J. Am. Chem. Soc. **1997,** *119,* 8523−8527

Multistate Gaussian Model for Electrostatic Solvation Free Energies

Gerhard Hummer,* Lawrence R. Pratt,* and Angel E. García

Contribution from the Theoretical Division, MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Received April 10, 1997. Revised Manuscript Received June 27, 1997[⊗]

Abstract: We develop and test a multistate Gaussian model for the distribution of electrostatic solvation energies of a solute in liquid water. The multistate Gaussian model depends on the discovery of simple indicators of structural substates that individually display Gaussian fluctuations of electrostatic interactions. The probability distribution of electrostatic interactions is then modeled as a superposition of Gaussian distributions of electrostatic interactions of the substates. We find that the number of hydrogen bonds to the solute is a suitable substate indicator that eliminates the chief failures of single Gaussian models for the distribution of electrostatic interactions and of quadratic models of the electrostatic contribution to the excess chemical potential. These results should improve calculations of ionic chemical processes in water, *i.e.*, acid−base chemistry, in particular those involving organic acids such as proteins and nucleic acids. The multistate Gaussian approach provides a specific and effective alternative to commonly discussed electrostriction and dielectric saturation modifications of dielectric continuum models. Moreover, the representation of complex energy distributions by a sum of simpler distributions based on structural substates is general and should be applicable in a variety of thermodynamic problems of solution chemistry.

Introduction

Dielectric continuum models[1−5] dominate our understanding of hydration of ionic and polar solutes in water. These models are physically simple and have been extensively studied.[6] They may be derived from basic starting points by taking a macroscopic limit in which most solvent molecular details vanish. Alternatively, dielectric models might be viewed as simplified implementations of second-order thermodynamic perturbation theory, or a Gaussian distribution model, for solute−solvent electrostatic interactions.[7] These molecular views of dielectric continuum models avoid and thus resolve the most serious limitation of dielectric models, namely, the definition of solute cavities on the basis of radii parameters. Those parameters should depend on temperature, pressure, composition of the solvent, and configuration of the solute. However, comparisons of computed solvation free energies with experiment over a limited range of conditions suggest that dielectric models can be helpfully accurate and that the values determined empirically for radii parameters are chemically reasonable. Dielectric models can break down more seriously when changes of conditions lead to changes in the correlations between near solvent molecules.

When tested on a molecular scale without the additional adjustment of radii, it becomes clear that dielectric models capture an essential basic ingredient of the solvation theory but are often not molecularly accurate. For instance, the free energies of ions in aqueous solution exhibit an approximately quadratic dependence on the ionic charge, in agreement with the Born continuum model of ionic solvation.[8] But higher order (up to eighth order) perturbation theory is required to capture fully the electrostatic hydration free energy of a water molecule in liquid water.[9] A physical conclusion from such results is that electrostatic hydrogen bonding interactions are difficult cases for dielectric continuum theories because the charges involved are relatively accessible. Even for classic atomic ions in water significant violations of linear response expectations can be observed for modest ionic charges: if the ion charge q is manipulated through the region $q \approx -0.2$ e, the structure of the first solvation shell makes a transition that is reflected in a nonlinear change of the solvation potential, a *kink*, with change of ionic charge.[8]

A physical picture that is consistent with the known successes and limitations of dielectric models is the following: The structure of the first hydration shell is most important. The structure of that first hydration shell can be viewed from the perspective of Stillinger−Weber "inherent structures" or substates.[10] These are potential energy basins of attraction for steepest-descent quenching of first hydration shell molecules. If those first hydration shell molecules stayed always in one basin, then a Gaussian model for thermal fluctuations would be reasonable. Empirical radii parameters reflect the characteristics of the most important basin. However, changing conditions may result in reweighting of slightly accessible basins or the opening of new basins. The Gaussian or dielectric models may fail to describe these possibilities well. We note that this picture is physically more definite than the commonly nonspecific discussions of electrostriction and dielectric saturation.

