

# Electrostatic Contributions to Residue-Specific Protonation Equilibria and Proton Binding Capacitance for a Small Protein<sup>†</sup>

Stina Lindman,<sup>\*,‡</sup> Sara Linse,<sup>\*,‡</sup> Frans A. A. Mulder,<sup>§</sup> and Ingemar André<sup>‡,||</sup>

Department of Biophysical Chemistry, Lund University, Chemical Center, SE-22100 Lund, Sweden, and Biophysical Chemistry, Groningen University, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Received August 1, 2006; Revised Manuscript Received September 19, 2006

**ABSTRACT:** Charge–charge interactions in proteins are important in a host of biological processes. Here we use <sup>13</sup>C NMR chemical shift data for individual aspartate and glutamate side chain carboxylate groups to accurately detect site-specific protonation equilibria in a variant of the B1 domain of protein G (PGB1-QDD). Carbon chemical shifts are dominated by changes in the electron distribution within the side chain and therefore excellent reporters of the charge state of individual groups, and the data are of high precision. We demonstrate that it is possible to detect local charge interactions within this small protein domain that stretch and skew the chemical shift titration curves away from “ideal” behavior and introduce a framework for the analysis of such convoluted data to study local charge–charge interactions and electrostatic coupling. It is found that, due to changes in electrostatic potential, the proton binding affinity,  $K_a$ , of each carboxyl group changes throughout the titration process and results in a linearly pH dependent  $pK_a$  value. This result could be readily explained by calculations of direct charge–charge interactions based on Coulomb’s law. In addition, the slope of  $pK_a$  versus pH was dependent on screening by salt, and this dependence allowed the selective study of charge–charge interactions. For PGB1-QDD, it was established that mainly differences in self-energy, and not direct charge–charge interactions, are responsible for shifted  $pK_a$  values within the protein environment.

Electrostatic interactions are fundamentally important in defining protein function. For example, charge–charge interactions modulate the binding of small ligands and proteins, provide essential interactions for chemistry in enzyme catalysis, are critical for stability, and engender protein solubility. The determination of  $pK_a$  values provides a sensitive method for experimentally accessing the electrostatic properties of a protein, and for testing electrostatic theories. Since the separation of contributions from self-energy and charge–charge interactions to the solvation energy is difficult to assess experimentally, experiments aimed at their study are of considerable interest.

Significant effort has been devoted to theoretical descriptions of proton binding phenomena in proteins. Continuum electrostatics within the framework of the Poisson–Boltzmann equation are typically used to calculate  $pK_a$  values of proteins (1, 2). The dielectric properties of the protein interior and the nature of the dielectric boundary with solvent are fundamentally unknown and represent one of the largest uncertainties in computational bioelectrostatics (3, 4). Many different approaches to increasing the accuracy of calculations of electrostatic interactions in proteins have been taken.

These include using multiple side chain conformations (5), inclusion of protein relaxation (6), and first-principles molecular mechanics (7). In spite of these advances, simulation of electrostatic properties of biomolecules remains a very challenging task.

Experimentally, the favored method for characterizing electrostatic interactions in proteins is NMR spectroscopy, where  $pK_a$  values for titrating residue side chains are determined from the measurement of chemical shifts of nuclei in the immediate vicinity as a function of pH. For aspartate (Asp) and glutamate (Glu), the most sensitive and specific way to determine  $pK_a$  values is to record the <sup>13</sup>C chemical shifts of carboxyl carbons (8–11). Since the electron density of the carboxyl carbon is significantly changed upon (de)protonation, the chemical shift will almost exclusively report on that specific titration event. Furthermore, the ionization-dependent chemical shift change is large, and this is necessary to yield very accurate data on charge–charge interactions and changes in the electrostatic potential with pH. As shown previously (8, 12, 13), proton data tend to be highly convoluted and therefore give much less accurate and reliable information about  $pK_a$  values in proteins.

The B1 domain of protein G (PGB1)<sup>1</sup> is used here as a model in investigating electrostatic interactions on the surface of proteins. PGB1 contains almost exclusively surface

<sup>†</sup> This work was supported by the Swedish Research Council (S. Linse).

<sup>\*</sup> To whom correspondence should be addressed: Lund University, Chemical Center, SE-22100 Lund, Sweden. Telephone: +46-46-2227092 or +46-46-2228246. Fax: +46-46-2224543. E-mail: stina.lindman@bpc.lu.se or sara.linse@bpc.lu.se.

<sup>‡</sup> Lund University.

<sup>§</sup> Groningen University.

<sup>||</sup> Present address: Department of Biochemistry, University of Washington, Box #357350, Seattle, WA 98195.

<sup>1</sup> Abbreviations: NMR, nuclear magnetic resonance; pI, isoelectric point; PGB1, protein G B1 domain; PGB1-QDD, PGB1-T2Q/N8D/N37D; wt, wild type.

exposed charged residues (14, 15) and exhibits distinct pH-dependent protein stability (16), resulting from electrostatic interactions (16, 17). Using a variant of PGB1 (T2Q/N8D/N37D) (17), residue-specific  $pK_a$  values were determined both from the carboxyl carbon and from the  $\beta$ - and  $\gamma$ -protons at low salt and at 0.5 M NaCl.

In this article, we present an array of techniques for analyzing titration data and dissecting electrostatic coupling in proteins. It is shown that the pH dependence of  $pK_a$  values of titrating groups can be determined directly from experimental data and that most groups display linear  $pK_a$  changes with pH. This linear pH dependence can be explained by direct charge–charge interactions as indicated by simulations based on Coulomb's law. The addition of salt leads to a reduced pH dependence of  $pK_a$  values due to electrostatic screening. The selective study of charge–charge interactions and the comparison of  $pK_a$  values at low and high salt provide information about the importance of Coulombic interactions in shifting  $pK_a$  values. These results will prove valuable in the development and calibration of electrostatic theories applied to proteins.

## MATERIALS AND METHODS

**NMR Spectroscopy.** Details about the cloning, expression, and purification of  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled PGB1-QDD have been reported previously (17). The side chain carboxyl groups and backbone carbonyl groups were assigned by correlating the side chain carboxyl  $^{13}\text{C}$  chemical shifts with the backbone amide proton chemical shift of the following residue as described by Tollinger et al. (10). The titration of acidic groups in PGB1-QDD was followed using a H(C)CO experiment where the side chain carboxyl carbon is correlated with the  $\text{H}^\beta$  or  $\text{H}^\gamma$  protons of aspartic or glutamic acid, respectively (18). At each pH value, a H(C)CO spectrum was recorded using 16 transients with 512 and 128 complex points in  $\omega_2$  ( $^1\text{H}$ ) and  $\omega_1$  ( $^{13}\text{C}$ ), respectively. All spectra were processed with a sine bell, typically shifted  $90^\circ$  in the indirect dimension, and the data size was doubled by zero-filling prior to Fourier transformation.

**NMR Sample Preparation and pH Titration.** PGB1-QDD labeled with  $^{15}\text{N}$  and  $^{13}\text{C}$  was dissolved in a 90%  $\text{H}_2\text{O}/10\%$   $\text{D}_2\text{O}$  mixture containing 0.1 mM  $\text{NaN}_3$  and 0.1 mM 2,2-dimethylsilapentane-5-sulfonic acid (DSS) to a protein concentration of 1.5 mM, and the initial NaCl concentration was 0 or 0.5 M. The initial pH was 5.0 for the sample without salt, and the pH of the sample containing 0.5 M NaCl was 4.7. In both cases, the sample was split in two so the same starting point could be used for upward and downward titrations. For the low-salt sample used for upward titrations, there was a systematic pH shift of  $-0.1$  unit for which data were corrected.

The pH was measured with a MP225 pH-meter equipped with a combination electrode [U402-M3-S7/200, calibrated with pH 4.01, 7.00, 9.21, and 11.0 standard solutions (Mettler Toledo)]. The pH was adjusted with 0.2 M HCl or NaOH, and the pH was measured before and after each titration point; in all cases, the variation before and after acquisition was no more than 0.05 pH unit. In the adjustment of pH during downward and upward titrations, the total added salt concentration amounted to 0.03 and 0.006 M, respectively.

