

Determination of the Effective Dielectric Constant from the Accurate Solution of the Poisson Equation

VLADISLAV VASILYEV

ANU Supercomputer Facility, Leonard Huxley Building 56, Canberra, ACT 0200, Australia
Biotechnology Research Institute, National Research Council, 6100 Royalmount Ave., Montreal,
Quebec, Canada, H4P 2R2

Received 17 September 2001; Accepted 30 April 2002

Abstract: Constant dielectric (CD) and distance-dependent dielectric (DDD) functions are the most popular and widespread in the Molecular Mechanics simulations of large molecular systems. In this article, we present a simple procedure to derive an effective dielectric constant, $\epsilon_{\text{out,eff}}$, for these two methods based on numerical solutions of the Poisson equation. It was found that because of the very approximate nature of the CD and DDD models there is no universal $\epsilon_{\text{out,eff}}$ which will work equally well for all molecular systems. For example, different MD trajectories of the same molecule can produce different optimal $\epsilon_{\text{out,eff}}$ s. The DDD function was found to yield better agreement with the numerical solutions of the Poisson equation than a CD model does. The reason is that a DDD function gives a better description of the electrostatic interactions at short distances between the atoms. Another interesting finding of this study is that under certain conditions $\epsilon_{\text{out,eff}}$ can take negative values for a system of two atoms at a limited distance range. However, in principle, there is nothing to prevent the $\epsilon_{\text{out,eff}}$ from taking negative values for specific conformations of some molecules.

© 2002 Wiley Periodicals, Inc. J Comput Chem 23: 1254–1265, 2002

Key words: effective dielectric constant; constant dielectric function; distant dependent dielectric function; the Poisson equation; boundary element method

Introduction

Electrostatic interactions and solvation effects play an important role in the structure and function of macromolecules in solution.^{1–5} There are two types of approaches for calculating the electrostatic energy of the macromolecules in solution in a more or less quantitative manner—microscopic and continuum methods. The major problem in applying microscopic approaches,^{6–10} such as slow growth simulations, thermodynamic integration, free energy perturbation, or linear response approximation, is that the presence of the explicit water molecules drastically increases the computational time for macromolecules such as proteins.

In continuum electrostatic models the solute is treated as a low-dielectric cavity embedded in a high-dielectric medium representing the solvent. The Poisson or Poisson–Boltzmann (PB) equation is then solved numerically using either finite difference¹¹ or boundary element (BEM)¹² methods. Although relatively inexpensive computationally, numerical continuum techniques are still too slow to be practical for the study of processes that require numerous evaluations of the solvation energy, so a number of approximate continuum approaches have been developed,^{13–24} among which the so-called generalized Born (GB) approxima-

tion^{13–22} seems to be the most promising (see ref. 4 for the review of GB).

At the low end of the continuum methods are very fast, simple, and popular solvation models such as constant dielectric (CD) and distance dependent dielectric (DDD) functions. Although CD and DDD functions are not based upon rigorous physics, as the Poisson equation is, they can be easily implemented into molecular mechanics programs and provide a fast evaluation of solvation effects, with the gradients and second derivatives of the electrostatic energy being readily available. At the same time there is a great demand for assessing the reliability of CD and DDD functions for large systems by comparison with the reference solvation energies. Because of the lack of experimental data on the solvation free energies for large molecules numerical solutions of the Poisson equation represent an ideal choice as reference values for the solvation free energies.^{14,25,26}

In this article we present a simple method to derive the effective dielectric constants for the CD and DDD models based on

Correspondence to: V. Vasilyev, ANU Supercomputer Facility, Australian National University, Leonard Huxley Bld., B56, Mills Road, Canberra, ACT 0200, Australia; e-mail: vv900@anuf.anu.edu.au

numerical solutions of the Poisson equation. GB-type models are the target of our next study, and are beyond the scope of the present article.

Formulation and Methods

CD and DDD functions

The solvent serves the dual function of solvating each individual charge of the molecule, which determines the self-energies of the charges, and screening the interaction between charge pairs. The electrostatic energy, $E_{\text{El/static}}$, of a system of charges can be calculated by integrating the energy density $u(\vec{x}) = \vec{D}^2/8\pi\epsilon(\vec{x})$ of the electric field over all space:^{14,27}

$$E_{\text{El/static}} = \int_{R^3} u(\vec{x})d^3x = \frac{1}{8\pi} \int_{R^3} \frac{\vec{D}^2(\vec{x})}{\epsilon(\vec{x})} d^3x \quad (1)$$

where R^3 denotes the infinite three-dimensional Cartesian space, $\vec{D}(\vec{x})$ is the dielectric displacement, and $\epsilon(\vec{x})$ is the position-dependent dielectric constant. For N charges $\vec{D}(\vec{x})$ is the superposition of their electric displacements $\vec{D}(\vec{x}) = \sum \vec{D}_i(\vec{x})$, and the electrostatic energy splits up into $N(N-1)/2$ crossterms with $\vec{D}_i\vec{D}_j$ and N quadratic terms with \vec{D}_i^2 .¹⁴

$$E_{\text{El/static}} = \frac{1}{4\pi} \sum_{i<j} \int_{R^3} \frac{\vec{D}_i(\vec{x})\vec{D}_j(\vec{x})}{\epsilon(\vec{x})} d^3x + \frac{1}{8\pi} \sum_i \int_{R^3} \frac{\vec{D}_i^2}{\epsilon(\vec{x})} d^3x \quad (2)$$

In the case of a homogeneous dielectric function, the first term yields the Coulomb interaction energy. The integrals in the second term diverge for the point charges but after making the assumption that each charge, q_i , is uniformly distributed on the surface of a sphere with radius R_i the second term, it gives us the sum of the Born self-energies. So, finally we have:¹⁴

$$E_{\text{El/static}} = \sum_{i<j} \frac{q_i q_j}{\epsilon r_{ij}} + \sum_i \frac{q_i^2}{2\epsilon R_i}, \text{ if } \epsilon(\vec{x}) = \epsilon = \text{const} \quad (3)$$

Thus, application of the CD function in molecular simulations involves the use of two basic approximations. First, one assumes a homogeneous dielectric constant instead of a position-dependent one. One may expect that this assumption is partly justified for small solutes whose atoms are highly exposed to solvent, so that there is a small departure of the field from the Coulomb one. However, this assumption is questionable for macromolecules. The self-solvation energies of the atoms in the molecule are omitted from consideration in the second approximation. The atoms' self-solvation energies depend on the geometry of the molecule, and can play an important role in conformational changes.

Contrary to the CD model, which can be derived from the integration of the Coulomb field, the DDD function represents an ad hoc assumption in employing the distant-dependent dielectric constant of the form ϵr_{ij} . Plausible explanations for the use of the DDD model can include reasoning that even powers of r make a computationally convenient expression for nonbonded energies²⁸

or that in the regions of the most important interactions from 2 to 5 Å the dielectric constant varies from 2 to 5.²⁸ Also, as in the case of the CD function, the atoms' self-solvation energies are omitted in the DDD approximation.

Another basic approximation of both solvation models is that they reduce a complex calculation of the molecular solvation free energy, which is not simply additive and nonpairwise by nature, and which is dependent on the particular charge distribution, molecular shape, etc., to the pairwise interactions between the atoms.

The key questions are how do these approximate models behave in the case of macromolecules and what are the optimal values of ϵ s in both models? To address this problem we choose the numerical solutions of the Poisson equation as a reference point to test both approximate methods.