This paper shows how to develop the physical idea expressed above. Attention is directed to the thermal probability distribution of electrostatic potential energies of the solute. Rather than approximate this distribution as a Gaussian, or largely Gaussian

---

[⊗] Abstract published in *Advance ACS Abstracts,* August 15, 1997.

(1) Born, M. *Z. Phys.* **1920**, *1*, 45.

(2) Latimer, W. M.; Pitzer, K. S.; Slansky, C. M. *J. Chem. Phys.* **1939**, *7*, 108.

(3) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509.

(4) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.

(5) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.

(6) Pratt, L. R.; Tawa, G. J.; Hummer, G.; García, A. E.; Corcelli, S. A. *Int. J. Quantum Chem.* **1997**, *64*, 121.

(7) Pratt, L. R.; Hummer, G.; García, A. E. *Biophys. Chem.* **1994**, *51*, 147.

(8) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1996**, *100*, 1206.

(9) Hummer, G.; Pratt, L. R.; García, A. E. *J. Phys. Chem.* **1995**, *99*, 14188.

(10) Stillinger, F. H.; Weber, T. A. *Phys. Rev. A* **1982**, *25*, 978.

S0002-7863(97)01148-7 CCC: $14.00

as would be the case upon inclusion of higher order perturbative corrections, we discriminate hydration structure on the basis of simple parameters diagnostic of hydration substates. We assume that the probability distribution of electrostatic potential energies is Gaussian for each substate. Therefore the full distribution is a superposition of Gaussians for the various substates. We show that identification of suitable substate diagnostic parameters can be simple, *e.g.*, they can be the number of hydrogen bonds made to the solute, and no explicit calculation of inherent structures is required. We show further that this approach can be highly accurate, eliminating some of the detailed numerical inaccuracies of the Gaussian fluctuation models.

Multistate Gaussian Model

The technical development begins from the fundamental relation between the distribution $p(u; \lambda = 0)$ of electrostatic energies u in the reference charge state $\lambda = 0$ and the part of the chemical potential due to electrostatic interactions $\Delta\mu(\lambda)$, which is the thermodynamic parameter sought,

$$e^{-\beta\Delta\mu(\lambda)} = \langle e^{-\beta\lambda u}\rangle_{\lambda=0} = \int du\, p(u; \lambda = 0) e^{-\beta\lambda u} \tag{1}$$

Here, $\beta = 1/k_B T$ is the inverse temperature and $\langle ... \rangle_{\lambda=0}$ denotes a thermal average in the reference state $\lambda = 0$. Direct use of eq 1 has recognized limitations. Though $p(u)$ is often substantially Gaussian—and that supports perturbation theory approximations—the formula eq 1 is sensitive to the tails of $p(u)$. That limits the applicability of eq 1 for calculations of even small changes in the charge state λ. In addition, the simple estimator $\ln\langle e^{-\beta\lambda u}\rangle_{\lambda=0} \approx \ln[M^{-1}\sum_{i=1}^{M} e^{-\beta\lambda u_i}]$ from M energies u_i observed in a simulation is highly biased for wide distributions $p(u)$ and requires extremely large numbers M.[11]

Methods that have been advanced to overcome those difficulties include perturbation expansions.[7-9,12-18] But direct extension of perturbation theory beyond fourth order has been impractical. In contrast to perturbation expansions, interpolative approximations polynomial in λ have been more successful. For the charging of water and ions in water, expansions to order six and higher were necessary to account for the simulation data.[8,9,16] Thus, perturbation theory would be unsatisfactory in such cases and additional calculations would be required. For atomic ions, a kink was observed for $d\Delta\mu(\lambda)/d\lambda$ as a function of charge λ. A different methodological approach was developed many years ago by Wheeler and Gordon[19] for utilizing power moments of $p(u)$ to model averages such as eq 1. That approach could be expected here to be of value for modeling the chemical potential $\Delta\mu(\lambda)$ but has the disadvantage that direct examination of $p(u)$ would not be informative unless *many* moments were available. In the present examples, it is typically impractical to calculate more than about four moments.