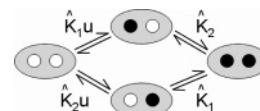


FIGURE 1: Microstates and microscopic binding constants of a two-site model. The state to the left with two white circles corresponds to a fully deprotonated state, while the state to the right is a fully protonated state.  $K_1$  and  $K_2$  are microscopic binding constants, and  $u$  is a pair interaction parameter that compensates for the reduced binding affinity of one site when the other site is deprotonated.

The pH was corrected by 0.04 to account for the 10%  $\text{D}_2\text{O}$  in the sample (19). Spectra were collected in steps of approximately 0.25 pH unit from pH 1.5 to 9.0.

**Data Analysis.** The observed  $^{13}\text{C}$  chemical shifts for the side chain carboxyl and  $^1\text{H}$  chemical shifts for  $\text{H}_\beta$  and  $\text{H}_\gamma$  protons were plotted and fitted with nonlinear least-squares regression analysis using MatLab 7 (The MathWorks, Natick, MA) or Kaleidagraph (Synergy Software). For the fitting procedure using MatLab, errors were MC-simulated using the standard deviation in spectra but do not include the uncertainty in pH which was estimated to be  $\pm 0.05$  pH unit. The chemical shift data were fitted to the Henderson–Hasselbalch equation without (eq 1) or including a Hill parameter ( $n_H$ ) (eq 2) in the model.

$$\delta_{\text{obs}} = \frac{\delta_{\text{HA}} + \delta_{\text{A}^-} \times 10^{\text{pH}-pK_a}}{1 + 10^{\text{pH}-pK_a}} \quad (1)$$

$$\delta_{\text{obs}} = \frac{\delta_{\text{HA}} + \delta_{\text{A}^-} \times 10^{n_H(\text{pH}-pK_a)}}{1 + 10^{n_H(\text{pH}-pK_a)}} \quad (2)$$

where  $\delta_{\text{obs}}$ ,  $\delta_{\text{HA}}$ , and  $\delta_{\text{A}^-}$  are the observed, protonated, and deprotonated chemical shifts, respectively.  $\Delta\delta$  in Tables 1, 2 and 3 corresponds to the chemical shift difference between the deprotonated and protonated forms.  $\text{pH} = -\log_{10}(a_{\text{H}^+})$ ;  $pK_a = -\log_{10} K_a$ , and the acid dissociation constant  $K_a = (a_{\text{H}^+}a_{\text{A}^-})/a_{\text{HA}}$  with the assumption  $\gamma_{\text{A}^-} = \gamma_{\text{HA}} = 1$ , where  $a_x$  and  $\gamma_x$  are the chemical activity and activity factor of species  $x$ , respectively. We refer, throughout the paper, to a two-state titration curve, eq 1, as being “ideal”. Multisite titration curves were modeled as

$$\delta_{\text{obs}} = \sum_i \frac{\delta_{\text{HA}}^i + \delta_{\text{A}^-}^i \times 10^{\text{pH}-pK_{ai}}}{1 + 10^{\text{pH}-pK_{ai}}} \quad (3)$$

where each term,  $i$ , represents one site. The improvement of a more complicated model to fit the experimental data was tested with the F-test (21).

In the alternative approach, data for two electrostatically coupled sites were fitted by a microscopic two-state model (22). In this situation, the protein has four microstates connected by four microscopic binding constants but can alternatively be completely described by two microscopic binding constants and a coupling parameter. The model is illustrated in Figure 1. The degree of protonation at site 1 as a function of pH,  $f_1(\text{pH})$ , is calculated from

$$f_1(\text{pH}) = \frac{\hat{K}_1 u a_{\text{H}^+} + \hat{K}_1 \hat{K}_2 u a_{\text{H}^+}^2}{1 + (\hat{K}_1 + \hat{K}_2) u a_{\text{H}^+} + \hat{K}_1 \hat{K}_2 u a_{\text{H}^+}^2} \quad (4)$$

where  $\hat{K}_1$  and  $\hat{K}_2$  are microscopic binding constants and  $u$  is a pair interaction parameter. The symbols are explained

further in Figure 1. In this formalism,  $pK_a = \log_{10}(\hat{K})$ . The chemical shift varies with  $f_1$  as

$$\delta_{\text{obs},1} = \delta_{\text{HA},1}f_1 + \delta_{\text{A}^-,1}(1 - f_1) \quad (5)$$

where  $\delta_{\text{obs},1}$ ,  $\delta_{\text{HA},1}$ , and  $\delta_{\text{A}^-,1}$  are the observed, protonated, and deprotonated chemical shifts of site 1, respectively.  $\hat{K}_1$  and  $\hat{K}_2$  can be equated to  $1/K_a$  for each dissociation event in the absence of electrostatic coupling.

**pH Dependence of  $pK_a$  Values.** Differences in electrostatic interaction energy at each titrating site as a function of pH are manifested as changes in proton binding affinities,  $K_a(\text{pH})$ , with pH. The fraction of charged residue at each pH was determined from

$$f_{\text{A}^-} = \frac{\delta_{\text{obs}} - \delta_{\text{HA}}}{\delta_{\text{A}^-} - \delta_{\text{HA}}} \quad (6)$$

where  $\delta_{\text{obs}}$ ,  $\delta_{\text{HA}}$ , and  $\delta_{\text{A}^-}$  are as defined above. Due to numerical complications for fractions close to 0 and 1, an  $f_{\text{A}^-}$  between 0.01 and 0.99 was used in the calculations.

The pH dependence of the  $pK_a$  values can be calculated as

$$pK_a(\text{pH}) = -\log_{10}\left(\frac{f_{\text{A}^-} \times 10^{-\text{pH}}}{1 - f_{\text{A}^-}}\right) \quad (7)$$

Alternatively, eq 7 can be written in terms of a pH-dependent free energy,  $\Delta G(\text{pH})$

$$\Delta G(\text{pH}) = -RT \ln\left(\frac{f_{\text{A}^-} \times 10^{-\text{pH}}}{1 - f_{\text{A}^-}}\right) \quad (8)$$

We define a parameter,  $\epsilon_H$ , that is a measure of the effect of screening by salt at each site as

$$\epsilon_H = \frac{\frac{\partial pK_a}{\partial \text{pH}}(I_2)}{\frac{\partial pK_a}{\partial \text{pH}}(I_1)} \quad (9)$$

where  $I_1$  and  $I_2$  are two ionic strengths. As shown below,  $\epsilon_H$  can be calculated from the Hill parameter.

**Physicochemical Basis of the Hill Parameter.** The Hill parameter in eq 2 is not a strictly defined quantity but has been used as a convenient way of fitting nonideal titration curves (20, 23). It can be given a physical meaning as follows. Let  $K_a$  be defined by

$$K_a = \frac{a_{\text{A}^-}a_{\text{H}^+}}{a_{\text{HA}}} = \frac{c_{\text{A}^-}\gamma_{\text{A}^-}a_{\text{H}^+}}{c_{\text{AH}}\gamma_{\text{HA}}} \quad (10)$$

where  $c_x$ ,  $a_x$ , and  $\gamma_x$  are the concentration, chemical activity, and activity factor of species  $x$ , respectively. The change in chemical potential upon deprotonation of a residue,  $\Delta\mu$ , is equal to  $-1/\beta \ln(\gamma_{\text{A}^-}/\gamma_{\text{HA}})$ , where  $\beta$  ( $1/kT$ ) is the Boltzmann factor. From eq 10, the fraction charged residue  $f_{\text{A}^-}$  can be found and equated with a similar expression from the modified Henderson–Hasselbalch expression (eq 2) as

$$f_{\text{A}^-} = \frac{1}{1 + 10^{\text{pH} - pK_a + \beta\Delta\mu/\ln 10}} = \frac{1}{1 + 10^{n_H(\text{pH} - pK_a)}} \quad (11)$$

The change in electrostatic interaction energy with pH can then be interpreted in terms of the Hill parameter

$$\frac{\partial(\beta\Delta\mu)}{\partial \text{pH}} = (n_H - 1) \times \ln 10 \quad (12)$$

