# Multiply-Protonated Protein Ions in the Gas Phase: Calculation of the Electrostatic Interactions between Charged Sites

**Maria Miteva,[†,‡] Plamen A. Demirev,[§] and Andrey D. Karshikoff*[,†]**

*Department of Bioscience at Novum, Karolinska Institute, 14157 Huddinge, Sweden, and Uppsala University, Division of Ion Physics, Box 534, 75121 Uppsala, Sweden*

*Received: July 10, 1997; In Final Form: September 3, 1997*[⊗]

A macroscopic approach for calculation of electrostatic interactions in proteins, initially developed for solution, was extended to describe isolated multiply-charged protein ions in the gas phase. It was combined with a Monte Carlo algorithm for determination of the most probable protonated sites for a different charge. Several quantities characterizing the behavior of protein ions in vacuum were calculated for nativelike lysozyme and ubiquitin. Among these are the apparent gas-phase basicity for various charge states, as well as the intrinsic basicities and the probabilities of protonation of individual sites. The contributions of the intramolecular solvation term and the peptide dipole-charge interactions to the intrinsic gas-phase basicity of each site were estimated. It was shown that the peptide dipoles may essentially influence the intrinsic gas-phase basicity of individual protonation sites. The approach can be successfully used to probe electrostatic interactions of gas-phase protein ions, provided information on the three-dimensional structure of these ions is available.

## Introduction

The three-dimensional structure of native proteins results from a nontrivial balance of different forces, mainly of a noncovalent nature. Among these forces, electrostatic interactions play an important role in a variety of functional properties of proteins. Their significance becomes evident at all pH-dependent phenomena involving proteins: protein−substrate/inhibitor interactions, pH-regulation of enzyme activity and stability, and many others. That is why the correct description of electrostatic interactions has turned out to be a high-priority task for a number of researchers.

The theory of electrostatic interactions in proteins was first developed for solutions based on a macroscopic approach. The protein molecule was represented as a material with a low dielectric constant immersed in the high-permittivity medium of the solvent. The Poisson−Boltzmann equation for this system was initially solved analytically for a spherical approximation of the protein molecule. The ionized side chains in this approximation are represented as point charges on the surface of the sphere.[1−4] Determination of the three-dimensional structure of a large number of proteins has stimulated the progress of the microscopic treatment of the problem.[5−7] The microscopic approach has been successfully applied for prediction of the solvation energy of small molecules,[8] calculation of intrinsic p*K* values of titratable groups in proteins,[7] analysis of mutational experiments involving charged side chains,[9] and many others. In parallel, the macroscopic approach has also undergone further elaboration. Warwicker and Watson[10] have proposed a numerical solution of the Poisson−Boltzmann equation, taking into consideration the actual shape of the protein macromolecule. Although the macroscopic approach has been criticized,[6,11] it is widely used for the analysis of electrostatic interactions in proteins in solution.[12−17]

Recently, electrospray ionization mass spectrometry has become a powerful tool for the analysis of proteins in the gas phase.[18] Identification, molecular mass, and primary structure elucidation of peptides and proteins in subnanomolar amounts can be routinely achieved by means of this technique.[19] An entirely new research field, addressing the nature of the interactions and high-order structures of isolated protein ions in a solvent-free environment, has emerged. Methods for probing the conformation space of protein ions in vacuum have been developed with direct implications for studies in protein folding and dynamics.[20] Among these techniques are determination of the nonreactive or reactive cross sections of protein ions with various small molecules in an ion drift tube[21−24] or in Fourier transform ion cyclotron resonance experiments.[25−31] Kaltashov and Fenselau[32] have used mass-analyzed ion kinetic-energy spectrometry to the measure intercharge repulsion and distances in multiprotonated peptide ions. McLafferty and coworkers[26] have examined various gas-phase conformations of protein cations (e.g., cytochrome *c*, myoglobin, RNAse, ubiquitin, etc.) by monitoring their hydrogen/deuterium exchange kinetics. These as well as other observations[33,34] reveal that gas-phase protein ions exist in different conformations depending on the charge state; protein ions electrosprayed from a native solution maintain the characteristics of a nativelike three-dimensional structure at a low charge state, whereas upon increasing the charge they tend to adopt a fully unfolded "stringlike" conformation. It becomes clear that the electrostatic interactions play a crucial role in the regulation of the conformational stability of proteins in vacuum.

In principle, there are no fundamental difficulties in the theoretical analysis of the electrostatic interactions in proteins in the gas phase. Moreover, the absence of solvent enables a direct comparison between the charge−charge interactions on one side and the intramolecular forces, such as hydrogen bonds and van der Waals interactions, stabilizing the nativelike protein structure on the other side. The main problem here is that the structure of the protein ions in the gas phase is not known. However, at least for low charge states, one can assume that the gas-phase protein conformation is similar to that in solution, hence known three-dimensional structures resolved by NMR or X-ray methods (if one assume that solution and crystal structures are essentially similar) can be used. On the basis of

[†] Karolinska Institute.
[‡] On leave from the Institute of Organic Chemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria.
[§] Uppsala University.
[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1997.

S1089-5647(97)02249-9 CCC: $14.00 

this assumption and using the Coulomb equation, Williams and co-workers have calculated the different properties of proteins and peptides in the gas phase, such as maximum charge state and gas-phase basicity of folded and unfolded structures[30,35−37] (see ref 31 for a review).

In this paper, we demonstrate that the theory of electrostatic interactions, developed for proteins in solution, can also be applied for proteins in the gas phase. The gas-phase basicities of nativelike lysozyme and ubiquitin were calculated applying a macroscopic approach. A Monte Carlo algorithm was developed for the determination of the most probable protonated sites for different charge states. The results were compared with available experimental gas-phase basicities for these proteins and discussed in the context of calculations based on a point-charge Coulomb interaction.

