The dependence of electrostatic solvation energy on dielectric constants in Poisson-Boltzmann calculations

Harianto Tjong and Huang-Xiang Zhou

Citation: The Journal of Chemical Physics **125**, 206101 (2006); doi: 10.1063/1.2393243
View online: http://dx.doi.org/10.1063/1.2393243
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/125/20?ver=pdfcov
Published by the AIP Publishing

The dependence of electrostatic solvation energy on dielectric constants in Poisson-Boltzmann calculations

Harianto Tjong and Huang-Xiang Zhou[a)]
Department of Physics and Institute of Molecular Biophysics and School of Computational Science, Florida State University, Tallahassee, Florida 32306

(Received 19 September 2006; accepted 18 October 2006; published online 22 November 2006)

[DOI: 10.1063/1.2393243]

The Poisson-Boltzmann (PB) equation is widely used for modeling electrostatic effects and solvation of biomolecules.[1–4] There are separate reasons for varying the solvent and solute dielectric constants (ε_s and ε_i) in PB calculations. The former case arises when the solvation energy of a solute in different solvents is of interest. The latter case arises because there is uncertainty on the value of the solute dielectric constant ε_i due to the empirical nature of the PB equation.[5] The purpose of this note is to present a simple formula that accurately predicts the electrostatic free energy for all combinations of ε_i and ε_s from the PB calculation on a single set of ε_i and ε_s values.

The formula has the form

$$\frac{\Delta G(\varepsilon_i,\varepsilon_s)}{1/\varepsilon_i - 1/\varepsilon_s} = \frac{\Delta G(\varepsilon_{ir},\varepsilon_{sr})}{1/\varepsilon_{ir} - 1/\varepsilon_{sr}} f(\varepsilon_i/\varepsilon_s), \tag{1}$$

$$f(\varepsilon_i,\varepsilon_s) = \frac{A + 2B\varepsilon_i/\varepsilon_s}{1 + 2\varepsilon_i/\varepsilon_s}, \tag{2}$$

where ε_{ir} and ε_{sr} are "reference" solute and solvent dielectric constants for which the PB equation is actually solved once, and A and B are fitting parameters. The factor $f(\varepsilon_i/\varepsilon_s)$ scales the solvation energy $\Delta G(\varepsilon_{ir},\varepsilon_{sr})$, calculated at the reference dielectric constants into the corresponding quantity at any desired dielectric constants. To find the fitting parameters, we solved the PB equation for eight different combinations of ε_i and ε_s, with the former at 1–4, and the latter at either 78.5 (for water) or 10.3 (for *n*-octanol), on a set of 55 proteins. These proteins have less than 10% sequence identities, better than 1.0 Å resolutions, and less than 250 residues. For the reference we chose $\varepsilon_{ir}=2$ and $\varepsilon_{sr}=78.5$. The Protein Data Bank name, the number of atoms (N_{atom}), the net charge (Q), and the fitting parameters of the 55 proteins are listed in Table I.

The fitting parameters have interesting properties. A shows very small variations among the 55 proteins. The average is 1.016, with a standard deviation of only 0.004. The scaling factor for the limiting situation $\varepsilon_i/\varepsilon_s \to 0$ is A; a near unity value reflects the closeness of the reference to the limiting situation. Nonetheless the variations of A appear to anticorrelate with the magnitude of the net charge and correlate with the number of atoms. A multilinear regression against $|Q|^{0.65}$ and N_{atom} gives the best fit as

TABLE I. Number of atoms, net charge, and fitting parameters of 55 proteins.

PDB	N_{atom}	Q	A	B
1a6m	2435	2	1.017	0.67
1aho	967	−2	1.017	0.66
1byi	3383	−4	1.018	0.63
1c75	987	−4	1.013	0.75
1c7k	1929	−5	1.019	0.61
1cex	2867	1	1.017	0.67
1eb6	2572	−15	1.012	0.75
1ejg	678	0	1.022	0.55
1etl	145	0	1.022	0.55
1exr	2240	−25	1.009	0.81
1f94	982	1	1.018	0.64
1f9y	2535	−5	1.017	0.66
1g4i	1842	−1	1.019	0.62
1g66	2794	−2	1.023	0.51
1gqv	2143	7	1.017	0.66
1hje	179	1	1.015	0.71
1iqz	1171	−17	1.005	0.89
1iua	1207	−1	1.018	0.65
1j0p	1597	8	1.013	0.74
1k4i	3253	−6	1.018	0.64
1kth	894	0	1.017	0.66
1l9l	1230	11	1.010	0.82
1mlq	1265	−7	1.013	0.75
1nls	3564	−7	1.021	0.55
1nwz	1912	−6	1.015	0.70
1od3	1900	−3	1.020	0.59
1ok0	1076	−5	1.016	0.69
1p9g	529	4	1.015	0.69
1pq7	3065	4	1.021	0.58
1r6j	1230	0	1.016	0.68
1ssx	2750	8	1.017	0.67
1tg0	1029	−12	1.009	0.83
1tqg	1660	−7	1.013	0.75
1tt8	2676	1	1.018	0.63
1u2h	1526	4	1.015	0.72
1ucs	997	0	1.017	0.66
1ufy	1926	−3	1.017	0.66
1unq	1966	−3	1.016	0.69
1vb0	921	3	1.016	0.70
1vbw	1058	8	1.012	0.77
1w0n	1856	−5	1.019	0.61
1wy3	560	1	1.015	0.71
1x6z	1741	0	1.016	0.68
1x8q	2815	−1	1.019	0.60
1xmk	1268	1	1.016	0.68