Obtaining the Effective Dielectric Constant

The total electrostatic energy of a molecule in Molecular Mechanics calculations using solvation free energy obtained from the numerical solution of Poisson equation can be expressed as a sum of the Coulomb and solvation energies:

$$E_{\text{El/static}} = E(\epsilon_{\text{in}})_{\text{Coul}} + E(\epsilon_{\text{in}}, \epsilon_{\text{out}})_{\text{Solv}} \quad (4)$$

or more specifically:

$$E_{\text{El/static}} = \sum_{i<j} \frac{q_i q_j}{\epsilon_{\text{in}} r_{ij}} + E(\epsilon_{\text{in}}, \epsilon_{\text{out}})_{\text{Solv}} \quad (4a)$$

where ϵ_{in} and ϵ_{out} the dielectric constants of solute and solvent, respectively.

Using accurate electrostatic solvation energies as reference values we can derive several approaches to obtaining the effective dielectric constant, $\epsilon_{\text{out,eff}}$, for the CD and DDD functions.

The simplest way to evaluate $\epsilon_{\text{out,eff}}$ is to set an approximate total electrostatic energy equal to the reference total electrostatic energy. So, for the CD function we have:

$$\sum_{i<j} \frac{q_i q_j}{\epsilon_{\text{out,eff}} r_{ij}} = \sum_{i<j} \frac{q_i q_j}{\epsilon_{\text{in}} r_{ij}} + E(\epsilon_{\text{in}}, \epsilon_{\text{out}})_{\text{Solv}} \quad (5)$$

It immediately follows that

$$\epsilon_{\text{out,eff}} = \frac{\sum_{i<j} \frac{q_i q_j}{r_{ij}}}{\sum_{i<j} \frac{q_i q_j}{\epsilon_{\text{in}} r_{ij}} + E(\epsilon_{\text{in}}, \epsilon_{\text{out}})_{\text{Solv}}} \quad (6)$$

By analogy we have for the DDD function:

$$\varepsilon_{\text{out,eff}} = \frac{\sum_{i<j} \frac{q_i q_j}{r_{ij}^2}}{\sum_{i<j} \frac{q_i q_j}{\varepsilon_{\text{in}} r_{ij}} + E(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{\text{Solv}}} \quad (7)$$

The reference solvation energy in eqs. (6) and (7) is always a negative value while the Coulomb terms can take either negative or positive values. As a result, we can obtain either negative or positive values for the $\varepsilon_{\text{out,eff}}$. In the present study, we used this approach in a modified form to calculate $\varepsilon_{\text{out,eff}}$ for a simple system of two atoms. In the modified expressions we subtracted the Born energies of the isolated atoms from the reference solvation energies, $E(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{\text{Solv}}$, to be consistent with the CD and DDD total electrostatic energies that do not contain self-solvation terms.

Another approach for calculating $\varepsilon_{\text{out,eff}}$ is to use the relative solvation energies rather than absolute ones. The rationale for deriving $\varepsilon_{\text{out,eff}}$ from the relative solvation energies is that for the Molecular Dynamics and optimization techniques a change in solvation energy is more important than its absolute value. To extract “solvation” energies from the CD and DDD models we can rewrite expressions for the total electrostatic energies using the Poisson solvation energy, CD, and DDD functions as follows:

$$E_{\text{El/static}} = \underbrace{\sum_{i<j} \frac{q_i q_j}{\varepsilon_{\text{in}} r_{ij}}}_{\text{Coulomb Energy}} + \underbrace{E(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{\text{Solv}}}_{\text{Solvation Energy}} \quad (8)$$

$$E_{\text{El/static/CD}} = \sum_{i<j} \frac{q_i q_j}{\varepsilon_{\text{out}} r_{ij}} = \underbrace{\sum_{i<j} \frac{q_i q_j}{\varepsilon_{\text{in}} r_{ij}}}_{\text{Coulomb Energy}} + \underbrace{\sum_{i<j} \left(\frac{1}{\varepsilon_{\text{out}}} - \frac{1}{\varepsilon_{\text{in}}} \right) \frac{q_i q_j}{r_{ij}}}_{\text{“Solvation” Energy}} \quad (9)$$

$$\begin{aligned} E_{\text{El/static/DDD}} &= \sum_{i<j} \frac{q_i q_j}{\varepsilon_{\text{out}}^2 r_{ij}^2} \\ &= \underbrace{\sum_{i<j} \frac{q_i q_j}{\varepsilon_{\text{in}} r_{ij}}}_{\text{Coulomb Energy}} + \underbrace{\sum_{i<j} \left(\frac{1}{\varepsilon_{\text{out}}} - \frac{r_{ij}}{\varepsilon_{\text{in}}} \right) \frac{q_i q_j}{r_{ij}^2}}_{\text{“Solvation” Energy}} \end{aligned} \quad (10)$$

There is an analogy between the Poisson solvation energy and the terms responsible for the solvation in the energy expressions for the CD and DDD functions.

Now, for N conformations of a given molecule we can construct an error function to calibrate an effective $\varepsilon_{\text{out,eff}}$ for the CD or DDD models against the reference solvation energies at given ε_{in} and ε_{out} :

$$\begin{aligned} \text{Err}(\varepsilon_{\text{out,eff}}) &= \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N (\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{ij,\text{Ref}} - \Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out,eff}})_{ij,\text{Approx}})^2}{N(N-1)/2} \end{aligned} \quad (11)$$

where $\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{ij,\text{Ref}}$ is a difference in the absolute electrostatic solvation energies between the i th and j th conformations of the molecule using the reference method, i.e.,

$$\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{ij,\text{Ref}} = \Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{i,\text{Ref}} - \Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{j,\text{Ref}} \quad (12)$$

$\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out,eff}})_{ij,\text{Approx}}$ has the same meaning as $\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{ij,\text{Ref}}$ but $\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out,eff}})_{i,\text{Approx}}$ and $\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out,eff}})_{j,\text{Approx}}$ are calculated using “solvation” energy terms for either the CD or DDD function.

Taking into account that at optimal value of $\varepsilon_{\text{out,eff}}$

$$\frac{\partial \text{Err}(\varepsilon_{\text{out,eff}})}{\partial \varepsilon_{\text{out,eff}}} = 0$$

$\varepsilon_{\text{out,eff}}$ can be expressed for the CD and DDD functions as

$$\varepsilon_{\text{out,eff}} = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \Delta E_{ij}^2}{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \Delta E_{ij} \left(\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{ij,\text{Ref}} + \frac{1}{\varepsilon_{\text{in}}} \Delta E_{ij} \right)} \quad \text{for CD} \quad (13)$$

$$\varepsilon_{\text{out,eff}} = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \Delta K_{ij}^2}{\sum_{i=1}^{N-1} \sum_{j=i+1}^N \Delta K_{ij} \left(\Delta\Delta G(\varepsilon_{\text{in}}, \varepsilon_{\text{out}})_{ij,\text{Ref}} + \frac{1}{\varepsilon_{\text{in}}} \Delta E_{ij} \right)} \quad \text{for DDD} \quad (14)$$

and

$$\Delta E_{ij} = E_i - E_j$$

$$\Delta K_{ij} = K_i - K_j$$

where E_i and K_i are Coulomb energies of the molecule at the dielectric constants 1 and r_{ij} , respectively.