Equation 12 shows that if a titration curve is described well by eq 2, the electrostatic energy at a site is changing linearly with pH. Furthermore, the reduction of repulsive charge–charge interactions due to salt screening can now be expressed in terms of the Hill parameter as

$$\epsilon_H = \frac{n_H(I_2) - 1}{n_H(I_1) - 1} \quad (13)$$

**Simulation of the Effect of Charge–Charge Interactions on Titration Curves.** The interaction of two charges in the protein was assumed to be described by Coulomb's law

$$U_{ij} = \frac{q_i q_j}{4\pi\epsilon\epsilon_0 r} \quad (14)$$

where  $q_i$  and  $q_j$  are the charges of residues  $i$  and  $j$ , respectively,  $r$  is the distance between them,  $\epsilon_0$  is the vacuum permittivity, and  $\epsilon$  is the dielectric constant. The dielectric constant was set to 78 in this simulation. The total electrostatic energy is equal to

$$U_{\text{tot}} = \sum_{i=1} \sum_{j < i} U_{ij} \quad (15)$$

The average number of protons bound to each site at a specific pH can be found by calculating the statistical mechanical partition function of the system as

$$\bar{x} = \sum x \exp(-\beta U)/Z \quad (16)$$

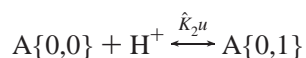
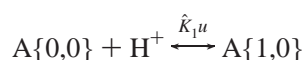
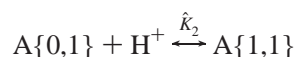
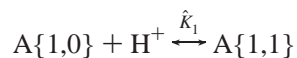
where  $Z$  is the partition function and  $\beta$  the Boltzmann factor. The protonation state ensemble is sampled with the standard Metropolis Monte Carlo method (24) using the following Boltzmann factor

$$\exp[-\Delta U_{\text{tot}}/kT \pm \ln 10(\text{pH} - pK_a^{\text{int}})] \quad (17)$$

where  $\Delta U_{\text{tot}}$  is the change in total energy due to the introduction or removal of a charge at a particular proton binding site.  $pK_a^{\text{int}}$  is equal to the  $pK_a$  value for the proton binding in water (the model compound value) plus the energy of transferring the charge to the proton binding site in a neutralized protein. The use of model  $pK_a$  values as an approximation of  $pK_a^{\text{int}}$  does not account for desolvation penalties, etc., and leads to a situation in which calculated  $pK_a$  values significantly deviate from experimentally determined ones. Since the shape of the titration curves is dependent on the  $pK_a$  values at which the residues titrate, the model was extended by allowing  $pK_a^{\text{int}}$  to change in the course of the simulation so that the difference between simulated and experimental  $pK_a$  values was minimized.  $pK_a^{\text{int}}$  values of acidic groups were randomly selected, and their values changed by 0.2 unit. The differences between the sums of all calculated and experimental fractions of acid for all residues at five different pH values (pH 3–5) were calculated, and this sum was minimized using the steepest descent

method. This generated new  $pK_a^{\text{int}}$  model values which were used as the basis for the simulation of the titration curves. The distances between charged residues were extracted from the wt PGB1 structure [PDB entry 1PGB (15)], and D8 and D37 were modeled into the structure using Swiss model (25). The correlation coefficient of calculated and experimental  $n_H$  values using this procedure was 0.93.

Titration curves were also modeled using a two- and three-site model. In the two-site model, the protein has four microstates with four microscopic binding constants. In the deprotonated state, A{0,0}, the microscopic  $K$  values are modified with a factor  $u$  due to electrostatic coupling. For this system, the following equilibria can be constructed:



where  $\hat{K}_1$  and  $\hat{K}_2$  are microscopic binding constants for sites 1 and 2, respectively. The degree of protonation at sites 1 [ $f_1(\text{pH})$ ] and 2 [ $f_2(\text{pH})$ ] as a function of pH can be calculated as

$$f_1(\text{pH}) = \frac{\hat{K}_1 u a_{H^+} + \hat{K}_1 \hat{K}_2 u a_{H^+}^2}{1 + (\hat{K}_1 + \hat{K}_2) u a_{H^+} + \hat{K}_1 \hat{K}_2 u a_{H^+}^2} \quad (18)$$

$$f_2(\text{pH}) = \frac{\hat{K}_2 u a_{H^+} + \hat{K}_1 \hat{K}_2 u a_{H^+}^2}{1 + (\hat{K}_1 + \hat{K}_2) u a_{H^+} + \hat{K}_1 \hat{K}_2 u a_{H^+}^2} \quad (19)$$

A three-site model can be constructed in the same manner as the two-site model but with three microscopic  $K$  values ( $\hat{K}_1$ ,  $\hat{K}_2$ , and  $\hat{K}_3$ ), three coupling factors ( $u_{12}$ ,  $u_{13}$ , and  $u_{23}$ ), and a triplet interaction factor ( $v_{123}$ ).  $K_1$ – $K_3$  are the macroscopic binding constants. The corresponding equations for the fraction protons bound to sites 1, 2, and 3 become

$$f_1(\text{pH}) = p(1,0,0)P_1 + [p(1,1,0) + p(1,0,1)]P_2 + P_3$$

$$f_2(\text{pH}) = p(0,1,0)P_1 + [p(1,1,0) + p(0,1,1)]P_2 + P_3$$

$$f_3(\text{pH}) = p(0,0,1)P_1 + [p(1,0,1) + p(0,1,1)]P_2 + P_3$$

where

$$P_n = K_n a_{H^+}^n / \sum_{n=0}^n K_n a_{H^+}^n$$

$$K_1 = (\hat{K}_1 u_{12} u_{13} + \hat{K}_2 u_{12} u_{23} + \hat{K}_3 u_{13} u_{23}) v_{123}$$

$$K_2 = (\hat{K}_1 \hat{K}_2 + \hat{K}_1 \hat{K}_3 + \hat{K}_2 \hat{K}_3) u_{12} u_{13} u_{23} v_{123}$$

$$K_3 = \hat{K}_1 \hat{K}_2 \hat{K}_3 u_{12} u_{13} u_{23} v_{123}$$

$$p(1,0,0) = \hat{K}_1 u_{12} u_{13} / (\hat{K}_1 u_{12} u_{13} + \hat{K}_2 u_{12} u_{23} + \hat{K}_3 u_{13} u_{23})$$

$$p(0,1,0) = \hat{K}_2 u_{12} u_{23} / (\hat{K}_1 u_{12} u_{13} + \hat{K}_2 u_{12} u_{23} + \hat{K}_3 u_{13} u_{23})$$

$$p(0,0,1) = \hat{K}_3 u_{13} u_{23} / (\hat{K}_1 u_{12} u_{13} + \hat{K}_2 u_{12} u_{23} + \hat{K}_3 u_{13} u_{23})$$

$$p(1,1,0) = \hat{K}_1 \hat{K}_2 / (\hat{K}_1 \hat{K}_2 + \hat{K}_1 \hat{K}_3 + \hat{K}_2 \hat{K}_3)$$

$$p(1,0,1) = \hat{K}_1 \hat{K}_3 / (\hat{K}_1 \hat{K}_2 + \hat{K}_1 \hat{K}_3 + \hat{K}_2 \hat{K}_3)$$

$$p(0,1,1) = \hat{K}_2 \hat{K}_3 / (\hat{K}_1 \hat{K}_2 + \hat{K}_1 \hat{K}_3 + \hat{K}_2 \hat{K}_3)$$

## RESULTS

The charge states of all carboxyl groups in this protein were determined as a function of pH from two-dimensional H(C)CO correlation spectra. The spectra are very well resolved, and all 12 side chain carboxyl groups and the carboxy terminus were monitored. Residue-specific  $pK_a$  values were determined from  $^{13}\text{C}$  chemical shifts of the carboxyl carbon and  $^1\text{H}$  chemical shifts of the  $\beta$ -protons (Asp) and  $\gamma$ -protons (Glu). The results are presented in Tables 1 and 3. For reference, model  $pK_a$  values for Asp, Glu, and the C-terminus are 4.0, 4.4, and 3.8, respectively (26). It is observed that a number of residues have  $pK_a$  values that deviate significantly from those obtained in water milieu. For example, D22, D47, and the C-terminus have significantly downshifted  $pK_a$  values of 3.0, 3.2, and 3.1, respectively, whereas D37 has a highly upshifted  $pK_a$  value of 6.5. It is apparent from Table 1 that at 0.5 M NaCl the  $pK_a$  values are closer to the model  $pK_a$  values for all carboxyl groups, but significant  $pK_a$  shifts are still present. This behavior is analyzed in more detail below.

