3-2002

An Integrated Effective Fragment—Polarizable Continuum Approach to Solvation: Theory and Application to Glycine

Pradipta Bandyopadhyay
Iowa State University

Mark S. Gordon
Iowa State University, mgordon@iastate.edu

Benedetta Mennucci
Universita di Pisa

Jacopo Tomasi
Universita di Pisa

An Integrated Effective Fragment—Polarizable Continuum Approach to Solvation: Theory and Application to Glycine

Abstract
A new discrete/continuum solvation model has been developed by combining the effective fragment potential (EFP) for the discrete part and the polarizable continuum model (PCM) for the continuum part. The usefulness of this model is demonstrated by applying it to the calculation of the relative energies of the neutral and zwitterionic forms of glycine. These calculations were performed by treating glycine with *ab initio* wave functions. Water clusters were treated with both *ab initio* and EFP methods for comparison purposes, and the effect of the continuum was accounted for by the PCM model. The energy barrier connecting the zwitterionic and neutral three-water clusters was also examined. The computationally efficient EFP/PCM model gives results that are in close agreement with the much more expensive full *ab initio*/PCM calculation. The use of methods that account for electron correlation is necessary to obtain accurate relative energies for the isomers of glycine.

Keywords
Ab initio calculations, Polyelectrolytes, Electron correlation calculations, Wave functions

Disciplines
Chemistry

Comments
The following article appeared in *Journal of Chemical Physics* 116 (2002): 5023, and may be found at doi:10.1063/1.1433503.

An integrated effective fragment—polarizable continuum approach to solvation: Theory and application to glycine

Pradipta Bandyopadhyay, Mark S. Gordon, Benedetta Mennucci, and Jacopo Tomasi

Citation: The Journal of Chemical Physics **116**, 5023 (2002); doi: 10.1063/1.1433503
View online: http://dx.doi.org/10.1063/1.1433503
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/116/12?ver=pdfcov
Published by the AIP Publishing

An integrated effective fragment—polarizable continuum approach to solvation: Theory and application to glycine

Pradipta Bandyopadhyay[a] and Mark S. Gordon[b]
Department of Chemistry, Iowa State University, Ames, Iowa 50010

Benedetta Mennucci and Jacopo Tomasi
Dipartimento di Chimica e Chimica Industriale, Universita di Pisa, Via Risorgimento 35, 56126 Pisa, Italy

(Received 24 September 2001; accepted 14 November 2001)

A new discrete/continuum solvation model has been developed by combining the effective fragment potential (EFP) for the discrete part and the polarizable continuum model (PCM) for the continuum part. The usefulness of this model is demonstrated by applying it to the calculation of the relative energies of the neutral and zwitterionic forms of glycine. These calculations were performed by treating glycine with *ab initio* wave functions. Water clusters were treated with both *ab initio* and EFP methods for comparison purposes, and the effect of the continuum was accounted for by the PCM model. The energy barrier connecting the zwitterionic and neutral three-water clusters was also examined. The computationally efficient EFP/PCM model gives results that are in close agreement with the much more expensive full *ab initio*/PCM calculation. The use of methods that account for electron correlation is necessary to obtain accurate relative energies for the isomers of glycine. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1433503]

I. INTRODUCTION

The modeling of solvation in chemical and biological sciences is an active field of research, even though the history of the literature on solvation is quite old. This suggests the importance and difficulty in developing an accurate and general solvation model. Most solvation models used in quantum chemistry are based on a continuum approach, in which the solvent is treated as a structureless medium.[1] These models solve Poisson's equation numerically or analytically depending upon the complexity of the electrostatic problem.[2] Continuum models have been quite successful in modeling solvation, though they are very sensitive to the choice of the parameters, such as the shape of the cavity used to embed the solute. The major drawback of continuum models is that the explicit structure of the solvent is largely neglected. This may be an important consideration when the interaction between the solute and the solvent is strong, as is the case when solvent–solvent or solvent–solute pairs can form hydrogen bonds. An alternative solvation approach is to treat solvent molecules explicitly. In such a model, the intermolecular interactions are explicitly accounted for. However, because of the presence of numerous minima in the configuration space, it is critical to adequately sample this space using such methods as molecular dynamics[3] or Monte Carlo.[4] This can be a time-consuming procedure, depending on the complexity of the intermolecular potential. When the solvent molecules are calculated using quantum mechanics, this approach is generally limited to the study of small molecular clusters.

A more general approach to modeling solvation is to combine an explicit treatment of a modest number of solvent molecules with a continuum to model the bulk. Such a combined method might be successful, even in cases where the solvent and the solute are strongly bound. Moreover, the structure of the solvent around the solute can provide insight into the solvation process. For these reasons the development of discrete/continuum solvation models is gaining popularity.[5] The major problem associated with such models is the quality and cost of the discrete part of the system. Since it is generally not possible to use many expensive explicit quantum mechanical solvent molecules, accurate yet inexpensive potentials are necessary. The effective fragment potential (EFP) method developed over the last few years[6] has been shown to possess both of these desirable characteristics. Because of its accuracy and firm foundation in quantum mechanics, this model gives results almost as good as the double zeta plus polarization restricted Hartree–Fock (RHF) method with a fraction of the computational cost. In addition, the EFP model can be improved systematically very much like quantum chemical methods.

We have recently[7] developed and illustrated the EFP/continuum approach using a simple Onsager method for the continuum. In general, however, it is desirable to interface the EFP Hamiltonian with more sophisticated continuum methods, in order to consistently obtain more reliable results. For this reason, in the present work an interface between the EFP method and the polarizable continuum model (PCM) is presented. This method is referred to as EFP/PCM. The accurate and successful PCM approach was originally developed in 1981,[8] revised in 1995,[9] and largely reformulated in 1998.[10] The common aspect maintained in all these formulations is that the Poisson equation is solved using a boundary element method (BEM) with the apparent surface charge (ASC) approach. More details of PCM are given in the meth-

[a]Present address: Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143.
[b]Electronic mail: mark@si.fi.ameslab.gov

odology section, where the most recent version of this model, the integral equation formalism (IEF),[10] is exploited.

However, before using any discrete/continuum model, it is necessary to articulate several points. The first question is, how many discrete solvent molecules are needed in a calculation to adequately represent the intermolecular interactions? One possibility is to represent the first solvation shell with discrete solvent molecules. However, the first solvation shell is generally not well defined. Moreover, the presence of several low-energy structures, even with a small number of discrete solvent molecules, makes it necessary to use some global optimization technique to get at least a representative number of isomers. It is also necessary to consider the entropic contributions from the different isomers. Since the EFP part of the EFP/PCM model provides the internal energy, while PCM provides the free energy of solvation, a comparison between the different structures must be done very carefully.

To illustrate the points made in the previous paragraph, in this paper the EFP/PCM model is applied to the question of relative stabilities of the neutral and zwitterionic forms of glycine, the simplest amino acid, using two different cluster sizes.

This paper is organized as follows: In Sec. II we present the theoretical and computational methods. In Sec. III, details of the calculation for glycine are described. Section IV consists of results and discussions. Conclusions and summary are presented in Sec. V.

II. THEORY AND COMPUTATIONAL METHODS

Detailed descriptions of the EFP and PCM methods are given in several papers.[6,8–10] Hence, in the present paper we present only brief descriptions of the two methods in Sec. II A. The working equations of the EFP/PCM model are presented in Sec. II B.