In the majority of force fields the Coulomb energy is calculated only for the atoms participating in $1 \geq 4$ interactions ($1-4$ interactions as well as beyond $1-4$ ones), with the $1-4$ ones being scaled by a factor. So, in such force fields E_i and K_i are expressed as:

$$E = \sum_{1>4} \frac{q_i q_j}{r_{ij}} + F \sum_{1-4} \frac{q_i q_j}{r_{ij}} \quad (15a)$$

$$K = \sum_{1>4} \frac{q_i q_j}{r_{ij}^2} + F \sum_{1-4} \frac{q_i q_j}{r_{ij}^2} \quad (15b)$$

where F is a scaling factor. So, changes in bond lengths and angle values are not reflected in the CD and DDD “solvation” energies what introduces another source of errors is the calculation of the

Table 1. $\epsilon_{\text{out,eff}}$ for *Ace-(Ala)₁₀-Nme* Calculated from 100 Snapshots Taken Along the MD Trajectories at 100-fs Intervals. Reference BEM solvation energies were calculated at $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 1/78.5$. No cutoff was used for the electrostatic interactions.

Trajectory <i>N</i>	CD function			DDD function		
	$\epsilon_{\text{out,eff}}^a$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(78.5)^b$	$\epsilon_{\text{out,eff}}^c$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(4)^b$
1	1.60	30.07	68.51	3.06	12.21	12.37
2	1.41	42.98	132.65	1.76	9.22	13.50
3	1.43	44.17	121.39	1.35	55.51	85.00
4	1.81	15.75	30.69	1.66	5.86	8.92
5	1.20	51.76	214.71	0.80	33.14	64.01
Mean	1.49	36.95	113.59	1.73	23.19	36.76

^a $\epsilon_{\text{out,eff}}$ is calculated using eq. (13).

^bError is calculated using error function (11).

^c $\epsilon_{\text{out,eff}}$ is calculated using eq. (14).

Coulomb energy in the AMBER-like²⁹ force fields. The use of a cutoff in the Molecular Mechanics simulations introduces an additional source of errors in calculating the “solvation” energies.

In the case of CD model we can also derive $\epsilon_{\text{out,eff}}$ by linear regression, i.e., for *N* conformations of a given molecule we need to minimize the function:

$$f(a, b) = \sum_{i=1}^N (\Delta G_{i,\text{Ref}} - (a\Delta G_{i,\text{CD}} + b))^2 \quad (16)$$

where $\Delta G_{i,\text{Ref}}$ is a reference solvation energy and $\Delta G_{i,\text{CD}}$ is a CD “solvation” one. Finally, we have

$$\epsilon_{\text{out,eff}} = \frac{1}{a + \frac{1}{\epsilon_{\text{in}}}} \text{ where}$$

$$a = \frac{\frac{1}{N} \sum \Delta G(\epsilon_{\text{in}}, \epsilon_{\text{out}})_{i,\text{Ref}} \sum E_i - \sum (E_i \Delta G(\epsilon_{\text{in}}, \epsilon_{\text{out}})_{i,\text{Ref}})}{\frac{1}{N} \sum E_i \sum E_i - \sum E_i^2} \quad (17)$$

where E_i is the Coulomb energy of the system when the dielectric constant is equal to 1. For our test systems, eq. (17) gave similar values of $\epsilon_{\text{out,eff}}$ to those calculated using eq. (13), so we reported results for the CD models using eq. (13).

Reference Electrostatic Solvation Free Energy Calculations

The reference electrostatic solvation free energies were calculated using a boundary element solution of the Poisson equation. The reaction field energies were computed using the BRI BEM program^{30,31} and the SIMS molecular surface program.³² The atomic radii used for reaction field energy calculations were taken from

Table 2. $\epsilon_{\text{out,eff}}$ for *Ace-(Ala)₁₀-Nme* Calculated from 100 Snapshots Taken Along the MD Trajectories at 100-fs Intervals. Reference BEM solvation energies were calculated at $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 2/78.5$.

Trajectory <i>N</i>	CD function			DDD function		
	$\epsilon_{\text{out,eff}}^a$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(78.5)^b$	$\epsilon_{\text{out,eff}}^c$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(4)^b$
1	3.13	7.28	16.93	5.72	2.97	3.11
2	2.77	10.28	32.68	3.40	2.21	2.29
3	2.79	10.54	29.98	2.63	13.58	15.67
4	3.53	3.75	7.51	3.21	1.40	1.49
5	2.37	12.44	52.99	1.58	8.31	12.86
Mean	2.92	8.86	28.02	3.31	5.69	7.08

^a $\epsilon_{\text{out,eff}}$ is calculated using eq. (13).

^bError is calculated using error function (11).

^c $\epsilon_{\text{out,eff}}$ is calculated using eq. (14).

Table 3. $\epsilon_{\text{out,eff}}$ for *Crambin* Calculated from 100 Snapshots Taken Along the MD Trajectories at 100-fs Intervals. Reference BEM solvation energies were calculated at $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 1/78.5$.

Trajectory <i>N</i>	CD function			DDD function		
	$\epsilon_{\text{out,eff}}^{\text{a}}$	$Err(\epsilon_{\text{out,eff}})^{\text{b}}$	$Err(78.5)^{\text{b}}$	$\epsilon_{\text{out,eff}}^{\text{c}}$	$Err(\epsilon_{\text{out,eff}})^{\text{b}}$	$Err(4)^{\text{b}}$
1	1.14	388.53	1229.75	1.49	142.85	296.91
2	1.87	421.66	814.45	1.42	77.95	266.95
3	3.25	313.23	395.87	1.28	156.28	299.69
4	1.22	534.47	1113.85	1.35	154.43	290.20
5	1.17	354.94	1182.49	1.20	186.28	366.59
Mean	1.73	402.57	947.28	1.3480	143.56	304.07

^a $\epsilon_{\text{out,eff}}$ is calculated using eq. (13).

^bError is calculated using error function (11).

^c $\epsilon_{\text{out,eff}}$ is calculated using eq. (14).

the AMBER 4.1 van der Waals radii²⁹ with two modifications: the radius of polar hydrogens was set to 1.0 Å, and the radius of the carboxylate oxygen was reduced to 1.5 Å. A dielectric constant of 78.5 was applied for the exterior medium (solvent water). We used either 1 or 2 for a dielectric constant for the solute interior.

Test Sets

Our test set for calculation of $\epsilon_{\text{out,eff}}$ included one peptide, Ace-Ala₁₀-Nme, and two globular proteins, crambin (1crn, 632 atoms) and insulin (2ins, 1480 atoms). Coordinates of 1crn and 2ins were taken from the Brookhaven protein data bank.³³

At a given combination of $\epsilon_{\text{in}}/\epsilon_{\text{out}}$ (1/78.5 or 2/78.5) for the reference solvation energy calculation we performed five 10 ps Molecular Dynamics simulations at 300 K with 1-fs time steps for each molecule using either the CD or DDD function with ϵ_{out} equal to 78.5 or $4r_{ij}$, respectively. BEM solvation energy was calculated at 100-fs intervals and after accumulation of 100 successive snapshots and corresponding BEM energies $\epsilon_{\text{out,eff}}$ was calculated for the CD or DDD function using either eq. (13) or (14).

In the case of Ace-Ala₁₀-Nme, a fully extended peptide conformation was used as an initial structure for the MD simulation, so $\epsilon_{\text{out,eff}}$ was calculated each time over a range of very different conformations, with the root-mean-square deviations (RMSD) from the time-averaged structures fluctuating between 3.0 and 4.0 Å, depending on the trajectory. In the case of globular proteins RMSDs were smaller, 1.5–1.9 and 1.7–2.3 Å for crambin and insulin, respectively.

