Thermal versus Guanidine-Induced Unfolding of Ubiquitin. An Analysis in Terms of the Contributions from Charge−Charge Interactions to Protein Stability[†]

Beatriz Ibarra-Molero,[‡,§] Vakhtang V. Loladze,[||,⊥] George I. Makhatadze,*[,||,⊥] and Jose M. Sanchez-Ruiz*[,§]

Facultad de Ciencias, Departamento de Quimica Fisica, Universidad de Granada, 18071-Granada, Spain, and Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061

Received March 12, 1999; Revised Manuscript Received April 26, 1999

ABSTRACT: We have characterized the guanidine-induced unfolding of both yeast and bovine ubiquitin at 25 °C and in the acidic pH range on the basis of fluorescence and circular dichroism measurements. Unfolding Gibbs energy changes calculated by linear extrapolation from high guanidine unfolding data are found to depend very weakly on pH. A simple explanation for this result involves the two following assumptions: (1) charged atoms of ionizable groups are exposed to the solvent in native ubiquitin (as supported by accessible surface area calculations), and Gibbs energy contributions associated with charge desolvation upon folding (a source of p*K* shifts) are small; (2) charge−charge interactions (another source of p*K* shifts upon folding) are screened out in concentrated guanidinium chloride solutions. We have also characterized the thermal unfolding of both proteins using differential scanning calorimetry. Unfolding Gibbs energy changes calculated from the calorimetric data do depend strongly on pH, a result that we attribute to the pH dependence of charge−charge interactions (not eliminated in the absence of guanidine). In fact, we find good agreement between the difference between the two series of experimental unfolding Gibbs energy changes (determined from high guanidine unfolding data by linear extrapolation and from thermal denaturation data in the absence of guanidine) and the theoretical estimates of the contribution from charge−charge interactions to the Gibbs energy change for ubiquitin unfolding obtained by using the solvent-accessibility-corrected Tanford−Kirkwood model, together with the Bashford−Karplus (reduced-set-of-sites) approximation. This contribution is found to be stabilizing at neutral pH, because most charged groups on the native protein interact mainly with groups of the opposite charge, a fact that, together with the absence of large charge-desolvation contributions, may explain the high stability of ubiquitin at neutral pH. In general, our analysis suggests the possibility of enhancing protein thermal stability by adequately redesigning the distribution of solvent-exposed, charged residues on the native protein surface.

The thermodynamic stability of proteins is commonly described by the unfolding Gibbs energy versus temperature profile, the so-called protein stability curve (*1*, *2*). It is a well-known fact that Gibbs energy changes of unfolding are small numbers, resulting from the cancellation of large contributions arising from the major "forces" that drive or oppose folding (hydrophobic effect, hydrogen bonding, van der Waals interactions, conformational entropy; see refs *3*−*8*). As an obvious consequence of this, even minor contributions to the unfolding Gibbs energy are important regarding protein stability. Thus, for instance, electrostatic contributions to unfolding Gibbs energies due to the interactions between the charges on ionizable groups are comparable to the total unfolding Gibbs energy (*9*, *10*); therefore, these electrostatic interactions might play a key role in determining operational thermal stability (as measured, for instance, by the denaturation temperature under given conditions).

Several features make ubiquitin an excellent model protein to investigate the contributions to protein stability that arise from charge−charge interactions: (1) It is a small protein, consisting of 76 amino acid residues with a molecular mass of 8433 Da and a three-dimensional structure known to high resolution from X-ray crystallography and NMR[1] (*11*, *12*). (2) Thermal unfolding of ubiquitin is reversible and conforms closely to the two-state, equilibrium model (*13*−*15*); therefore, the thermal stability of this protein is described by a

[†] This work was made possible by the support of a NATO Collaborative Research Grant (Grant CRG-970228 to G.I.M. and J.M.S.-R.). Research in individual labs was supported by the Robert A. Welch Foundation (Grant D-1403 to G.I.M.), the National Institutes of Health (Grant GM54537 to G.I.M.), the Spanish Ministry of Education and Culture (Grant PB96-1439 from the DGES to J.M.S.-R.), and the "FUNDACION RAMON ARECES" (Grant to J.M.S.-R.).

* To whom correspondence should be addressed. Fax: 34-958-272879 (J.M.S.-R.), 717-5317022 (G.I.M.). E-Mail: sanchezr@goliat.ugr.es (J.M.S.-R.).

[‡] B.I.-M. was a recipient of a predoctoral fellowship from the Spanish Ministry of Education and Culture during the period 1995−1998. Present address: Department of Chemistry, Penn State University, University Park, PA 16802.

[§] Universidad de Granada.

[||] Address after July 1, 1999: Department of Biochemistry and Molecular Biology, Penn State College of Medicine, Hershey, PA 17033-0850.

[⊥] Texas Tech University.

[1] Abbreviations: ASA, accessible surface area; CD, circular dichroism; DSC, differential scanning calorimetry; LEM, linear extrapolation method; NMR, nuclear magnetic resonance; TK−BK, the Tanford−Kirkwood model, together with the Bashford−Karplus approximation.

10.1021/bi9905819 CCC: $18.00 © 1999 American Chemical Society
Published on Web 06/03/1999

thermodynamic stability curve, which can be derived from experimental thermal unfolding data by using a straightforward two-state analysis. (3) Accessible surface area calculations show that the charged atoms of ionizable groups are highly exposed to the solvent in native ubiquitin [in fact, all ionizable groups in ubiquitin belong to class III in the classification of Rashin and Honig (*16*)]; this suggests that the Gibbs energy contributions associated with the change of the environment of charges upon folding ("desolvation" contributions: ref *9*) can be neglected, at least as a first approximation, thus allowing us to focus on the charge−charge interactions, a point of view which will be supported to some extent by the experimental results reported in this work. (4) Ubiquitin is very stable at neutral pH, the denaturation temperature being around or higher than 100 °C (*14*); however, the thermal stability of ubiquitin depends strongly on pH, which suggests that electrostatic factors may play a prominent role in the high stability observed at neutral pH.

In this work we will investigate the electrostatic contributions to ubiquitin thermal stability using experimental as well as theoretical approaches. On the experimental side, we will explore our previous suggestion (*10*) and estimate contributions due to charge−charge interactions from the comparison between the unfolding Gibbs energy values derived from guanidine-induced denaturation studies by using the linear extrapolation method (LEM)[1] and those obtained from the analysis of differential scanning calorimetry (DSC) thermograms corresponding to the thermal denaturation of the protein. These experimental studies will be carried out as a function of pH with two forms of ubiquitin that differ in three nonionizable amino acids and which have essentially the same structure: bovine and yeast ubiquitin (*17*). Guanidine-induced denaturation experiments will be based on circular dichroism and fluorescence measurements. On the theoretical side, we will estimate the energies associated with charge−charge interactions on the basis of the solvent-accessibility-corrected, Tanford−Kirkwood model (*18*, *19*), together with the Bashford−Karplus ("reduced-set-of-sites") approximation (*20*).

MATERIALS AND METHODS

Ubiquitin from bovine blood was purchased from Sigma, and guanidinium chloride was ultrapure grade from Pierce. Both were used without further purification. Yeast ubiquitin was overexpressed in *E. coli* and purified as described (*14*). Deionized water was used throughout.

Aqueous stock solutions of ubiquitin were prepared by exhaustive dialysis at 4 °C against 10 mM sodium acetate buffer (for pHs within the range 3.5−4.5) or 10 mM glycine (for pHs within the range 1.5−3.5). Protein concentrations were determined spectrophotometrically using an extinction coefficient of 0.149 mg^{-1}·mL·cm^{-1} calculated according to Gill and von Hippel (*21*). Stock solutions of 8 M guanidinium chloride in acetate or glycine buffers were prepared by mixing two guanidine solutions in the acidic and basic forms of the buffer used, so that the desired pH was obtained. Ubiquitin solutions at different guanidine concentrations were prepared by mixing appropriate volumes of the dialysis buffer, and stock solutions of the protein and 8 M guanidine. Guanidine concentrations were determined from refraction index measurements (*22*) using an Atago R 5000 hand refractometer.

Differential scanning calorimetry experiments were performed using a VP-DSC microcalorimeter from MicroCal or a DS-93 calorimeter (John Hopkins University). The samples were degassed at room temperature prior to calorimetric experiments. Calorimetric cells (operating volume ∼0.5 mL) were kept under an excess pressure of 30 psi to prevent degassing during the scan. The scan rate was always 1.5 K/min, and, in most cases, the protein concentrations in the solutions used in the calorimetric experiments were in the range 1−2 mg/mL. Further experimental details regarding the DSC experiments and the calorimetric data analysis can be found elsewhere (*15*). Far-UV circular dichroism (CD) spectra of ubiquitin from 260 to 200 nm were collected at 25 °C on a Jasco J-20 spectropolarimeter using rectangular quartz cells with 1 mm path length as described (*23*). Protein concentrations used in CD experiments were on the order of 0.1 mg/mL. The fluorescence of the protein was measured by using Perkin-Elmer LS-5 or LS50B instruments, with excitation and emission wavelengths of 279 and 310 nm, respectively. Protein concentrations used in fluorescence measurements were about 0.04 mg/mL.