## Method of Calculation

**Gas-Phase Basicity.** The gas-phase basicity (GB) of a molecule, M, with a single protonation site is defined as the negative of the Gibbs free energy of the hypothetical reaction $M + H^+ \rightleftarrows MH^+$, i.e.,

$$GB = -\Delta G = RT \ln K \tag{1}$$

where $K$ is the equilibrium constant. In the case of multiply-charged ions, the protonation sites may have different gas-phase basicities, so the apparent gas-phase basicity, $GB_{app}$, of these ions is determined by the site with the lowest GB value. The task to calculate $GB_{app}$ of a protein is then reduced to calculation of the $GB_i$ value of each protonation site $i$. For this purpose, eq 1 can be used, however, the presence of other charges and the dielectric properties of the macroion must be taken into account. Here, one can employ the strategy developed for the calculation of the p$K$ of titratable groups (i.e., the protonation sites) in proteins in solution. For simplicity, we will follow the scheme presented in ref 38, although other more comprehensive considerations are available.[6,17] For the apparent gas-phase basicity of an individual protonation site, $GB_{app,i}$, one can write eq 2. By analogy with Tanford's definition of intrinsic

$$GB_{app,i} = GB_{int,i} + \Delta GB_{c-c,i} \tag{2}$$

p$K$,[4] the intrinsic gas-phase basicity, $GB_{int,i}$, is the gas-phase basicity of a site in a hypothetical protein without any other protonated (charged) sites. The contribution of all other protonated sites of a protein on the gas-phase basicity of site $i$ is accounted for by the term $\Delta GB_{c-c,i}$. The above expression does not differ from that used by Williams and co-workers.[30,35−37] These authors[35] have calculated the values of $GB_{int,i}$ for various types of site (arginines, lysines, etc.) as a sum of the GB values obtained for model compounds and a correction term of 15 kcal/mol, which accounts for the intramolecular solvation. This approximation presumes that all sites are equally solvated in the protein moiety. The degree of exposure or burial of the different protonation sites cannot accurately be estimated, so that a uniform correction is quite acceptable. However, the protonation sites are situated in different local environments, which can essentially, and in different proportions, influence the magnitude of the correction term. Besides the intramolecular solvation, another factor contributing to the diversity of the intrinsic gas-phase basicity is the electrostatic influence of the protein partial charges, which do not belong to the protonation sites. It has been shown in a number of studies that the partial charges determining the backbone dipoles play an important role in the ionization behavior of proteins in solution.[10,39−43] Exploiting the eventual similarity of the protein structures in solution and in the gas phase,[22,27] we assume that the backbone conformation in vacuum is identical with that in solution and include the peptide dipoles in the calculations. Thus, for the intrinsic gas-phase basicity of a site $i$ one can write eq 3, where

**TABLE 1: Free Amino Acids: Gas-phase Basicity and Charged Atoms of the Protonation Sites Used in the Calculations**

| group | GB (ref) | charged atoms[a] | charge |
|---|---|---|---|
| arg | 235.8 (49) | NH1; NH2 | 0.5 |
| lys | 226.0 (50) | NZ | 1.0 |
| his | 224.1 (29) | ND1; NE2 | 0.5 |
| pro | 221.9 (50) | N | 1.0 |
| trp | 217.6 (29) | NE1 | 1.0 |
| tyr | 214.5 (29) | OH | 1.0 |
| asn | 212.0 (29) | OD1; ND2 | 0.5 |
| gln | 210.6 (29) | OE1; NE2 | 0.5 |
| thr | 210.8 (29) | OG1 | 1.0 |
| ser | 209.0 (29) | OG | 1.0 |
| N-terminus | 206.0 (35) | N | 1.0 |

[a] Atom name designation is according to PDB nomenclature.[46]

$$GB_{int,i} = GB_{aa,i} + \Delta GB_{sol,i} + \Delta GB_{dip,i} \tag{3}$$

$GB_{aa,i}$ is the gas-phase basicity experimentally determined for the different types of amino acid (Table 1), $\Delta GB_{sol,i}$ is the contribution of the intramolecular solvation of site $i$, and $\Delta GB_{dip,i}$ represents the influence of the electrostatic interactions of this site with the peptide dipoles.

If one assumes that the major contribution to the perturbation of the gas-phase basicity of a given site upon intramolecular solvation is of an electrostatic nature, then $\Delta GB_{sol,i}$ can be evaluated in terms of the Born theory. The Born solvation energy results from the interaction of a charge with the polarization that this charge induces in the environment. This energy differs for a protonation site in a free amino acid and in a protein. The difference between these energies, which is the energy of transfer of a free amino acid to its position in the protein, gives the desired correction given in eq 4, where the

$$\Delta GB_{sol,i} = -\Delta\Delta G_{Born,i} = \frac{C}{2}\sum_j q_j[\varphi_{j,prot}(r_j) - \varphi_{j,mod}(r_j)] \tag{4}$$

sum is over the possible charge locations within the protonation site, $q_j$, is the charge (in atomic units, values given in Table 1) with coordinates $r_j$, $\varphi_j$ is the potential created by this charge, and $C \approx 332$ is a numerical constant chosen so that $\Delta GB_{sol,i}$ is in kcal/mol. The above equation is valid if the deprotonated form of the site does not have any charges. More accurate calculations require inclusion of the partial charges of an amino acid in both the protonated and deprotonated state. A comprehensive expression for $\Delta G_{Born}$ for this case can be found in ref 44.

The contribution of the permanent partial charges can be calculated by eq 5, where the first sum is over the charge

$$\Delta GB_{dip,i} = -C\sum_j q_j \sum_k \varphi_k(r_j) \tag{5}$$

locations $j$ of the protonation site $i$, the second sum is over all partial charges determining peptide dipoles, and $\varphi_k(r_j)$ is the potential created by charge $k$ at the position of the charge $q_j$.

The last term in eq 2, $\Delta GB_{c-c,i}$, can be evaluated using formula 5, where $\varphi_k(r_j)$ is now the potential created by the charges of the protonation sites.

**Electrostatic Calculations.** The problem of the calculation of the apparent gas-phase basicity of the individual protonation sites reduces to calculation of the electrostatic potential and performing the arithmetic in eqs 4 and 5. In principle, the

microscopic approach gives the most rigorous description of the electrostatic interactions in proteins on a molecular level. However, the lack of structural information may induce uncertainty that can overweigh the accuracy achieved by a precise consideration of the protein crystal structure. Therefore, we think that the macroscopic approach is most appropriate for the problem.