A. The effective fragment potential (EFP) method

The EFP method was developed to represent weak intermolecular interactions between a quantum mechanical solute molecule and a solvent molecule represented by an effective fragment potential. Interactions among the different fragments are also included in the model. These interactions are evaluated by considering the fundamental intermolecular interactions. In the case of solute–fragment interactions, one-electron terms are added to the *ab initio* Hamiltonian of the solute, representing electrostatic (Coulombic), polarization, and exchange-repulsion plus charge-transfer interactions. For the μth fragment molecule, the effective fragment potential is given by

$$V_{\mathrm{el}}(\mu,s)=\sum_{k=1}^{K} V_k^{\mathrm{Elec}}(\mu,s)+\sum_{l=1}^{L} V_l^{\mathrm{pol}}(\mu,s)+\sum_{m=1}^{M} V_m^{\mathrm{rep}}(\mu,s), \tag{1}$$

where s represents the electronic coordinates. The three terms on the right-hand side of Eq. (1) represent the Coulomb, polarization, and exchange-repulsion plus charge transfer interactions, respectively. The electrostatic potential is expressed in terms of the distributed multipolar analysis (DMA)[11] of the solvent charge distributions. K in the first term of Eq. (1) is the number of expansion points (e.g., each atom center and bond midpoint in water). Each expansion is carried through octopole terms at each point. The entire expansion is multiplied by a distant-dependent cutoff term to account for overlapping charge densities. The exchange repulsion plus charge transfer interactions are modeled by simple Gaussian functions located at the fragment centers, denoted by m in the third term of Eq. (1), where M is the number of expansion centers (each atom center and the center of mass of water). The polarization of the solvent molecules by the electric field of the solute molecules [the second term in Eq. (1)] is treated self-consistently. In this process the dipole polarizability tensor of the solvent molecules is expanded into its component bond and lone pair localized orbital dipole polarizability tensors centered at the centroids of the L localized valence molecular orbitals.[12]

B. Polarizable continuum model (PCM)

The PCM describes the solvent effects on quantum mechanical solutes by treating the solvent as a continuum. This method is based on the fact that the reaction potential generated in the solvent by the presence of a solute charge density can be described, at any point in the space, in terms of an apparent surface charge (ASC) distribution spread over the solute cavity surface. These ASC distributions can be calculated from the total electrostatic potential Φ_{tot}, which has contributions Φ^{γ} from the solute charge distribution and Φ^{σ} from the ASC distribution. Φ^{σ} accounts for the self-polarization of the ASCs and necessitates an iterative calculation:

$$\Phi_{\mathrm{tot}}=\Phi^{\gamma}+\Phi^{\sigma}. \tag{2}$$

The solvent electrostatic potential can be calculated by integrating the ASCs over the solute cavity as shown in Eq. (3),

$$\Phi^{\sigma}(\mathbf{r})=\int \frac{\sigma(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|}d^2s, \tag{3}$$

where $\mathbf{s}$ and $\sigma(\mathbf{s})$ denote the cavity surface and the ASCs at the cavity surface, respectively. The notation d^2s in Eq. (3) indicates a two-dimensional surface integral.

In the practical implementation of the PCM method, the solute is embedded in a cavity formed by interlocking spheres centered on atoms and atomic groups. The cavity surface is partitioned into small areas, called tessera. The continuous ASC distribution is represented by discrete point charges placed in each tessera. The solvent potential is determined as a summation over the potentials coming from the discrete charges, as shown in Eq. (4),

$$\Phi^{\sigma}(\mathbf{r})=\sum_k \frac{q_k}{|\mathbf{r}-\mathbf{s}_k|},\qquad q_k=a_k\sigma(\mathbf{s}_k), \tag{4}$$

where $\sigma(\mathbf{s}_k)$ is the value of the ASC distribution at a representative point s_k within the tessera k and a_k is the area of the tessera k. Note that $\sigma(\mathbf{s})$ represents the continuous ASC,

while $\{q\}$ represents the finite set of point charges (each of these placed in each tessera), i.e., the discrete representation of the ASC used in the calculation. ASC charges q can be determined either iteratively or by a matrix inversion procedure,[9] which considers the iterations implicitly. The matrix equation is shown below:

$$\mathbf{Dq} = -\mathbf{b}, \tag{5}$$

where the vector **q** collects the ASC in each tessera and **b** contains the solute electrostatic potential. The **D** matrix depends on the solute cavity partition (position and area of the tesserae) and on the solvent dielectric constant; detailed expressions for the components of **D** can be found in Ref. 10.

Once the calculation of Φ^{σ} is completed, it is introduced into the Schrödinger equation of the solute system,

$$(H^0 + \Phi^{\sigma})|\Psi'\rangle = E'|\Psi'\rangle, \tag{6}$$

where H^0 is the gas phase Hamiltonian and Ψ' and E' are the final wave function and the energy, respectively, in the presence of the solvent. It can be shown[9] that the variational minimization of Eq. (6) leads to the following free energy quantity:

$$\begin{aligned} G &= \langle \Psi' | H^0 + \frac{1}{2}\Phi^{\sigma} | \Psi' \rangle \\ &= \langle \Psi' | H^0 | \Psi' \rangle + \frac{1}{2} \sum_i \left(q_i^e \langle \Psi' | \frac{1}{|\mathbf{r}-\mathbf{s}_i|} | \Psi' \rangle \right. \\ &\quad \left. + q_i^N \langle \Psi' | \frac{1}{|\mathbf{r}-\mathbf{s}_i|} | \Psi' \rangle \right), \end{aligned} \tag{7}$$

where the last two terms on the right-hand side represent the interactions between the electronic part of the solute charge density and the ASC due to the polarization of the solvent [generated by the electronic (q_i^e) and nuclear (q_i^N) parts of the solute]. Index i runs over all tesserae and $\mathbf{s}_i$ are the positions of the ASCs. The addition of the solute nuclear contribution to Eq. (7) gives the following equation:

$$\begin{aligned} G_{\mathrm{tot}} &= \langle \Psi' | H^0 | \Psi' \rangle + \frac{1}{2} \sum_i q_i \langle \Psi' | \frac{1}{|\mathbf{r}-\mathbf{s}_i|} | \Psi' \rangle \\ &\quad + \frac{1}{2} \sum_i (q_i^e + q_i^N) \sum_m \frac{Z_m}{|\mathbf{R}_m - \mathbf{s}_i|} + V_{NN}, \end{aligned} \tag{8}$$

where the second and third terms of Eq. (7) are collected into the second term of Eq. (8). The third term represents the interaction between the nuclear charges of the solute and the ASC due to the electronic and nuclear part of the solute. The last term denotes the nuclear repulsion energy.

The second and third terms of Eq. (8) can be written as

$$\frac{1}{2} U_{\gamma\sigma} = \frac{1}{2} \sum_i \int \frac{\gamma(\mathbf{r}) q_i(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} d\mathbf{r}, \tag{9}$$

where

$$\begin{aligned} \gamma(\mathbf{r}) &= \gamma^N(\mathbf{r}) + \rho(\mathbf{r}) \\ &= \sum_m^N Z_m \delta(\mathbf{R}_m - \mathbf{r}) + \sum_{\mu\nu} P_{\mu\nu} \chi_{\mu}(\mathbf{r}) \chi_{\nu}(\mathbf{r}) \end{aligned}$$

is the solute charge distribution. Z_m denote the nuclear charges, $\mathbf{R}_m$ denote the positions of the nuclei, $P_{\mu\nu}$ is an element of the density matrix, $\{\chi_{\mu}\}$ is the basis set used in the *ab initio* calculation. Equation (9) gives all solute–solvent interaction terms.