On a broader level, however, the idea of modeling the distribution $p(u)$, perhaps including physical intuition in addition to the information available, is sure to be useful. This point of view has recently been helpful in understanding and computing hydrophobic effects[20-22] and in extending thermodynamic perturbation theory calculations.[23-26]

Here, we will attempt to represent the observed complicated properties of $p(u)$ considering a combination of simpler states. We attempt to find states $\{n\}$ with simple densities $p_n(u)$ such that the weighted combination of states will reproduce the behavior of the overall system,

$$p(u) = \sum_n w_n p_n(u) \tag{2}$$

with weights $w_n \geq 0$, $\sum_n w_n = 1$ and normalized densities $p_n(u) \geq 0$, $\int du\, p_n(u) = 1$. For electrostatic systems, we will seek $p_n(u)$'s of Gaussian form, representing the overall system as a linear combination of Gaussian subsystems, each showing linear response to electrostatic interactions. A natural, nonunique choice of the states $\{n\}$ is sorting with respect to the number and type of neighboring solvent molecules.[20-22,27] This is based on the assumption that for each number of neighboring molecules the dielectric response of the solvent will be approximately linear, whereas the characteristics of that linear response may vary considerably between solute states with different numbers of neighbors.

Representing $p(u)$ by a sum of Gaussian densities can give nontrivial results for the chemical potential, as can be seen by substituting eq 2 into eq 1,

$$\Delta\mu(\lambda) = -\beta^{-1} \ln \sum_n w_n e^{-\beta\lambda m_n + \beta^2\lambda^2\sigma_n^2/2} \tag{3}$$

where m_n and σ_n^2 are the mean and variance of the Gaussian p_n, respectively.

Results and Discussion

(a) Hydration of Water. We will now explore the validity of a multistate Gaussian model, considering two test cases: (a) the uncharging of a water molecule in water and (b) the charging of an ion in water. Accurate reference data are available for both systems. In the case of uncharging water, we performed a Monte Carlo (MC) simulation of 512 SPC water molecules[28] using Ewald-summation electrostatics[29] at a temperature of 298 K and a density of $\rho = 0.99707$ g/cm^3 (see ref 9 for further details). The electrostatic interaction energies u of each water molecule (solute) with all other water molecules (solvent) were calculated for 8000 configurations taken from 400 000 MC passes. The Ewald interaction potential was approximated by a tenth-order kubic-harmonic expansion.[9,30-32] The probability density of those energies u is shown in Figure 1. Also shown is a Gaussian probability density with identical mean and variance, which is somewhat wider in the tails, indicative of a negative kurtosis of $p(u)$.

---

[11] Wood, R. H.; Mühlbauer, W. C. F.; Thompson, P. T. *J. Phys. Chem.* **1991**, *95*, 6670.

[12] Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

[13] Hwang, J. K.; Warshel, A. *J. Am. Chem. Soc.* **1987**, *109*, 715.

[14] Levy, R. M.; Belhadj, M.; Kitchen, D. B. *J. Chem. Phys.* **1991**, *95*, 3627.

[15] Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 577.

[16] Hummer, G.; Szabo, A. *J. Chem. Phys.* **1996**, *105*, 2004.

[17] Liu, H.; Mark, A. E.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *100*, 9485.

[18] Archontis, G.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 11246.

[19] Wheeler, J. C.; Gordon, R. G. In *The Padé Approximant in Theoretical Physics*; Baker, G. A., Jr., Gammel, J. L., Eds.; Academic Press: New York, 1970; pp 99−128.

[20] Hummer, G.; Garde, S.; García, A. E.; Pohorille, A.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951.

[21] Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8800.

[22] Garde, S.; Hummer, G.; García, A. E.; Paulaitis, M. E.; Pratt, L. R. *Phys. Rev. Lett.* **1996**, *77*, 4966.