NMR experiments were carried out using a 500 MHz VARIAN UNITY plus instrument. Protein was dissolved in 10 mM phosphate buffer in D_2O, pH 8.4, and, subsequently, the pH was adjusted by addition of a solution of DCl in D_2O (the pH was calculated by adding 0.4 to the pH-meter reading). The final protein concentration was 800 μM. TMS in D_2O was used as standard for chemical shifts. Monodimensional ^{1}H NMR spectra were obtained at 25 °C, using the sequence PRESAT, with a spectral width of 8000 Hz and a pulse angle of 90°. Spectra were obtained at several pH values within the range 2−8, and the chemical shift of the C_2−H in His-68 was assigned on the basis of previous work (*24*, *25*). The chemical shift versus pH profile could be analyzed in terms of an ideal titration curve, and from this analysis, a pK value of 5.92 ± 0.02 was obtained for His-68.

Analyses of experimental data were carried out using programs written in the MLAB environment (Civilized Software Inc.). Accessible surface area (ASA) calculations were performed using the Shrake−Rupley algorithm (*26*) with a radius of 1.4 Å for the solvent probe and the Chothia set for the protein atoms (*27*).

Electrostatic calculations based on the Tanford−Kirkwood model (*18*, *19*), together with the Bashford−Karplus approximation for fractional protonation calculation (*20*), were carried out using programs written by us. The calculation procedure (which will be subsequently referred to simply as the "TK-BK procedure") is briefly described below.

The interaction energies between unit positive charges placed in the protonation sites of ionizable groups i and j were calculated as

$$E_{ij} = \epsilon^2\left(\frac{A_{ij} - B_{ij}}{2b} - \frac{C_{ij}}{2a}\right)(1 - \mathrm{SA}_{ij}) \qquad (1)$$

where ϵ is the unit charge, A_{ij}, B_{ij} and C_{ij} (defined by Tanford and Kirkwood: ref *18*) are functions of the positions of the charges, the dielectric constants of the protein and solvent, and (in the case of C_{ij}) the ionic strength, and b stands for the radius of the sphere representing the protein, but solvent ions cannot penetrate a larger sphere of radius a. In our

calculations, we have taken $a = b + 1.4$ (1.4 Å being the mean of the radii of the cation and anion for a typical salt: NaCl), and b was derived from the volume of the protein (calculated from the molecular mass using a typical value for the specific volume: 0.72 mL/g). Finally, SA_{ij} is the mean of the accessibilities to solvent for groups i and j. The accessibility of a group was calculated as the ratio between the side-chain ASA values of the group in the native protein and in a Gly-X-Gly tripeptide in which the conformation of the side chain was the same as in the native protein. In most calculations, we employed 78.5 and 4 as values for the dielectric constants of solvent and the protein, respectively; however, calculations with other values for the protein dielectric constant were also performed and will be described further below. In all calculations, charges were assumed to be placed on the protein surface.

To calculate the fractional protonations of ionizable groups and other quantities of interest, we have followed the general approach described by Bashford and Karplus (*20*). This approach is briefly outlined here, since some of the equations and definitions involved will be required for subsequent discussions in this paper, as well as for the theoretical analysis given in the Appendix. For a protein with n ionizable groups, there are 2^n protonation states, which are represented by the 2^n values of the protonation vector χ (a vector whose elements, x_i, are 0 or 1, depending on the protonation state of the corresponding groups). The Gibbs energy of a given protonation state in the native protein is

$$\Delta G_N(\chi) = -RT\,(\ln 10)\sum_{i=1}^{n}(q_i + x_i)\mathrm{p}K_{\mathrm{int},i} + \frac{1}{2}\sum_{i,j=1}^{n}E_{ij}(q_i + x_i)(q_j + x_j) \quad (2)$$

where q_i is the charge of group i in the unprotonated state and $\mathrm{p}K_{\mathrm{int},i}$ (the intrinsic pK) is the pK of group i if all other groups had zero charge. Equation 2 seems to differ somewhat from eq 3 in the work of Bashford and Karplus (*20*). The difference, however, is just a matter of the reference level chosen; thus, according to our equation, $\Delta G_N(\chi) = 0$ when the charges of all ionizable groups are set to zero (when $q_i + x_i = 0$ for all i). That is, we are taking the protein in the neutral state (with all ionizable groups in their uncharged form) as the reference state for Gibbs energy calculations.

The fraction of native protein that is present as a given protonation state is given by

$$\rho_N(\chi) = \frac{1}{Z_N}\exp\left[-\frac{\Delta G_N(\chi)}{RT} - \nu(\chi)(\ln 10)\mathrm{pH}\right] \quad (3)$$

with $\nu(\chi)$ the number of protonated ionizable groups in that protonation state and Z_N the protonation partition function of the native protein:

$$Z_N = \sum_{\chi}\exp\left[-\frac{\Delta G_N(\chi)}{RT} - \nu(\chi)(\ln 10)\mathrm{pH}\right] \quad (4)$$

where the sum is over the 2^n values of the protonation vector χ.

The fractional protonation on group i in the native protein (θ_i), the charge−charge interaction energy between two ionizable groups ($\langle W_{ij}\rangle$), and the total charge−charge interaction energy in the native protein ($\langle W_{q-q}\rangle$) are averages over all protonation states. For instance, the latter quantity is given by

$$\langle W_{q-q}\rangle = \sum_{\chi}\left[\sum_{i,j=1}^{n}\frac{1}{2}E_{ij}(q_i + x_i)(q_j + x_j)\right]\rho_N(\chi) = \frac{1}{2}\sum_{i,j=1}^{n}\langle W_{ij}\rangle \quad (5)$$

Equations 3 and 4 refer to the native protein. It will be assumed in this work that ionizable groups in native ubiquitin are exposed to the solvent and that the charge−charge interaction is the main factor responsible for the different behavior of ionizable groups in the native and unfolded proteins (a justification for this assumption is given further below). Accordingly, we assume that the intrinsic pK values are equal to the model compound values for both the native and unfolded protein, and write for the unfolded protein:

$$\Delta G_U(\chi) = -RT\,(\ln 10)\sum_{i=1}^{n}(q_i + x_i)\mathrm{p}K_{\mathrm{int},i} \quad (6)$$

which is similar to eq 2, the only difference being that the charge−charge interaction term is not included here. [The values we have used for the model compound pK values are the same as those employed by the program TITRA (*28*).] Then, the fraction of unfolded protein present as a given protonation state [$\rho_U(\chi)$] and the protonation partition function of the unfolded protein (Z_U) can be calculated as indicated by eqs 3 and 4, but using ΔG_U (eq 6) instead of ΔG_N (eq 2). Finally, since the neutral forms of the native and unfolded protein are taken as reference levels for the partition functions Z_N and Z_U (respectively), the contribution from charge−charge interactions to the unfolding Gibbs energy is (*9*)

$$\Delta G_{q-q} = -RT\ln\left(\frac{Z_U}{Z_N}\right) \quad (7)$$

The application of the above equations involves summations over 2^n protonation states, a task that may become impossible, even for moderate values of n. We have, therefore, used the Bashford−Karplus, "reduced-set-of-sites" approximation (*20*), with upper and lower limits for the fractional protonations of 0.999 and 0.001, respectively. Since our calculations were carried out for pH <7, our choice of limits simply implied that lysines, arginines, and Tyr-59 were always fixed in the protonated forms and that most calculations took into account all values of a protonation subvector representing all carboxylic acid groups and His-68.

RESULTS AND DISCUSSION

Guanidine-Induced Denaturation of Ubiquitin Studied by Fluorescence Measurements. Figure 1 shows profiles of fluorescence intensity versus denaturant concentration for the guanidine-induced unfolding of bovine ubiquitin at 25 °C and at several pHs in the range 1.7−4.3. All profiles appear to be biphasic, showing a major transition centered around 4 M guanidine and what would seem to be a minor transition at very low guanidine concentrations. A detailed analysis

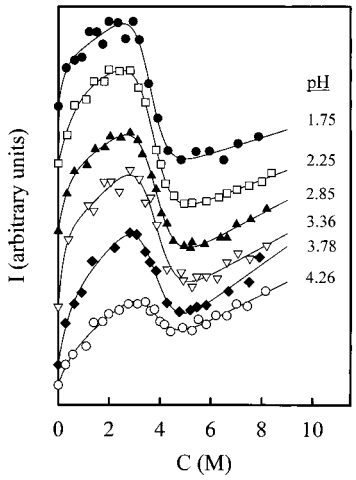

FIGURE 1: Profiles of fluorescence intensity versus denaturant concentration for the guanidine-induced unfolding of bovine ubiquitin at 25 °C and at different pH values. The continuous lines represent the best nonlinear, least-squares fits of eq 9 (together with eqs 10 and 11) to the experimental data. The profiles have been displaced in the vertical axis for the sake of clarity.