Because of the linear nature of the electrostatic equation, it is possible to write Eq. (9) as

$$\begin{aligned} \frac{1}{2} U_{\gamma\sigma} &= \frac{1}{2} \sum_i \int \frac{\gamma(\mathbf{r})[q_i^e(\mathbf{s}) + q_i^N(\mathbf{s})]}{|\mathbf{r}-\mathbf{s}|} d\mathbf{r} \\ &= \frac{1}{2}(U_{ee} + U_{eN} + U_{Ne} + U_{NN}), \end{aligned} \tag{10}$$

where we have divided both the solute charge distribution and the ASC into electronic and nuclear parts. The first index of U on the right-hand side of Eq. (10) refers to the component of γ, the second one to the component of the ASC. For instance, U_{Ne} implies the interaction between the solute nuclear charges and ASC due to the electronic part of the solute. Equation (10) is the starting point for the derivation of the equations of the EFP/PCM model. However, first it is useful to write each of the terms of Eq. (10) in a different manner for later use, in order to be consistent with the notation of Cammi and Tomasi.[9] First we express U_{ee} as

$$U_{ee} = \mathrm{tr}\,\mathbf{PX}(\mathbf{P}), \tag{11a}$$

where **P** is the density matrix and X has matrix elements:

$$X_{\mu\nu} = \sum_i V_{\mu\nu}^i q_i^e .$$

$V_{\mu\nu}^i$ is an electrostatic potential matrix element in the AO basis, q_i^e is the ASC due to the electronic part of the solute, and i runs over all tessera points. In a similar manner,

$$U_{eN} = \mathrm{tr}\,\mathbf{PJ}, \tag{11b}$$

with small $J_{\mu\nu} = \Sigma_i V_{\mu\nu}^i q_i^N$, where q_i^N is the ASC due to the nuclear part of the solute:

$$U_{Ne} = \mathrm{tr}\,\mathbf{PY}, \tag{11c}$$

with $Y_{\mu\nu} = \Sigma_i V_i^N q_i^e$, where V_i^N is the electrostatic potential of the nuclei charges, and finally,

$$U_{\mathrm{NN}} = \sum_i V_i^N q_i^N . \tag{11d}$$

Using the above notation, Eq. (8) can be rewritten as

$$\begin{aligned} G &= [\mathrm{tr}\,\mathbf{Ph} + \tfrac{1}{2}\mathbf{PG}(\mathbf{P})] \\ &\quad + \tfrac{1}{2}[\mathrm{tr}\,\mathbf{PX}(\mathbf{P}) + \mathrm{tr}\,\mathbf{PJ} + \mathrm{tr}\,\mathbf{PY} + U_{\mathrm{NN}}] + V_{\mathrm{NN}}. \end{aligned} \tag{12}$$

The terms in the first set of square brackets denote the usual matrix expression for the electronic energy *in vacuo*. **h** and **G**(**P**) are the matrix representations of the one-electron and two-electron operators used in Hartree–Fock theory. The next four terms inside the second set of square brackets correspond to the second and third terms of Eq. (8).

C. EFP/PCM model

Before deriving the EFP/PCM equations, it is necessary to consider the additional terms arising from the effective

fragments. First, there are two sources of charge distribution from the fragments, the Coulomb part, and the polarization part. The Coulomb part of the EFP is described by the distributed multipoles, which will contribute to the ASC. For the polarization part, the induced dipole moments of the fragments will contribute to the ASC. Since the Coulomb contribution does not depend on the *ab initio* density, the apparent surface charges due to the Coulomb contribution can be calculated before initiating the SCF process. On the contrary, the ASCs due to the polarization contribution to the EFP depend on the *ab initio* density. Hence, these have to be calculated during the SCF process. Last, the apparent surface charges, once produced, can polarize the fragments. So, the electric field that generates the induced dipole moment in the fragments has a contribution from the apparent surface charges in addition to the contributions from the other terms in the EFP. In the following, the working equations of the EFP/PCM model are derived. The summations in the following equations are over all tessera points on the cavity surface. Recall that in the EFP/PCM approach the molecular cavity is obtained in terms of a generalized solute represented by the solute itself plus the solvent molecules introduced by the EFP model; the technique exploited is completely equivalent to that formulated for the standard PCM/solute-only calculations.[2]

The solute (*ab initio*+EFP)–PCM solvent interaction energy can be written as an extension of Eq. (10),

$$\frac{1}{2}U_{\gamma\sigma}=\frac{1}{2}\sum\int\frac{(\gamma_e+\gamma_N+\gamma_{\mathrm{stat}}^{\mathrm{EFP}}+\gamma_{\mathrm{pol}}^{EFP})(q^e+q^N+q_{\mathrm{stat}}^{\mathrm{EFP}}+q_{\mathrm{pol}}^{\mathrm{EFP}})}{|\mathbf{r}-\mathbf{s}|}d\mathbf{r}, \tag{13}$$

where $\gamma_{\mathrm{stat}}^{\mathrm{EFP}}$ and $\gamma_{\mathrm{pol}}^{\mathrm{EFP}}$ are the charge distributions of the electrostatic part (static multipoles) and polarization part (induced dipoles) of the fragments, and $q_{\mathrm{stat}}^{\mathrm{EFP}}$ and $q_{\mathrm{pol}}^{\mathrm{EFP}}$ are the ASC due to the electrostatic and polarization part of the fragments, respectively. The electrostatic component of the free energy of the solute (*ab initio*+EFP)–solvent system can be written in a manner similar to Eq. (8),

$$\begin{aligned}G=&\left[\langle\Psi'|H^0+\sum_{k=1}^{K}V_k^{\mathrm{Elec}}+\sum_{l=1}^{L}V_l^{\mathrm{pol}}+\sum_{m=1}^{M}V_m^{\mathrm{rep}}|\Psi'\rangle\right.\\&\left.+\frac{1}{2}(U_{ee}+U_{eN}+U_{Ne}+U_{\mathrm{NN}})+V'_{\mathrm{NN}}\right]\\&+\sum q_{\mathrm{stat}}^{\mathrm{EFP}}\langle\Psi'|\frac{1}{|\mathbf{r}-\mathbf{r}_i|}|\Psi'\rangle\\&+\sum q_{\mathrm{pol}}^{\mathrm{EFP}}\langle\Psi'|\frac{1}{|\mathbf{r}-\mathbf{r}_i|}|\Psi'\rangle+\frac{1}{2}\sum q_{\mathrm{stat}}^{\mathrm{EFP}}V_{\mathrm{stat}}^{\mathrm{EFP}}\\&+\frac{1}{2}\sum q_{\mathrm{pol}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{EFP}}+\sum q_{\mathrm{stat}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{EFP}}+\sum q^NV_{\mathrm{stat}}^{\mathrm{EFP}}\\&+\sum q^NV_{\mathrm{pol}}^{\mathrm{EFP}},\end{aligned} \tag{14}$$

where all terms within the square brackets are due to the EFP and PCM methods separately along with the gas phase terms. H^0 is the gas phase Hamiltonian. V_k^{Elec}, V_l^{pol}, and V_m^{rep} are the operators for the electrostatic, polarization, and repulsion terms of the fragments. The next four terms are defined in Eq. (10). $V'_{\mathrm{NN}}=V_{\mathrm{NN}}+V_{\mathrm{NF}}+V_{\mathrm{FF}}$, where V_{NF} and V_{FF} are the *ab initio* nuclei–fragment and fragment–fragment interactions, respectively. The remaining terms are due to the coupling between the fragment and the continuum and are explained in the next paragraph.

The first two of these terms contain the interactions between the electronic part of the *ab initio* charge density and the ASC due to the static multipoles ($q_{\mathrm{stat}}^{\mathrm{EFP}}$) and the induced dipoles ($q_{\mathrm{pol}}^{\mathrm{EFP}}$) of the fragments. V in the remaining terms denotes the electrostatic potential; the subscript and superscript in each of these terms indicates the source of the potential. For instance, $V_{\mathrm{stat}}^{\mathrm{EFP}}$ is the electrostatic part of the EFP. The first three of these terms represent the (a) interactions between the static multipoles of the fragments and the ASC due to these static multipoles ($q_{\mathrm{stat}}^{\mathrm{EFP}}V_{\mathrm{stat}}^{\mathrm{EFP}}$), (b) induced dipole of the fragments and the ASC due to the induced dipole of the fragments ($q_{\mathrm{pol}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{EFP}}$), and (c) interaction between the induced dipoles of the fragments and ASC due to the static multipoles of the fragments ($q_{\mathrm{stat}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{EFP}}$). The last two terms are the interactions between the ASC due to the nuclear charges of the *ab initio* part with the static multipoles of the fragments and with the induced dipoles of the fragments, respectively. Note that the self-interaction terms (for which the source of electrostatic potential and ASCs are the same, e.g., $q_{\mathrm{pol}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{FP}}$) are preceded by a factor of $\frac{1}{2}$, while the other terms have a factor of 1. This is because both of the two equivalent terms are included in the latter. For example, for the term $\frac{1}{2}q_{\mathrm{stat}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{EFP}}$ there is an equivalent term $\frac{1}{2}q_{\mathrm{pol}}^{\mathrm{EFP}}V_{\mathrm{stat}}^{\mathrm{EFP}}$, and these two terms are combined to form $q_{\mathrm{stat}}^{\mathrm{EFP}}V_{\mathrm{pol}}^{\mathrm{EFP}}$.