The electrostatic potential, $\varphi_i(r)$, is determined by the Poisson equation eq 6, where $\rho(r)$ represents the set of point charges creating the electrostatic field and $\epsilon(r)$ is the dielectric constant as a function of $r$: $\epsilon(r) = \epsilon_i$ in the protein interior and $\epsilon(r) = \epsilon_0$ outside the protein. This equation is solved numerically using

$$\nabla(\epsilon(r)\nabla\varphi(r)) = -4\pi\rho(r) \qquad (6)$$

the finite difference method.[10,12,15] To obtain $\Delta GB_{sol,i}$, eq 6 must be solved for each individual site $i$ in the protein and as a free amino acid, one site at a time. Thus, for $N$ protonation sites, $2N$ finite difference calculations are performed. The potential $\varphi_i(r)$ used in the calculation of $\Delta GB_{dip,i}$, as well as in the calculation of the charge−charge interactions between the protonation sites, is calculated independently for each protonation site, thus, additional $N$ finite difference calculations are performed.

**Determination of the Charge Distribution.** The major problem in calculation of $\Delta GB_{c-c,i}$ is that the distribution of the protonated sites in a protein ion is not known. For a protein with $N$ protonation sites, $n$ of which are protonated, the number of the possible charge configurations, $R = N!/n!(N - n)!$, may be enormously large and prohibitive for any calculations, even for medium proteins. To solve the problem, Schnier et al.[35] have developed a "pseudorandom walk" algorithm for finding the lowest free-energy charge configuration. In this work, we propose an alternative procedure based on the Metropolis algorithm[45] for Monte Carlo sampling of charge configurations. For a given charge state, $n$, the algorithm starts from a randomly chosen charge distribution. The free energy (the electrostatic term only) for this distribution is calculated and the protonation state of one site is altered. To keep the charge state of the protein unchanged, the protonation state of another, randomly chosen, site is simultaneously altered. If a given site is altered from its charged to its neutral form, the concomitant alteration is performed within the subset of neutral sites and *vice versa*. The free energy of the new charge distribution is calculated and compared with that of the previous one. If the change of the free energy is negative, $\Delta G_{el} \leq 0$, the new charge configuration is "accepted". For $\Delta G_{el} > 0$, the new configuration is "accepted" if $\exp(\Delta G_{el}/RT) > \eta$, where $\eta$ is a random number between 0 and 1. Otherwise, the old configuration remains "accepted". This procedure is repeated in a loop, consisting of a consecutive alteration of the protonation states of all sites. To avoid trapping in a local minimum, an annealing procedure was applied. A nonrelated Monte Carlo step was determined by four loops, each executed for different temperatures: $T_1 = 2360$ K, $T_2 = 1180$ K, $T_3 = 590$ K, and $T_4 = 295$ K. The first loop begins with a randomly chosen charge configuration, whereas each following one uses the configuration obtained by the former one. The charge configuration obtained by one nonrelated Monte Carlo step was considered as a possible configuration, $x$, and the gas-phase basicities, $GB_i$, for the protonated sites were calculated. The minimum $GB_i$ value determined the gas-phase basicity of the protein for this charge configuration. The apparent GB of the protein was calculated by

$$\langle GB_{app}\rangle = \frac{1}{M}\sum_{\{x\}} GB_{min}(x)$$

where $\boldsymbol{x} = (x_1,..., x_i,..., x_N)$ is a vector of $N$ element ($N$ is the number of the protonation sites) corresponding to one charge distribution obtained by a noncorrelated Monte Carlo step. The element $x_i$ is 0 or 1 for neutral or protonated site $i$, respectively. The sum is taken over all sampled configurations, $\{\boldsymbol{x}\}$, and $M$ is the number of the noncorrelated Monte Carlo steps.

In this work, the calculations have been performed with $M = 10^4$. Further increasing of $M$ to $10^5$ did not affect the results substantially. The standard deviation of $GB_{app}$ obtained after $10^4$ Monte Carlo steps was 6.9 kcal/mol.

**Application to Lysozyme and Ubiquitin.** According to the macroscopic model, the protein molecule was represented as a material with dielectric constant $\epsilon = 4$. Calculations with $\epsilon = 2$ were also performed. Equation 6 was solved using the scheme given in ref 38, where the procedure for the p$K$ calculations was substituted by the algorithm for Monte Carlo simulation described above. The X-ray structures of hen−egg lysozyme (Protein Data Bank[46] (PDB) entry 1LZT) and ubiquitin (PDB entry 1UBQ) were used in the calculations. As described above, $\Delta GB_{sol,i}$ is calculated for each individual site $i$ in the protein and as a free amino acid, one site at a time. For each site in the protein, finite difference calculations were performed in a cubic grid with $64 \times 64 \times 64$ nodes centered on the corresponding site. The final solution of eq 6 was obtained after focusing[12] with a lattice spacing of 0.4 Å. The same calculations were repeated for the free amino acids after removing all other protein atoms. The potentials $\varphi_i(r)$, used in the calculation of $\Delta GB_{dip,i}$ as well as in the calculation of the charge−charge interactions between the protonation sites (eq 5), were calculated independently in a grid centered at the geometrical center of the proteins. Initially the lattice spacing was 1.4 Å for lysozyme and 1.2 Å for ubiquitin, followed by focusing with a lattice spacing of 0.7 Å for lysozyme and 0.6 Å for ubiquitin. The protonation sites, their GB values in free amino acids, as well as the ionizable atoms are listed in Table 1. The partial charges (atoms N, C, O, CA) forming the peptide dipoles are taken from CHARMM parameter set 19.[47]

## Results and Discussion

**Intrinsic Gas-Phase Basicity.** One main difference between the calculations performed by Williams and co-workers[35] and those in this work is in the estimation of the intrinsic gas-phase basicity. These authors have used a constant correction term of 15 kcal/mol to account for the effect of intramolecular solvation of the different protonation sites. Here, the influence of the protein molecule (excluding other protonation sites) on the GB values is evaluated by eq 3, which reflects the local electrostatic environment ($\Delta GB_{dip}$) and the extent of intramolecular solvation of each individual site ($\Delta GB_{sol}$). In Tables 2 and 3, the values of the intrinsic gas-phase basicities, as well as the perturbation due to the intramolecular solvation and peptide dipoles, are listed for lysozyme and ubiquitin, respectively. As shown, $\Delta GB_{sol}$ varies between 24.3 and 5.2 kcal/mol for lysozyme and 20.1 and 8.2 kcal/mol for ubiquitin, i.e., the individual corrections due to the intramolecular solvation vary around the value used by Schnier et al.[35]