Results and Discussion

Tables 1–6 show effective dielectric constants for the CD and DDD models obtained from the MD trajectories of three molecules using BEM solvation energies as reference ones at two different combinations of $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 1/78.5$ and $2/78.5$. No cutoff was used for calculation of the electrostatic energies. $\epsilon_{\text{out,eff}}$ s in all cases were calculated using eqs. (13) and (14), and corresponding errors in predicting the relative solvation energies were estimated according to the error function (11).

Table 4. $\epsilon_{\text{out,eff}}$ for *Crambin* Calculated from 100 Snapshots Taken Along the MD Trajectories at 100-fs Intervals. Reference BEM solvation energies were calculated at $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 2/78.5$.

Trajectory <i>N</i>	CD function			DDD function		
	$\epsilon_{\text{out,eff}}^{\text{a}}$	$Err(\epsilon_{\text{out,eff}})^{\text{b}}$	$Err(78.5)^{\text{b}}$	$\epsilon_{\text{out,eff}}^{\text{c}}$	$Err(\epsilon_{\text{out,eff}})^{\text{b}}$	$Err(4)^{\text{b}}$
1	2.26	93.49	299.83	2.92	34.13	41.57
2	3.65	99.31	197.64	2.79	18.67	29.43
3	6.14	74.91	96.42	2.53	37.21	48.12
4	2.42	127.22	269.27	2.66	37.34	46.32
5	2.33	84.73	288.66	2.37	44.58	60.30
Mean	3.3600	95.9320	230.3640	2.6540	34.3860	45.1480

^a $\epsilon_{\text{out,eff}}$ is calculated using eq. (13).

^bError is calculated using error function (11).

^c $\epsilon_{\text{out,eff}}$ is calculated using eq. (14).

Table 5. $\epsilon_{\text{out,eff}}$ for Insuline Calculated from 100 Snapshots Taken Along the MD Trajectories at 100-fs Intervals. Reference BEM solvation energies were calculated at $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 1/78.5$.

Trajectory N	CD function			DDD function		
	$\epsilon_{\text{out,eff}}^a$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(78.5)^b$	$\epsilon_{\text{out,eff}}^c$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(4)^b$
1	s	3338.39	3730.84	10.18	1186.14	1236.44
2	1.81	1304.15	3343.87	2.58	1120.68	1151.46
3	1.92	5143.64	6230.22	1.54	1612.12	1860.39
4	1.79	1403.76	3150.79	2.88	564.82	575.12
5	1.62	1322.09	3728.54	5.00	1134.29	1146.32
Mean	2.06	2502.41	4036.85	4.4360	1123.6100	1193.9460

^a $\epsilon_{\text{out,eff}}$ is calculated using eq. (13).^bError is calculated using error function (11).^c $\epsilon_{\text{out,eff}}$ is calculated using eq. (14).

It is very interesting to note that the CD constant never reaches the value of bulk water and takes values typical for organic solvents, with the $\epsilon_{\text{out,eff}}$ being lower for the combination of $\epsilon_{\text{in}}/\epsilon_{\text{out}}$ equal to $1/78.5$. In the case of the DDD model we do not see any unexpected results and with the exception of insuline at $2/78.5$ $\epsilon_{\text{out,eff}}$ is in the range $1.3*r_{ij} - 4.4*r_{ij}$. Finally, it should be emphasized that in general the DDD model shows a better agreement with the reference BEM values than the CD one. Scatter plots of the relative approximate “solvation” energies vs. the relative reference ones at a combination of $\epsilon_{\text{in}}/\epsilon_{\text{out}}$ equal to $1/78.5$ are shown in Figures 1–6. Only the trajectory giving the smallest error was chosen for each molecule. One can see again that the DDD model gives better agreement in predicting the relative BEM solvation energies than the CD one. For comparison Figures 7 and 8 show scatter plots for the two trajectories of insuline that give the largest errors in reproducing the relative BEM solvation energies using CD and DDD models, respectively. Again, we see better behavior in the case of the DDD function.

Dependence of the CD and DDD models on the cutoff is shown in Tables 7 and 8 for crambin and insuline. As might be expected, the errors in predicting the relative BEM solvation energies are

larger in the case of calculations with a cutoff. The only exception is for crambin using the CD function. Most likely it is caused by a fortunate cancellation of the errors.

Thus, the results indicate that the DDD function gives the best agreement with the BEM method in reproducing the relative electrostatic solvation energies. The findings of other studies^{25,26,34} vary from one investigator to another. Edinger et al.²⁵ compared solvation free energies obtained from the DDD function ($\epsilon = r_{ij}$), Generalized Born approximation, and accurate solutions of the Poisson equation for a large data set of peptide structures, ranging from a single amino acid to a peptide sequence of length nine. They concluded that the DDD model yields qualitatively erroneous results.²⁵ Given and Gilson³⁴ examined five different dielectric models, CD ($\epsilon = 1, 4, \text{ or } 80$), DDD ($\epsilon = 4r_{ij}$), and GB for selecting the correct ligand conformations for three complexes with known crystal structure. The DDD model was found to be the best electrostatic screening function of those tested.³⁴ David et al.²⁶ concluded that although using the DDD model is probably preferable to complete neglect of the electrostatic influence of water, it is a very crude approximation whose chief virtue is computational speed.

Table 6. $\epsilon_{\text{out,eff}}$ for Insuline Calculated from 100 Snapshots Taken Along the MD Trajectories at 100-fs Intervals. Reference BEM solvation energies were calculated at $\epsilon_{\text{in}}/\epsilon_{\text{out}}, 2/78.5$.

Trajectory N	CD function			DDD function		
	$\epsilon_{\text{out,eff}}^a$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(78.5)^b$	$\epsilon_{\text{out,eff}}^c$	$Err(\epsilon_{\text{out,eff}})^b$	$Err(4)^b$
1	6.03	807.53	907.16	17.23	288.04	368.53
2	3.56	314.95	820.23	4.94	270.14	273.82
3	3.82	1231.28	1494.01	3.02	385.77	395.96
4	3.51	335.17	769.16	5.40	134.61	139.18
5	3.17	319.15	922.44	8.84	269.65	360.05
Mean	4.0180	601.6160	982.6000	7.8860	269.6420	307.5080

^a $\epsilon_{\text{out,eff}}$ is calculated using eq. (13).^bError is calculated using error function (11).^c $\epsilon_{\text{out,eff}}$ is calculated using eq. (14).

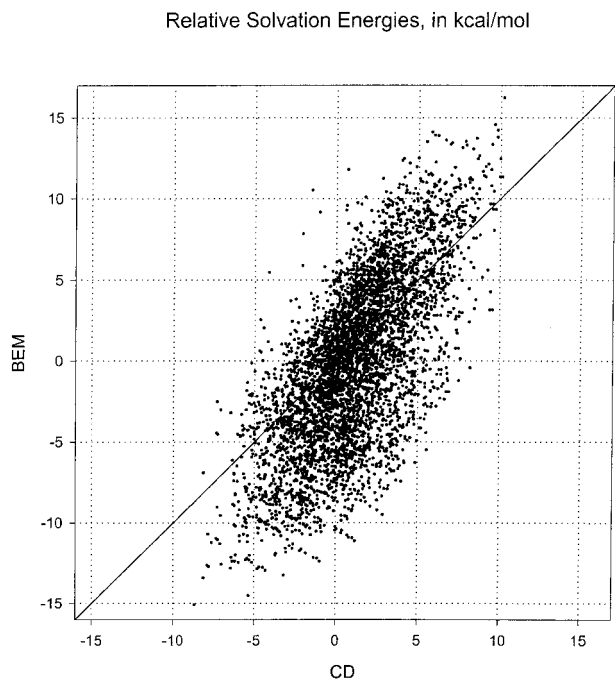


Figure 1. Scatter plot of the relative CD “solvation” energies (in kcal/mol) vs. BEM ones for Ace-(Ala)₁₀-NMe. The trajectory giving the smallest error in the relative solvation energies was chosen, $r^2 = 0.0460$.