Equation (14) can also be written in a manner similar to Eq. (12):

$$\begin{aligned}G=&\left[\mathrm{tr}\,\mathbf{P}\tilde{\mathbf{h}}+\tfrac{1}{2}\mathbf{P}\tilde{\mathbf{G}}(\mathbf{P})+V_{\mathrm{NN}}+\tfrac{1}{2}U_{\mathrm{NN}}\right]+\mathrm{tr}\,\mathbf{P}(\mathbf{J}_{\mathrm{stat}}^{\mathrm{EFP}}+\mathbf{J}_{\mathrm{pol}}^{\mathrm{EFP}})\\&+\tfrac{1}{2}(U_{\mathrm{stat,stat}}+U_{\mathrm{pol,pol}})+U_{\mathrm{pol,stat}}+U_{\mathrm{stat},N}+U_{\mathrm{pol},N},\end{aligned} \tag{15}$$

where $\tilde{\mathbf{h}}=\mathbf{h}+\mathbf{J}+\mathbf{Y}$ and $\tilde{\mathbf{G}}(\mathbf{P})=\mathbf{G}(\mathbf{P})+\mathbf{X}(\mathbf{P})$ [see Eq. (12)]. All terms inside the square brackets occur in Eq. (12) and are defined there. All other terms are due to the coupling between the EFP and PCM methods. The two contributions to J are $(J_{\mathrm{stat}}^{\mathrm{EFP}})_{\mu\nu}=\Sigma q_{\mathrm{stat}}V_{\mu\nu}^e$ and $(J_{\mathrm{pol}}^{\mathrm{EFP}})=\Sigma q_{\mathrm{pol}}^{\mathrm{EFP}}V_{\mu\nu}^e$. $V_{\mu\nu}^e=\langle\mu|1/|\mathbf{r}-\mathbf{s}_i||\nu\rangle$, the electrostatic potential integrals in the atomic orbital basis. Equation (15) can be written as

$$G=\left(\mathrm{tr}\,\mathbf{P}\tilde{\mathbf{h}}'+\tfrac{1}{2}\mathbf{P}\tilde{\mathbf{G}}(\mathbf{P})\right)+V_{\mathrm{NN}}+\tfrac{1}{2}U_{\mathrm{NN}}+\tfrac{1}{2}(U_{\mathrm{stat,stat}}+U_{\mathrm{pol,pol}})+U_{\mathrm{pol,stat}}+U_{\mathrm{stat,N}}+U_{\mathrm{pol,N}}, \tag{16}$$

where $\tilde{\mathbf{h}}'=\mathbf{h}+\mathbf{J}_{\mathrm{stat}}^{\mathrm{EFP}}+\mathbf{J}_{\mathrm{pol}}^{\mathrm{EFP}}$. All but the first two terms of Eq. (16) are additive, i.e., they can be evaluated after the determination of the wave function. Note that the induced dipole of the fragments depends on the wave function. However, since this is a second-order effect, the variation of the induced dipoles with respect to the wave function parameters is not considered.

From Eq. (16), by using the stationary condition $\delta G=0$ for an arbitrary variation of the molecular orbital (with the normalization of the wave function), one can obtain the matrix HF-like equation:

$$\mathbf{F}'\mathbf{C}=\mathbf{SCE}', \tag{17}$$

where $\mathbf{F}'=\tilde{\mathbf{h}}+\tilde{\mathbf{G}}$. Equation (17) is the final working equation of the EFP/PCM model. These equations have been coded and implemented in the electronic structure code GAMESS.[13]

III. APPLICATION TO GLYCINE

A. Preliminary discussion

Glycine, the simplest amino acid, has been studied widely in aqueous solution with both supermolecule-continuum and continuum only calculations. In the supermolecule-continuum calculations, previous works explored several different choices for the discrete part. Rzepa *et al.*[14] used 7 and 15 discrete water molecules and found that AM1 and PM3 semiempirical methods give a poor estimate of the neutral (*N*)-zwitterion (*Z*) energy difference. This was attributed to the poor parameterization of the nitrogen atom. Tortonda *et al.*[15] used one discrete water molecule in their HF and MP2 calculation and concluded that addition of one water does not favor the zwitterion formation in solution. Jensen and Gordon[16] concluded that at the MP2 level of theory, two water molecules are sufficient to stabilize a zwitterionic local minimum, with the neutral isomer still lower in energy. Bandyopadhyay and Gordon used eight water molecules to illustrate the EFP/Onsager model.[7] However, in most of the previous works, the configuration space of the discrete waters was not adequately sampled. This is a serious issue considering the existence of a number of structures that have similar energies.

There have also been several continuum only calculations at the Hartree–Fock (HF), density functional theory (DFT), and MP2 levels of theory.[17] Tomasi *et al.*,[17(a)] using HF with small basis sets, showed that electrostatic solute–solvent interactions can stabilize the zwitterionic form with respect to the neutral one. Many different continuum models have been applied to the same system with different results. The most recent ones, however, seem to confirm that the use of an accurate solvation model (including solvent effects on both the electronic and geometrical description of the solute) and of a correlated quantum mechanical method are necessary to obtain closer agreement with the experimental result that the free energy of the process $Z(\mathrm{aq})\rightarrow N(\mathrm{aq})$ is 7.3 kcal/mol.[18] On the other hand, in the supermolecule-continuum calculation, it is not clear what number of discrete water molecules is needed to converge to the experimental results.

It is therefore interesting that in a recent paper, Kassab *et al.* employed[19] the continuum model of Claverie and co-workers[20] using the electrostatic potential derived from density functional theory (B3LYP/6-31G*) to calculate the energy difference between *N* and *Z* using both unhydrated and trihydrated species. These authors found that for the unhydrated species, both the enthalpy and the free energy have the wrong signs, i.e., they favor *N*. On the other hand with the trihydrated structures they found a free energy difference of 5.4 kcal/mol, close to the experimental value of 7.3 kcal/mol favoring *Z*. Of course, this close agreement with experiment with just three water molecules raises two questions.

(1) Does the three water cluster system provide converged results? That is, will the results change if we use more waters and sample the structures carefully?

(2) Are the cluster structures used for a comparison between *N* and *Z*, the lowest-energy structures for both *N* and *Z* clusters? If not, the reported relative energy difference between these two clusters must be called into question.

Because of the importance of understanding the role of clusters of waters in stabilizing the zwitterion, we have investigated both the above-mentioned questions in this work. The first step in this study was to calculate the glycine$(H_2O)_3$ system, systematically sampling the configuration space, in order to determine if there are structures with lower energies than those found by Kassab *et al.* Then to understand the effect of cluster size on the relative energies, clusters with eight water molecules (in which the first solvation shell is essentially filled) were investigated. The configuration space of the eight water cluster was sampled using the Monte Carlo simulated annealing (MCSA) method.[21]

The three three-water structures considered by Kassab *et al.*, the *cis*-Z cluster (CZ), the transition state structure (TS), and the *cis*-N (CN) are shown in Figs. 1(a), 1(b), and 1(c), respectively. The *trans*-N cluster (TN) found in the present work is shown in Fig. 1(d).

CN is the neutral structure used by Kassab *et al.* to determine the *N*-*Z* energy difference. However, at the RHF/6-31++G(d,p) level of theory, TN is 6.5 kcal/mol lower in energy than the *cis* isomer. On the other hand, there does not appear to be a $Z(H_2O)_3$ isomer that is lower than the structure considered by Kassab *et al.* The reaction path connecting the two neutral clusters, CN and TN, is discussed below.