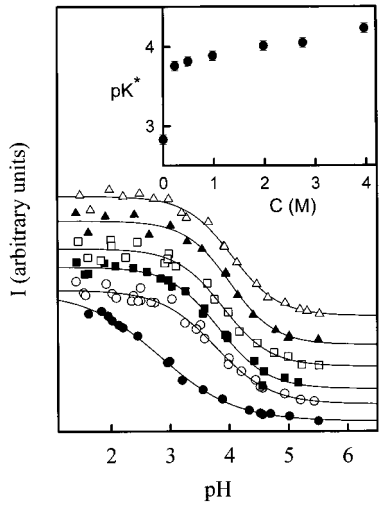

FIGURE 2: Profiles of fluorescence intensity versus pH for bovine ubiquitin at 25 °C. The different profiles correspond to different guanidine concentrations: ●, 0 M; ○, 0.23 M; ■, 0.49 M; □, 0.98 M; ▲, 1.98; Δ, 2.76 M. The continuous lines represent the best nonlinear, least-squares fits of eq 8 to the experimental data. From these fits, values of p*K** (the p*K* of a hypothetical single group responsible for the pH-dependent quenching of Tyr-59 fluorescence) were obtained; these p*K** values are plotted versus guanidine concentration in the inset.

(described below) shows, however, that the minor "transition" is not due to an unfolding process.

Ubiquitin fluorescence is due to tyrosine-59 and depends strongly on pH, a fact that has been explained in terms of quenching by the deprotonated form of a nearby carboxylic acid group (*29*). Figure 2 shows the profiles of tyrosine fluorescence intensity versus pH for several guanidine concentrations. We have analyzed these profiles by using the equation:

$$I = \frac{I_A + I_B 10^{(pK^* - pH)}}{1 + 10^{(pK^* - pH)}} \quad (8)$$

where p*K** is the p*K* value of the group responsible for the fluorescence quenching and I_A and I_B stand for the low-pH and high-pH limits of the fluorescence value. The calculated p*K** values (see Figure 2) change abruptly at low guanidine concentrations, a result that appears to be a nonspecific salt effect on the p*K* of the group responsible for quenching, since we have obtained very similar results with sodium chloride (results not shown).

It is clear now that the low guanidine "transition" observed in Figure 1 does not reflect an unfolding event, but it is result of the strong salt concentration dependence of p*K** at low guanidine concentration. Thus, at a given pH, the population of the quenching form of the amino acid (deprotonated carboxylate, most likely) and, consequently, the tyrosine fluorescence change abruptly with denaturant concentration. Accordingly, in our analysis of the profiles of Figure 1, we have included the sharp increase in fluorescence observed at low guanidine concentration in the fluorescence base line corresponding to the native state. The profiles of fluorescence intensity versus guanidine concentration have been analyzed according to the two-state equation:

$$I = \frac{I_N + I_D \exp[m_{1/2}(C - C_{1/2})/RT]}{1 + \exp[m_{1/2}(C - C_{1/2})/RT]} \quad (9)$$

where $C_{1/2}$ is the guanidine concentration (C) at which the unfolding Gibbs (ΔG) energy is zero, $m_{1/2}$ is the slope of ΔG versus C at the $C_{1/2}$ concentration, and the fluorescence intensity values for the native and denatured states (I_N and I_D in eq 9) are given by

$$I_D = a + b \times C \quad (10)$$

$$I_N = d + (f + g \times C) \exp(-hC) \quad (11)$$

That is, we have assumed a linear dependence of I_D with C, while in the case of I_N, we have included an exponential dependence in order to account phenomenologically for the sharp increase in fluorescence observed at low guanidine concentration. Fittings of eq 9 (together with eqs 10 and 11) to the experimental fluorescence data were always excellent (see Figure 1), and the values of $C_{1/2}$, $m_{1/2}$, and ΔG_W derived from these fittings are shown in Figure 3.

Guanidine-Induced Denaturation of Ubiquitin Studied by Circular Dichroism. Profiles of ellipticity at 222 nm versus denaturant concentration for the guanidine-induced denaturation of ubiquitin at 25 °C were obtained at pHs 2, 3, and 4 for the bovine protein, and at pHs 2, 3, 4, and 5.5 for the protein from yeast. These profiles (Figure 4) showed a single transition, and no cooperative change in ellipticity was detected at low denaturant concentration, thus confirming that the low guanidine "transition" seen in the fluorescence profiles (Figure 1) is not due to an unfolding event. Ellipticity versus denaturant concentration profiles were analyzed on the basis of the two-state model, but since the experimental uncertainty associated with each data value seems in this case to be significant in comparison to the amplitude of the unfolding transition (see Figure 4), we deemed it convenient to carry out a global analysis of the several profiles corresponding to different pH values. Thus, the two-state equation for the ellipticity at 222 nm was written as

$$\theta_{222} = \frac{\theta_N + \theta_D \exp[-(\Delta G_W(pH) - m_{1/2}C)/RT]}{1 + \exp[-(\Delta G_W(pH) - m_{1/2}C)/RT]} \quad (12)$$

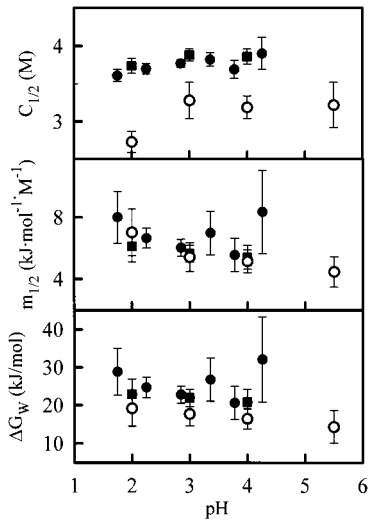

FIGURE 3: Thermodynamic parameters of unfolding for yeast ubiquitin (open symbols) and bovine ubiquitin (filled symbols), as determined from guanidine-induced denaturation experiments. These experiments were based on circular dichroism measurements in the case of the yeast protein and on fluorescence (circles) and circular dichroism (squares) measurements in the case of bovine ubiquitin. Note that the ΔG_W values shown in this figure have been obtained using the linear extrapolation method.

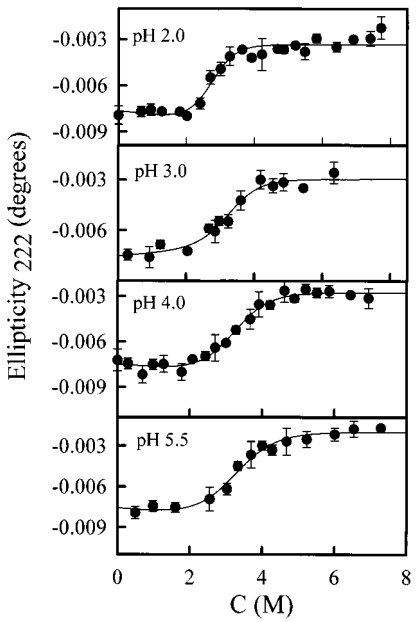

FIGURE 4: Profiles of ellipticity versus denaturant concentration for the guanidine-induced unfolding of yeast ubiquitin at 25 °C and at the pHs shown. Each experimental value shown is the average of three measurements. The lines represent the best global fit of eq 12 (together with eqs 13−15) to all the experimental profiles (see text for details).

with linear denaturant concentration dependencies for the native and unfolded state ellipticities:

$$\theta_N = \alpha_N + \beta_N C \quad (13)$$

$$\theta_D = \alpha_D + \beta_D C \quad (14)$$

And we assumed that the pH dependence of the unfolding Gibbs energy at zero denaturant concentration is linear:

$$\Delta G_W(\mathrm{pH}) = \Delta G_{W,ref} + (\ln 10) RT \Delta\nu(\mathrm{pH} - \mathrm{pH}_{ref}) \quad (15)$$

where pH_{ref} is a reference pH value, $\Delta G_{W,ref}$ is the unfolding Gibbs energy at zero denaturant concentration and at the reference pH, and $\Delta\nu$ is the number of protons taken from the solvent upon protein unfolding. The use of a linear pH dependence (that is, the use of a pH-independent $\Delta\nu$ value) is in this case justified by our analysis of the fluorescence unfolding curves of Figure 1, which indicates that ΔG_W for ubiquitin, as obtained from guanidine-induced unfolding data by linear extrapolation, barely changes with pH within experimental uncertainty (Figure 3).