Such controversy in the results is not surprising taking into account the very approximate nature of the CD and DDD models. For example, our data (Tables 1–6) show that even different MD trajec-

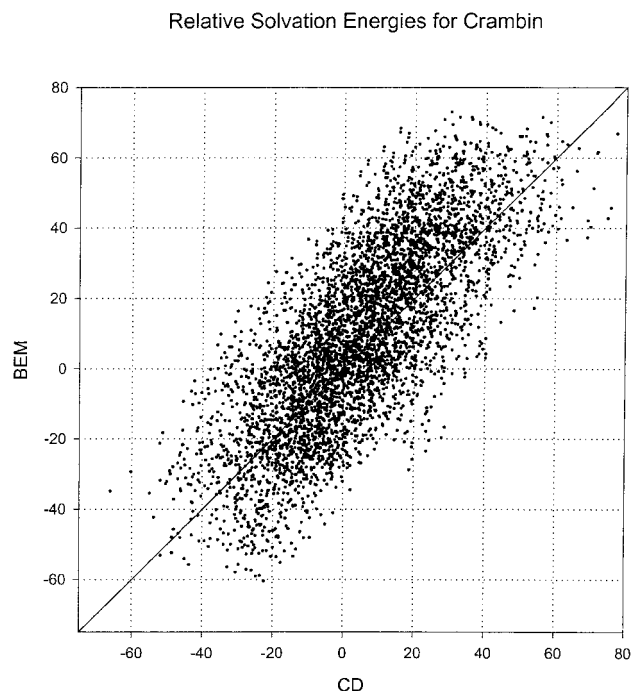


Figure 3. Scatter plot of the relative CD “solvation” energies (in kcal/mol) vs. BEM ones for crambin. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.576$.

tories of the same molecule can produce different errors in the relative solvation energies and different optimal $\epsilon_{\text{out,eff}}$ s. The data also indicate that there is no universal $\epsilon_{\text{out,eff}}$ which will work equally well for

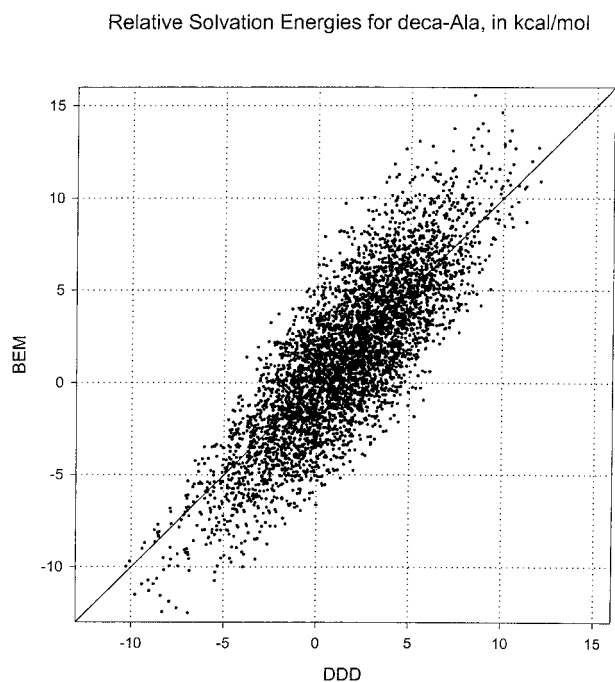


Figure 2. Scatter plot of the relative DDD “solvation” energies (in kcal/mol) vs. BEM ones for Ace-(Ala)₁₀-NMe. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.672$.

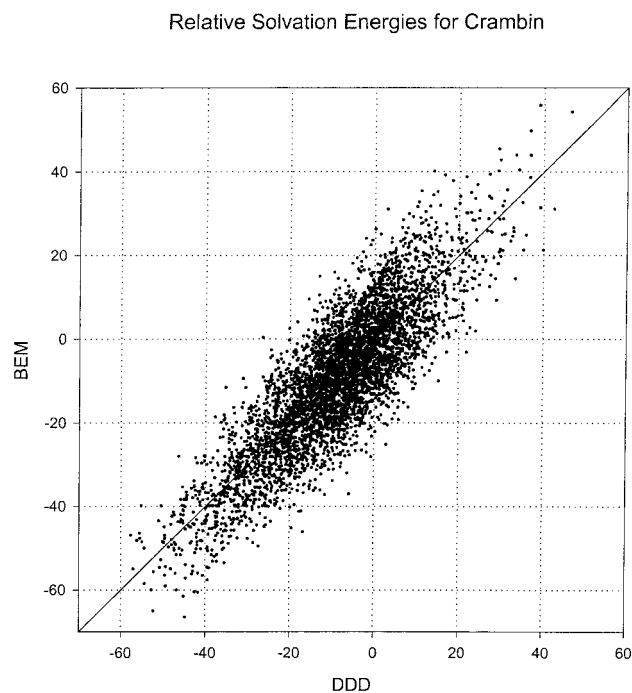


Figure 4. Scatter plot of the relative DDD “solvation” energies (in kcal/mol) vs. BEM ones for crambin. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.716$.

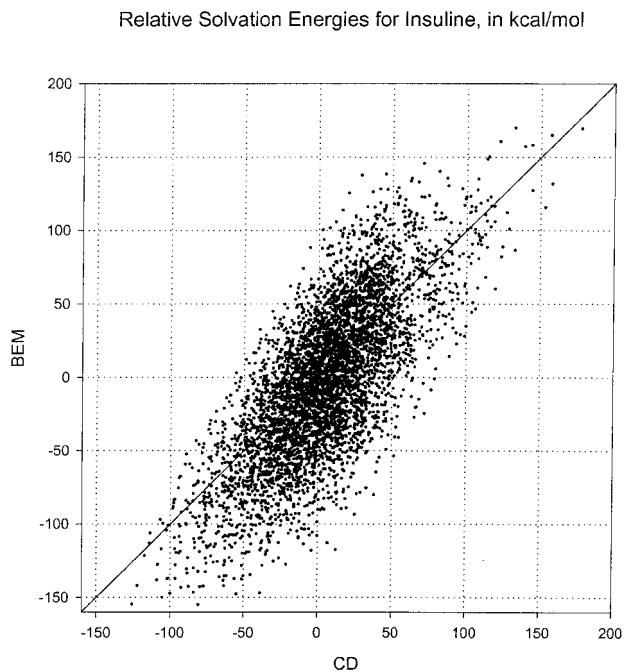


Figure 5. Scatter plot of the relative CD “solvation” energies (in kcal/mol) vs. BEM ones for insulin. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.536$.

all molecular systems. It is worth noting in this respect that none of the mentioned studies^{25,26,34} pursued an improvement in the accuracy of the CD or DDD model by reparametrizing for the systems studied.