For the eight water clusters, the Monte Carlo simulated annealing method detects several structures close in energy. As discussed in detail below, once the configuration spaces are adequately sampled, it is found that the clusters containing three and eight waters, each surrounded by a continuum, give similar results for the relative energies.

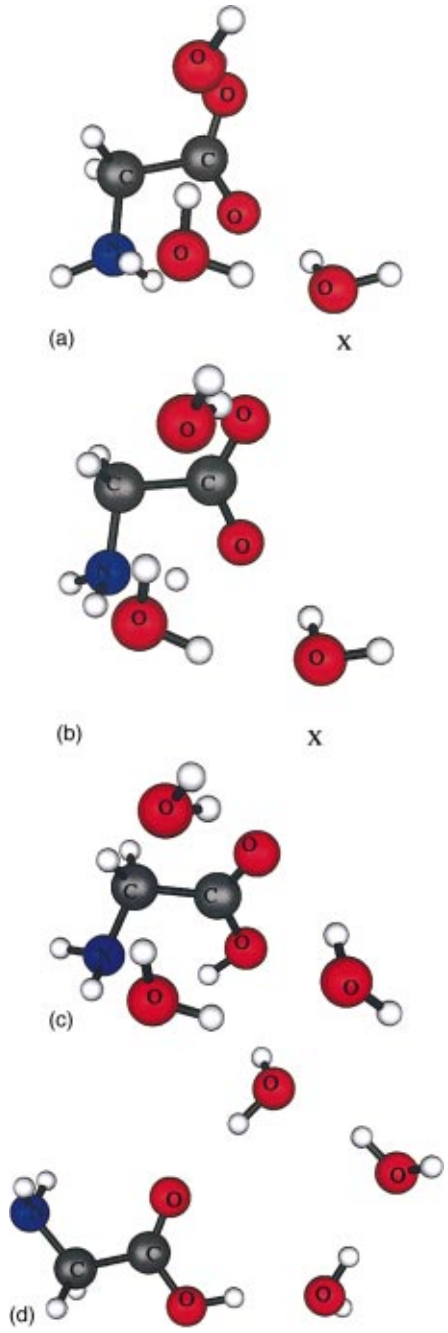

FIG. 1. (Color) Optimized structures of (a) *cis*-zwitterion cluster; (b) transition state; (c) *cis*-neutral cluster; (d) *trans*-neutral cluster; with three waters at the RHF/6-31++G(d,p) level with effective fragment waters. Symbol X in (a) and (b) represent the water molecule with a different orientation than that of Kassab *et al.* (see the text for more information).

B. Details of the calculations

Three-water cluster: First, full geometry optimizations were performed on the CN and CZ structures found by Kassab *et al.*, including the transition state connecting them. The next step was to determine the lowest-energy structure on the $N(H_2O)_3$ potential energy surface. In addition, the barrier connecting $Z(H_2O)_3$ with the $N(H_2O)_3$ lowest-energy structure was determined by locating the saddle points connecting these structures. All stationary points were characterized by calculating the Hessian matrix of energy second derivatives and determining the number of imaginary frequencies. Minimum energy path (MEP)[22] calculations were used to confirm the minima connected by each saddle point. All of these calculations were performed at the RHF/6-31++G(d,p) level of theory, with discrete water molecules treated both quantum mechanically and by the EFP method. To assess the effect of the continuum, each cluster was placed in a cavity using PCM to calculate energies at the cluster optimized geometries. The effect of correlation on the stationary points was taken into account by reoptimizing all the structures using second-order Møller–Plesset (MP2) perturbation theory.[23]

8-water cluster: The Monte Carlo simulated annealing method including the minimization approach of Li and Scheraga[24] was used to determine the low-energy structures of the $N(H_2O)_8$ and $Z(H_2O)_8$ clusters using EFP water. Both the *ab initio* glycine molecule and the EFP waters were moved during the MC process. For the water molecules, both single fragment movement and multifragment movements were tried during the MC process.

The energy difference between a bare cluster and the cluster embedded in the continuum is defined as the *solvation energy* of the cluster. Note that only the electrostatic part of the solvation energy was considered in the present work.

The zero-point energies (ZPE) for all the structures are calculated in this work using both EFP and *ab initio* waters. Note that the ZPE corrections are very similar for structures with EFP and with *ab initio* waters. The maximum difference in ΔZPE between the EFP and *ab initio* results is only 1.2 kcal/mol.

All the calculations were performed using the GAMESS program package.[13]

IV. RESULTS AND DISCUSSION

The optimized structures of CZ and CN and the transition state connecting them are shown in Fig. 1. Structures 1(a) and 1(b) are the same as those found by Kassab *et al.*, except for the orientation of the water labeled X in the figure. However, the different orientation of water X does not change the relative energies of these species. The relative energies presented in Table I, are exactly the same as those reported by Kassab *et al.* at the HF/6-31++G(d,p) level of theory. Table I also shows the relative energies of these three structures calculated with EFP waters and the effect of the continuum on these structures obtained using both *ab initio* +PCM and EFP/PCM methods.

It is seen in Table I that CN is more stable than CZ by 9.1 kcal/mol (*ab initio*) and 9.0 kcal/mol (EFP). Note that the zero-point vibrational energy (ZPE) favors CN by about 2 kcal/mol relative to CZ. The transition state connecting them is located at 15.6 (15.8) kcal/mol higher energy than CN at the *ab initio* (EFP) level of theory. So the EFP calculation is in excellent agreement with the more expensive *ab initio* calculation. The effect of the continuum was determined by single-point PCM calculations on these three cluster structures. The relative energies in bulk solution are indicated in parentheses in Table I, and the solvation energies are given in square brackets. The continuum has a profound effect on the *Z* cluster, as can be seen by the large solvation energy of 27.5

TABLE I. Relative RHF/6-31++G(d,p) energies (in kcal/mol) of *cis*-$N(H_2O)_3$, *cis*-$Z(H_2O)_3$, and the transition state connecting them. Values in parentheses refer to PCM calculations while those in square brackets give PCM solvation free energies of each molecular system. Δ(ZPE) refers to the difference in vibrational zero-point energies for cluster calculations.

Structure	Method of calculation			
	Ab initio (continuum) [solvation energy]	Δ(ZPE)	EFP (continuum) [solvation energy]	Δ(ZPE)
CN	0.0 (0.0) [13.8]	0.2	0.0 (0.0) [12.0]	0.0
TS	15.6 (9.5) [19.9]	−2.0	15.8 (6.8) [21.0]	−1.8
CZ	9.1 (−4.6) [27.5]	1.7	9.0 (−6.9) [28.0]	2.9

kcal/mol compared with 13.8 kcal/mol for the *N* cluster. The corresponding EFP values are 28.0 and 12.0 kcal/mol, in excellent agreement with the *ab initio* results. This large differential solvation energy changes the relative CN and CZ stabilities. In bulk solution, CZ is more stable than CN by 4.6 (6.9) kcal/mol at the *ab initio* (EFP) level. Kassab *et al.* found this value to be 5.4 kcal/mol (after adding thermal and entropic contributions) using the continuum model of Claverie. They compared this value favorably with the experimental free energy of 7.3 kcal/mol for the process $Z(\mathrm{aq}) \rightarrow N(\mathrm{aq})$. However, as discussed below, CN is not the lowest-energy structure of $N(H_2O)_3$, so the actual energy and free-energy differences for three-water clusters, with and without the continuum, must be reevaluated. Now, it is interesting that for bare glycine, *cis* is the lowest-energy neutral structure in the continuum according to Tortonda *et al.*, while in the gas phase *trans* is the lowest-energy neutral structure.[15] Apparently, the relative energies reverse when neutral glycine is placed in a three-water cluster.