The inspection of $\Delta GB_{dip}$ values from Tables 2 and 3 indicates that the majority of the protonation sites that belong to long side chains are in favorable interactions with the peptide dipoles. For the shorter side chains, the influence of the peptide dipoles becomes essentially unfavorable. For the proline residues, it is larger in magnitude than the solvation term. As a result, the arginines and the lysines are characterized by a relatively high $GB_{int}$, whereas the prolines have the lowest $GB_{int}$ values (Tables 2 and 3). Thus, accounting for $GB_{int}$ only, it follows that due to the electrostatic influence of the peptide

**TABLE 2: Intrinsic Gas-Phase Basicities, $\Delta GB_{sol}$ (kcal/mol), and the Terms Arising from Intramolecular Solvation, $\Delta GB_{sol}$, and Peptide Dipole Interactions, $\Delta GB_{dip}$, of the Individual Protonation Sites in Lysozyme**

| | $\epsilon = 4$ | | | $\epsilon = 2$ | | |
|---|---|---|---|---|---|---|
| site | $\Delta GB_{sol}$ | $\Delta GB_{dip}$ | $GB_{int}$ | $\Delta GB_{sol}$ | $\Delta GB_{dip}$ | $GB_{int}$ |
| Arg21 | 14.3 | 8.3 | 258.4 | 8.3 | 12.3 | 256.4 |
| Arg112 | 14.4 | 5.7 | 255.8 | 8.5 | 8.4 | 252.7 |
| Lys96 | 19.2 | 9.4 | 254.6 | 12.1 | 17.0 | 255.0 |
| Arg73 | 13.0 | 4.2 | 253.1 | 7.3 | 6.7 | 249.9 |
| Lys97 | 13.8 | 12.8 | 252.6 | 7.6 | 20.0 | 253.7 |
| Arg68 | 14.0 | 2.4 | 252.2 | 8.2 | 3.7 | 247.7 |
| Arg45 | 14.8 | 1.2 | 251.8 | 8.8 | 1.7 | 246.3 |
| Arg114 | 11.2 | 4.0 | 251.0 | 6.0 | 5.6 | 247.4 |
| Arg14 | 13.8 | −3.4 | 246.2 | 8.0 | −5.3 | 238.5 |
| Arg61 | 15.1 | −5.0 | 245.9 | 9.2 | −9.4 | 235.6 |
| Lys33 | 16.5 | 2.3 | 244.7 | 9.9 | 3.3 | 239.2 |
| Arg125 | 11.9 | −3.0 | 244.7 | 6.4 | −3.5 | 238.8 |
| Lys116 | 12.9 | 4.1 | 243.0 | 7.0 | 5.3 | 238.3 |
| Arg5 | 13.1 | −6.3 | 242.6 | 7.4 | −9.5 | 233.7 |
| Ser100 | 20.7 | 12.1 | 241.8 | 13.0 | 20.3 | 242.2 |
| Tyr20 | 19.3 | 7.8 | 241.6 | 12.5 | 12.7 | 239.7 |
| Trp108 | 19.8 | 3.6 | 241.0 | 14.0 | 5.4 | 237.0 |
| Trp28 | 19.3 | 3.6 | 240.5 | 13.6 | 6.4 | 237.6 |
| Gln57 | 20.8 | 8.3 | 239.6 | 14.0 | 15.2 | 239.8 |
| Asn103 | 18.7 | 8.6 | 239.3 | 11.7 | 13.5 | 237.2 |
| Ser36 | 24.4 | 5.4 | 238.8 | 16.8 | 9.9 | 235.6 |
| His15 | 17.5 | −4.0 | 237.6 | 11.0 | −6.7 | 228.4 |
| Lys1 | 15.8 | −5.5 | 236.3 | 9.3 | −8.4 | 226.9 |
| Lys13 | 13.6 | −3.6 | 235.9 | 7.5 | −5.4 | 228.1 |
| Trp111 | 19.0 | −0.8 | 235.8 | 13.2 | −2.8 | 228.0 |
| Arg128 | 5.2 | −5.2 | 235.8 | 2.2 | −7.4 | 230.6 |
| Tyr23 | 19.3 | 1.8 | 235.6 | 12.6 | 1.1 | 228.2 |
| Tyr53 | 21.5 | −1.6 | 234.3 | 14.9 | −3.8 | 225.6 |
| Asn37 | 15.2 | 6.3 | 233.6 | 8.7 | 10.5 | 231.1 |
| Trp63 | 18.1 | −2.2 | 233.5 | 12.2 | −6.1 | 223.7 |
| Thr43 | 17.1 | 4.0 | 231.9 | 10.0 | 6.7 | 227.5 |
| Asn77 | 11.8 | 8.0 | 231.8 | 5.9 | 11.5 | 229.4 |
| Ser91 | 25.2 | −2.5 | 231.8 | 17.7 | −4.6 | 222.0 |
| Thr51 | 21.4 | −0.6 | 231.6 | 14.6 | −1.8 | 223.5 |
| Trp123 | 17.6 | −3.6 | 231.6 | 11.8 | −6.4 | 223.0 |
| Ser72 | 22.2 | −0.1 | 231.1 | 14.7 | −1.9 | 221.8 |
| Asn113 | 14.4 | 3.7 | 230.1 | 8.0 | 4.4 | 224.4 |
| Asn44 | 16.4 | 1.1 | 229.5 | 9.4 | 1.6 | 223.0 |
| Trp62 | 15.6 | −4.4 | 228.7 | 10.0 | −9.4 | 218.2 |
| Asn106 | 16.7 | −0.3 | 228.4 | 9.9 | −2.8 | 219.0 |
| Thr118 | 19.7 | −2.8 | 227.7 | 12.8 | −5.5 | 218.1 |
| Asn39 | 17.3 | −1.7 | 227.6 | 10.5 | −3.8 | 218.7 |
| Thr69 | 21.5 | −4.7 | 227.6 | 14.8 | −10.2 | 215.4 |
| Asn27 | 20.9 | −5.6 | 227.3 | 13.8 | −11.2 | 214.6 |
| Thr40 | 21.2 | −4.7 | 227.3 | 14.2 | −8.2 | 216.8 |
| Asn93 | 16.6 | −1.9 | 226.8 | 9.9 | −3.1 | 218.8 |
| Asn74 | 18.8 | −4.1 | 226.7 | 12.0 | −10.2 | 213.8 |
| Asn46 | 18.3 | −4.9 | 225.4 | 11.5 | −9.2 | 214.3 |
| Gln41 | 14.3 | −0.3 | 224.6 | 8.1 | −0.6 | 218.1 |
| Asn19 | 15.5 | −4.0 | 223.5 | 8.8 | −8.4 | 212.5 |
| Ser50 | 22.3 | −10.2 | 221.1 | 15.0 | −20.1 | 203.9 |
| Ser60 | 23.5 | −11.5 | 221.0 | 16.2 | −23.6 | 201.6 |
| Asn59 | 20.5 | −11.9 | 220.6 | 13.6 | −23.9 | 201.7 |
| N-terminus | 21.1 | −7.8 | 219.3 | 12.9 | −14.1 | 204.8 |
| Ser86 | 18.9 | −9.6 | 218.3 | 11.4 | −16.0 | 204.4 |
| Asn65 | 16.8 | −11.5 | 217.3 | 9.9 | −22.9 | 199.0 |
| Thr89 | 19.3 | −13.3 | 216.8 | 12.1 | −23.7 | 199.2 |
| Ser85 | 19.3 | −14.4 | 213.9 | 11.6 | −25.2 | 195.5 |
| Ser24 | 22.0 | −18.7 | 212.4 | 14.4 | −35.4 | 188.0 |
| Ser81 | 17.0 | −14.8 | 211.3 | 9.8 | −26.4 | 192.5 |
| Gln121 | 11.5 | −11.0 | 211.2 | 6.0 | −17.0 | 199.6 |
| Thr47 | 11.6 | −14.6 | 207.7 | 6.0 | −26.0 | 190.8 |
| Pro70 | 16.9 | −51.0 | 187.8 | 10.2 | −91.4 | 140.7 |
| Pro79 | 18.6 | −52.9 | 187.4 | 11.6 | −94.8 | 138.7 |