For each protein (bovine or yeast ubiquitin), global, nonlinear least-squares fittings of eq 12 (together with eqs 13−15) to the several unfolding profiles were carried out imposing that the values of $\Delta G_{W,ref}$ and $\Delta\nu$ must be the same for all profiles, while each profile was allowed to have its own native and unfolded base lines (θ_N and θ_D of eqs 13 and 14) and its own $m_{1/2}$ value. These global fits were excellent (see, for instance, the several profiles for the yeast protein in Figure 4), and the unfolding parameters derived from them are shown in Figure 3.

General Comments on the Values for the Ubiquitin Unfolding Parameters Derived from Guanidine-Induced Denaturation Data. The values of ΔG_W (linear extrapolation), $C_{1/2}$, and $m_{1/2}$ derived from guanidine-induced unfolding profiles are summarized in Figure 3. Several points are worth emphasizing: (1) There is an excellent agreement between the data for bovine ubiquitin obtained from fluorescence and circular dichroism profiles. (2) Stability, *as characterized by guanidine-induced denaturation*, shows a very weak dependence on pH within the studied pH range; thus, both $C_{1/2}$ and ΔG_W obtained by linear extrapolation change little with pH. This appears to be an unusual result since, for other proteins, the ΔG_W values determined from guanidine-induced denaturation do seem to depend significantly on pH (*30*). An explanation for the lack of strong pH dependence found with ubiquitin will be proposed further below. (3) Ubiquitin from yeast is less stable than bovine ubiquitin, as shown by lower $C_{1/2}$ and ΔG_W values. (4) There is agreement within experimental uncertainty, however, between the $m_{1/2}$ values determined for both proteins. (5) The $m_{1/2}$ value seems to increase as the pH is lowered. If $m_{1/2}$ values are interpreted in terms of unfolding changes in accessible surface area (*31*), a possible explanation for this pH dependence could be that the denatured state becomes more accessible to the denaturant as the pH is lowered, as a result of the "expansion" (*30*) caused by the repulsion between positives charges. Another plausible explanation could be that, since guanidine is a salt, its interaction with proteins may depend not only on the accessible surface area but also on the number and sign of the charges exposed. On the other hand, an explanation in terms of the potential nonlinearity of the ΔG versus C dependence is not plausible in this case, since, for each protein, the different $m_{1/2}$ values in Figure 3 correspond to practically the same guanidine concentration (the $C_{1/2}$ values). Finally, whichever physical factor is responsible for the pH dependence of $m_{1/2}$, it is also likely to be responsible for the pH dependence of the linear extrapolation ΔG_W values; we note again, however, that the latter dependence is weak and certainly minor when compared with the pH dependence of the ΔG_W values determined from differential scanning calorimetry (see below).

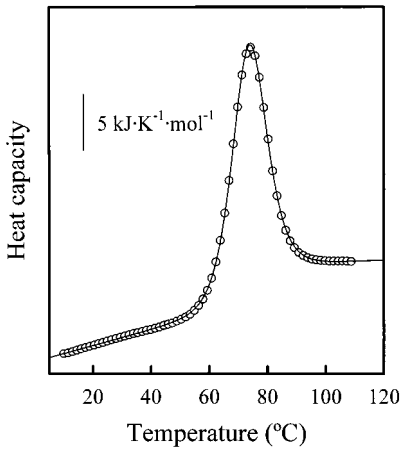

FIGURE 5: Representative DSC profile and two-state fit for ubiquitin unfolding under the conditions employed in this work. The experimental data (○) were obtained with bovine ubiquitin in 10 mM glycine buffer, pH 3. Protein concentration was 1.4 mg/mL, and the scanning rate was 1.5 K/min. The line represents the best fit of the two-state model to the data. See Ibarra-Molero et al. (*15*) for further details on DSC data analysis.

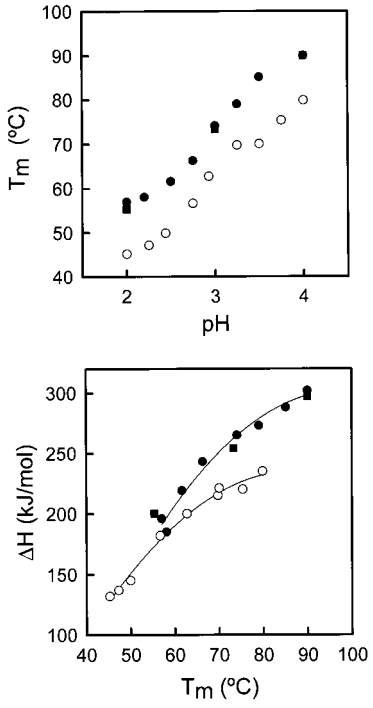

FIGURE 6: Profiles of denaturation temperature versus pH (upper panel) and unfolding enthalpy versus denaturation temperature (lower panel) for yeast ubiquitin (open symbols) and bovine ubiquitin (closed symbols). The T_m and ΔH values have been calculated from the two-state analysis of DSC experiments performed as part of the present work (○, yeast ubiquitin; ■, bovine ubiquitin), but we have also included the data on bovine ubiquitin (●) previously reported by Wintrode et al. (*13*). The lines shown in the lower panel represent the best fit of second-order polynomials to the ΔH versus T_m data.

Thermal Denaturation of Ubiquitin in the Absence of Guanidine As Studied by Differential Scanning Calorimetry. Differential scanning calorimetry (DSC) experiments were carried for both the bovine and the yeast protein, in the absence of guanidine and at several pHs in the range 2−4 (additional experiments in the presence of guanidine were performed with the bovine protein). The DSC transitions were analyzed on the basis of the two-state equilibrium model; excellent fits were obtained in all cases (see, for instance, Figure 5). The values for the denaturation temperature and the unfolding enthalpy derived from the two-state analysis are shown in Figure 6. It must be noted that the T_m and ΔH values for bovine ubiquitin are in good agreement with those previously determined by Wintrode et al. (*13*) under the same solvent conditions and that the plots of ΔH versus temperature are nonlinear for both proteins, reflecting temperature-dependent heat capacity changes for unfolding (*13*). Also, the lower stability of the yeast protein is apparent in the denaturation temperature values.

Unfolding Gibbs energies at 25 °C were calculated by direct integration of the Gibbs−Helmholtz equation:

$$\Delta G_W = -298 \int_{T_m}^{298} \frac{\Delta H}{T^2}\, dT \qquad (16)$$

where the temperature dependencies of the unfolding enthalpies were described by the second-order polynomials[2] that fit the experimental ΔH versus T data (see Figure 6). The resulting ΔG_W versus pH data are shown in Figure 7.

Effect of Guanidine on the Thermal Denaturation and Stability of Bovine Ubiquitin at pH 2 As Studied by DSC. We have carried out DSC experiments on the thermal unfolding of bovine ubiquitin at pH 2 and in the presence of several guanidine concentrations within the range 0−4

[2] Second-order polynomials are required here, since the plots of ΔH vs T are clearly nonlinear (Figure 6). This nonlinearity reflects a temperature-dependent unfolding heat capacity change, as previously reported by Wintrode et al. (*13*) and as suggested by the native and unfolded base lines of our DSC thermograms (see, for instance, Figure 5). In fact, the ΔC_p values calculated as $d\Delta H/dT$ from the dependencies shown in Figure 6 are in good agreement (better than 1 kJ·mol^{-1}·K^{-1}) with the values determined by extrapolating to the T_m the native an unfolded base lines of our DSC experiments. Obviously, the calculation of ΔG_W described in the text takes into account the temperature dependence of ΔC_p as given by the (nonlinear) experimental ΔH vs T profiles, a procedure which involves an assumption: the temperature dependence of ΔH given by the second-order polynomial that fits the experimental data is *assumed* to hold outside the temperature range in which experimental ΔH values are available. We also calculated ΔG_W values at 25 °C from the individual DSC profiles on the basis of

$$\Delta G_W = \Delta H(T_m) + \int_{T_m}^{298} \Delta C_p\, dT - 298\left[\frac{\Delta H(T_m)}{T_m} + \int_{T_m}^{298} \frac{\Delta C_p}{T}\, dT\right]$$

using the ΔC_p vs T dependence given in Table 3 of Wintrode et al. (*13*), which was calculated by these authors *assuming* that the heat capacity of the unfolded state is equal to the theoretical value derived from model-compound heat capacities. The ΔG_W values obtained in this manner (results not shown) were in good agreement with those shown in Figure 7 (average difference between the two series of "DSC-values" was 1.3 kJ/mol). Finally, the situation is somewhat different for the ΔG values shown in Figure 8. We include there data corresponding to different guanidine concentrations, and reliable information about the temperature dependence of ΔC_p in guanidine solutions is not available. Therefore, we used constant-ΔC_p Gibbs−Hemholtz extrapolation (with the average ΔC_p values from all the DSC experiments in guanidine solutions: 3.76 ± 0.80 kJ·mol^{-1}·K^{-1}) to calculate ΔG values in the presence of guanidine and, for the sake of consistency, also for the ΔG_W value (ΔG value at 0 M guanidine) given in Figure 8. As a result, there is a small difference (about 3 kJ/mol) between the ΔG_W values (derived from DSC) for bovine ubiquitin unfolding at pH 2 shown in Figures 7 and 8. A similar small difference between ΔG_W values obtained by constant-ΔC_p and nonconstant-ΔC_p extrapolation was recently reported for HEW lysozyme (see Figure 5 in reference *10*). We believe that these small discrepancies simply reflect the unavoidable uncertainty associated with Gibbs−Hemholtz extrapolation. The uncertainty is, nevertheless, fairly small and does not affect any of the general results and conclusions of the present work.

