NOE data demonstrating a compact unfolded state from an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* 289:619–638.

Neri, D., M. Billeter, G. Wider, and K. Wuthrich. 1992. NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science.* 257:1559–1563.

Nozaki, Y., and C. Tanford. 1967a. Acid-base titrations in concentrated guanidine hydrochloride. Dissociation constants of the guanidinium ion and of some amino acids. *J. Am. Chem. Soc.* 89:736–742.

Nozaki, Y., and C. Tanford. 1967b. Proteins as random coils. II. Hydrogen ion titration curves of ribonuclease in 6 M. guanidine hydrochloride. *J. Am. Chem. Soc.* 89:742–749.

Oliverberg, M., V. L. Arcus, and A. R. Fersht. 1995. pKa values of carboxyl groups in the native and denatured states of barnase: the pKa values of the denatured state are on average 0.4 units lower than those of model compounds. *Biochemistry.* 34:9424–9433.

Pace, C. N., R. W. Alston, and K. L. Shaw. 2000. Charge–charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci.* 9:1395–1398.

Pace, C. N., D. V. Laurents, and J. A. Thomson. 1990. pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry.* 29:2564–2573.

Schwalbe, H., K. M. Fiebig, M. Buck, J. A. Jones, S. B. Grimshaw, A. Spencer, S. J. Glaser, L. J. Smith, and C. M. Dobson. 1997. Structural and dynamical properties of a denatured proteins. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry.* 36:8977–8991.

Shortle, D. 1996. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J.* 10:27–34.

Shortle, D., and M. S. Akerman. 2001. Persistence of native-like topology in a denatured protein in 8 M urea. *Science.* 293:487–489.

Sinclair, J. F., and D. Shortle. 1999. Analysis of long-range interactions in a model denatured state of staphylococcal nuclease based on correlated changes in backbone dynamics. *Protein Sci.* 8:991–1000.

Stigter, D., D. O. Alonso, and K. A. Dill. 1991. Protein stability: electrostatics and compact denatured states. *Proc. Natl. Acad. Sci. U.S.A.* 88:4176–4180.

Swint-Kruse, L., and A. D. Robertson. 1995. Hydrogen bonds and the pH dependence of Ovomucoid third domain stability. *Biochemistry.* 34: 4724–4732.

Tan, Y. J., M. Oliverberg, B. Davis, and A. R. Fersht. 1995. Pertubed pKa-values in the denatured states of proteins. *J. Mol. Biol.* 254: 980–992.

Tanford, C. 1970. Protein denaturation. Part C. *Adv. Protein Chem.* 24: 1–95.

Vijayakumar, M., and H.-X. Zhou. 2000. Prediction of residue–residue pair frequencies in proteins. *J. Phys. Chem. B.* 104:9755–9764.

Whitten, S. T., and B. Garcia-Moreno. 2000. pH dependence of stability of staphylococcal nuclease: evidence of substantial electrostatic interactions in the denatured state. *Biochemistry.* 39:14292–14304.

Wong, K.-B., J. Clarke, C. J. Bond, J. L. Neira, S. M. V. Freund, A. R. Fersht, and V. Daggett. 2000. Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding. *J. Mol. Biol.* 296:1257–1282.

Wrabl, J., and D. Shortle. 1999. A model of the changes in denatured state structure underlying *m* value effects in staphylococcal nuclease. *Nature Struct. Biol.* 6:876–883.

Yi, Q., M. L. Scalley-Kim, E. J. Aim, and D. Baker. 2000. NMR characterization of residual structure in the denatured state of protein L. *J. Mol. Biol.* 299:1341–1351.

Zhou, H.-X. 2002a. A Gaussian-chain model for treating residual charge–charge interactions in the unfolded state of proteins. *Proc. Natl. Acad. Sci. U.S.A.* 99:3569–3574.

Zhou, H.-X. 2002b. Residual electrostatic effects in the unfolded state of the N-terminal domain of L9 can be attributed to nonspecific nonlocal charge–charge interactions. *Biochemistry.* 41:6538–6533.

Zhou, H.-X., and K. A. Dill. 2001. Stabilization of proteins in confined spaces. *Biochemistry.* 40:11289–11293.

Zhou, H.-X., and M. Vijayakumar. 1997. Modeling of conformational fluctuations in pKa predictions. *J. Mol. Biol.* 267:1002–1011.