dipoles, most probable protonated sites in the gas phase are those protonated in solution.

On the basis of the model calculations and the experimental measurements of $GB_{app}$ of the individual charge states of cytochrome *c*, a dielectric constant of 2 has been evaluated.[30] We repeated the calculations with a dielectric constant $\epsilon = 2$.

**TABLE 3: Intrinsic Gas-Phase Basicities, $\Delta GB_{sol}$ (kcal/mol) and the Terms Arising from Intramolecular Solvation, $\Delta GB_{sol}$, and Peptide Dipole Interactions, $\Delta GB_{dip}$, of the Individual Protonation Sites in Ubiquitin**

| | $\epsilon = 4$ | | | $\epsilon = 2$ | | |
|---|---|---|---|---|---|---|
| site | $\Delta GB_{sol}$ | $\Delta GB_{dip}$ | $GB_{int}$ | $\Delta GB_{sol}$ | $\Delta GB_{dip}$ | $GB_{int}$ |
| Arg72 | 14.7 | 6.6 | 257.0 | 9.1 | 11.9 | 256.8 |
| Arg42 | 13.0 | 1.2 | 250.0 | 7.6 | 1.4 | 244.9 |
| Lys27 | 19.9 | 1.1 | 247.0 | 13.8 | 3.0 | 242.8 |
| Lys48 | 12.5 | 8.2 | 246.7 | 7.1 | 14.3 | 247.4 |
| Lys33 | 14.6 | 6.0 | 246.5 | 8.8 | 9.4 | 244.1 |
| Lys11 | 14.1 | 4.8 | 244.8 | 8.4 | 6.4 | 240.8 |
| Arg74 | 8.2 | −0.2 | 243.8 | 4.1 | −0.8 | 239.1 |
| Lys29 | 15.9 | 0.5 | 242.4 | 9.9 | 2.0 | 237.9 |
| Arg54 | 10.1 | −5.8 | 240.1 | 5.3 | −8.1 | 233.0 |
| Lys6 | 11.4 | 2.2 | 239.6 | 6.2 | 2.7 | 234.8 |
| His68 | 14.7 | 0.2 | 239.0 | 9.0 | −0.6 | 232.5 |
| Lys63 | 9.6 | −2.7 | 232.8 | 4.9 | −4.2 | 226.7 |
| Gln41 | 19.3 | 1.0 | 230.9 | 13.5 | 1.4 | 225.5 |
| Thr66 | 16.9 | 1.7 | 229.5 | 10.5 | 2.4 | 223.7 |
| Thr12 | 16.9 | 1.1 | 228.9 | 10.6 | 0.0 | 221.4 |
| Ser65 | 20.1 | −0.2 | 228.8 | 13.2 | −0.4 | 221.8 |
| Asn60 | 11.9 | 3.0 | 226.9 | 6.4 | 5.2 | 223.6 |
| Thr14 | 15.5 | 0.2 | 226.5 | 9.2 | 0.6 | 219.4 |
| Tyr59 | 18.3 | −6.3 | 226.5 | 12.5 | −11.4 | 215.6 |
| Gln49 | 15.1 | 0.6 | 226.3 | 9.1 | 0.8 | 220.5 |
| Gln40 | 15.8 | −1.9 | 224.5 | 9.9 | −4.3 | 216.2 |
| Thr9 | 14.6 | −1.7 | 223.7 | 8.7 | −6.5 | 213.0 |
| Gln2 | 14.0 | −1.4 | 223.2 | 8.2 | −2.6 | 216.2 |
| Gln31 | 14.1 | −2.1 | 222.6 | 8.3 | −4.7 | 214.2 |
| Thr7 | 18.1 | −6.5 | 222.4 | 12.0 | −15.5 | 207.3 |
| N-terminus | 19.0 | −2.8 | 222.1 | 11.6 | −3.2 | 214.4 |
| Asn25 | 16.0 | −8.1 | 219.9 | 9.7 | −13.8 | 207.9 |
| Gln62 | 11.0 | −1.6 | 219.9 | 5.8 | −2.7 | 213.7 |
| Ser57 | 16.9 | −11.2 | 214.7 | 10.2 | −19.4 | 199.8 |
| Thr55 | 17.3 | −13.8 | 214.2 | 11.1 | −25.0 | 196.9 |
| Thr22 | 17.5 | −16.1 | 212.2 | 11.2 | −29.0 | 193.0 |
| Ser20 | 15.1 | −15.4 | 208.7 | 8.6 | −26.2 | 191.5 |
| Pro37 | 17.2 | −44.7 | 194.4 | 11.0 | −80.5 | 152.4 |
| Pro38 | 17.6 | −52.5 | 188.0 | 11.4 | −95.3 | 138.0 |
| Pro19 | 17.3 | −61.8 | 177.4 | 11.2 | −110.4 | 122.7 |

As expected, the contribution of the intramolecular solvation reduces whereas the influence of the dipole moments increases. However, the classification of the protonation sites with respect to $GB_{int}$ was essentially not changed.