*Analysis of Titration Data and Extraction of  $pK_a$  Values.* An example of the experimental data obtained for PGB1-QDD is given in Figure 2a. The curve shown for D8 represents the Henderson–Hasselbalch model, eq 1, fit to the carboxyl carbon chemical shift for this residue measured for 29 pH values. The fit to experimental data is far from perfect, which is expected for a protein with simultaneously titrating groups in the vicinity. Indeed, eq 1 fails to describe the experimental data for seven residues and the C-terminus in PGB1-QDD. It is common practice in such cases to introduce a cooperativity parameter, usually termed the Hill parameter,  $n_H$  (eq 2), to obtain an improved fit to the experimental data (20, 23), which often appear stretched. The results of including a Hill parameter in the fitting of the experimental  $^1\text{H}$  and  $^{13}\text{C}$  data are presented in Table 1. The fits are significantly improved using eq 2, as shown in Figure 2b for D8, although systematic deviations are still observed in some cases. As shown by Figure 2, these small systematic deviations between the fit and the experimental data points can be detected only with high-accuracy titration data. For example, for a number of residues, there is a significant deviation between experimental data and the fit to eq 2 between pH 5 and 6. A multisite model, eq 3, was therefore tested in fitting the data. As judged from the F-test, this two-site model yields significantly better fits for residues D8, D36, D40, and E42 and the C-terminus, when compared to eq 2. The improved fit for D8 is apparent from Figure 2c. Interestingly, the  $pK_a$  values determined by the different fits do not vary much for D36 and the C-terminus but slightly



Table 1: pH Titrations<sup>a</sup>

residue	low salt concentration			0.5 M salt			$\epsilon_H$	$\epsilon_{pK_a}$
	$pK_a$ for $^{13}C$	Hill parameter	$\Delta\delta$ (ppm)	$pK_a$ for $^{13}C$	Hill parameter	$\Delta\delta$ (ppm)		
D8	4.91 $\pm$ 0.02	0.67 $\pm$ 0.02	3.90 $\pm$ 0.02	4.69 $\pm$ 0.01	0.82 $\pm$ 0.02	3.90 $\pm$ 0.02	0.55	0.76
E15	4.63 $\pm$ 0.01	0.88 $\pm$ 0.02	4.14 $\pm$ 0.02	4.65 $\pm$ 0.01	0.94 $\pm$ 0.02	4.16 $\pm$ 0.02	0.50	
E19	3.90 $\pm$ 0.02	1.02 $\pm$ 0.03	3.54 $\pm$ 0.03	4.03 $\pm$ 0.02	0.97 $\pm$ 0.03	3.71 $\pm$ 0.03		0.74
D22	2.99 $\pm$ 0.05	1.05 $\pm$ 0.09	1.56 $\pm$ 0.06	3.24 $\pm$ 0.04	1.13 $\pm$ 0.09	1.58 $\pm$ 0.04		0.75
E27	4.81 $\pm$ 0.01	0.83 $\pm$ 0.02	4.26 $\pm$ 0.02	4.88 $\pm$ 0.01	0.92 $\pm$ 0.02	4.27 $\pm$ 0.02	0.47	
D36	4.20 $\pm$ 0.02	0.82 $\pm$ 0.03	3.20 $\pm$ 0.03	4.14 $\pm$ 0.02	0.89 $\pm$ 0.03	3.23 $\pm$ 0.03	0.61	0.70
D37	6.51 $\pm$ 0.02	1.11 $\pm$ 0.04	3.60 $\pm$ 0.02	6.16 $\pm$ 0.01	1.13 $\pm$ 0.03	3.45 $\pm$ 0.02		0.86
D40	4.12 $\pm$ 0.02	0.72 $\pm$ 0.02	4.15 $\pm$ 0.04	4.07 $\pm$ 0.01	0.88 $\pm$ 0.02	4.08 $\pm$ 0.03	0.43	0.58
E42	4.81 $\pm$ 0.01	0.73 $\pm$ 0.02	4.34 $\pm$ 0.02	4.55 $\pm$ 0.01	0.88 $\pm$ 0.02	4.37 $\pm$ 0.02	0.44	0.37
D46	3.84 $\pm$ 0.02	0.90 $\pm$ 0.03	2.99 $\pm$ 0.03	3.87 $\pm$ 0.02	0.95 $\pm$ 0.03	3.03 $\pm$ 0.03		0.81
D47	3.15 $\pm$ 0.02	0.92 $\pm$ 0.03	3.96 $\pm$ 0.05	3.35 $\pm$ 0.02	0.97 $\pm$ 0.03	3.84 $\pm$ 0.04	0.38	0.76
E56	3.81 $\pm$ 0.05	0.56 $\pm$ 0.02	2.99 $\pm$ 0.06	3.88 $\pm$ 0.03	0.63 $\pm$ 0.02	3.08 $\pm$ 0.05	0.84	0.88
C-terminus	3.13 $\pm$ 0.04	0.69 $\pm$ 0.03	3.59 $\pm$ 0.08	3.23 $\pm$ 0.03	0.76 $\pm$ 0.03	3.59 $\pm$ 0.06	0.77	0.85

<sup>a</sup>  $pK_a$  values and Hill parameters ( $n_H$ ) derived by fitting eq 2 to titration data for  $^{13}C$  chemical shifts. The shift differences over the pH range are also indicated.  $\epsilon_H$  and  $\epsilon_{pK_a}$  are calculated from eqs 13 and 20, respectively, and report on the electrostatic interactions in the system. For some residues, the  $pK_a$  or  $n_H$  change between  $I_1$  and  $I_2$  was insignificant, and hence,  $\epsilon_H$  or  $\epsilon_{pK_a}$  could not be obtained.

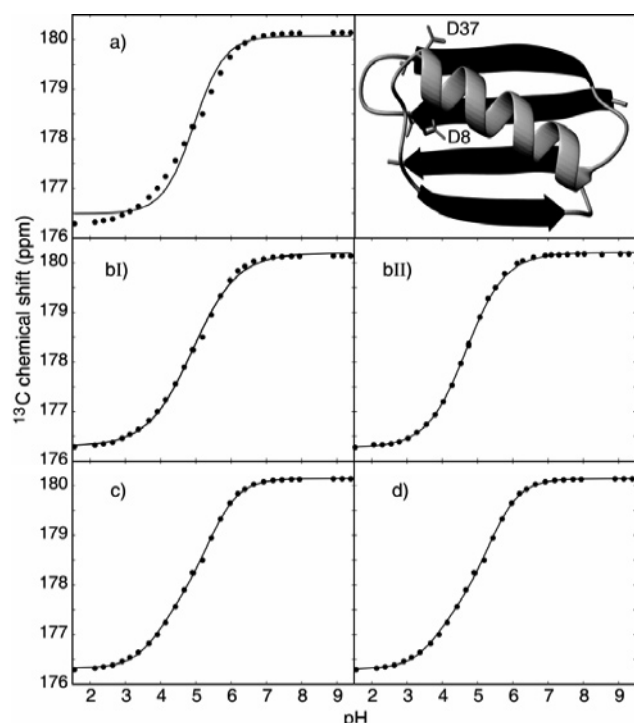


FIGURE 2: Titration curves for D8 with fitted curves using (a) eq 1, (b) eq 2 (I, low salt; and II, 0.5 M NaCl), (c) eq 3, and (d) eq 5. At the top right is the structure for PGB1-QDD (the mutated charges are explicitly shown) generated using PDB entry 1PGB (15). D8 and D37 were modeled into the structure using Swiss model (25), and the figure was generated using MOLMOL (30).

more for D8, D40, and E42 (Figure 2 and Tables 1 and 2). For D8, D40, and E42, the chemical shift differences are large for both sites, and hence, the  $pK_a$  value determined from eq 2 is more an average of the biphasic behavior. There are several possible explanations for the multiphasic titration curves. In  $^1H$  NMR, these effects are usually attributed to changes in the local electron distribution upon titration of various nearby groups. Using  $^{13}C$  carboxyl groups as reporters should significantly minimize this problem due to the large chemical shift changes upon proton binding, as compared to long-range electrostatic coupling. The multiphasic titration curves could report on a pH-dependent conformational change such as a change in the hydrogen bonding pattern.