The most stable structure of $N(H_2O)_3$, as determined by performing Monte Carlo simulations starting from several different geometries and glycine isomers, is shown in Fig. 1(d). In this structure, the O–H group is *trans* with respect to the NH_2 group. The three waters form a hydrogen bonded network involving the oxygen of the carbonyl group and the hydrogen of the hydroxyl group, and each water molecule acts simultaneously as proton acceptor and a proton donor. This structure, TN, is more stable than *cis*-$N(H_2O)_3$, CN, by 6.5 kcal/mol at the RHF/6-31++G(d,p) level of theory. So, the proper comparison between *Z* and *N* in aqueous solution, if one is using three water clusters, should include the structure shown in Fig. 1(d). Even with three waters, the existence of several isomers of $N(H_2O)_3$ makes the determination of the lowest-energy structure complicated. However, we are confident that TN is one of the lowest-energy structures.

Table II shows the energy difference between CN and TN at the EFP, RHF/6-31++G(d,p), and MP2/6-31++G(d,p) levels of theory. The effect of bulk solvent is also shown for the EFP and HF structures. The HF and EFP methods predict TN to be more stable than CN by 6.5 and 5.7 kcal/mol, respectively. More importantly, the neutral cluster (TN) is more stable than the zwitterion cluster (CZ) by 15.6 (14.7) kcal/mol at the HF (EFP) level of theory. As noted above, the ZPE correlation favors TN relative to CZ by ~2 kcal/mol. The PCM continuum stabilizes TN more than CZ, but even so, TN is more stable than CZ in solution by 2.2 kcal/mol (2.6 kcal/mol using EFP waters). So, the continuum does not stabilize the zwitterion relative to the neutral three-water cluster sufficiently to make the zwitterion the global minimum at the Hartree–Fock or EFP levels of theory.

Next, consider how the relative energies change when the three structures are optimized in solution using the PCM model. The optimizations were performed at the RHF/6-31G(d) level of theory, because it was found that with diffuse functions on all atoms and polarization functions on hydrogens, the optimizations do not converge. Once the optimization was completed, single-point RHF/6-31++G(d,p) energies were calculated, and the results are given in Table II. At this level of theory, CZ and TN are essentially isoenergetic, and both are more than 6 kcal/mol more stable than the *cis-N* cluster. The results are consistent with those quoted above and shown in Table II: The continuum preferentially stabilizes the zwitterion, and this effect is amplified when the geometry is optimized in the presence of the continuum.

To determine the effect of correlation on these three three-water clusters, they were reoptimized as the MP2/6-31

TABLE II. Relative RHF/6-31++G(d,p) and MP2/6-31++G(d,p) energies among *cis*-$N(H_2O)_3$, *trans*-$N(H_2O)_3$ and $Z(H_2O)_3$ (in kcal/mol). The numbers in square brackets are obtained by optimizing these three clusters in the PCM continuum with the 6-31G(d) basis set and then calculating the energies with the 6-31++G(d,p) basis set.

Method	*E*(CN)-*E*(TN) (continuum)	*E*(*Z*)-*E*(TN) (continuum)	Solvation energy [$E_{\mathrm{trans}}(N)$]
RHF/6-31++G(d,p) (EFP waters)	5.7/5.6[a](9.5)	14.7/17.4[a](2.6)	15.9
RHF/6-31++G(d,p) (Full quantum)	6.5/6.9[a](6.8) [6.5]	15.6/17.7[a](2.2) [0.2]	14.1
MP2/6-31++G(d,p) (Full quantum)	3.4/3.8[a](4.4)	4.3/5.7[a](−5.4)	18.8

[a]Including Δ(ZPE).

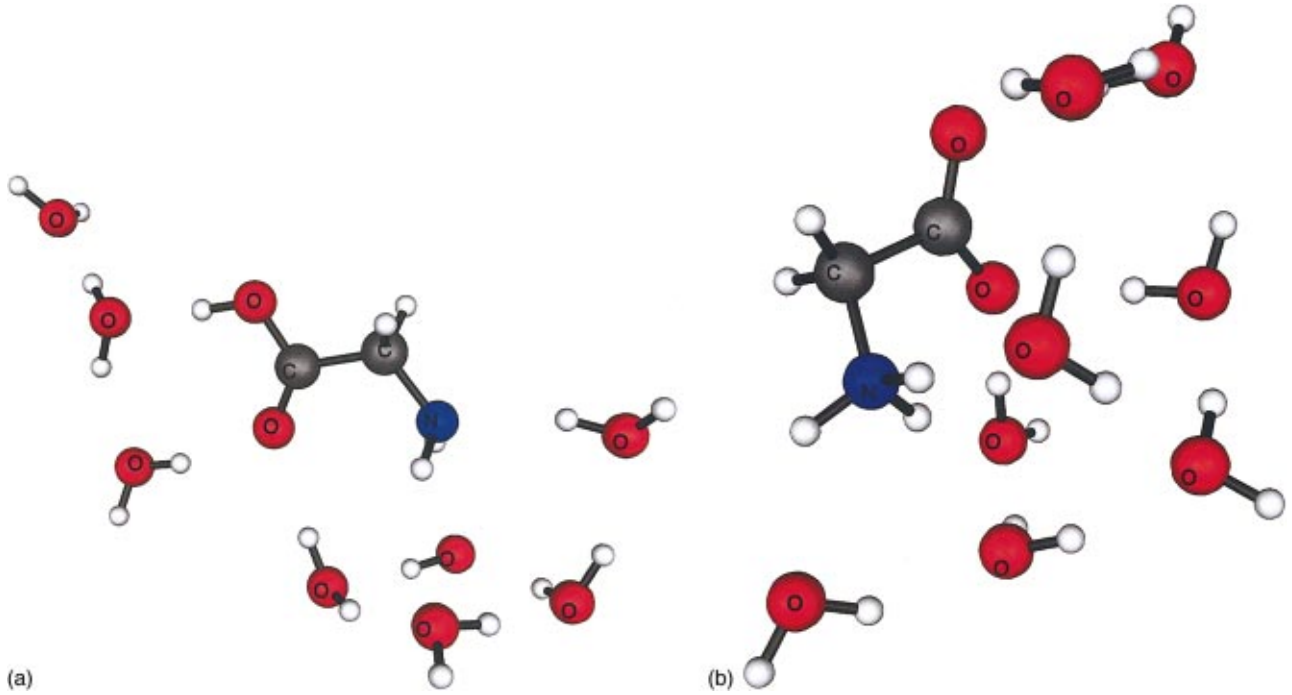

FIG. 2. (Color) Optimized structures of (a) $N(H_2O)_8$, (b) $Z(H_2O)_8$, with a RHF/6-31++G(d,p) level with effective fragment waters.

++G(d,p) level. The results are shown in Table II. As expected, there is a significant correlation effect, since MP2 preferentially stabilizes the zwitterion, reducing the TN-CZ energy difference from 15.6 kcal/mol at the HF level of theory to 4.3 kcal/mol when correlation is included. Even so, the neutral structure is still favored in the absence of the continuum. To assess the combined effects of correlation and bulk, MP2+PCM calculations have been performed using the algorithm of Cammi *et al.*[25] As shown in Table II, CZ is now more stable than TN by 5.4 kcal/mol. So, a quantitative calculation of the energy difference between the *N* and *Z* clusters requires a correlated level of electronic structure theory.

8-water case: The lowest-energy structures found for $N(H_2O)_8$ and $Z(H_2O)_8$ are shown in Figs. 2(a) and 2(b), respectively, and their relative energies at the EFP and HF level and with the PCM model are shown in Table III. With EFP waters, the *Z* cluster is more stable by 3.8 kcal/mol while full quantum calculation at the RHF level makes the *Z* cluster more stale by 1.2 kcal/mol. However, including vibrational zero-point energies reverses the energy ordering, leaving the neutral species slightly more stable. Adding PCM with EFP waters and with the full quantum cluster makes the *Z* cluster more stable by 1.2 and 0.2 kcal/mol, respectively. These results suggest that at the HF level these two clusters are essentially isoenergetic, with the EFP and all *ab initio* results once again in good agreement with each other. MP2 calculations on these clusters again illustrate the importance of electron correlation. As for the three water clusters, the zwitterion is preferentially stabilized, with the Z cluster now 11.2 kcal/mol lower in energy than the N cluster. Since the MC simulations were not exhaustive, the basis set may not be converged, and since the continuum effect was not present during the simulations, these results cannot be considered to be quantitative. However, they do emphasize the importance of electron correlation.