**Apparent Gas-Phase Basicity.** The gas-phase basicities calculated for the different charge states of lysozyme and ubiquitin are shown in Figure 1. Experimental data for electrosprayed lysozyme and ubiquitin are also shown (data for ubiquitin, Figure 1B, are for protein ions electrosprayed from a methanol/water solution[48]). As far as the native protein gives charge states between 6 and 8,[28] the measured GB values for these states may be close to the values for the nativelike structure. The values of $GB_{app}$, calculated using a dielectric constant $\epsilon_i = 4$, essentially do not differ from those obtained with $\epsilon_i = 2$. In both cases, $GB_{app}$ values were found to lie below the experimental data. Since the calculations are based on the presumption of thermodynamic equilibrium (eq 1), an underestimate of the calculated $GB_{app}$ value compared to the experimental results can be expected.[29,30,35] It is notable that the calculated $GB_{app}$ values are closer to the experimental data for the lower than for higher charge states. One can relate this observation to the increase of the electrostatic repulsion between the protonated sites at higher charge states which produces conformational changes[35] and ion "swelling" (increasing of the mutual distance between the charges). As discussed, in these calculations we used a fixed protein structure, possible conformational changes were not taken into account. We assume that the smooth increase of the discrepancy between the calculated and the experimental $GB_{app}$ values reflects the energy costs associated with the conformational changes upon changing the

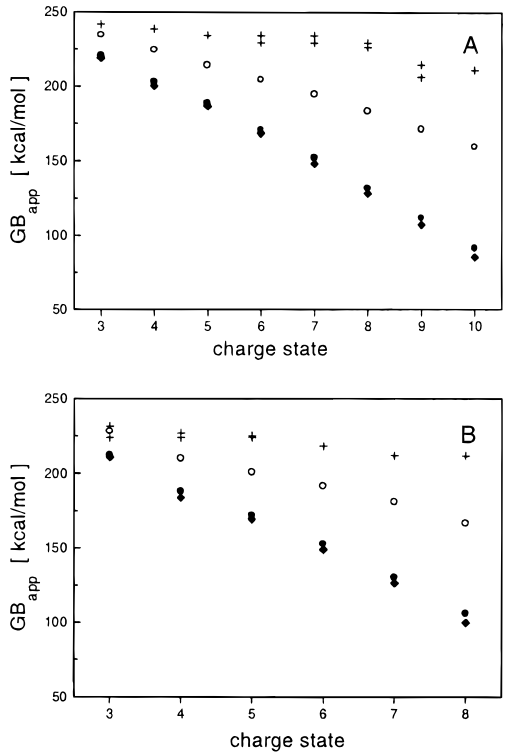


**Figure 1.** Apparent gas-phase basicity as a function of the charge state of disulfide-intact lysozyme (A) and ubiquitin (B): (●) values calculated with $\epsilon_i = 4$; (♦) values calculated with $\epsilon_i = 2$; (⊙) values calculated using the Coulomb equation; (+) experimental data taken from refs 36 (for lysozyme) and 48 (for ubiquitin).

charge state. This interpretation is qualitatively in agreement with the experimental observation of Valentine et al.[23]

The $GB_{app}$ values calculated from the Coulomb formula with $\epsilon_r = 2$ are still below, but essentially closer to, experimental data (see also ref 36). This seeming improvement is, however, based on a representation that does not correspond to the physical reality. The use of the Coulomb equation resumes that the electrostatic interactions between protonation sites are realized in a homogeneous material with dielectric properties equivalent to those of the protein molecule, which, in fact, is a very constricted zone.

The electrostatic term of the free energy, calculated as a function of the charge states, showed minimum values at charge states 8 and 9 for lysozyme and between 6 and 8 for ubiquitin. This result corresponds to the charge state distributions detected in electrospray mass spectra of the nativelike molecules. For disulfide-intact lysozyme, these values are between 8 and 11,[36] whereas for ubiquitin, they are between 6 and 8.[28] The minima obtained from the calculations based on the Coulomb equation showed an essential shift toward high charge states, 14 for lysozyme and 12 for ubiquitin, which are typical for unfolded electrosprayed proteins. These results are evidence that the model for the calculation of the electrostatic interactions in gas-phase proteins, proposed in this work, is realistic.

**Preferred Protonation Sites.** The probability, $P_i$, that a given site $i$ is protonated was calculated from

$$P_i = \frac{1}{M}\sum_{\{x\}} x_i$$

where $x_i$ (1 or 0) is the element of a vector $\boldsymbol{x}$ describing the protonation state of site $i$. In Tables 4 and 5, the sites with probabilities greater than 0.1 are listed. The sites with the highest probability to be protonated for a given charge state are given in bold. As shown, the sites that are preferably protonated at low charge states remain protonated at high charge states. Also, for both lysozyme and ubiquitin, the most preferred protonation sites belong to long side chains, such as arginines and lysines, independent of the charge state. Thus, both charge−charge and charge−peptide dipole interactions lead to one and the same effect. This is in accord with the commonly accepted opinion that protonated side chains in the gas phase and in solution coincide.