Onufriev et al. have shown that protein titration curves can generally be described by a set of noninteracting sites with quasi- $pK_a$  values (27). They used their model to fit experimental data of a diprotic acid with excellent results. However, the large number of sites in a protein renders it unfeasible for fitting experimental data for a protein to such a model. A more comprehensive way of fitting experimental data is to use a microscopic model of the titration process involving the interaction between charged residues (Figure 1). The number of parameters in such a model will increase rapidly with the number of sites included so that only a two-site model can be confidently fitted to the experimental data. A two-site model involves two microscopic  $pK_a$  values and one coupling parameter describing the electrostatic interaction between the sites (see eqs 4 and 5 and Figure 1). Importantly, the  $pK_a$  values in this model can be interpreted as those found in a protein without charge–charge interactions. Because the electrostatic interaction at one site is a sum of charge–charge interactions, the second  $pK_a$  value should be seen as an effective value. Fitting eq 5 to data yields an excellent fit to data with significance similar to that of eq 3 (Figure 2d). Two  $pK_a$  values were extracted for each residue together with a pair interaction parameter, which gives a measure of the electrostatic coupling at a site. These data are compiled in Table 2. Although the interpretation of the fitting parameters is complex, the titration data are accurately reproduced by the explicit model.

**Electrostatic Coupling.** Many of the residues display stretched titration curves, indicating significant electrostatic coupling in the system. The proton binding capability for each residue as a function of pH is displayed in the derivatives of the normalized titration curve. The derivative can be interpreted as a measure of the change in charge at each pH value, called charge capacitance (28), and is useful in analyzing nonideal and asymmetric titration curves. Analogous to the charge capacitance measure we introduce here, a quantity named binding capacity, which is the second derivative of binding potentials, was introduced by Di Cera et al. as a general way of describing coupled equilibria in ligand binding (29). They showed that the pH dependence of the binding capacity can be used to study cooperativity.

Panels a and b of Figure 3 show the titration curves for all acidic residues in PGB1-QDD at low salt concentrations.

Table 2: pH Titrations at Low Salt Concentrations

residue	$pK_a-1^a$	$pK_a-2^a$	$\Delta\delta-1$ (ppm) <sup>a</sup>	$\Delta\delta-2$ (ppm) <sup>a</sup>	$pK_a-1^b$	$pK_a-2^b$	$RT \ln(u)^b$ (kJ/mol)
D8	$4.00 \pm 0.09$	$5.39 \pm 0.05$	$1.43 \pm 0.08$	$2.40 \pm 0.08$	$4.41 \pm 0.06$	$4.2 \pm 0.1$	$4.5 \pm 0.3$
D36	$4.11 \pm 0.03$	$6.18 \pm 0.27$	$2.87 \pm 0.04$	$0.31 \pm 0.04$	$4.16 \pm 0.02$	$5.1 \pm 0.1$	$6 \pm 1$
D40	$3.83 \pm 0.05$	$5.12 \pm 0.13$	$2.9 \pm 0.1$	$1.1 \pm 0.1$	$3.96 \pm 0.03$	$4.3 \pm 0.1$	$3.7 \pm 0.3$
E42	$4.20 \pm 0.08$	$5.38 \pm 0.07$	$2.1 \pm 0.1$	$2.2 \pm 0.1$	$4.49 \pm 0.04$	$4.5 \pm 0.1$	$3.66 \pm 0.2$
C-terminus	$3.08 \pm 0.03$	$5.09 \pm 0.19$	$2.89 \pm 0.04$	$0.43 \pm 0.04$	$3.14 \pm 0.03$	$4.0 \pm 0.1$	$6.1 \pm 0.7$

<sup>a</sup>  $pK_a$  values from fitting eq 3 to data of some residues that exhibited a statistically significant (according to the F-test) better fit compared to fitting with eq 2. <sup>b</sup>  $pK_a$  values from fitting eq 5 to data of some residues that exhibited a statistically significant (according to the F-test) better fit compared to fitting with eq 2.

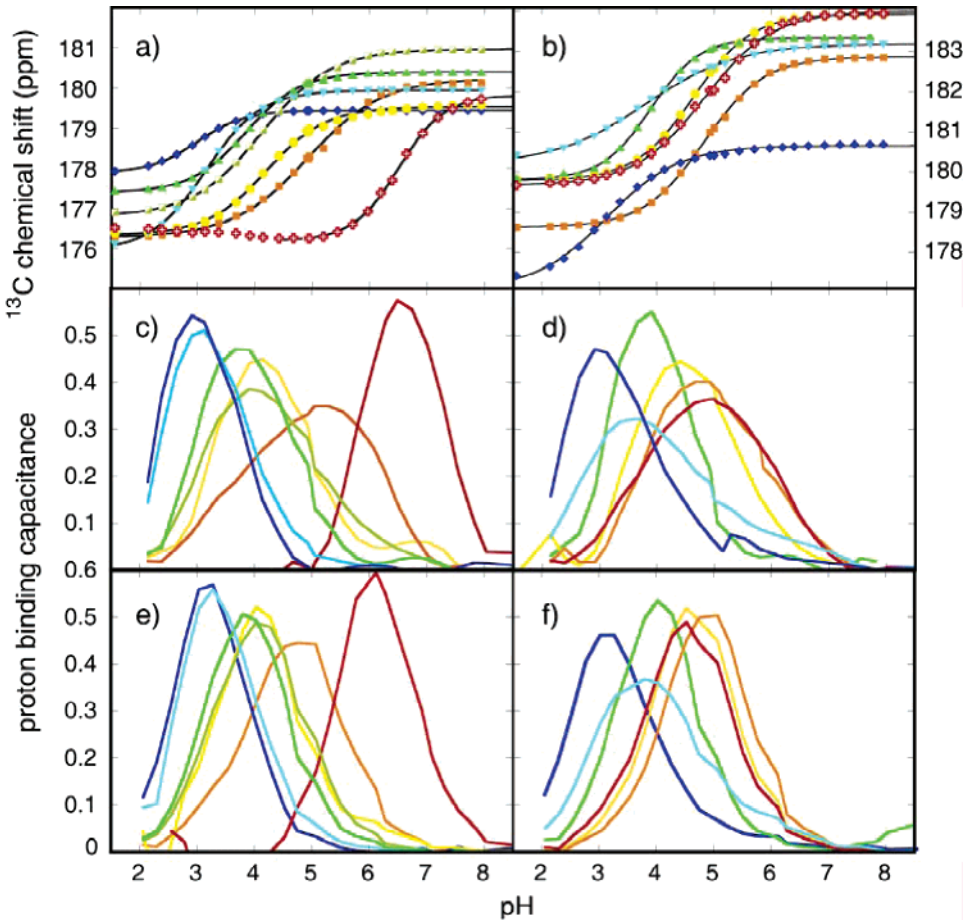


FIGURE 3: Titration curves of the seven Asp residues (a) and the five Glu residues and the C-terminus (b) in PGB1-QDD at low salt concentrations. (c–f) Capacitance curves for all carboxyl groups and the C-terminus in PGB1-QDD: (c) Asp at a low salt concentration, (d) Glu and the C-terminus at a low salt concentration, (e) Asp in 0.5 M salt, and (f) Glu and the C-terminus in 0.5 M salt. The capacitance of proton binding was calculated from the data points by taking the centered differences between subsequent titration points of the normalized titration curves: D8 (filled orange squares), D22 (filled blue diamonds), D36 (filled yellow circles), D37 (red pluses), D40 (empty green-yellow squares), D46 (filled green up triangles), D47 (filled turquoise down triangles), E15 (solid yellow circles), E19 (solid green up triangles), E27 (solid orange squares), E42 (red pluses), E56 (solid turquoise down triangles), and C-terminus (solid blue diamonds). Equation 2 was fit to data (—) in panels a and b.