[23] Amadei, A.; Apol, M. E. F.; Di Nola, A.; Berendsen, H. J. C. *J. Chem. Phys.* **1996**, *104*, 1560.

[24] Apol, M. E. F.; Amadei, A.; Berendsen, H. J. C. *J. Chem. Phys.* **1996**, *104*, 6665.

[25] Apol, M. E. F.; Amadei, A.; Berendsen, H. J. C. *Chem. Phys. Lett.* **1996**, *256*, 172.

[26] Amadei, A.; Roccatano, D.; Apol, M. E. F.; Berendsen, H. J. C.; Di Nola, A. *J. Chem. Phys.* **1996**, *105*, 7022.

[27] Extension of the cavity distribution modeling from hard core model solutes to continuous van der Waals interaction could be based upon this idea also; the distribution of numbers of solvent molecule centers in a defined solute excluded volume *jointly* with the embedding van der Waals energy should be determined to obtain results for hydrophobic free energies.

[28] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces: Proceedings of the 14th Jerusalem Symposium on Quantum Chemistry and Biochemistry*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981; pp 331−342.

[29] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, UK, 1987.

[30] von der Lage, F. C.; Bethe, H. A. *Phys. Rev.* **1947**, *71*, 612.

[31] Adams, D. J.; Dubey, G. S. *J. Comput. Phys.* **1987**, *72*, 156.

[32] Hummer, G.; Soumpasis, D. M. *J. Chem. Phys.* **1993**, *98*, 581.

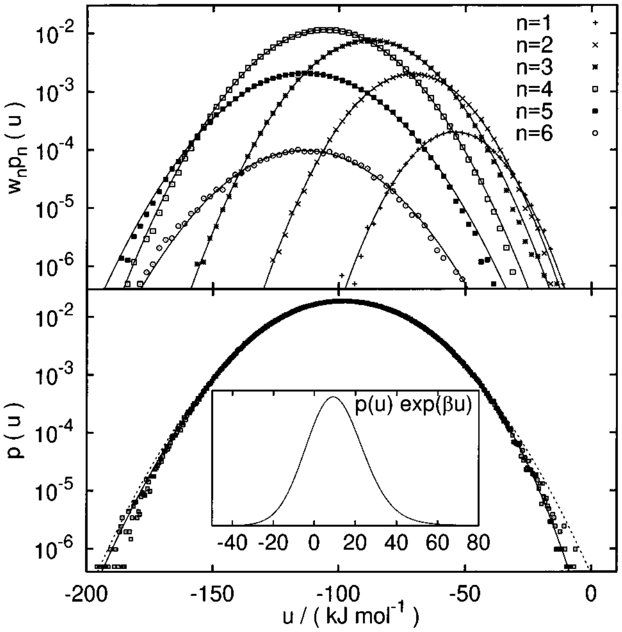

Figure 1. Probability density of the electrostatic energy *u* of a water molecule in water on a logarithmic scale (bottom panel). The square symbols show histogram data. The dashed line is the Gaussian probability density with estimated mean and variance. The solid line is the result of adding Gaussian probability densities for $n = 1$ to 6 neighbor atoms, as defined by an OH distance of 0.24 nm. Note that for this definition of *u* as the electrostatic interaction energy of the solute water molecule with the solvent, we have to change the sign of λ in eqs 1 and 3. Therefore, the upper tail of $p(u)$ is most important. This is illustrated by the inset showing $p(u)\exp(\beta u)$ that is significantly larger than zero in a region from about -30 to 60 kJ mol^{-1}, where essentially no histogram data are available. (The multistate Gaussian model is used for $p(u)$.) The top panel shows the substate contributions $w_n\, p_n(u)$ for $n = 1$ to 6 neighboring molecules as symbols together with the corresponding Gaussian approximations (lines). Gaussians with different mean values m_n and variances σ_n^2 are excellent approximations to the substate distributions of electrostatic energies *u*.