The relatively small number of protonation sites of ubiquitin allows us to calculate the probability for protonation, accounting for all possible charge distributions for charge states $n \leq 8$ and using the Boltzmann weighted sum:

$$P_i = \frac{1}{Z}\sum_{\{x\}} x_i e^{-G_{el}(x)/RT}$$

**TABLE 4: Probabilities of Protonated Sites in Disulfide-Intact Lysozyme for Different Charge States, *n*[a]**

| site/*n* | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 |
|---|---|---|---|---|---|---|---|---|
| Arg128 | **1.00** | **1.00** | **1.00** | **1.00** | **0.99** | **0.93** | **0.86** | **0.73** |
| Arg73 | **0.96** | **0.93** | **0.91** | **0.89** | **0.85** | **0.86** | **0.85** | 0.31 |
| Arg21 | **0.91** | **0.90** | **0.88** | **0.86** | **0.84** | **0.79** | **0.77** | **0.83** |
| Arg114 | **0.90** | **0.88** | **0.84** | **0.82** | **0.77** | **0.66** | **0.61** | **0.57** |
| Arg125 | **0.81** | **0.79** | **0.78** | **0.82** | **0.68** | **0.59** | 0.27 | 0.19 |
| Lys116 | **0.69** | **0.61** | **0.50** | 0.35 | 0.22 | 0.11 | | |
| Asn77 | **0.67** | **0.58** | **0.47** | 0.27 | | | | |
| Arg68 | **0.58** | **0.58** | **0.59** | **0.60** | **0.61** | **0.63** | **0.65** | **0.65** |
| Lys96 | **0.55** | **0.52** | 0.44 | **0.46** | **0.42** | 0.28 | 0.19 | |
| Arg14 | **0.47** | **0.43** | 0.38 | 0.27 | 0.21 | 0.13 | 0.10 | |
| Arg112 | **0.41** | 0.41 | **0.46** | **0.49** | **0.42** | 0.33 | 0.19 | 0.22 |
| Arg45 | **0.41** | 0.40 | 0.40 | 0.38 | 0.38 | 0.36 | 0.34 | 0.26 |
| Arg61 | 0.39 | 0.28 | 0.18 | 0.12 | 0.12 | | | |
| Lys1 | 0.39 | 0.34 | 0.24 | 0.14 | | | | |
| Lys13 | 0.32 | 0.21 | 0.11 | | | | | |
| Lys97 | 0.29 | 0.26 | 0.23 | 0.17 | 0.14 | | | |
| Gln41 | 0.22 | 0.16 | | | | | | |
| Arg5 | 0.20 | 0.19 | 0.17 | 0.12 | | | | |
| Thr47 | 0.17 | | | | | | | |
| Lys33 | 0.11 | 0.11 | 0.12 | | | | | |

[a] The values of the most probable sites for a given charge state are in bold.

**TABLE 5: Probabilities of Protonated Sites in Nativelike Ubiquitin for Different Charge States, *n*[a]**

| site/*n* | 9 | 8 | 7 | 6 | 5 | 4 |
|---|---|---|---|---|---|---|
| | | | Monte Carlo | | | |
| Arg74 | **1.00** | **1.00** | **1.00** | **1.00** | **1.00** | **0.93** |
| Arg54 | **0.99** | **0.99** | **0.97** | **0.91** | **0.75** | 0.35 |
| Lys63 | **0.93** | **0.89** | **0.90** | **0.93** | **0.92** | **0.76** |
| Lys48 | **0.92** | **0.88** | **0.83** | **0.64** | **0.52** | **0.67** |
| Lys6 | **0.82** | **0.77** | **0.61** | 0.42 | 0.14 | |
| Arg72 | **0.75** | **0.67** | **0.57** | **0.44** | 0.24 | 0.10 |
| Lys11 | **0.72** | **0.67** | **0.58** | **0.54** | **0.56** | **0.57** |
| Arg42 | **0.69** | 0.48 | 0.39 | 0.38 | 0.39 | 0.13 |
| Lys29 | **0.68** | **0.61** | 0.40 | 0.21 | | |
| Lys33 | 0.50 | 0.44 | 0.41 | 0.40 | 0.37 | 0.38 |
| Asn60 | 0.21 | 0.10 | | | | |
| Gln31 | 0.14 | | | | | |
| Thr9 | 0.12 | | | | | |
| | | | Boltzmann Statistics | | | |
| Arg74 | | **1.00** | **1.00** | **1.00** | **1.00** | **1.00** |
| Arg54 | | **1.00** | **1.00** | **1.00** | **1.00** | |
| Lys63 | | **1.00** | **1.00** | **1.00** | **1.00** | **0.97** |
| Lys48 | | **1.00** | **1.00** | **0.91** | | **0.94** |
| Lys6 | | **1.00** | **1.00** | 0.12 | | |
| Arg72 | | **1.00** | **1.00** | **0.68** | | |
| Lys11 | | **1.00** | | **0.86** | **0.87** | **0.93** |
| Arg42 | | | | 0.29 | **0.98** | |
| Lys29 | | **1.00** | | | | |
| Lys33 | | | **1.00** | 0.14 | 0.13 | |

[a] The values of the most probable sites for a given charge state are in bold.

where $Z = \sum \exp(-G_{el}(\boldsymbol{x})/RT)$ and $G_{el}(\mathbf{x})$ is the electrostatic term of the free energy of the charge distribution $\boldsymbol{x}$. The sum is taken over all possible charge configurations $\{\boldsymbol{x}\}$. The charge distributions, calculated by means of the Boltzmann statistics, for the different charge states are distinctly defined (Table 5). A single charge distribution, with probabilities essentially close to unity, was found for charge states 7 and 8. There are few sites with a somewhat lower probability of protonation at charge states 4−6, however, their probabilities are still essentially higher than those of the rest of the sites. The important result here is that the charge distribution calculated using the Boltzmann statistics and Monte Carlo simulation are essentially identical (see Table 5). The only difference found between the two sets of distributions was for charge state 7, where according to the Boltzmann statistics Lys33 was in its protonated form, whereas Monte Carlo calculations suggested a probability of 0.41 for this site and a slightly higher probability (0.58) for Lys11. The identity of the charge distributions calculated by the two techniques indicates that the computational approach proposed in this work predicts the ionization behavior of proteins in the gas phase sufficiently well.