The derivatives of the normalized titration data, coined capacitance curves, for low and high salt concentrations are shown in panels c–f of Figure 3. High charge capacitance will result in a narrower curve as a function of pH, while a titration process with a low charge capacitance yields a stretched shape compared to the ideal model. Because the titration curves are normalized, the area below the curves in panels c–f of Figure 3 is constant, and a maximum is found when  $pK_a = (\ln 10 \times n_H/4)$ . Highly upshifted (D37) and highly downshifted (D22 and D47) residues titrate in a pH interval where few or no other residues change charge state and therefore have charge capacitances close to 0.58, the

value expected when  $n_H = 1$ . This is shown in panels a and b of Figure 4 for D22 and D37, respectively. E56 and D8, on the other hand, exhibit stretched titration curves, and hence, the capacitance for binding a proton is lower at each pH point. It is observed that the derivatives are asymmetric, indicating that the charge capacitance changes during (de-)protonation (Figure 4c,d). A residue that starts to ionize at pH values lower than those of most other residues, such as E56, displays a lower capacitance at the high-pH end of the titration process where titration processes of other residues influence the ionization. The opposite behavior is obtained for D8, which is the second-to-last acidic residue to give up

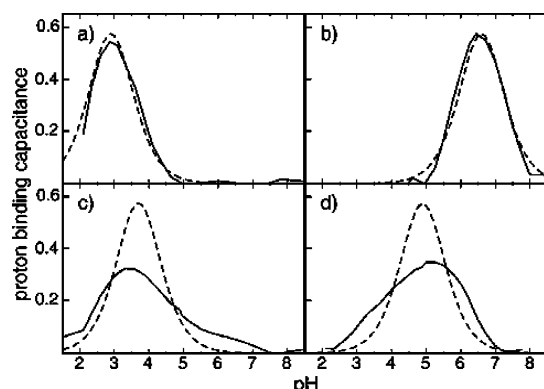


FIGURE 4: Capacitance curves under low-salt conditions. Capacitance curve for (a) D22 (—) and for a case where  $pK_a = 3.0$  and  $n_H = 1.0$  (---), (b) D37 (—) and a case where  $pK_a = 6.5$  and  $n_H = 1.0$  (---), (c) E56 (—) and a case where  $pK_a = 3.8$  and  $n_H = 1.0$  (---), and (d) D8 (—) and a case where  $pK_a = 4.9$  and  $n_H = 1.0$  (---).

its proton with an increase in pH. The curve can be explained by the influence of (un)charging all other acidic residues in the vicinity that affect the equilibrium of D8 directly through Coulombic forces. Evidently, the capacitance of proton binding provides an intuitive physical picture of electrostatic coupling in the system. For many residues, the observed titration curves and derivatives deviate significantly from ideal behavior. At low pH, few groups titrate and ionization equilibria become close to ideal. In the pH interval of 3.5–5.5, a large number of groups titrate, leading to significant coupling as witnessed by low charge capacitance and asymmetric titration curves. Surprisingly, E19 titrates in this interval and displays an almost ideal charge capacitance, indicating insignificant charge–charge repulsion at this site. As the pH becomes more neutral, very few residues titrate and titration processes again become close to ideal and described by eq 1. Titration curves do not display Hill parameters well above 1.0 since there are no positively charged residues titrating in this interval (see Table 1). The observation that the  $n_H$  of D37 equals 1.1 is currently unexplained but may be a result of its high  $pK_a$ . Since PGB1-QDD contains no histidines, the only basic group that is expected to titrate in this region is the amino terminus, but this interaction has not been verified.

**Screening of the Electrostatic Coupling by Salt.** All residues, except E19, display stretched titration curves, indicating that there is a coupled network of charges titrating in the same range resulting in lowered charge capacitance at each site. The addition of salt inhibits electrostatic coupling, yielding sharper titrations and higher capacitance. The asymmetric titration behavior seen in the derivative of the titration curve is also diminished by the addition of salt, and the data for D8 (Figure 2b, I and II) can now be better fitted by eq 2. The exception is the C-terminal residue, E56, which displays a stretched titration curve also in the presence of high salt concentrations. This can be explained by the inevitable coupling between the carboxyl groups in the main and side chains.

The effects of screening by salt on the Hill parameter, expressed by  $\epsilon_H$  (calculated from eq 13), are given in Table 1. For the majority of the carboxyl groups,  $\epsilon_H$  has a value of 0.3–0.5, indicating that most of the electrostatic coupling is lost with added salt. E56 side chain and main chain carboxylate have increased values of  $\sim 0.8$ . This indicates

that salt is not able to screen out the interaction between these two carboxyl groups due to their proximity.

**Electrostatic Coupling Inferred from  $^1H$  and  $^{13}C$  Data.** Table 3 lists the  $pK_a$  values obtained from the  $^1H$  chemical shifts measured in the H(C)CO spectra. As observed previously by several groups (8, 11, 13), different  $pK_a$  values are obtained for the two prochiral protons in the same amino acid side chain. For example, the  $pK_a$  values for D8 are reported to be 5.8 and 4.7 from the  $^1H$  data, whereas a value 4.9 is obtained for  $^{13}C$ . Here, the two protons give  $n_H$  values of 1.0 and 0.6, whereas carbon yields a value of 0.7. The values obtained for protons are less reliable because these display much smaller chemical shift differences than carbon. In addition, the amount of cooperativity, borne out by the Hill parameter, varies to a much greater extent (from 0.4 to 1.5) for  $^1H$  as opposed to  $^{13}C$  (from 0.6 to 1.1). As there is no reason to assume that  $^1H$  and  $^{13}C$  sense cooperative electrostatic interactions differently, this observation is a direct result of the difficulty of deciphering convoluted  $^1H$  titration data. These observations substantiate the need to utilize carbon chemical shift data for the accurate investigation of electrostatic interactions in proteins.

**Electrostatic Coupling Visualized by the pH Dependence of Individual  $pK_a$  Values.** The electrostatic environment at each titrating site in a protein changes as a function of pH due to the change in charge at other titrating sites. This leads to changes in electrostatic interaction energy at each titrating site as a function of pH and is manifested as a change in proton binding affinity,  $K_a(pH)$ , with pH. Without invoking any theoretical model, we can determine pH-dependent  $pK_a$  values directly from our raw experimental data. The input to this operation is the chemical shifts as a function of pH, and the only necessary parameters are the chemical shifts of the fully deprotonated and protonated species, which are used to derive the fraction of each particular residue in its ionized form,  $f_A^-$ , as a function of pH. The pH dependence of the  $pK_a$  values can then be calculated according to eq 7 or 8. For an ideal, noninteracting Asp or Glu residue, the  $pK_a$  values are independent of pH. However, for an interacting proton binding site, the  $pK_a$  value will change as a function of pH.

These data show that the  $pK_a$  for D22 and D37 is nearly independent of proton activity, whereas for D8 and E56, it is clearly pH dependent. This is illustrated in Figure 5a with a dashed line generated for each residue based on the curve fit using eq 2. The  $pK_a$  values of most residues are increasing linearly with pH in the examined interval ( $0.01 \leq f_A^- \leq 0.99$ ). The slope is proportional to the level of electrostatic coupling, with larger slopes indicating stronger coupling. Figure 5b shows a comparison of  $pK_a$  values at low and high salt concentrations for D8. The slope of the curve decreases with added salt, showing that the level of electrostatic coupling is decreased. When charge–charge interactions are totally screened out by salt, the slope of the pH dependence of the  $pK_a$  values should approach zero. Thus, the slope of the  $pK_a$  value curve can be used as a measure of the strength of electrostatic interactions.

**Origin of the Linear pH Dependence of Electrostatic Interactions.** All carboxyl groups were observed to display a linear pH dependence of the free energy of protonation, which is manifested as linearly pH dependent  $pK_a$  values. This is caused by repulsive charge–charge interactions

Table 3: Comparison of Apparent  $pK_a$  Values and Hill Parameters ( $n_H$ ) from the  $\beta$ - or  $\gamma$ -Protons<sup>a</sup>

residue	$pK_a$ for $^1\text{H}$ -1	Hill parameter for $^1\text{H}$ -1	$\Delta\delta$ (ppm) for $^1\text{H}$ -1	$pK_a$ for $^1\text{H}$ -2	Hill parameter for $^1\text{H}$ -2	$\Delta\delta$ (ppm) for $^1\text{H}$ -2
D8	5.8	1.0	-0.10	4.7	0.6	-0.49
E15	4.3	0.6	-0.39	4.6	1.2	-0.38
E19	3.7	1.2	-0.20	4.0	1.2	-0.43
D22	3.1	1.2	-0.34	3.0	1.0	-0.32
E27	4.8	1.1	-0.15	4.3	0.8	0.13
D36	4.0	1.0	-0.25	3.5	0.6	-0.26
D37	6.6	1.1	-0.17	6.1	0.4	-0.39
D40	3.9	0.6	-0.22	4.5	0.6	-0.22
E42	4.8	0.6	-0.24	5.1	0.8	-0.25
D46	3.8	1.1	-0.35	3.7	0.8	-0.20
D47	3.0	1.3	-0.18	3.1	1.5	-0.36
E56	3.9	0.5	-0.32			

<sup>a</sup> Equation 2 was fit to titration data for  $^1\text{H}$  chemical shifts. The reported  $pK_a$  values, Hill parameters, and chemical shift differences are determined within less than  $\pm 0.1$  pH unit.