The "neighborhood" of a water molecule can be defined by using a hydrogen bonding criterion. We consider two molecules to be neighboring if any of their oxygen–hydrogen distances is below 0.24 nm, which is the first minimum in the OH pair-correlation function. For that criterion, we find an average number of neighbors of 3.64 with $w_n > 10^{-3}$ for $1 \leq n \leq 6$ and $n = 4$ being the most probable number of neighbors. We calculated the probability density of energies $p_n(u)$ by binning the energies *u* of a particular molecule in histograms according to the number *n* of its neighbors. The mean m_n and variance σ_n^2 of those probability densities give the Gaussian approximations $p_n(u) = (2\pi\sigma_n^2)^{-1}\exp[-(u - m_n)^2/2\sigma_n^2]$. The weighted sum of those Gaussian densities, $\sum_{n=1}^{6} w_n p_n(u)$, is an excellent approximation to the probability density $p(u)$, as shown in Figure 1. In particular, the tails of the distribution are better represented than in the simple approximation by one Gaussian. This is important because in the average eq 1, the tails dominate due to the Boltzmann factor. As illustrated in Figure 1 (top panel), Gaussians are indeed excellent approximations to the individual energy distributions $p_n(u)$ for different numbers of neighbors *n*.

This is even more evident in the calculated change of the chemical potential upon change of the charge state of a solute water molecule. The states $\lambda = 0$ and 1 correspond to a fully charged and uncharged solute water molecule, where the partial charges of that molecule are varied linearly with λ. In Figure 2, we compare the chemical-potential change calculated from a single Gaussian and from the sum over Gaussian probability densities for neighbors $n = 1$ to 6. The single-Gaussian approximation gives poor behavior, as observed previously,[9,33] and deviates from the reference curve for $\lambda > 0.25$ already. However, the chemical potential calculated from a weighted sum of Gaussian probability distributions deviates only for large perturbations $\lambda > 0.8$. The chemical-potential difference between a charged and an uncharged water molecule is accurate to within 5%. This is a surprising result, because $\Delta\mu(\lambda)$ is nonquadratic, requiring an eighth-order polynomial to fit the simulation data for chemical-potential derivatives.[8,9,16] It shows that even for large perturbations involving changes of the chemical potential of about $14k_BT$ the energy distribution $p(u)$ contains sufficient information, and that a multistate Gaussian theory is an accurate way of extracting that information to calculate chemical potential differences.

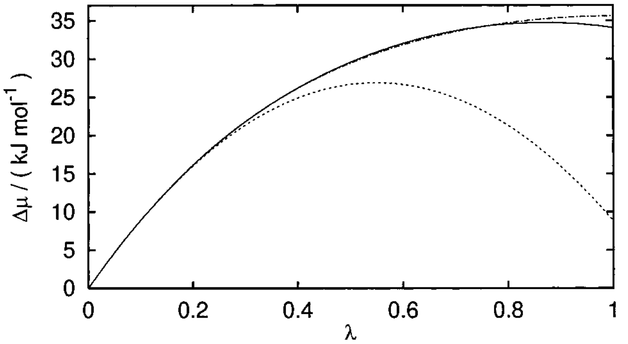

Figure 2. Chemical potential of a water molecule as a function of its charge state. $\lambda = 0$ and 1 correspond to fully charged and uncharged states, where the partial charges are varied linearly with λ. The solid curve shows the result of expanding about $\lambda = 0$ by adding Gaussian probability densities for $n = 1$ to 6 neighbor atoms. The dashed curve is the result of using a single Gaussian probability density. As a reference, the polynomial approximation of order eight to the chemical potential $\Delta\mu(\lambda)$ from ref 9 is shown with a dot-dashed line.