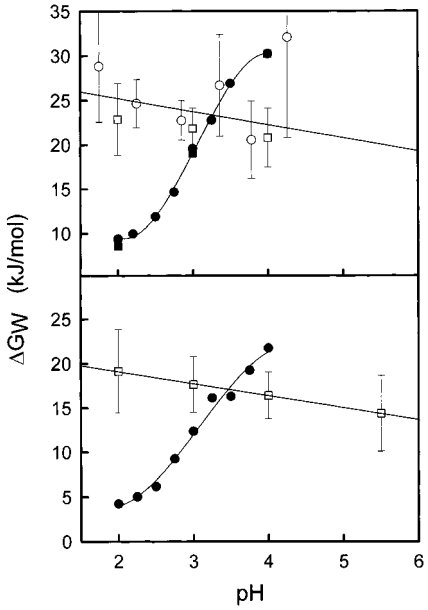

FIGURE 7: Profiles of unfolding Gibbs energy versus pH at 25 °C for yeast ubiquitin (lower panel) and bovine ubiquitin (upper panel). Closed symbols represent the ΔG_W values calculated from DSC experiments; in the upper panel, the closed squares represent the data obtained in this work, and the close circles are the data previously reported by Wintrode et al. (*13*) under the same buffer conditions. Open symbols are the ΔG_W values calculated from guanidine-induced unfolding profiles (open squares, circular dichroism profiles; open circles, fluorescence profiles). Lines do not carry any meaning, but are drawn to guide the eye.

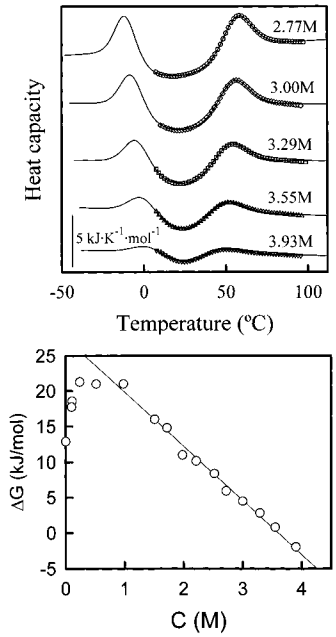

FIGURE 8: Upper panel: Examples of DSC profiles for the unfolding of bovine ubiquitin at pH 2 and in the presence of several guanidine concentrations (the numbers alongside the profiles in the figure) showing evidence of both heat and cold denaturation transitions. The profiles have been displaced in the vertical axis for the sake of clarity. The symbols are the experimental heat capacity data, and the continuous lines represent the best fits of the two-state model; see Ibarra-Molero et al. (*15*) for details on the data analysis procedure; the complete peak due to cold denaturation is shown for illustration purposes only. Lower panel: Profile of unfolding Gibbs energy versus guanidine concentration for bovine ubiquitin at pH 2; the ΔG values were obtained from two-state analyses of DSC thermograms, such as those shown in the upper panel. The straight line is the best linear fit to the ΔG values above 1 M guanidine.

M (see Figure 8). Evidence of cold denaturation is apparent in the thermograms corresponding to the higher guanidine concentrations employed. We have recently reported and discussed (*15*) ubiquitin cold denaturation in the presence of guanidine (although not at pH 2). Here, we just wish to comment on the unfolding ΔG versus guanidine concentration profile obtained from the thermograms in the upper panel of Figure 8 (see ref *15* for details on the data analysis). This ΔG vs C profile is shown in the lower panel of Figure 8 and demonstrates clearly that the guanidine concentration dependence of the unfolding ΔG is linear to a very good approximation within the range 1−4 M and that deviations from linearity only occur below 1 M guanidine approximately (see refs *14* and *32* for an interpretation of these deviations in terms of anion interactions with the positively charged protein). Similar results have been previously described with HEW lysozyme (*10*) and thioredoxin (*33*). Of course, it would have been of interest to obtain the ΔG vs C profile at a pH value near the isoelectric point, since our data (see below) suggest a clear positive deviation from linearity in this case; DSC experiments near the isoelectric pH are, however, problematic due to the very high stability of ubiquitin under these conditions.

General Comments on the Difference between the Stability of Ubiquitin Determined from DSC and the Stability Determined from Guanidine-Induced Denaturation Using Linear Extrapolation. It is clear that the stability of ubiquitin, *as determined by differential scanning calorimetry*, is affected by pH, as is demonstrated by the strong pH dependencies of both T_m and ΔG_W. This contrasts sharply with the stability *as determined from guanidine-induced denaturation*, which shows only a weak pH dependence. Furthermore, the difference between the two ΔG_W versus pH profiles (see Figure 7) appears to change sign at about pH 3.5−4. We tentatively suggest the following interpretation for these results:

(1) The charged atoms of ionizable groups in native ubiquitin are largely exposed to the solvent [accessible surface area calculations show that all ionizable groups in ubiquitin belong to class III in the classification of Rashin and Honig (*16*)]; as a result, desolvation penalties upon folding may perhaps be neglected as a first approximation, and the electrostatic contribution to the unfolding Gibbs energy is mostly associated with the charge−charge interactions (see refs *9* and *10* for a discussion on desolvation and charge−charge contributions to unfolding Gibbs energies).

(2) These charge−charge interactions are eliminated in high guanidine concentration, due again to the large degree of exposure to solvent of the charged atoms in the native protein, together with the salt character of the denaturant. Note, however, that guanidine could be expected to be more effective than most salts at eliminating charge−charge interactions, because the "screening effect" may be enhanced by the binding of guanidinium ions to the protein surface (*34*) and, at acidic pH, by the binding of chloride ions (*14*).

(3) ΔG_W values determined from guanidine denaturation data by linear extrapolation from high guanidine concentration do not contain the contribution from the charge−charge interactions and are expected to show only a weak pH dependence, since we have assumed that charge−charge

interactions (one source of p*K* shifts upon folding) are screened out in concentrated guanidine and that desolvation contributions (other source of p*K* shifts) are negligible for ubiquitin. [Accordingly, the fact that, for other proteins, the ΔG_W values determined from guanidine-induced denaturation do seem to depend significantly on pH (*30*) could be taken as evidence that Gibbs energy contributions arising from desolvation of charges upon folding are *not* negligible in those proteins.]

(4) The charge−charge contribution is present in the ΔG_W values determined by DSC in the absence of guanidine. Thus, these ΔG_W values are strongly pH-dependent.[3]

(5) The difference between both series of ΔG_W data (obtained from guanidine denaturation experiments and from DSC) should therefore provide an estimate of the contribution to protein stability that arises from interactions between charged ionizable groups.

(6) The sign of this contribution will change with pH (*9*, *10*) because, at pH values near the isoelectric pH (about 7 for both proteins), charge−charge interactions are expected to be stabilizing (groups of a given sign are, on average, surrounded by groups of the opposite sign, giving rise to an overall attractive and stabilizing interaction), while they will be destabilizing at acidic pH (due to the neutralization of negatively charged groups and the concomitant repulsion between the remaining positive charges). This sign change is clearly suggested by our experimental results (Figure 7).

A Theoretical Estimate of the Contribution from Charge−Charge Interactions to Ubiquitin Stability. Electrostatic calculations for proteins have been previously carried out on the basis of various theoretical approaches, including the Tanford−Kirkwood model (*18*, *19*, *35*), the numerical solving of the Poisson−Boltzmann equation (*36*, *9*, *37*), the generalized Born model (*38*, *39*), cavity-function theory (*40*), and all-atoms models and dipolar models (*41*, *42*). Here, we have employed the Tanford−Kirkwood model with the solvent accessibility correction of Gurd (*18*, *19*), because it is simple and yet it appears to account for the overall features of the titration behavior of several proteins. Also, the use of the Tanford−Kirkwood model would seem acceptable in this case, since this model does not take into account solvation terms, which we have assumed in fact to be negligible, since charged atoms of ionizable residues in native ubiquitin are exposed to the solvent. Interestingly, it has been remarked (*43*) that for ionizable groups that reside on the protein surface "many fundamentally flawed models work as well as physically sensible ones" and that "good results are obtained as long as a large shielding for charge−charge interactions is built into the model".