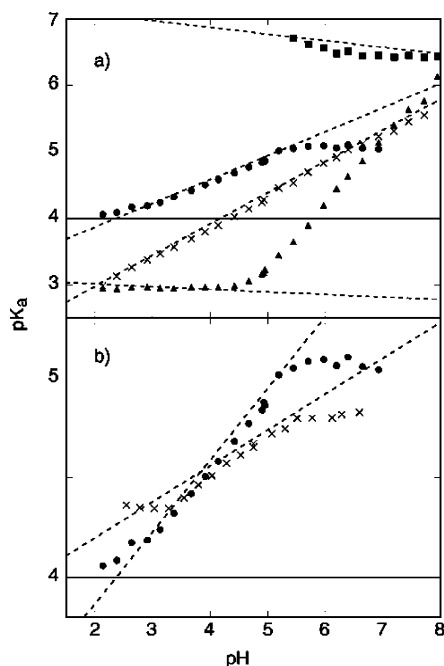


FIGURE 5: (a)  $pK_a$  values calculated from titration data at low salt concentrations as a function of pH: D8 (●), E56 (×), D37 (■), and D22 (▲).  $pK_a$  values were calculated from a curve fit with a Hill parameter for the respective residue (eq 2) (---). (b) Comparison of  $pK_a$  values at low salt concentrations and 0.5 M NaCl for D8: low salt (●), high salt (×), and curve fits to experimental data with a Hill parameter (eq 2) (---).

whose magnitudes change as a function of pH. However, it is not immediately evident why there is a linear pH dependence. To investigate whether Coulombic interactions can explain the linear pH dependence, a Monte Carlo simulation of electrostatic interactions was conducted to produce synthetic data. In this simulation, only direct charge–charge interactions were taken into account and interactions between charges were calculated using Coulomb’s law. The effective  $pK_a$  model values determined using this procedure were then used in the simulation of titration. If D37 (whose non-Coulombic shift could not be accurately determined since it was highly upshifted) and the C-terminus (which is likely to be flexible) are removed, the correlation coefficient between calculated and experimental  $n_H$  values is 0.93 using this approach (data not shown). This is a surprisingly good correlation given the primitive model and indicates that the shape of the titration curves can be

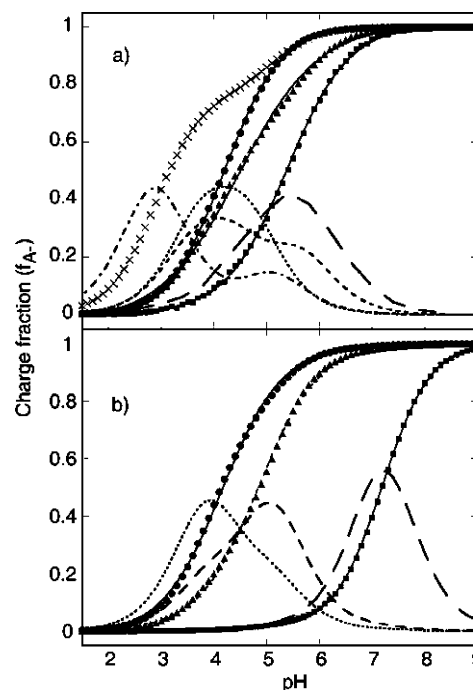


FIGURE 6: Simulated curves. Equation 2 was fit to data with parameters from each fit indicated in parentheses. (a) MC simulated titration curve for D8 assuming only Coulombic interactions (■) ( $pK_a = 5.4$ ;  $n_H = 0.74$ ) and simulated curve using a two-site model where both  $pK_a$  values are set to 4.0 with an interaction parameter of 0.4 (●) ( $pK_a = 4.2$ ;  $n_H = 0.82$ ), a three-site model where  $pK_a = 4.0$  and the interaction parameter = 0.4 (▲) ( $pK_a = 4.5$ ;  $n_H = 0.58$ ), and a two-site model where  $pK_a = 3.0$  and the interaction parameter is increased to 1.5 (×). (b) MC simulations of a three-site model with titration curves for the different sites:  $pK_a = 4.0$  for S1 (●) ( $pK_a = 4.2$ ;  $n_H = 0.76$ ),  $pK_a = 4.4$  for S2 (▲) ( $pK_a = 4.8$ ;  $n_H = 0.76$ ), and  $pK_a = 6.0$  for S3 (■) ( $pK_a = 7.2$ ;  $n_H = 0.94$ ). The capacitance curves are also shown in the picture.

accurately explained by Coulombic interactions. Equation 2 fits excellently to all points in the synthetic data, and hence, the free energy of protonation changes linearly with pH just as for the experimental data (Figure 6a). From the simulation, it is evident that the repulsive charge–charge interactions cause a linear pH dependence of the free energy of protonation, but why is this so?

To examine whether the linear pH dependence is caused by a multitude of coupled residues or can occur also in a simple system, a two-site model was evaluated. Simulated titration curves are shown in Figure 6a. For moderate values



of the interaction parameter, eq 2 fits well to the generated curves showing that the free energy of protonation depends linearly on pH. Only if the interaction energy is sufficiently high does a nonlinear dependence arise. When the difference in  $pK_a$  values between the two sites is increased, the two sites do not interact and ideal titration curves are observed. Hence, the linear pH dependence of the free energy of protonation is found also in the most simple model system of two interacting sites if the interaction energy is sufficiently low. The same trends are observed for a three-site model (Figure 6b). In Figure 6b, the simulations were performed for a three-site model where two of the sites have similar  $pK_a$  values while the  $pK_a$  value of the third site is 2 units higher. It is evident that the sites titrating close to each other in pH are electrostatically coupled and display extended titration curves while the highly shifted site titrates in a manner independent of the two other sites and shows a nearly pH independent  $pK_a$  value.

## DISCUSSION AND CONCLUSION

$pK_a$  values of titrating residues provide information about the solvation energy of charges at different sites in a protein. The solvation energy is a sum of different contributions, including direct charge–charge interactions. The latter are the primary source of pH-dependent electrostatic interactions that provide the protein with plasticity so that it can bind proteins, membranes, and ligands. Short-range electrostatic interactions between charges of opposite sign, termed salt bridges, have been shown to be important for the function of proteins and to affect stability. It is impossible from  $pK_a$  values alone to determine the contribution to shifts from charge–charge interactions. Analysis of titration curves provides a means of selectively investigating charge–charge interactions. In the presence of significant charge–charge interaction between residues, titration curves should deviate from the ideal case described by eq 1. Indeed, most residues exhibit a considerable change in interaction energy as a function of pH, and much of this pH dependence is lost with the addition of salt. The  $\epsilon_H$  values report that for most residues roughly 30–50% of the direct charge–charge interactions are remaining in 0.5 M NaCl. It should be pointed out that this is an approximate number since we are not determining the absolute electrostatic free energy but the change in this energy with pH. At the same time, most residues exhibit small  $pK_a$  shifts as a consequence of the addition of salt; on average,  $pK_a$  values shift 0.15 pH unit. This clearly indicates that charge–charge interactions are not the primary source of  $pK_a$  shifts for these residues. In particular, D37 is highly upshifted in low salt (6.5), but the  $pK_a$  value is reduced only 0.3 unit by the addition of 0.5 M salt. Just as the  $\epsilon_H$  parameter is a measure of the repulsive interactions that are lost with the addition of salt, we can introduce a new parameter,  $\epsilon_{pK_a}$ , as

$$\epsilon_{pK_a} = \frac{pK_{a,model} - pK_{a,I_2}}{pK_{a,model} - pK_{a,I_1}} \quad (20)$$

where  $pK_a$  values  $I_1$  and  $I_2$  correspond to different ionic strengths (Table 1). This parameter can, in combination with  $\epsilon_H$ , report on how much of the  $pK_a$  shift is caused by charge–charge interactions and how much is a result of non-pH-

dependent effects. Hence,  $\epsilon_{pK_a}$  could provide useful information about hydrogen bonding or buried charges in a protein. The calculated  $\epsilon_{pK_a}$  values in Table 1 show that the relative changes in  $pK_a$  values are much smaller than the relative changes reported by  $\epsilon_H$ . There is one important difference between  $\epsilon_H$  and  $\epsilon_{pK_a}$ :  $\epsilon_H$  values report on screening of repulsive interactions alone, whereas  $\epsilon_{pK_a}$  values report on screening of both repulsive and attractive interactions. If  $pK_a$  value shifts would arise mainly due to charge–charge interactions,  $\epsilon_H$  and  $\epsilon_{pK_a}$  would be expected to have similar magnitudes. Thus, the difference is a clear indicator that self-energy terms are dominant in shifting  $pK_a$  values in PGB1-QDD.