(b) Hydration of Ions. The second system used to investigate the applicability of a multistate Gaussian theory is that of an ion in water. We focus on an ion with methane Lennard-Jones parameters,[34] as studied previously.[8] We carried out a MC simulation of an uncharged methane particle Me in 128 SPC water molecules at 298 K (see ref 8 for details), analyzing 100 000 configurations taken from 10^6 MC passes. Figure 3 shows the probability density of the electrostatic energy *u* of charging that particle from zero charge to the elementary charge e. $p(u)$ is approximately Gaussian at the resolution available for this system.

The definition of "neighborhood" of a solvent molecule is somewhat more difficult for an uncharged solute than for water because the first peaks in the solute–water pair-correlation functions (for O and H) are broad. We decided to consider a water molecule as neighboring if either of the two Me–H distances was below 0.32 nm (where the Me–H pair correlation function reaches a value of one for the first time). That stringent criterion results in an average number of neighboring solvent molecules of 2.24 with weights $w_n > 10^{-3}$ for $0 \leq n \leq 7$. The weighted sum of Gaussian probability distributions according to that criterion gives a $p(u)$ distribution that is positively skewed, but otherwise close to the single-Gaussian approximation. Alternative criteria could also involve the oxygen position or the directionality of the hydrogen bond.

---

(33) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1994**, *116*, 3949.

(34) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.

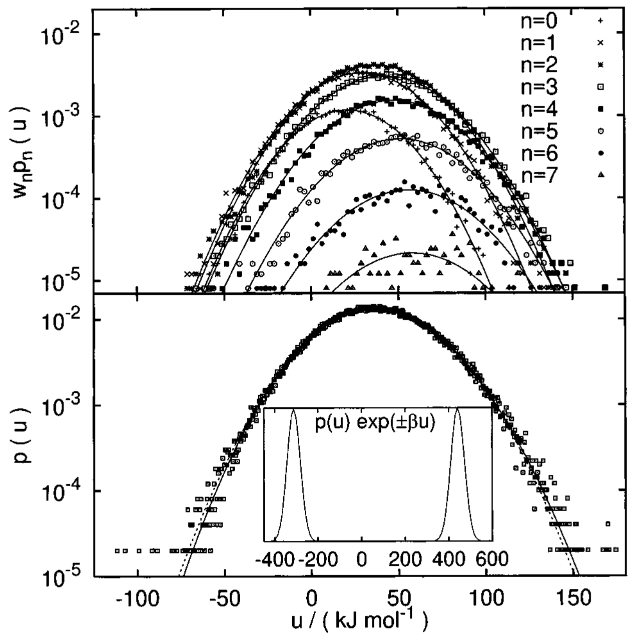

Figure 3. Probability density of the electrostatic energy *u* of a methane-size ion in water on a logarithmic scale, determined from a simulation of an uncharged ion in water (bottom panel). The dashed line is the Gaussian probability density with estimated mean and variance. The solid line is the result of adding Gaussian probability densities for $n = 0$ to 7 neighbor atoms, as defined by an ion−hydrogen distance of 0.32 nm. The inset shows $p(u)\ \exp(\pm\beta u)$, scaled appropriately for negative and positive ions to fit on the same graph. (The multistate Gaussian model is used for $p(u)$.) The two curves are significantly larger than zero below about −200 and above 350 kJ mol^{-1}, respectively, in regions far from where *u* histogram data have been collected. The top panel shows the substate contributions $w_n\ p_n(u)$ for $n = 0$ to 7 neighboring water molecules as symbols together with the corresponding Gaussian approximations (lines). As for water (Figure 1), Gaussians are excellent approximations to the substate distributions of electrostatic energies *u*.

The chemical potential of charging the methane-sized particle from 0 to ±e is shown in Figure 4 along with the first and second derivative with respect to the dimensionless charge $\lambda = q/\mathrm{e}$. The multistate Gaussian theory results in a qualitative improvement over a single-Gaussian approximation, with both the anion and cation chemical potentials in better agreement with the reference data from ref 8. However, quantitative agreement with the reference data is observed only for the cation data for the chemical potential and its derivatives. For the anion, the agreement is only qualitative. The first derivative of $\Delta\mu$, *i.e.*, the average potential at the ion position, is curved upwards, but the kink at $q \approx -0.2$ e is weaker than in the simulation data.