Note that the energy difference between TN and CZ three-water clusters (after geometry optimization in solution) is the same as the energy difference between TN and CZ with eight-water clusters (with single-point PCM calculations). The very similar difference (0.2 kcal/mol favoring *N*) is certainly fortuitous. However, these values are likely to be similar even when the eight-water cluster is optimized in

TABLE III. Relative energies of the lowest energy N-$(H_2O)_8$ and Z-$(H_2O)_8$ based on RHF/6-31++G(d,p) structures.

Structure	RHF+EFP (PCM)	Full quantum RHF(PCM)	MP2/6-31++G(d,p)
$N(H_2O)_8$	0.0/0.0[a](0.0)	0.00/0.0[a](0.0)	0.0
$Z(H_2O)_8$	−3.8/0.3[a](−1.2)	−1.2/2.1[a](−0.2)	−11.2

[a]Including Δ(ZPE).

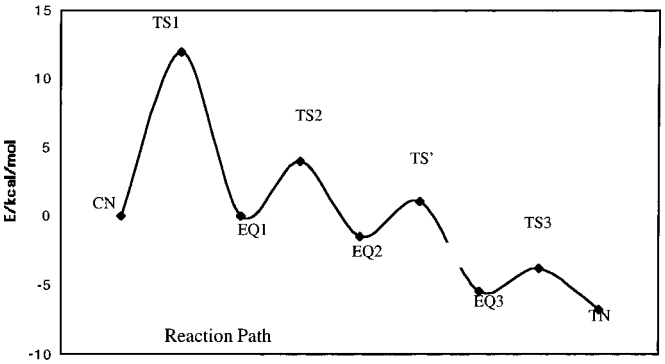

FIG. 3. Schematic reaction path connecting *cis*-neutral cluster to the *trans*-neutral cluster at the RHF/6-31++G(d,p) level with effective fragment waters. All structures are shown in Fig. 4.

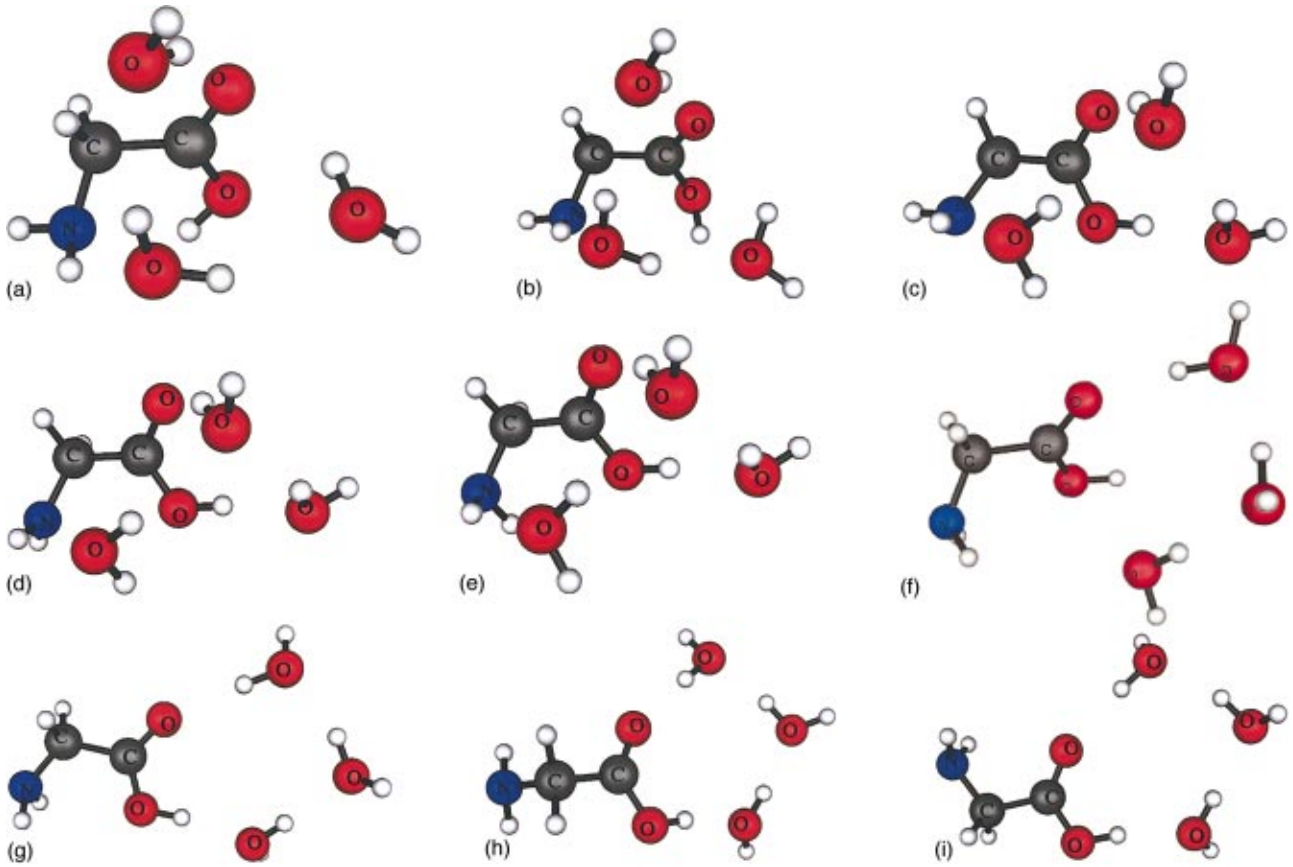

FIG. 4. (Color) Structures of the stationary points shown in Fig. 3. (a) CN (cis-neutral cluster), (b) TS1, (c) EQ1, (d) TS2, (e) EQ2, (f) TS′, (g) EQ3, (h) TS3, (i) TN (*trans*-neutral cluster).

solution, since the continuum has a small effect on the eight water clusters.

Estimate of the CN↔TN barrier: One possible reaction path connecting CN and TN is shown in Fig. 3 with the corresponding structures and energies shown in Fig. 4 and Table IV, respectively. There appear to be three local minima on this path (EQ1, EQ2, EQ3), so there must be four saddle points connecting CN and TN. The orientations of the three waters do not change significantly upon going from CN to EQ2 vis TS1 and TS2. However, there are substantial changes in the orientations of the three waters between EQ2 and EQ3. From EQ3 to TN, the orientation of waters remains almost unchanged. All transition states have been found and verified by minimum energy paths, except the one connecting EQ2 to EQ3. There is one transition state structure between EQ2 and EQ3 with an energy that is 3.0 kcal/mol above EQ2. A minimum energy path calculation starting from this HF TS (denoted TS' in Table IV) goes to EQ3 in one direction and to a new structure, EQ2', in the reverse direction. The structure of EQ2' is very close to that of EQ2. A linear least motion (LLM) path calculation connecting EQ2 and EQ2′ suggests that the upper bound to this TS is about 2 kcal/mol. The EFP and HF potential energy profiles (Table IV) are in excellent agreement with each other and

TABLE IV. Relative energies of the stationary points connecting CN and TN. Numbers in parentheses show the relative energies in the presence of the PCM continuum; the numbers within square brackets show the PCM solvation energies.