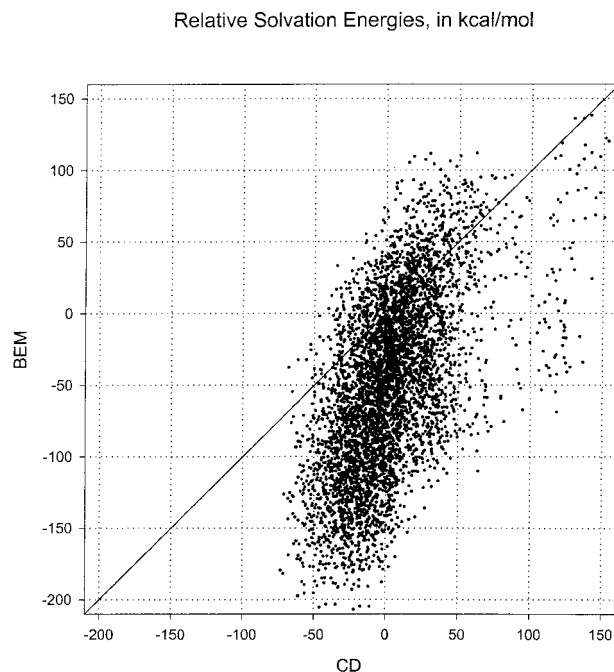


Figure 7. Scatter plot of the relative CD “solvation” energies (in kcal/mol) vs. BEM ones for crambin. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.333$.

Two Charges Model

To gain greater insight into behavior of approximate models we calculated the total electrostatic energies using different ap-

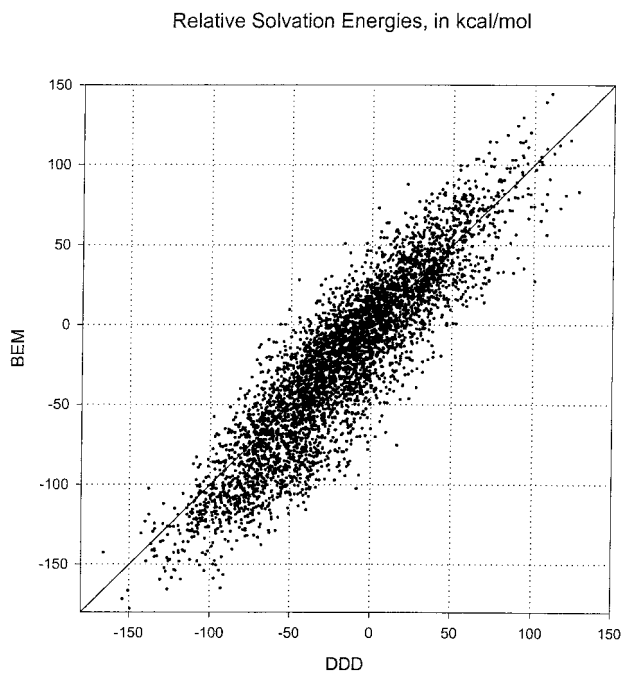


Figure 6. Scatter plot of the relative DDD “solvation” energies (in kcal/mol) vs. BEM ones for insulin. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.804$.

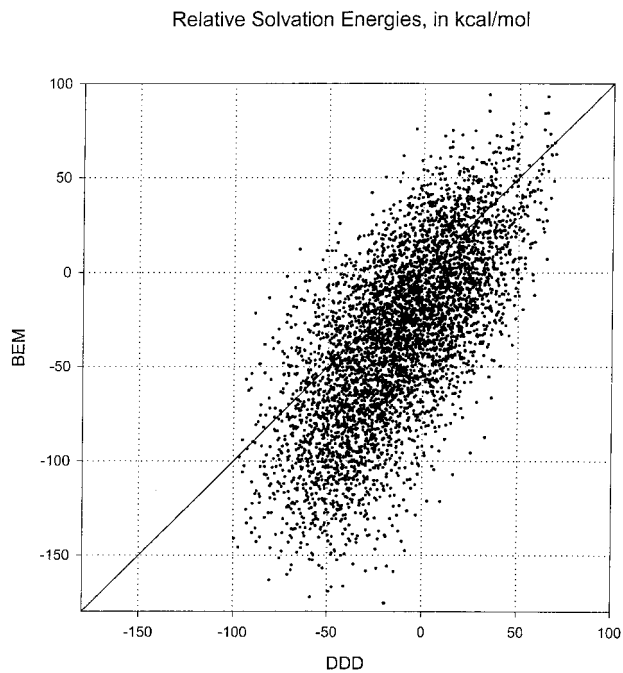


Figure 8. Scatter plot of the relative DDD “solvation” energies (in kcal/mol) vs. BEM ones for insulin. The trajectory giving the smallest error in the relative solvation energies was chosen. $r^2 = 0.461$.

Table 7. No Cutoff vs. 8-Å Cutoff for the CD Function.

Molecule	$\epsilon_{in}/\epsilon_{out}$	Cutoff	Mean $\epsilon_{out,eff}$ ^a	Mean $Err(\epsilon_{out,eff})$ ^b	Mean $Err(78.5)$ ^b
Crambin	1/78.5	No	1.73	402.57	947.28
		8 Å	1.44	371.02	1168.54
	2/78.5	No	3.36	95.93	230.36
		8 Å	3.14	79.03	236.36
Insuline	1/78.5	No	2.06	2502.41	4036.85
		8 Å	2.16	3484.45	6437.13
	2/78.5	No	4.02	601.62	982.60
		8 Å	3.21	718.83	1870.73

Mean $\epsilon_{out,eff}$ s and corresponding errors are from five MD trajectories at different $\epsilon_{in}/\epsilon_{out}$.

^a $\epsilon_{out,eff}$ is calculated using eq. (13).

^bError are calculated using error function (11).

proaches for the simplest model consisting of two atoms having the same radii and the opposite charges of 0.5 and -0.5 . All the calculations were done at the combination of $\epsilon_{in}/\epsilon_{out}$, 1/78.5. Two different radii sets were used for the calculation of BEM solvation energies, with the first one, 2 Å, falling within the radius of the carbon atom. This model corresponds to the case of small or unfolded molecules with atoms highly exposed to the solvent. In this instance, effective Born radii of the atoms are around the same as Born radii of the isolated atoms (in our case Born radii are equal to the van der Waals radii of the atoms). In another calculation we used radii equal to 3 Å for both atoms, which corresponds to a folded globular molecule where the effective Born radii of the atoms are larger than their van der Waals ones. Total electrostatic energies (Coulomb plus solvation) according to different models for a system of two charges vs. interatomic distance are shown in Figure 9. $\epsilon_{out,eff}$ s at each point were calculated using modified eqs. (6) and (7). To be more consistent with the CD and DDD models, we subtracted Born energies of isolated atoms from the BEM solvation energies, $E(\epsilon_{in}, \epsilon_{out})_{Solv}$, in modified equations.

What we can see from Figure 9 is that the CD model at $\epsilon_{out} = 78.5$ gives a wrong description of electrostatic interaction at close distances between the atoms. The discrepancy is more pronounced in the case of atomic radii of 3 Å, which corresponds to the situation of a folded molecule. Contrary to the CD function, the

DDD model yields a good agreement with BEM complemented electrostatic energies.

At a glance the CD model with $\epsilon_{out} = 2$ gives a very poor description of electrostatic interaction over a wide range of interatomic distances. However, contrary to the CD function with $\epsilon_{out} = 78.5$, which practically completely screens electrostatic interaction, at $\epsilon_{out} = 2$ the CD model better reproduces a change in total electrostatic energy that can even be seen visually. This is the reason why we got low values of $\epsilon_{out,eff}$ for our test systems after fitting them against the relative BEM solvation energies.

Can $\epsilon_{out,eff}$ Be Negative?

