Comment on the Validation of Continuum Electrostatics Models

MARCO SCARSI, AMEDEO CAFLISCH

Department of Biochemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received 4 January 1999; accepted 1 June 1999

ABSTRACT: A validation based on solvation energies (vacuum to water transfer) is not sufficient to justify the use of approximated models of electrostatics to rank ligand/protein complexes. A full validation should be based on energies in solution, i.e., solvation plus vacuum Coulomb energies, because of the anticorrelation between solvation and vacuum energies. The energy in solution is the relevant quantity in simulations of biological macromolecules and complexes. © 1999 John Wiley & Sons, Inc. J Comput Chem 20: 1533–1536, 1999

Keywords: continuum electrostatics; solvation energy; Poisson equation; thrombin; docking

Fast methodologies are needed to estimate solvation effects for docking flexible ligands in receptor binding sites. Approximated continuum electrostatic approaches are often used, but sometimes a good agreement with the exact solution of the Poisson equation (finite-difference method[1]) can originate uniquely from an inappropriate validation.

A comprehensive continuum treatment of the electrostatic effects of a high-dielectric solvent in the calculation of ligand-receptor binding energies should include the evaluation of the screened intermolecular interaction, as well as the receptor and ligand desolvations. The former is the intermolecular energy of the complex in solution, whereas the desolvation of a solute molecule is defined as the solute–solvent energy that is lost when part of the surrounding solvent is replaced by molecules with lower polarizability. Desolvation is the only energy term that disfavors the association between two opposite charges in solution.

In this comment, we focus on the screened interaction between two molecules (e.g., a receptor and a ligand) in a solvent. It can be calculated from the potential solution of the Poisson equation, assuming that one of the two molecules (in the present case the ligand) is an uncharged low dielectric cavity:

$$E^{int}_{solution} = \sum_{i=1}^{N} q_i \phi(\vec{r}_i) \qquad (1)$$

where N is the number of ligand atoms, q_i is the partial charge of the ligand atom i, and $\phi(\vec{r}_i)$ is the electrostatic potential generated by the charges of the receptor at the position of atom i. This approach requires the solution of the Poisson equation

Correspondence to: A. Caflisch; e-mail: caflisch@bioc.unizh.ch
Contract/grant sponsor: Swiss National Science Foundation; contract/grant number: 31-53604.98

CCC 0192-8651 / 99 / 141533-04

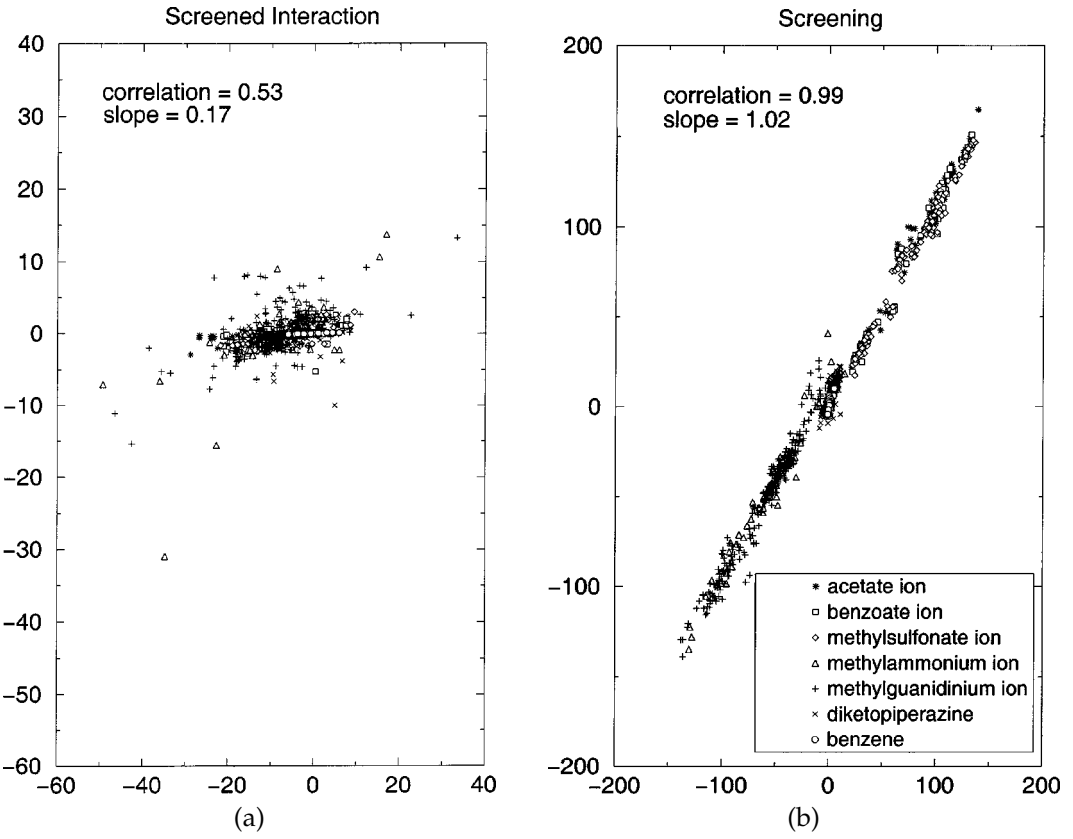

FIGURE 1. (a) Screened interaction calculated according to eq. (1) (*x*-axis) and to eq. (2) (*y*-axis) for 1025 ligand–thrombin complexes. (b) Screening energy calculated according to eq. (6) (*x*-axis) and to eq. (7) (*y*-axis) for 1025 ligand–thrombin complexes. All values are in kcal/mol.