**Conclusions.** In this work, we showed that the theory of electrostatic interactions in proteins in solution can successfully be applied to protein ions in the gas phase. The advantage of the approach presented here can be summarized in two points. First, the concept of $GB_{int}$ was extended from a constant correction to an expression, accounting for the intramolecular solvation of the individual protonation sites and for the electrostatic influence of the peptide dipoles. The latter was found to have an important role in determination of the $GB_{int}$ values, and hence in determination of the preferred protonated sites (arginines and lysines). Second, the Monte Carlo procedure proposed here gives charge distributions identical or very close to the exact solutions for small proteins, such as ubiquitin. This allows calculations of charge−charge interactions between the protonation site and, hence, enables the evaluation the apparent gas-phase basicities, as well as other quantities of gas-phase protein ions.

It must be pointed out that the calculations were based on a fixed, X-ray, structure of the protein molecule. Although there are experimental evidence that, for low charge state, the gas-phase protein ions resemble the nativelike structure, the question of how and to what degree these structures are close to each other has no answer. Thus, we are aware that the identification of a compact gas-phase structure of a protein ion with the corresponding three-dimensional X-ray structure in the crystalline form is a weak approximation. Nevertheless, we believe that the approach proposed here can be used in combination with other methods for the theoretical analysis of the protein structure, such as molecular dynamics, and can contribute to the prediction of the ionization behavior of proteins in the gas phase.

**Acknowledgment.** This work was partially supported by Grant No. I-AA/LS 06301-316 from the Swedish Natural Research Council. A. Karshikoff thanks Dr. Lennart Nilsson for the helpful discussions.

## References and Notes

(1) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351−361.
(2) Kirkwood, J. G.; Westheimer, F. H. *J. Chem. Phys.* **1938**, *6*, 506.
(3) Westheimer, F. H.; Kirkwood, J. G. *J. Chem. Phys.* **1938**, *6*, 513.
(4) Tanford, Ch.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, *79*, 5333.
(5) Warshel, A.; Russell, S. T. Q. *Rev. Biophys.* **1984**, *17*, 283.
(6) Warshel, A.; Åqvist, *J. Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267.
(7) Warshel, A.; Russell, S. T.; Churg, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4785.
(8) Lee, F. S.; Chu, Z. T.; Warshel, A. *J. Comput. Chem.* **1993**, *14*, 161.
(9) Hwang, J.-K.; Warshel, A. *Nature* **1988**, *334*, 270.
(10) Warwicker, J.; Watson, N C. *J. Mol. Biol.* **1982**, *157*, 671.
(11) van Duijnen, P. T.; De Vries, A. H. *Int. J. Quantum Chem.* **1995**, *29*, 523.
(12) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. *Proteins* **1986**, *1*, 47.
(13) Gilson, M.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1987**, *9*, 327.
(14) Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1990**, *94*, 7684.
(15) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435.
(16) Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95*, 9556.
(17) Schaefer, M.; Sommer, M.; Karplus, M. *J. Chem. Phys. B* **1997**, *101*, 1663.
(18) Fenn, J.; Mann, M.; Meng, Ch.; Wong, Sh.; Whitehouse, Cr. *Science* **1989**, *246*, 64.
(19) Senko, M.; McLafferty, F. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 763.
(20) Wolynes, P. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2426.
(21) Wyttenbach, T.; von Helden, G.; Bowers, M. *J. Am. Chem. Soc.* **1996**, *118*, 8355.
(22) Collins, B. A.; Douglas, D. J. *J. Am. Chem. Soc.* **1996**, *118*, 4488.
(23) Valentine, S. J.; Anderson, J. G.; Ellington, A. D.; Clemmer, D. E. *J. Phys. Chem.* **1997**, *101*, 3891.
(24) Shelimov, K.; Jarrold, M. *J. Am. Chem. Soc.* **1996**, *118*, 10313.
(25) Zhang, X.; Cassady, C. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1211.
(26) Suckau, D.; Shi, Y.; Beu, St.; Senko, M.; Quinn, J.; Wampler, F., III; McLafferty, F. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 790.
(27) Wood, T. D.; Russell, A. C.; Wampler, F. M., III; Little, D. P.; O'Connor, P. B.; McLafferty, F. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 2451.
(28) Cassady, C. J.; Wronka, J.; Kruppa, G. H.; Laukien, F. H. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 394.
(29) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695.
(30) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Chem. Soc.* **1995**, *117*, 6747.
(31) Williams, E. R. *J. Mass Spectrom.* **1997**, *31*, 831.
(32) Kaltashov, I.; Fenselau, C. *Proteins* **1997**, *27*, 165.
(33) Sullivan, P.; Axelsson, J.; Altmann, S.; Quist, A.; Sundqvist, B.; Reimann, C. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 329.
(34) Jockush, R.; Schnier, P.; Price, W.; Strittmater, E.; Demirev, P.; Williams E. *Anal. Chem.* **1997**, *69*, 1119.
(35) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086.
(36) Gross, D. S.; Schnier, P. D.; Rodriguez-Cruz, S. E.; Fagerquist, C. K.; Williams, E. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 3143.
(37) Schnier, P. D.; Price, W. D.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 972.
(38) Karshikoff, A. *Protein Eng.* **1995**, *8*, 243.
(39) Hol, W. G. J.; van Duijnen, P. T.; Berendsen, H. J. C. *Nature* **1978**, *273*, 443.
(40) Wada, A. *Adv. Biophys.* **1976**, *9*, 1.
(41) Hol, W. G. J. *J. Mol. Struct. (THEOCHEM)* **1985**, *123*, 27.
(42) Karshikoff, A.; Reinemer, P.; Huber, R.; Ladenstein, R. *Eur. J. Biochem.* **1993**, *215*, 663.
(43) Spassov, V. Z.; Ladenstein, R.; Karshikoff, A. D. *Protein Sci.* **1997**, *6*, 1190.
(44) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219.
(45) Metropolis, N.; Rosenbluth, A W.; Rosenbluth, M. N.; Teller, A. H. *J. Chem. Phys.* **1953**, *21*, 1087.
(46) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. D., Jr.; Brice, M. D.; Rogers, J. R.; Kennard, O.; Shinaniushi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.
(47) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(48) Zhang, X.; Cassady, C. J. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1211.
(49) Wu, Z.; Fenselau, C. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 403.
(50) Gorman, G. S.; Spier, J. P.; Turner, C. A.; Amster, I. J. *J. Am. Chem. Soc.* **1992**, *114*, 3986.