As we have seen above, the effective dielectric constant in CD and DDD models accounts for many effects of the total electrostatic energy of the solute in solvent, i.e., this sole parameter effectively includes nonadditivity of the electrostatic solvation free energy and different polarizabilities of solvent and solute among others. As a result, the effective dielectric constant of water can take values which correspond to those of nonpolar organic compounds, as happens with the CD function. Now we can further address the problem and ask whether $\epsilon_{out,eff}$ can take unphysical values. In the general case it will depend on at least two main factors, namely,

Table 8. No Cutoff vs. 8-Å Cutoff for the DDD Function.

Molecule	$\epsilon_{in}/\epsilon_{out}$	Cutoff	Mean $\epsilon_{out,eff}$ ^a	Mean $Err(\epsilon_{out,eff})$ ^b	Mean $Err(4)$ ^b
Crambin	1/78.5	No	1.35	143.56	304.07
		8 Å	1.22	420.19	651.69
	2/78.5	No	2.65	34.39	45.15
		8 Å	2.39	104.11	132.12
Insuline	1/78.5	No	4.44	1123.61	1193.95
		8 Å	9.91	3203.40	3386.34
	2/78.5	No	7.89	269.64	307.51
		8 Å	12.42	785.75	832.52

Mean $\epsilon_{out,eff}$ s and corresponding errors are from five MD trajectories at different $\epsilon_{in}/\epsilon_{out}$.

^a $\epsilon_{out,eff}$ is calculated using eq. (14).

^bError are calculated using error function (11).

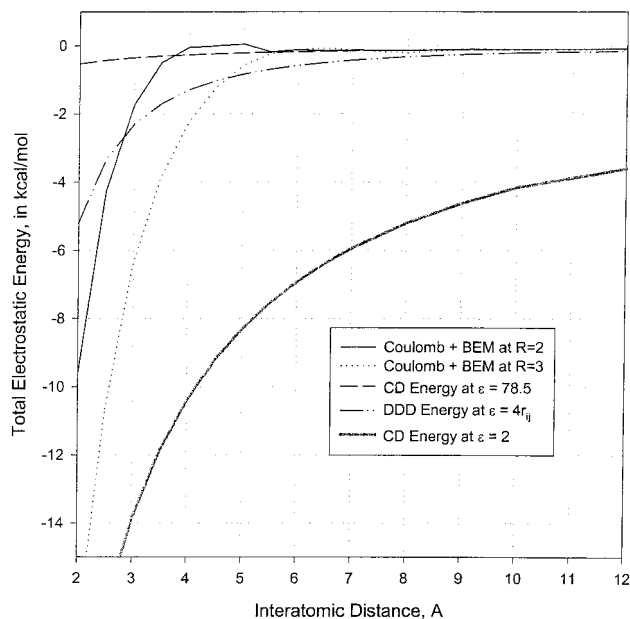


Figure 9. Total electrostatic energies (in kcal/mol) according to different models for a system of two charges vs interatomic distance. To be more consistent with the CD and DDD total electrostatic energies the Born energies of isolated atoms are subtracted from BEM solvation energies.

how shall we calculate $\epsilon_{\text{out,eff}}$ and what shall we take as reference values.

If we look again at Figure 9 we notice that the curve of the total electrostatic energy for one of our reference systems with atomic radii of 2 Å slightly crosses the zero energy level and takes small positive values. The “local” values of $\epsilon_{\text{out,eff}}$ at each interatomic distance calculated using eqs. (6) and (7) after subtracting Born energies of isolated atoms from the total BEM energies are shown in Table 9. We see very interesting behavior of $\epsilon_{\text{out,eff}}$ in the interatomic range from 4 to 5 Å for the atoms with a radius of 2 Å. First, $\epsilon_{\text{out,eff}}$ takes very high values in both the CD and DDD models and then becomes negative. To gain greater insight into such “unphysical” behavior of $\epsilon_{\text{out,eff}}$ one needs to take a look at how the surface is calculated in our BEM method. It makes use of molecular surface (see Fig. 10) because we feel that it is superior to both van der Waals and solvent-accessible surfaces in capturing the desolvation processes. As the distance between the two sphere surfaces goes below one solvent diameter, the molecular surface begins to balloon out along the approach axis (see Figs. 10 and 11), which results in a larger total molecular surface area than that for the isolated atoms (Fig. 11). This leads to slightly unfavorable interaction energy according to BEM and, as a result, to a negative $\epsilon_{\text{out,eff}}$. The last two columns of Table 9 show $\epsilon_{\text{out,eff}}$ s calculated using a van der Waals surface. In this case $\epsilon_{\text{out,eff}}$ s do not take negative values.

In our case a negative $\epsilon_{\text{out,eff}}$ might arise from using the molecular surface for the BEM solvation energy calculations.

Table 9. Effective Dielectric Constants for the CD and DDD Models for Two Charges within Spheres of Different Radii (R , in Å).

Interatomic distance	$\epsilon_{\text{out,eff}}$					
	Molecular surface ^a				VDW surface	
	CD model		DDD model		CD model	DDD model
	$R = 2 \text{ \AA}$	$R = 3 \text{ \AA}$	$R = 2 \text{ \AA}$	$R = 3 \text{ \AA}$	$R = 2 \text{ \AA}$	$R = 2 \text{ \AA}$
1.0	1.83	1.46	1.83	1.46	1.83	1.83
1.5	2.69	1.83	1.80	1.22	2.73	1.82
2.0	4.30	2.36	2.15	1.18	4.23	2.11
2.5	7.80	3.12	3.12	1.25	7.78	3.11
3.0	15.99	4.34	5.33	1.45	16.37	5.46
3.5	48.78	6.17	13.94	1.76	28.71	8.20
4.0	410.09	9.22	102.52	2.31	34.57	8.64
4.5	4745.29	15.08	1054.51	3.35	41.57	9.24
5.0	-322.50	29.21	-64.50	5.84	88.11	17.62
5.5	84.52	66.04	15.37	12.01	80.04	14.55
6.0	125.35	125.35	20.89	20.89	86.28	14.38
6.5	87.64	194.25	13.48	29.88	94.10	14.48
7.0	96.28	96.28	13.75	13.75	96.28	13.75
8.0	73.98	49.36	9.25	6.17	73.98	9.25
9.0	86.28	78.90	9.59	8.77	86.28	9.59
10.0	79.66	72.68	7.97	7.27	79.66	7.97
15.0	83.67	64.25	5.58	4.28	83.67	5.58
20.0	98.57	66.84	4.93	3.34	98.57	4.93

Modified expressions (6) and (7) were used for calculating $\epsilon_{\text{out,eff}}$ at each interatomic distance.

^aSee definition of molecular surface in Figure 10.

Another source of a negative $\epsilon_{\text{out,eff}}$ might come from numerical error in solving the Poisson equation. On the other hand, the negative $\epsilon_{\text{out,eff}}$ values can be understood from the physical point of view. At a certain critical distance between two charged atoms of a particular radius the structure of the hydration shells is disrupted as a result of a solvation-desolvation process that leads to unfavorable solvation energy. In any case, even if the phenomenon of a negative $\epsilon_{\text{out,eff}}$ for two ions in solution takes place in real systems it can only occur in a small distance range and at certain radii of solute and solvent atoms.

Table 9 also shows that an ideal dielectric function must account for both short- and long-range interactions, giving low dielectric constants at short distances and screening enough above 5–7 Å. Another plausible candidate for this role (in addition to the CD and DDD functions) is a sigmoidal distant-dependent dielectric function.^{21,35–37} However, it is beyond the scope of the present study.