for every ligand-receptor conformation and, consequently, is very time consuming. For efficiency reasons, it has been assumed that the receptor is the only low-dielectric medium that needs to be considered and the additional contribution of a ligand to the polarization and the potential is negligible.[2, 3] According to this assumption, the electrostatic potential of the isolated receptor ($\tilde{\phi}$) is calculated only once, and the screened interaction energy between the receptor and a bound ligand is obtained as:

$$\tilde{E}^{int}_{solution} = \sum_{i=1}^{N} q_i \tilde{\phi}(\vec{r}_i) \qquad (2)$$

This assumes that different positions of the ligands will not significantly perturb the potential.[2, 3] Because the solvent effects are used as a correction to a vacuum energy function, one can rewrite eqs. (1) and (2) as:

$$E^{int}_{solution} = E^{int}_{vacuum} + E^{int}_{solvation} \qquad (3)$$

$$\tilde{E}^{int}_{solution} = E^{int}_{vacuum} + \tilde{E}^{int}_{solvation} \qquad (4)$$

where $E^{int}_{solution}$ and $\tilde{E}^{int}_{solution}$ are the exact and the approximate screened interaction, and $E^{int}_{solvation}$ and $\tilde{E}^{int}_{solvation}$ are the exact and the approximate screening. The screening is the change in intermolecular electrostatic energy upon solvation of the complex. It is a solvation energy lacking the intramolecular contributions. With a solute dielectric constant of 1, one has:

$$E^{int}_{vacuum} = \sum_{i \in ligand} \sum_{j \in receptor} \frac{q_i q_j}{r_{ij}} = \sum_{i=1}^{N} q_i \phi_{vacuum}(\vec{r}_i) \qquad (5)$$

$$E^{int}_{solvation} = \sum_{i=1}^{N} q_i \left(\phi(\vec{r}_i) - \phi_{vacuum}(\vec{r}_i)\right) \qquad (6)$$

$$\tilde{E}^{int}_{solvation} = \sum_{i=1}^{N} q_i \left(\tilde{\phi}(\vec{r}_i) - \phi_{vacuum}(\vec{r}_i)\right) \qquad (7)$$

where $\phi_{vacuum}(\vec{r}_i)$ is the electrostatic potential of the isolated receptor *in vacuo*. It can be evaluated analytically exactly on every position, $\vec{r}_i$, or on the same grid used for $\phi(\vec{r}_i)$. Using a grid spacing of 0.5 Å, the two approaches practically coincide (correlation 0.9999, slope 1.0022) and we adopted the first one.