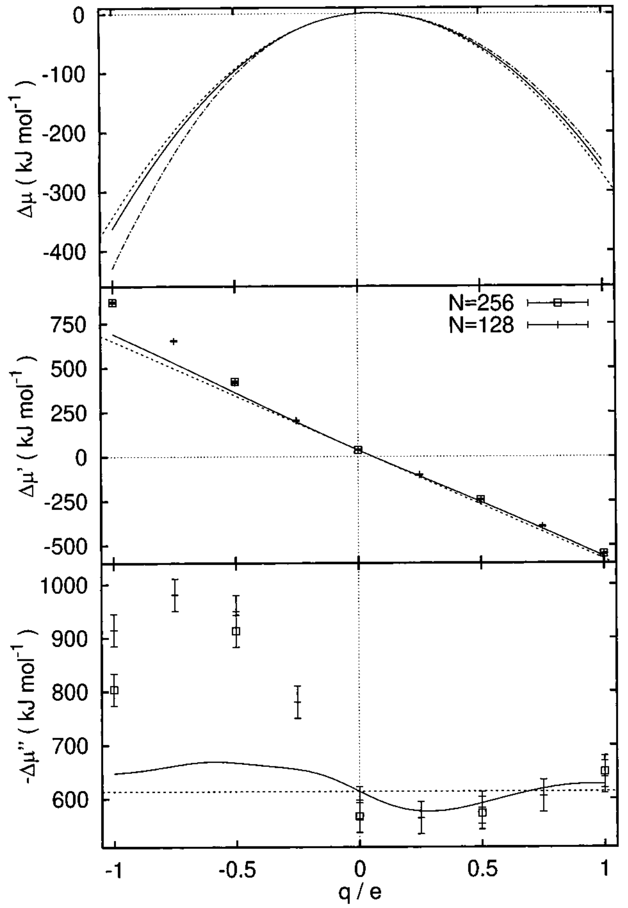

Figure 4. Chemical potential and its derivatives with respect to charge of a methane-size ion in water as a function of its charge state *q*. The chemical potential (top) and its first (middle) and second derivative (bottom) with respect to the charge *q*/*e* are shown. The results of a single Gaussian and of adding Gaussian probability densities for $n = 0$ to 7 neighbor molecules are shown with dashed and solid lines, respectively. For reference, the polynomial approximation p_8 of order eight (dot-dashed line) as well as simulation data for the first and second derivative from simulations of charged ions and $N = 128$ (plus) and 256 (square) water molecules are shown, taken from ref 8.

These discrepancies might be understood in part from the increased statistical errors of the ion data compared to the water case where every water molecule could be treated individually as a solute resulting in great sampling thoroughness. But the primary causes are the structural differences in the first hydration layer between the uncharged particle and the anion. The first peak of the anion−hydrogen pair-correlation function is in a region where the corresponding correlation for the uncharged particle is zero.[8] As a consequence, none of the configurations representative of anionic hydration are observed in the simulation of the uncharged Me particle. It may be possible to find a better reference system than the uncharged solute or a better indicator of significant structural substates. A negative charge state $q \approx -0.2$ e, for instance, should be a better reference state because the anionic hydration structure would be better represented. Consideration of the fully charged states of the ions should further increase the accuracy, in part because the problem of extrapolating from an uncharged reference state would then become a problem of interpolating between two states, which is generally simpler.[8] Moreover, the separation into well-defined substates would be simplified because of the more ordered structure in the first hydration layer of a charged ion compared to an uncharged, spherical Lennard-Jones particle. Substate criteria could again be the number of hydrogen bonds or the number of solvent molecules in the first hydration layer. Nevertheless, the multistate Gaussian model can give correct indication of the changes when going from positive to negative ionic charges. This multistate Gaussian model suggests the two different quadratic regimes of the chemical potential for anions and cations.