We used the energies of interaction between unit positive charges calculated by using the Tanford−Kirkwood model as the input for a procedure based on the Bashford−Karplus, "reduced-set-of-sites" approximation (*20*), which yielded the fractional protonation of each ionizable group, as well as the average values of the pairwise charge−charge interaction energies and the total charge−charge contribution to the unfolding Gibbs energy (see Materials and Methods for details). The use of the Bashford−Karplus procedure is feasible in this case due to the fact that ubiquitin is a small protein with a comparatively small number of ionizable residues. We note, nevertheless, that almost identical results were obtained in preliminary calculations using the program TITRA (*28*), which is also based upon the solvent-accessibility-corrected Tanford−Kirkwood model, but employs the (less rigorous) mean-field approximation of Tanford and Roxby (*44*).

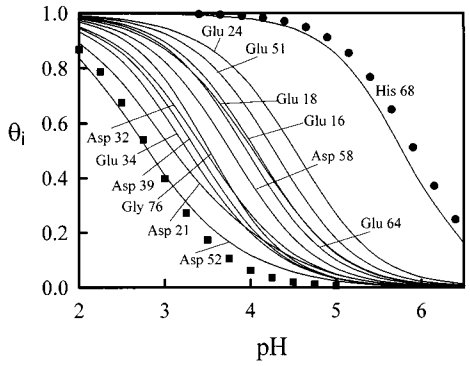

FIGURE 9: Profiles of fractional protonation versus pH for the ubiquitin residues that undergo protonation in the acidic pH range. The profile labeled Gly 76 corresponds to the carboxyl terminal. The fractional protonation values were calculated using the TK−BK procedure. The closed circles depict the ionization behavior of His 68 as determined experimentally from NMR measurements. The closed squares show the ionization behavior of the hypothetical single group responsible for the pH-dependent quenching of the Tyr 59 fluorescence in ubiquitin (see Figure 2).

Figure 9 shows the calculated profiles of fractional protonation versus pH for all the groups that undergo ionization in the acidic pH range. Also, in Figure 10 we compare the theoretical ΔG_{q-q} versus pH profile with the profile calculated as the difference ΔG_W(DSC) − ΔG_W(guanidine+LEM). The agreement between both profiles is excellent, which, to some extent, supports our interpretation of the difference ΔG_W(DSC) − ΔG_W(guanidine+LEM) as an estimate of the contribution from charge−charge interactions to the unfolding Gibbs energy (see preceding section). It is also encouraging (see Figure 9) that there is a good agreement between the ionization profile predicted for His-68 and that obtained by NMR, as well as the fact that the pH dependence of Tyr-59 fluorescence agrees with the ionization profile predicted for Asp-52, one of the two carboxylate groups which are close to Tyr-59.

It is worth noting that our theoretical results do not depend strongly on the value used for the dielectric constant of the protein and, in fact, we obtained similar ΔG_{q-q} versus pH profiles using values of ϵ_P as low as 2 and as high as 20. On the other hand, the shielding of the charge−charge interactions introduced by the solvent-accessibility correction does make a significant difference. Thus, leaving out the accessibility correction leads to significantly larger (in absolute value) ΔG_{q-q} values; nevertheless, the overall qualitative features of the ΔG_{q-q} versus pH profile remain the same:

[3] Clearly, we may also expect the charge−charge contribution to be present in the ΔG_W values determined from urea-induced unfolding experiments, since urea is not a salt. This is supported by our preliminary experiments on the urea-induced unfolding of yeast ubiquitin. Thus, LEM analysis of CD unfolding profiles at pH 2 and 5 yields the following results: $C_{1/2} = 2.1$ M, $m_{1/2} = 3.6$ kJ·mol^{-1}·M^{-1}, $\Delta G_W = 7.5$ kJ/mol (pH 2); $C_{1/2} = 5.9$ M, $m_{1/2} = 4.5$ kJ·mol^{-1}·M^{-1}, $\Delta G_W = 26.4$ kJ/mol (pH 5). Note that ΔG_W and $C_{1/2}$ change strongly with pH and that these ΔG_W values are in good agreement with those determined by DSC (Figure 7).

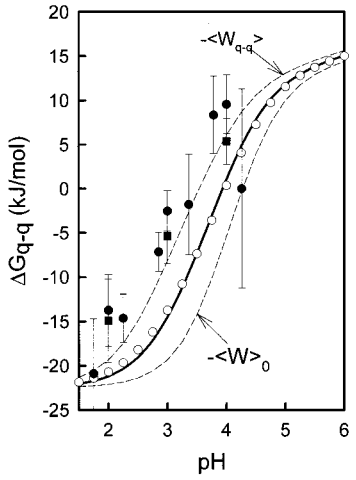

FIGURE 10: Comparison between theoretical and experimental estimates of the contribution to ubiquitin unfolding Gibbs energy that arises from charge−charge interactions. The closed symbols are the experimental estimates calculated as ΔG_W(DSC) − ΔG_W(guanidine-induced unfolding + linear extrapolation); circles and squares are, respectively, the values of this difference for bovine ubiquitin and yeast ubiquitin. The error bars have been calculated assuming the main source of error is the ΔG_W values obtained from guanidine-induced unfolding profiles. The open circles are the theoretical estimates of ΔG_{q-q} derived using the TK−BK procedure. $\langle W_{q-q}\rangle$ is the sum of charge−charge interactions energies in the native protein, a value which is, in fact, an average over all protein protonation states. $\langle W\rangle_0$ is the same average, but obtained using the distribution of protonation states corresponding to the unfolded protein. The continuous line shown in the figure is the mean of $-\langle W_{q-q}\rangle$ and $-\langle W\rangle_0$. See the Appendix for a discussion on the meaning of these quantities.

ΔG_{q-q} is negative at very acidic pH, but becomes positive at pH values near the isoelectric one.

Analysis of the Charge−Charge Interactions in Ubiquitin at the Individual Residue Level. We proceed in this section to interpret the contribution from charge−charge interactions to ubiquitin stability, in terms of contributions from the individual ionizable groups. Before doing so, however, it is important to realize that ΔG_{q-q} is not strictly equal to the energy required to "break" all pairwise charge−charge interactions in the native protein (as one might perhaps naively expect). This point becomes obvious if we mentally carry out the unfolding process at a given pH in three steps: (1) The pairwise charge−charge interactions in the native protein are "disconnected" without changing the protonation states distribution (that is, the relative amounts of the different protonation states are kept at the values corresponding to the native protein). (2) The protein is unfolded, maintaining the native distribution of protonation states. (3) The relative amounts of the different protonation states are allowed to change ("relax") from the values given by the native distribution (values affected by the charge−charge interactions) to the values corresponding to the "unfolded" distribution (values not affected by the charge−charge interactions). Step 1 involves breaking all charge−charge interactions in the native protein, but the total contribution of charge−charge interactions to the unfolding ΔG is given by the sum of the Gibbs energy changes associated with steps 1 and 3. Thus, if we refer to step 3 as a "relaxation" step (for lack of a better name), we have

$$\Delta G_{q-q} = -\langle W_{q-q}\rangle + \Delta G_{\text{rel}} \qquad (17)$$

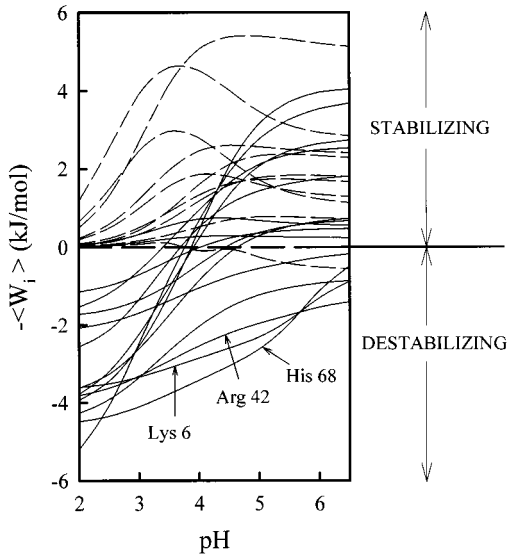

FIGURE 11: pH dependence of the energies due to the charge−charge interaction of each ionizable group in ubiquitin with the rest of the ionizable groups (see eq 18). The values have been calculated using the TK−BK procedure. Residues that bear or develop a negative charge (Asp, Glu, Tyr, and carboxyl terminal) are shown with dashed lines. Groups that bear or develop a positive charge (Lys, Arg, His, and amino terminal) are shown with continuous lines.