Adaptive Fitting

As we have seen, the parameter $\epsilon_{\text{out,eff}}$ in the CD and DDD models absorbs all the solvation effects and varies from system to system and even for the different MD trajectories of the same molecule. It would be unlikely to have a single value that works best for all molecules and for all packing configurations. Thus, $\epsilon_{\text{out,eff}}$ can be refitted periodically as the molecular conformation changes during a molecular dynamics or Monte Carlo simulation.

Taking into account that in the case of optimization, MD, or Monte Carlo simulations we are more interested in the relative change of the solvation energy rather than in its absolute value, the form of the error function (11) suggested here may be a promising candidate for refitting $\epsilon_{\text{out,eff}}$. Recalibration of the $\epsilon_{\text{out,eff}}$ can be done “on the fly” at the beginning or in the course of the simulation.

Conclusions

We have described an approach for evaluating an effective dielectric constant for both the CD and DDD models making use of BEM solvation free energies as reference ones, with the emphasis on reproducing the relative changes in the solvation energies during the simulation. Such an approach seems to be more appro-

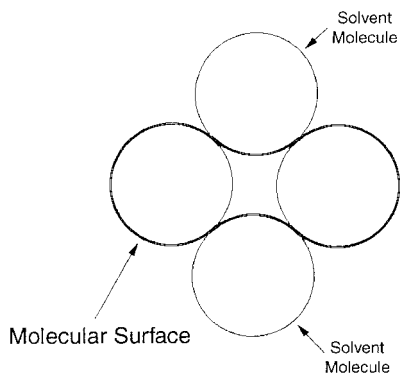


Figure 10. Definition of molecular surface.

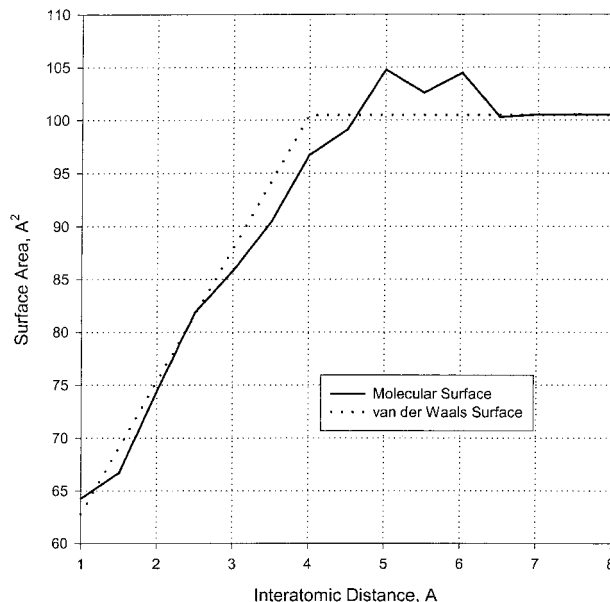


Figure 11. Molecular and van der Waals surfaces vs. interatomic distance for two atoms with radii of 2 Å

priate for CD and DDD approximate methods intended to be used in Monte Carlo, MD, or optimization techniques. An adaptive fitting procedure may be used for further improvement in the description of the electrostatic energy for a particular molecule during the simulation. Fitting against the relative electrostatic solvation energies is a quite general approach, and it can be applied for calibration of any other approximate method.

The DDD function was found to be in better agreement with BEM solvation energies than a CD one because of a better description of the electrostatic interactions at short distances between the charges.

One interesting finding of this study was that, under certain conditions, the effective dielectric constant in both the CD and DDD models can take negative values. This was demonstrated for a system of two charged atoms at a particular distance range. However, in principle, there is nothing to prevent the $\epsilon_{\text{out,eff}}$ from taking negative values for specific conformations of some molecules.

Acknowledgments

The author thanks Dr. Enrico Purisima for many useful discussions and for careful reading of the manuscript. The author also thanks Tara Sprules for help in preparation of the manuscript. This is National Research Council of Canada publication no. 44846.

References

1. Honig, B.; Nicholls, A. *Science* 1995, 268, 1144.
2. Honig, B.; Sharp, K.; Yang, A.-S. *J Phys Chem* 1993, 97, 1101.
3. Levy, R. M.; Gallicchio, E. *Annu Rev Chem* 1998, 49, 531.
4. Bashford, D.; Case, D. A. *Annu Rev Chem* 2000, 51, 129.

5. Orozco, M.; Luque, F. J. *Chem Rev* 2000, 100, 4187.
6. Karplus, M.; Petsko, G. A. *Nature* 1990, 347, 631.
7. Kollman, P. A. *Chem Rev* 1993, 93, 2395.
8. Van Gunsteren, W. F.; Berendsen, H. J. C. *Angew Chem Int Ed Engl* 1990, 29, 992.
9. Jorgensen, W. L. *Acc Chem Res* 1989, 22, 184.
10. Aqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng* 1994, 7, 385.
11. Nicholls, A.; Honig, B. *J Comput Chem* 1991, 12, 435.
12. Zauhar, R. J.; Morgan, R. S. *J Comput Chem* 1988, 9, 171.
13. Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J Am Chem Soc* 1990, 112, 6127.
14. Schaefer, M.; Karplus, M. *J Phys Chem* 1996, 100, 1578.
15. Scarsi, M.; Apostolakis, J.; Caffisch, A. *J Phys Chem A* 1997, 101, 8098.
16. Onufriev, A.; Bashford, D.; Case, D. A. *J Phys Chem B* 2000, 104, 3712.
17. Srinivasan, J.; Trevathan, M. W.; Beroza, P.; Case, D. A. *Theor Chem Acc* 1999, 101, 426.
18. Harada, M.; Watanabe, I.; Watarai, H. *Chem Phys Lett* 1999, 301, 270.
19. Giesen, D. J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor Chem Acc* 1997, 98, 85.
20. Babu, C. S.; Lim, C. *J Chem Phys* 2001, 114, 889.
21. Guarnieri, F.; Schmidt, A. B.; Mehler, E. L. *Int J Quant Chem* 1998, 69, 57.
22. Jayaram, B.; Sprous, D.; Beveridge, D. L. *J Phys Chem B* 1998, 102, 9571.
23. Augspurger, J. D.; Scheraga, H. A. *J Comp Chem* 1996, 17, 1549.
24. Stouten, P. F. W.; Frommel, C.; Nakamura, H.; Sander, C. *Mol Simulat* 1993, 10, 97.
25. Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J Phys Chem B* 1997, 101, 1190.
26. Laurent, D.; Ray, L.; Gilson, M. K. *J Comp Chem* 2000, 21, 295.
27. Jackson, J. D. *Classical Electrodynamics*; John Wiley & Sons: New York, 1975.
28. McCammon, J. A.; Wolynes, P. G.; Karplus, M. *Biochemistry* 1979, 18, 927.
29. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J Am Chem Soc* 1995, 117, 5179.
30. Purisima, E. O. *J Comput Chem* 1998, 19, 1494.
31. Purisima, E. O.; Nilar, S. H. *J Comput Chem* 1995, 16, 681.
32. Vorobjev, Y. N.; Hermans, J. *Biophys J* 1997, 73, 722.
33. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res* 2000, 28, 235.
34. Given, J. A.; Gilson, M. K. *Proteins* 1998, 33, 475.
35. Ramstein, J.; Lavery, R. *Proc Natl Acad Sci USA* 1988, 85, 7231.
36. Mehler, E. L. *Protein Eng* 1990, 3, 415.
37. Mehler, E. L. *J Phys Chem* 1996, 100, 16006.