Conclusions

We found quantitative agreement between the multistate Gaussian model and reference data for uncharging of a water molecule in water and for charging a cation and qualitative agreement for charging an anion in water. The multistate Gaussian model produces the correct nonlinear behavior directly from the systematic decomposition of the electrostatic energy distributions into a sum of Gaussians based on structural substates. The observed agreement is not the consequence of a numerical fit. These results are encouraging in several ways. (1) They indicate that we can indeed find simple criteria identifying a small set of states, each exhibiting approximately linear response, with the whole producing a nonlinear result for the system thermodynamics. This has implications regarding the development of improved dielectric continuum models that contain molecular information about the solvent.[6,33] A weighted mixture of a few continuum-dielectric systems can give the complicated behavior stemming from nonlinear dielectric response. This is a specific and effective alternative to the ideas of electrostriction and dielectric saturation for these problems. (2) On the practical side, the multistate Gaussian model provides a way of maximizing the use of data of a single simulation for the calculation of free energies. As shown in the water case, large energetic perturbations can be manageable with this approach. (3) A decomposition of a complicated probability density into a weighted sum of simpler ones is a general concept. A possible area of application includes glasses and amorphous materials, where one expects considerable differences of the energy distributions of a particle depending upon its local environment, but where the probability densities should be simpler when sorted according to coordination numbers. For instance, Sciortino, Geiger, and Stanley[35] investigated the correlation of the diffusivity of water molecules with binding energies and coordination numbers in quenched configurations of water.

A particularly interesting area of application is the analysis of simulation data for proteins and nucleic acids. Dielectric continuum models with empirical atomic radii and dielectric constants are used with some success for the calculation of pK_a values of ionizable groups.[5] Recently, advances were made in calculating pK_a values of ionizable groups from computer simulations.[36−40] However, that analysis is still awkward and tedious, often involving several independent free-energy simulations for each ionizable group. That effort can be greatly reduced if accurate free energies of protonation can be calculated from a small number of equilibrium simulations. Equilibrium simulations have the great advantage that one can extract additional structural, dynamic, and energetic information. The accuracy of such an analysis should increase considerably if the multistate Gaussian model is used.

The problem of multiple protonation sites could be treated more reliably without simulation of each of the possible charge states if an accurate extrapolation method such as the multistatate Gaussian theory is used. That analysis would then eliminate a number of energetically unfavorable protonation states such that only the relevant states need to be considered for a more detailed analysis. High accuracy will in general require interpolation between at least two reference states (*e.g.*, charged and uncharged), because extrapolation from one state alone is difficult when the free energy changes by several hundred kilojoules per mole.[8]

A multistate Gaussian analyis of amino acid protonation states might require sorting of configurations according to the changes in the local environment of the ionizable group. For instance, the properties should be different if another amino acid side chain of a protein is interacting directly with the group under consideration during parts of the simulation. This suggests a classification of states based on the number and type of interacting amino acids, where each state can exhibit approximately Gaussian statistics of electrostatic energy fluctuations.

JA971148U

---

(35) Sciortino, F.; Geiger, A.; Stanley, H. E. *Nature* **1991**, *354*, 218.

(36) Warshel, A.; Sussman, F.; King, G. *Biochemistry* **1986**, *25*, 8368.

(37) Merz, K. M. *J. Am. Chem. Soc.* **1991**, *113*, 3572.

(38) Del Buono, G. S.; Figueirido, F. E.; Levy, R. M. *Proteins Struct. Funct. Genet.* **1994**, *20*, 85.

(39) Saito, M. *J. Phys. Chem.* **1995**, *99*, 17043.

(40) Figueirido, F.; Del Buono, G. S.; Levy, R. M. *J. Phys. Chem.* **1996**, *100*, 6389.