The theoretical relation between ΔG_{q-q} and $-\langle W_{q-q}\rangle$ is further explored in the Appendix. We show there that $-\langle W_{q-q}\rangle$ is an upper limit to the value of ΔG_{q-q} (that is, the relaxation contribution is always destabilizing and $\Delta G_{q-q} \leq -\langle W_{q-q}\rangle$). In the case of ubiquitin, the theoretical calculations shown in Figure 10 indicate, nevertheless, that the difference between ΔG_{q-q} and $-\langle W_{q-q}\rangle$, while significant, is not very large and that, as a rough first approximation, we can write:

$$\Delta G_{q-q} \approx -\langle W_{q-q}\rangle = -\frac{1}{2}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\langle W_{ij}\rangle\right) = -\frac{1}{2}\sum_{i=1}^{n}\langle W_i\rangle \qquad (18)$$

This approximation allows us to interpret the ΔG_{q-q} values in terms of contributions arising from the individual ionizable groups, as expressed in the last equality of the above expression, where $\langle W_i\rangle$ is the energy due to the charge−charge interaction of group i with the rest of the ionizable groups in the protein.

Figure 11 shows a plot of $-\langle W_i\rangle$ versus pH for all ionizable groups in ubiquitin. These individual $-\langle W_i\rangle$ values have been calculated on the basis of the solvent-accessibility-corrected Tanford−Kirkwood model (see Materials and Methods) and must be taken with due caution. Nevertheless, the overall qualitative picture is clear enough and should be independent of specific models. Thus, at very acidic pH, the values of $-\langle W_i\rangle$ for carboxylic acid groups are zero (reflecting a zero charge on these groups), while for positively charged groups (His, Lys, and Arg) $-\langle W_i\rangle$ is negative (that is, destabilizing), reflecting the repulsive interaction between identical charges. On the other hand, at values of pH above 5, carboxylic acid groups bear negative charge, and most $-\langle W_i\rangle$ values are positive (stabilizing), indicating that each charged group is predominantly interacting with charges of the opposite sign, thus giving rise to an overall attractive interaction.

General Comments on Ubiquitin Thermal Stability. Ubiquitin is a very stable protein at neutral pH (denaturation

temperature around or higher than 100 °C; ref *14*). Electrostatic factors must play an important role in determining ubiquitin stability, as shown by the strong pH dependence of the denaturation temperature and the unfolding Gibbs energy values determined from differential scanning calorimetry experiments (see Figures 6 and 7). In particular, the experimental results and theoretical calculations reported in this work suggest that the high stability of ubiquitin at neutral pH may be due, at least in part, to the following two causes:

(1) Gibbs energy contributions arising from the desolvation of charges upon folding could be small in the case of ubiquitin, since the charged atoms of ionizable residues are very exposed to the solvent in the native structure. This is important for stabilization, because desolvation contributions are expected to be destabilizing (*9*).

(2) At neutral pH, the Gibbs energy contribution that arises from charge−charge interactions is stabilizing. This is due to the fact that the charged residues are distributed on the protein surface in such a way that most charges of a given sign (but not all, see below) interact mainly with charges of the opposite sign, thus giving rise to an overall attractive interaction.

The second point above suggests the possibility of enhancing protein thermal stability by adequately redesigning the distribution of solvent-exposed, charged residues on the protein surface. In connection with this, it is interesting to note that the positively charged residues Lys-6, His-68, and Arg-42 in ubiquitin are predicted to be involved in destabilizing charge−charge interactions in a wide pH range (see Figure 11). These residues are located close to each other on the solvent-exposed side of the β-sheet (but on different strands) and make obvious targets for charge-reversal mutations that could perhaps lead to a significant stability enhancement. We are currently testing this possibility.

ACKNOWLEDGMENT

NMR experiments were kindly carried out by Dr. Xiaomin Yang in the laboratory of Dr. Maria Luisa Tasayco (City College of New York); we are very grateful to these researchers, as well as to the NMR facility from the Hunter College of the City University of New York. We also thank Dr. Peter Fojan (University of Aalborg, Denmark) for carrying out for us preliminary electrostatic calculations with the program TITRA (*28*).

APPENDIX

We explore here in some detail the theoretical relationship between the contribution from charge−charge contributions to the unfolding Gibbs energy (ΔG_{q-q}) and the sum of the charge−charge interaction energies in the native protein ($\langle W_{q-q}\rangle$).

We begin by noticing the similarity between eq 2 (for the native protein) and eq 6 (for the unfolded protein). The only difference is the charge−charge interaction term, which may be considered as a "perturbation". Therefore, we will use a common procedure in Statistical Thermodynamics (see chapters 19 and 20 in reference *45*) and introduce a parameter, λ, to smoothly connect the perturbation. Thus, we write:

$$\rho(\chi,\lambda) = \frac{\exp[-((\Delta G_0(\chi,\mathrm{pH}) + \lambda W(\chi))/RT]}{Z(\lambda)} \quad (19)$$

$$Z(\lambda) = \sum_{\chi} \exp[-(\Delta G_0(\chi,\mathrm{pH}) + \lambda W(\chi))/RT] \quad (20)$$

where

$$\Delta G_0(\chi,\mathrm{pH}) = -RT\,(\ln 10)[\nu(\chi)\mathrm{pH} - \sum_{i=1}^{n}(q_i + x_i)\mathrm{p}K_{\mathrm{int},i}] \quad (21)$$

$$W(\chi) = \frac{1}{2}\sum_{i,j=1}^{n} E_{ij}(q_i + x_i)(q_j + x_j) \quad (22)$$

Clearly, when $\lambda = 0$ the "perturbation" is not connected, and eqs 19 and 20 yield the unfolded distribution of protonation states (ρ_U) and the protonation partition function of the unfolded protein (Z_U). On the other hand, when $\lambda = 1$ the "perturbation" is fully connected, and eqs 19 and 20 yield ρ_N and Z_N (eqs 3 and 4).

Straightforward differentiation yields the λ dependence of $Z(\lambda)$:

$$\frac{\partial \ln Z(\lambda)}{\partial\lambda} = -\frac{\sum_{\chi} W(\chi)\rho(\chi,\lambda)}{RT} = -\frac{\langle W\rangle_\lambda}{RT} \quad (23)$$

where the subscript λ in $\langle W\rangle_\lambda$ means that the average of the sum of the charge−charge interaction energies is taken over all protonation states, but using the distribution corresponding to the given λ value [that is, $\rho(\chi,\lambda)$]. For instance, for $\lambda = 1$, we have $\rho(\chi,\lambda) = \rho_N(\chi)$, and, therefore, $\langle W\rangle_1$ is equal to $\langle W_{q-q}\rangle$, the sum of the charge−charge interaction energies in the native protein. On the other hand, for $\lambda = 0$, we have $\rho(\chi,\lambda) = \rho_U(\chi)$, and $\langle W\rangle_0$ is the sum of the charge−charge interaction energies, but averaged with the protonation-state distribution of the "unperturbed" system (that is, $\langle W\rangle_0$ could be interpreted as the first-order estimate of $\langle W_{q-q}\rangle$).

Equation 23 allows us to arrive at a very convenient expression for the contribution from charge−charge interactions to the unfolding Gibbs energy (given by eq 7):

$$\Delta G_{q-q} = -RT\ln\left(\frac{Z_U}{Z_N}\right) = RT\,(\ln Z(1) - \ln Z(0)) = RT\int_0^1 \frac{\partial \ln Z(\lambda)}{\partial\lambda}\,\mathrm{d}\lambda = -\int_0^1 \langle W\rangle_\lambda\,\mathrm{d}\lambda \quad (24)$$

The λ dependence of $\langle W\rangle_\lambda$ is easily found by differentiation:

$$\frac{\partial\langle W\rangle_\lambda}{\partial\lambda} = -\frac{\langle W^2\rangle_\lambda - \langle W\rangle_\lambda^2}{RT} = -\frac{\langle (W - \langle W\rangle_\lambda)^2\rangle_\lambda}{RT} \quad (25)$$

where the average of the square of $W - \langle W\rangle_\lambda$ is simply the second moment of the distribution that describes the fluctuations in the sum of charge−charge interaction energies (calculated for a given λ value). The relevant point here is, of course, that second moments of distributions must necessarily be positive numbers. Therefore, $\langle W\rangle_\lambda$ is a monotonically decreasing function of λ, and, consequently, the following inequalities necessarily hold:

$$\langle W\rangle_0 \geq \langle W\rangle_\lambda \geq \langle W\rangle_1 = \langle W_{q-q}\rangle \quad (26)$$

Changing signs and integrating from $\lambda = 0$ to $\lambda = 1$, we have

$$-\int_0^1 \langle W\rangle_0 \, d\lambda \leq -\int_0^1 \langle W\rangle_\lambda \, d\lambda \leq -\int_0^1 \langle W_{q-q}\rangle \, d\lambda \quad (27)$$

and using eq 24 and the fact that both $\langle W\rangle_0$ and $\langle W_{q-q}\rangle$ are λ-independent values, we arrive at

$$-\langle W\rangle_0 \leq \Delta G_{q-q} \leq -\langle W_{q-q}\rangle \quad (28)$$

which shows that ΔG_{q-q} must be in the interval $(-\langle W\rangle_0, -\langle W_{q-q}\rangle)$. This is clearly illustrated in Figure 10, where we have included the calculated pH dependencies of the three quantities involved in eq 28. The most important point here is, of course, that $\Delta G_{q-q} \leq -\langle W_{q-q}\rangle$; that is, if we interpret the contribution from charge−charge interaction in terms of eq 17, it is clear that the relaxation term (ΔG_{rel}) is always destabilizing.