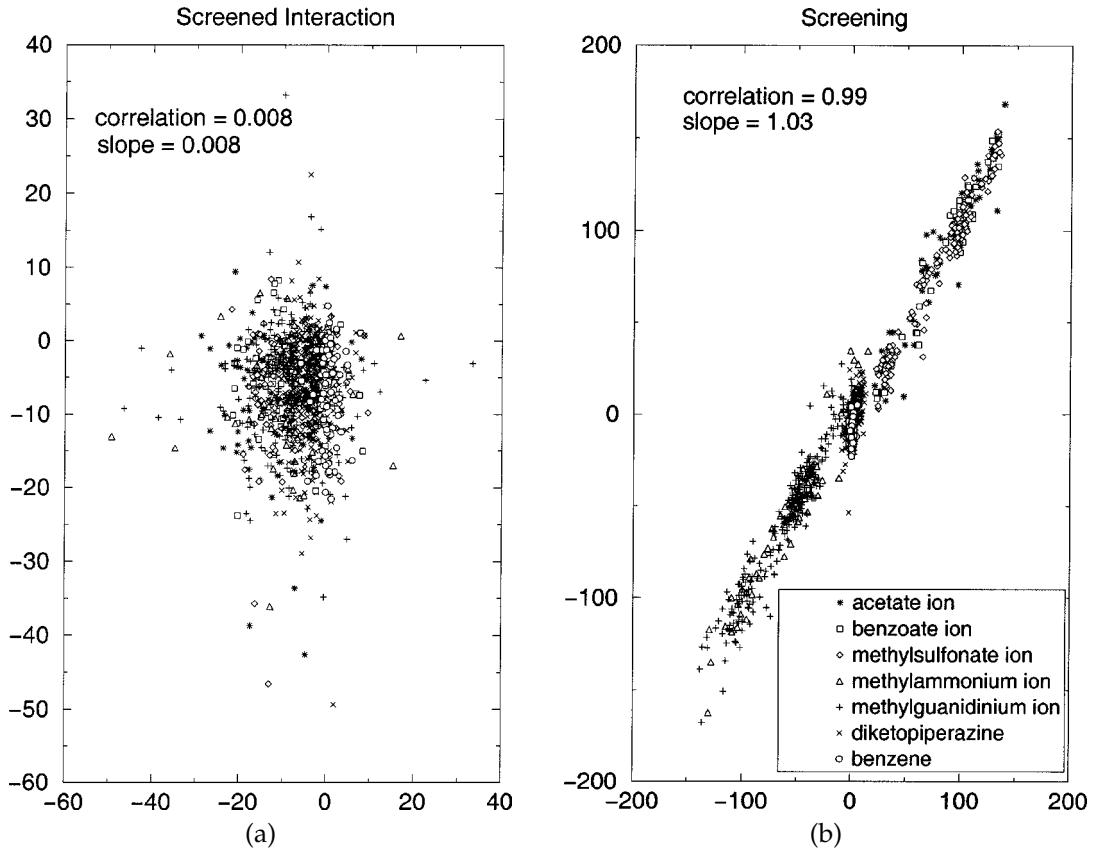

FIGURE 2. (a) Screened interaction calculated according to eq. (1) (*x*-axis) and to the "scrambled" model (*y*-axis) for 1025 ligand–thrombin complexes. (b) Screening energy calculated according to eq. (6) (*x*-axis) and to the "scrambled" model (*y*-axis) for 1025 ligand–thrombin complexes. All values are in kcal/mol.

One is tempted to validate the solvation model by comparing the approximate with the exact screening ($\widetilde{E}^{int}_{solvation}$ with $E^{int}_{solvation}$). Such validation is not correct because the energy function used for docking ultimately contains only $\widetilde{E}^{int}_{solution}$, the screened interaction. Solvation energies of molecular complexes often show a strong anticorrelation with the vacuum energies of the complexes themselves.[4–6] In eqs. (6) and (7), $\sum_i q_i \phi_{vacuum}(\vec{r}_i)$ can, in general, be much higher in module than $\sum_i q_i \phi(\vec{r}_i)$ or $\sum_i q_i \tilde{\phi}(\vec{r}_i)$. Consequently, the correlation between $\widetilde{E}^{int}_{solvation}$ and $E^{int}_{solvation}$ can result mainly from the correlation of $\sum_i q_i \phi_{vacuum}(\vec{r}_i)$ with itself.

The UHBD program[7] was used to calculate the approximate and the exact screened interaction (and, consequently, also the corresponding screening) for 1025 complexes of the rigid structure of thrombin with small organic compounds (acetate ion, benzoate ion, methylsulfonate ion, methylammonium ion, methylguanidinium ion, 2,5-diketopiperazine, and benzene). These compounds were distributed on the surface of the thrombin active site from P3 to P2′. A solute dielectric constant of 1 and a solvent dielectric constant of 78.5 were employed in all calculations together with a grid spacing of 0.5 Å. The comparisons between the screened interactions and between the screenings are shown in Figure 1a and b, respectively. As is evident from Figure 1a, the approximate method does not show good agreement with the exact solution of the Poisson equation. Nevertheless, a validation based on comparison between screenings (Fig. 1b) gives the opposite impression (see also Fig. 2 of ref. 3).

As a paradox, we generated another set of screened interactions for the same complexes by randomly mixing the list of 1025 exact screened interactions. As expected, the "scrambled" model showed no agreement with the exact solution of the Poisson equation (Fig. 2a). Yet, the screening resulting from the "scrambled" model still showed a very good correlation with the exact screening (Fig. 2b) because of the aforementioned correlation of E^{int}_{vacuum} with itself.

This indicates that approximate continuum electrostatic models should be validated on the energies in solution rather than on the solvation energies.[4] The latter quantities often tend to show a much better agreement with the exact solution of the Poisson equation than the former ones, but only energies in solution are useful for discriminating between favorable and unfavorable binding modes.

Acknowledgment

We thank Prof. J. A. McCammon for providing the UHBD program, which was used for all finite-difference calculations.

References

1. Warwicker, J.; Watson, H. C. J Mol Biol 1982, 157, 671.
2. Shoichet, B. K.; Stroud, R. M.; Santi, D. V.; Kuntz, I. D.; Perry, K. M. Science 1993, 259, 1445.
3. Vieth, M; Hirst, J. D.; Kolinski, A.; Brooks, C. L. III J Comput Chem 1998, 19, 1612.
4. Scarsi, M.; Apostolakis, J.; Caflisch, A. J Phys Chem A 1997, 101, 8098.
5. Fujinaga, M.; Huang, K.; Bateman, K. S.; James, M. N. G. J Mol Biol 1998, 284, 1683.
6. Vorobjev, Y. N.; Almagro, J. C.; Hermans, J. Prot Struct Function Genet 1998, 32, 399.
7. Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. Comp Phys Commun 1991, 62, 187.