TABLE I. (*Continued.*)

PDB	N_{atom}	Q	A	B
1yk4	770	−8	1.009	0.83
1zzk	1252	1	1.016	0.69
2a6z	3432	−3	1.018	0.65
2bf9	560	−2	1.012	0.77
2chh	1624	−3	1.019	0.61
2cws	3400	−3	1.020	0.59
2erl	573	−6	1.010	0.81
2fdn	731	−8	1.010	0.80
2fwh	1830	−6	1.014	0.72
3lzt	1960	8	1.016	0.68

$$A = -1.63 \times 10^{-3}|Q|^{0.65} + 2.18 \times 10^{-6}N_{atom} + 1.016, \quad (3)$$

with $R^2=0.68$. A similar regression analysis yields

$$B = 3.31 \times 10^{-2}|Q|^{0.65} - 4.77 \times 10^{-5}N_{atom} + 0.683, \quad (4)$$

with $R^2=0.65$. Note that the correlation trends of A and B are opposite to each other.

Using A and B values fixed by Eqs. (3) and (4) and the PB results for the reference dielectric constants, the solvation energy given by Eq. (1) agrees very well with the actual PB results for the other seven sets of dielectric constants. With $(\varepsilon_i, \varepsilon_s)=(1, 78.5)$, (3, 78.5), (4, 78.5), (1, 10.3), (2, 10.3), (3, 10.3), and (4, 10.3), the average unsigned relative errors are 0.2%, 0.1%, 0.2%, 0.6%, 1.2%, 1.5%, and 1.7%, respectively. The magnitude of the errors corresponds to the deviation of $\varepsilon_i/\varepsilon_s$ from the reference value $\varepsilon_{ir}/\varepsilon_{sr}$. To show that Eq. (1) is indeed predictive, we repeated the procedure using only 29 of the 55 proteins. Equation (1) was then applied to the remaining 26 proteins. The corresponding errors are 0.2%, 0.1%, 0.2%, 0.6%, 1.2%, 1.7%, and 2.0%, showing at most a slight deterioration. Equation (1) also works well when salt effects are included. For example, for ionic strengths between 0.1 M and 0.5 M, the average unsigned relative errors are 0.2%, 0.2%, and 0.4%, respectively, for the $(\varepsilon_i, \varepsilon_s)=(1, 78.5)$, (3, 78.5), and (4, 78.5).

For a spherical model (with radius R) having all the charge located at the center, the Born formula,[6]

$$\Delta G(\varepsilon_i, \varepsilon_s) = -(1/\varepsilon_i - 1/\varepsilon_s)Q^2/R, \quad (5)$$

corresponds to $A=1$ and $B=0$. For spherical-model proteins with arbitrary distributions of point charges, calculations with the Tanford and Kirkwood formula[7] lead to fitted A and B values that behave similarly to those for the actual proteins. Namely, the values of A show only small fluctuations above unity and both $-A$ and B show correlation with the magnitude of the net charge and anticorrelation with the number of point charges. On the other hand, the values of the coefficients in Eqs. (3) and (4) reflect the general properties of proteins with up to 250 residues. The coefficients will likely vary somewhat, for example, for nucleic acids.

Based on the Tanford and Kirkwood formula[7] for spherical-model proteins, Sigalov *et al.*[8] suggested a scaling formula similar to Eq. (1) but with the coefficient of $\varepsilon_i/\varepsilon_s$ set to 0.57, instead of our value of 2, in their development of a generalized Born method. We found less success with their coefficient in reproducing the dependence of the solvation energy on dielectric constants.

In summary, we obtained a simple, accurate formula for predicting the dependence of the electrostatic free energy on solute and solvent dielectric constants. This formula will facilitate the application of the Poisson-Boltzmann equation and be useful for the development of generalized Born methods.

This work was supported in part through Grant No. GM058187 by the National Institutes of Health.

[a]Electronic mail: zhou@sb.fsu.edu

[1] K. A. Sharp and B. Honig, Annu. Rev. Biophys. Biophys. Chem. **19**, 301 (1990).
[2] J. D. Madura, J. M. Briggs, R. C. Wade *et al.*, Comput. Phys. Commun. **91**, 57 (1995).
[3] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, and J. A. McCammon, Proc. Natl. Acad. Sci. U.S.A. **98**, 10037 (2001).
[4] N. A. Baker, Curr. Opin. Struct. Biol. **15**, 137 (2005).
[5] J. Antosiewicz, J. A. McCammon, and M. K. Gilson, J. Mol. Biol. **238**, 415 (1994).
[6] M. Born, Z. Phys. **1**, 45 (1920).
[7] C. Tanford and J. G. Kirkwood, J. Am. Chem. Soc. **79**, 5333 (1957).
[8] G. Sigalov, P. Scheffel, and A. Onufriev, J. Chem. Phys. **122**, 094511 (2005).