An estimate of ΔG_{q-q} more precise than that provided by the interval $(-\langle W\rangle_0, -\langle W_{q-q}\rangle)$ can be derived, but on the basis of an approximation, as is described below. We first differentiate in eq 25 to arrive at

$$\frac{\partial^2 \langle W\rangle_\lambda}{\partial \lambda^2} = \frac{\langle (W - \langle W\rangle_\lambda)^3\rangle_\lambda}{(RT)^2} \quad (29)$$

where the average of the third power of $W - \langle W\rangle_\lambda$ is the third moment of the distribution that describes the fluctuations in the sum of charge−charge interaction energies. Assume now that, at a certain pH, several ionizable groups are partially protonated. Of course, the fractional protonations of ionizable groups are just averages over all the protonation states and, according to the ergodic hypothesis, equal to the corresponding time averages. Actually, those groups fluctuate between the protonated and the nonprotonated forms (forms which have different charges), thus contributing to the fluctuation in the sum of charge−charge interaction energies. Some groups might be considered to fluctuate in a roughly independent manner in the mean field created by the other charges, while for other groups the fluctuations may be strongly coupled. In any case, if at a given pH value a sufficiently large number of ionizable groups are fluctuating significantly, we may consider that the fluctuations in the sum of charge−charge interaction energies (W) are the result of a large number of comparatively small effects operating more or less independently. Then, the central limit theorem suggests that the distribution function for the fluctuations in W will approach the normal distribution law, which is a symmetric distribution with the third moment equal to zero. A third moment equal to zero for all λ implies that the second derivative in eq 29 is zero, that the first derivative in eq 25 is constant, and, therefore, that $\langle W\rangle_\lambda$ changes linearly with λ:

$$\langle W\rangle_\lambda = \langle W\rangle_0 + \lambda(\langle W_{q-q}\rangle - \langle W\rangle_0) \quad (30)$$

and substitution in eq 24 followed by integration yields

$$\Delta G_{q-q} = -\frac{1}{2}(\langle W_{q-q}\rangle + \langle W\rangle_0) \quad (31)$$

According to this equation, the ΔG_{q-q} versus pH profile is expected to be the mean of the profiles defined by $-\langle W_{q-q}\rangle$ and $-\langle W\rangle_0$.

We have not carried out yet a systematic investigation into the range of validity of the "Gaussian approximation" described above, but in the case of ubiquitin at least, eq 31 provides an excellent estimate of ΔG_{q-q} (see Figure 10).

NOTE ADDED IN PROOF

After this work was completed, we became aware of a recent mutational analysis into the electrostatic contributions to protein stability, which focuses on the effect of inter-helix ion pairs (*46*). We thank Dr. Robert Baldwin for making a preprint of this work available to us.

REFERENCES

1. Becktel, W. J., and Schellman. J. A. (1987) *Biopolymers 26*, 1859−1877.
2. Sanchez-Ruiz, J. M. (1995) in *Subcelllar Biochemistry, Volume 24: Proteins: Structure, Function and Engineering* (Biswas, B. B., and Roy, S., Eds.) pp 133−176, Plenum, New York.
3. Dill, K. A. (1990) *Biochemistry 29*, 7133−7155.
4. Yang, A.-S., Sharp, K. A., and Honig, B. (1992) *J. Mol. Biol. 227*, 889−900.
5. Matthews, B. W. (1995) *Adv. Protein Chem. 46*, 249−278.
6. Makhatadze, G. I., and Privalov, P. L. (1995) *Adv. Protein Chem. 47*, 307−425.
7. Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) *FASEB J. 10*, 75−83.
8. Luque, I., and Freire, E. (1998) *Methods Enzymol. 295*, 100−127.
9. Yang, A.-S., and Honig, B. (1993) *J. Mol. Biol. 231*, 459−474.
10. Ibarra-Molero, B., and Sanchez-Ruiz, J. M. (1996) *Biochemistry 35*, 14689−14702.
11. Vijay-Kumar, S., Bugg, C. E., and Cook, W. J. (1987) *J. Mol. Biol. 194*, 531−544.
12. Weber, P. L., Brown, S. C., and Mueller, L. (1987) *Biochemistry 26*, 7282−7290.
13. Wintrode, P. L., Makhatadze, G. I., and Privalov, P. L. (1994) *Proteins: Struct., Funct., Genet. 18*, 246−253.
14. Makhatadze, G. I., Lopez, M. M., Richardson, J. M., and Thomas, S. T. (1998) *Protein Sci. 7*, 689−697.
15. Ibarra-Molero, B., Makhatadze, G. I., and Sanchez-Ruiz, J. M. (1999) *Biochim. Biophys. Acta 1429*, 384−390.
16. Rashin, A. A., and Honig, B. (1984) *J. Mol. Biol. 173*, 515−521.
17. Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., Vierstra, R. D., Hatfield, P. M., and Cook, W. J. (1987) *J. Biol. Chem. 262*, 6396−6399.
18. Tanford, C., and Kirkwood, J. G. (1957) *J. Am. Chem. Soc. 79*, 5333−5339.
19. Matthew, J. B., and Gurd, F. R. N. (1986) *Methods Enzymol. 130*, 413−453.
20. Bashford, D., and Karplus, M. (1991) *J. Phys. Chem. 95*, 9556−9561.
21. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
22. Pace, C. N., Shirley, B. A., and Thomson, J. A. (1989) in *Protein structure, a practical approach* (Creighton, T. E., Ed.) pp 331−330, IRL Press at Oxford University Press, Oxford.
23. Gribenko, A. V., and Makhatadze, G. I. (1998) *J. Mol. Biol. 283*, 679−694.
24. Harding, M. M., Williams, D. H., and Woolfson, D. N. (1991) *Biochemistry 30*, 3120−3128.
25. Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1993) *Biochemistry 32*, 7054−7063.
26. Shrake, A., and Rupley, J. A. (1973) *J. Mol. Biol. 79*, 351−372.
27. Chothia, C. (1976) *J. Mol. Biol. 105*, 1−12.

28. Neves-Petersen, M. T., Martel, P., Petersen, E. I., Drabløs, F., and Petersen, S. B. (1997) *Methods Enzymol. 284*, 130−153.
29. Jenson, J., Goldstein, G., and Breslow, E. (1980) *Biochim. Biophys. Acta 624*, 378−385.
30. Pace, C. N., Laurents, D. V., and Thomson, J. A. (1990) *Biochemistry 29*, 2564−2572.
31. Myers, J. K., Pace, N. C., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
32. Makhatadze, G. I. (1999) *J. Phys. Chem.* (in press).
33. Santoro, M. M., and Bolen, D. W. (1992) *Biochemistry 31*, 4901−4907.
34. Makhatadze, G. I., and Privalov, P. I. (1992) *J. Mol. Biol. 226*, 491−505.
35. Garcia-Moreno, B. (1994) *Methods Enzymol. 240*, 645−667.
36. Warwicker, J., and Watson, H. C. (1982) *J. Mol. Biol. 157*, 671−679.
37. Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1996) *Biochemistry 35*, 7819−7833.
38. Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) *J. Am. Chem. Soc. 112*, 6127−6129.
39. Ghosh, A., Rapp, C. S., and Friesner, R. A. (1998) *J. Phys. Chem. B 102*, 10983−10990.
40. Zhou, Y. (1998) *J. Phys. Chem. B 102*, 10615−10621.
41. Del Buono, G. S., Figueirido, F. E., and Levy, R. M. (1994) *Proteins: Struct., Funct., Genet. 20*, 85−97.
42. Sham, Y. Y., Chu, Z. T., and Warshel, A. (1997) *J. Phys. Chem. B 101*, 4458−4472.
43. Warshel, A., and Papazyan, A. (1998) *Curr. Opin. Struct. Biol. 8*, 211−217.
44. Tanford, C., and Roxby, R. (1972) *Biochemistry 11*, 2192−2198.
45. Callen, B. C. (1985) *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed., Wiley, New York.
46. Ramos, C. H. I., Kay, M. S., and Baldwin, R. L. (1999) *Biochemistry* (in press).

BI9905819