Structure	Method		
	RHF (PCM)	EFP (PCM)	MP2
Cis-$N(H_2O)_3$	0.0/0.0[a] (0.0) [13.8]	0.0/0.0[a] (0.0) [12.0]	0.0/0.0[a]
TS1	13.2/11.3[a] (12.0) [14.8]	13.0/11.2[a] (8.5) [16.7]	16.2/15.2[a]
EQ1	1.4/0.8[a] (0.0) [15.2]	1.5/1.2[a] (3.5) [17.1]	3.3/2.8[a]
TS2	4.4/2.5[a] (4.0) [14.2]	4.5/3.0[a] (0.6) [15.9]	6.9/5.2[a]
EQ2	−1.0/−1.6[a] (−1.5) [14.2]	−1.1/−1.3[a] (−5.0) [16.1]	1.1/0.4[a]
TS′	2.0/0.7[a] (1.07) [14,7]	2.9/1.7[a] (−1.2) [16.0]	
EQ3	−5.3/−5.7[a] (−5.5) [14.0]	−4.4/−4.2[a] (−8.2) [15.8]	−2.5/−2.9[a]
TS3	−3.8/−4.0[a] (−3.8 [13.8]	−2.7/−2.3[a] (−6.6) [15.7]	
Trans-$N(H_2O)_3$	−6.5/−6.9[a] (−6.8) [14.1]	−5.7/−5.6[a] (−9.5) [16.6]	−3.4/−3.8[a]

[a]Including Δ(ZPE).

suggest that the first step (TS1) is the overall bottleneck, with all other stationary points being much lower in energy. The continuum has only a small effect on this profile, because the solvation energies of the structures are similar. The numbers in square brackets are the solvation energies.

Finally, all the structures were reoptimized at the MP2 level of theory. The TS3 transition state was not found at the MP2 level of theory. This may mean that this structure and EQ3 are no longer stationary points on the MP2 potential energy surface. MP2 energies for this reaction path are shown in Table IV. The results of the calculations suggest that on the glycine+3-water cluster potential energy surface, only *Z* is significantly influenced by electron correlation. Kassab *et al.* came to a similar conclusion.

V. SUMMARY AND CONCLUSION

A unifying model for solvation, which treats the system of interest in three layers, quantum solute, explicit solvent molecules (described by the EFP method) and the bulk solvent as continuum (described by the PCM) has been developed. This model, EPM/PCM, calculates solvation energies that are in good agreement with the computationally more demanding full quantum calculations. The usefulness of the EFP/PCM model is illustrated by (a) calculating the energy difference between the neutral and zwitterionic forms of glycine, the simplest amino acid, with two different cluster sizes, three waters, and eight waters and (b) by calculating the energy barrier in going from *cis*-$N(H_2O)_3$ to *trans*-$N(H_2O)_3$ for the three-water case.

A proper identification of the lowest-energy structures of the clusters is key to obtaining reliable energetics in a supermolecule-continuum calculation. For several discrete water molecules, global optimization techniques like Monte Carlo simulated annealing can play a very important role in determining low lying structures. With the lowest-energy structures, the three-water and eight-water clusters give similar results when embedded in the continuum. Geometry optimization of the cluster structures in the continuum for the three-water case makes the result even closer to the eight-water case, where the first solvation shell is essentially filled and the continuum has a smaller effect. For the neutral zwitterion equilibrium, the use of a correlated level of electronic structure theory is necessary to produce results in agreement with the experiment.

Stationary points connecting *cis*-$N(H_2O)_3$ and *trans*-$N(H_2O)_3$ were determined with EFP, RHF, and MP2 levels of theory using the 6-31++G(d,p) basis set. The first step of this reaction path is the rate determining one, with other barriers much lower. A bulk solvent has little effect on this reaction path since the solvation energies of the structures are similar.

ACKNOWLEDGMENTS

The authors would like to thank Professor X. Song and Professor R. Cammi for helpful discussions. This work was supported by a Department of Defense software development (CHSSI) grant. The calculations were performed in part on an IBM Power3 cluster provided by IBM Shared University Research and Department of Energy grants, and in part at the Maui High Performance Computation Center under a DoD Grand Challenge grant.

[1] C. J. F. Bottcher, *Theory of Electric Polarization* (Elsevier, Amsterdam, 1973), Vol. I.
[2] J. Tomasi and M. Persico, Chem. Rev. **94**, 2027 (1994).
[3] R. M. Levy, D. B. Kitchen, J. T. Blair, and K. Krogh-Jesperson, J. Phys. Chem. **94**, 4470 (1990).
[4] P. Bandyopadhyay, S. Ten-no, and S. Iwata, Mol. Phys. **96**, 349 (1999).
[5] (a) S. T. Russel and A. Warshel, J. Mol. Biol. **185**, 389 (1985); (b) A. H. De Vries, P. Th. Van Duijnen, A. H. Juffer, J. A. C. Rullman, J. P. Dijkman, H. Merenga, and B. T. Thole, J. Comput. Chem. **16**, 37 (1995).
[6] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, J. Phys. Chem. A **105**, 293 (2001).
[7] P. Bandyopadhyay and M. S. Gordon, J. Chem. Phys. **113**, 1104 (2000).
[8] S. Miertus, E. Srocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).
[9] R. Cammi and J. Tomasi, J. Comput. Chem. **16**, 1449 (1995).
[10] (a) E. Cances and B. Mennucci, J. Math. Chem. **23**, 309 (1998); (b) E. Cances, B. Mennucci, and J. Tomasi, J. Chem. Phys. **107**, 3032 (1997); (c) B. Mennucci, E. Cances, and J. Tomasi, J. Phys. Chem. B **101**, 10 5061997)
[11] A. J. Stone, Chem. Phys. Lett. **83**, 233 (1981).
[12] D. R. Garmer and W. J. Stevens, J. Phys. Chem. **93**, 8263 (1989).
[13] (a) M. W. Schmidt, K. K. Baldridge, J. A. Boatz *et al.*, J. Comput. Chem. **14**, 1347 (1993); (b) G. D. Fletcher, M. W. Schmidt, and M. S. Gordon, Adv. Chem. Phys. **110**, 267 (1999).
[14] H. S. Rzepa and M. Y. Yi, J. Chem. Soc., Perkin Trans. 2 **4**, 531 (1991).
[15] F. R. Tortonda, J. L. Pascual-Ahuir, E. Silla, and I. Tunon, Chem. Phys. Lett. **260**, 21 (1996).
[16] J. H. Jensen and M. S. Gordon, J. Am. Chem. Soc. **117**, 8159 (1995).
[17] (a) R. Bonaccorsi, P. Palla, and J. Tomasi, J. Am. Chem. Soc. **106**, 1945 (1984); (b) T. N. Truong and E. V. Stefanovich, J. Chem. Phys. **103**, 3709 (1995); (c) J. Andzelm, C. Kolmel, and A. Klamt, *ibid.* **103**, 9312 (1995); (d) C. Adamo, V. Dillet, and V. Barone, Chem. Phys. Lett. **263**, 113 (1996); (e) F. R. Tortonda, J.-L. Pascual Ahuir, E. Silla, I. Tunon, and F. J. Ramirez, J. Chem. Phys. **109**, 592 (2000); (f) I. Tunon, E. Silla, and M. F. Ruiz-Lopez, Chem. Phys. Lett. **321**, 433 (2000); (g) L. Gontrani, B. Mennucci, and J. Tomasi, J. Mol. Struct.: THEOCHEM **500**, 113 (2000).
[18] G. Wada, E. Tamura, M. Okina, and M. Nakamura, Bull. Chem. Soc. Jpn. **55**, 3064 (1982).
[19] E. Kassab, J. Langlet, E. Evleth, and Y. Akacem, J. Mol. Struct.: THEOCHEM **531**, 267 (2000).
[20] P. Claverie, J. P. Daudey, J. Langlet, B. Pullman, D. Piazzola, and M. J. Huron, J. Phys. Chem. **82**, 405 (1978).
[21] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecci, Science **220**, 671 (1983).
[22] C. Gonsalez and H. B. Schegel, J. Phys. Chem. **94**, 5520 (1990).
[23] C. Moller and M. S. Pleset, Phys. Rev. **46**, 618 (1934).
[24] Z. Li and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 6611 (1987).
[25] R. Cammi, B. Mennucci, and J. Tomasi, J. Phys. Chem. A **103**, 9100 (1999).