Theoretical calculations of  $pK_a$  values of proteins can provide valuable insights, but currently, many fundamental aspects of proteins, such as the dielectric properties, are still unknown which make accurate predictions complicated. Hence, accurate experimental  $pK_a$  value determinations are highly valuable, and the methods presented here can be used for calibration of electrostatic calculations. In particular, we suggest that not only  $pK_a$  values themselves but also their pH dependence could be used as parameters to judge the accuracy of a computational method. The correlation of calculated and experimental Hill parameters was studied by Georgescu et al. in the benchmarking of their computational method (5), but as we show here, data derived from proton chemical shifts are not generally reliable and can obscure the result. The detailed experimental data presented here will be valuable for investigators aiming to develop calculation schemes for predicting electrostatic effects from protein structure.

In conclusion, we provide an extensive analysis of high-resolution NMR titration data by which charge–charge interactions and electrostatic coupling in a protein can be selectively studied. All these analyses are performed with NMR chemical shifts as the input data. By taking the derivative of these titration curves, we show that the capacitance of binding a proton can change throughout the titration. The changing charge capacitance is further manifested in pH-dependent  $pK_a$  values. Furthermore, the extent of electrostatic coupling in a protein that is removed by salt can be captured. From this, it can be experimentally derived that differences in self-energy are mainly responsible for the atypical  $pK_a$  values within this small protein.

## ACKNOWLEDGMENT

We thank Prof. Bo Jönsson and Prof. Bengt Jönsson for valuable discussions about electrostatics and Mikael Lund for helpful insight into  $pK_a$  value calculations. Help with protein expression and purification by Hanna Nilsson is gratefully acknowledged.

## REFERENCES

1. Warwicker, J., and Watson, H. C. (1982) Calculation of the electric potential in the active site cleft due to  $\alpha$ -helix dipoles, *J. Mol. Biol.* 157, 671–9.
2. Gilson, M. K., Rashin, A., Fine, R., and Honig, B. (1985) On the calculation of electrostatic interactions in proteins, *J. Mol. Biol.* 184, 503–16.
3. Schutz, C. N., and Warshel, A. (2001) What are the dielectric “constants” of proteins and how to validate electrostatic models? *Proteins* 44, 400–17.

4. Simonson, T. (2001) Macromolecular electrostatics: Continuum models and their growing pains, *Curr. Opin. Struct. Biol.* **11**, 243–52.
5. Georgescu, R. E., Alexov, E. G., and Gunner, M. R. (2002) Combining conformational flexibility and continuum electrostatics for calculating pK<sub>a</sub>s in proteins, *Biophys. J.* **83**, 1731–48.
6. Sham, Y. Y., Chu, Z. T., and Warshel, A. (1997) Consistent calculations of pK<sub>a</sub>'s of ionizable residues in proteins: Semi-microscopic and microscopic approaches, *J. Phys. Chem. B* **101**, 4458–72.
7. Ivanov, I., Chen, B., Raugei, S., and Klein, M. L. (2006) Relative pK<sub>a</sub> values from first-principles molecular dynamics: The case of histidine deprotonation, *J. Phys. Chem. B* **110**, 6365–71.
8. Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y., and Nakamura, H. (1994) Individual ionization constants of all the carboxyl groups in ribonuclease HI from *Escherichia coli* determined by NMR, *Biochemistry* **33**, 5275–84.
9. Tollinger, M., Crowhurst, K. A., Kay, L. E., and Forman-Kay, J. D. (2003) Site-specific contributions to the pH dependence of protein stability, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 4545–50.
10. Tollinger, M., Forman-Kay, J. D., and Kay, L. E. (2002) Measurement of side-chain carboxyl pK<sub>a</sub> values of glutamate and aspartate residues in an unfolded protein by multinuclear NMR spectroscopy, *J. Am. Chem. Soc.* **124**, 5714–7.
11. Chen, H. A., Pfuhl, M., McAlister, M. S. B., and Driscoll, P. C. (2000) Determination of pK<sub>a</sub> values of carboxyl groups in the N-terminal domain of rat CD2: Anomalous pK<sub>a</sub> of a glutamate on the ligand-binding surface, *Biochemistry* **39**, 6814–24.
12. Oliveberg, M., Arcus, V. L., and Fersht, A. R. (1995) pK<sub>a</sub> values of carboxyl groups in the native and denatured states of barnase: The pK<sub>a</sub> values of the denatured state are on average 0.4 units lower than those of model compounds, *Biochemistry* **34**, 9424–33.
13. Forman-Kay, J. D., Clore, G. M., and Gronenborn, A. M. (1992) Relationship between electrostatics and redox function in human thioredoxin: Characterization of pH titration shifts using two-dimensional homo- and heteronuclear NMR, *Biochemistry* **31**, 3442–52.
14. Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., and Clore, G. M. (1991) A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G, *Science* **253**, 657–61.
15. Gallagher, T., Alexander, P., Bryan, P., and Gilliland, G. L. (1994) Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR, *Biochemistry* **33**, 4721–9.
16. Alexander, P., Fahnestock, S., Lee, T., Orban, J., and Bryan, P. (1992) Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: Why small proteins tend to have high denaturation temperatures, *Biochemistry* **31**, 3597–603.
17. Lindman, S., Xue, W.-F., Szczepankiewicz, O., Bauer, M. C., Nilsson, H., and Linse, S. (2006) Salting the Charged Surface: pH and Salt Dependence of Protein G B1 Stability, *Biophys. J.* **90**, 2911–21.
18. Yamazaki, T., Yoshida, M., and Nagayama, K. (1993) Complete assignments of magnetic resonances of ribonuclease H from *Escherichia coli* by double- and triple-resonance 2D and 3D NMR spectroscopies, *Biochemistry* **32**, 5656–69.
19. Glasoe, P. K., and Long, F. A. (1960) Use of glass electrodes to measure acidities in deuterium oxide, *J. Phys. Chem.* **64**, 188–90.
20. Markley, J. L. (1975) Observation of Histidine Residues in Proteins by Means of NMR Spectroscopy, *Acc. Chem. Res.* **8**, 70–80.
21. Bevington, P. R., and Robinson, D. K. (2003) *Data Reduction and Error Analysis*, McGraw-Hill, New York.
22. Borkovec, M., Jönsson, B., and Koper, G. J. M. (2001) Ionization Processes and Proton Binding in Polyprotic Systems: Small Molecules, Proteins, Interfaces, and Polyelectrolytes, in *Surface and Colloid Science* (Matijevic, E., Ed.) pp 169–99, Kluwer Academic/Plenum Publishers, New York.
23. Markley, J. L. (1973) Nuclear Magnetic-Resonance Studies of Trypsin-Inhibitors. Histidines of Virgin and Modified Soybean Trypsin-Inhibitor (Kunitz), *Biochemistry* **12**, 2245–50.
24. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953) Equation of State Calculations by Fast Computing Machines, *J. Chem. Phys.* **21**, 1087–92.
25. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling, *Electrophoresis* **18**, 2714–23.
26. Nozaki, Y., and Tanford, C. (1967) Examination of titration behavior, *Methods Enzymol.* **11**, 715–34.
27. Onufriev, A., Case, D. A., and Ullmann, G. M. (2001) A novel view of pH titration in biomolecules, *Biochemistry* **40**, 3413–9.
28. Lund, M., and Jönsson, B. (2005) On the charge regulation of proteins, *Biochemistry* **44**, 5722–7.
29. Di Cera, E., Gill, S. J., and Wyman, J. (1988) Binding capacity: Cooperativity and buffering in biopolymers, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 449–52.
30. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics* **14**, 51–5.

BI061555V