Title

An Estimation of Hybrid Quantum Mechanical Molecular Mechanical Polarization Energies for Small Molecules Using Polarizable Force-Field Approaches

Journal

Journal of Chemical Theory and Computation, 13(2)

ISSN

1549-9618

Authors

Huang, Jing
Mei, Ye
König, Gerhard
et al.

Publication Date

2017-02-14

DOI

10.1021/acs.jctc.6b01125

Peer reviewed

HHS Public Access

Author manuscript

J Chem Theory Comput. Author manuscript; available in PMC 2017 February 18.

Published in final edited form as:

J Chem Theory Comput. 2017 February 14; 13(2): 679–695. doi:10.1021/acs.jctc.6b01125.

An Estimation of Hybrid Quantum Mechanical Molecular Mechanical (QM/MM) Polarization Energies for Small Molecules Using Polarizable Force-Field Approaches

Jing Huang[1,2], **Ye Mei**[3,4], **Gerhard König**[5], **Andrew C. Simmonett**[2], **Frank C. Pickard IV**[2], **Qin Wu**[6], **Lee-Ping Wang**[7], **Alexander D. MacKerell Jr.**[1], **Bernard R. Brooks**[2], and **Yihan Shao**[8,9,a]

[1]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 Penn Street, Baltimore, Maryland 21201, United States

[2]Laboratory of Computational Biology, National Institutes of Health, National Heart, Lung and Blood Institute, 5635 Fishers Lane, T-900 Suite, Rockville, Maryland 20852, United States

[3]State Key Laboratory of Precision Spectroscopy, School of Physics and Materials Science, East China Normal University, Shanghai 200062, China

[4]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

[5]Max-Planck-Institut für Kohlenforschung, 45470 Mülheim an der Ruhr, NRW, Germany, EU

[6]Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, New York 11973, United States

[7]Department of Chemistry, University of California, 1 Shields Ave., Davis, California 95616, United States

[8]Q-Chem Inc., 6601 Owens Drive, Suite 105, Pleasanton, California 94588, United States

[9]Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019, United States

Abstract

In this work, we report two polarizable molecular mechanics (polMM) force field models for estimating the polarization energy in hybrid quantum mechanical molecular mechanical (QM/MM) calculations. These two models, named the potential of atomic charges (PAC) and potential of atomic dipoles (PAD), are formulated from the *ab initio* quantum mechanical (QM) response kernels for the prediction of the QM density response to an external molecular mechanical (MM) environment (as described by external point charges). The PAC model is similar to fluctuating charge (FQ) models because the energy depends on external electrostatic potential values at QM atomic sites; the PAD energy depends on external electrostatic field values at QM atomic sites, resembling induced dipole (ID) models. To demonstrate their uses, we apply the PAC

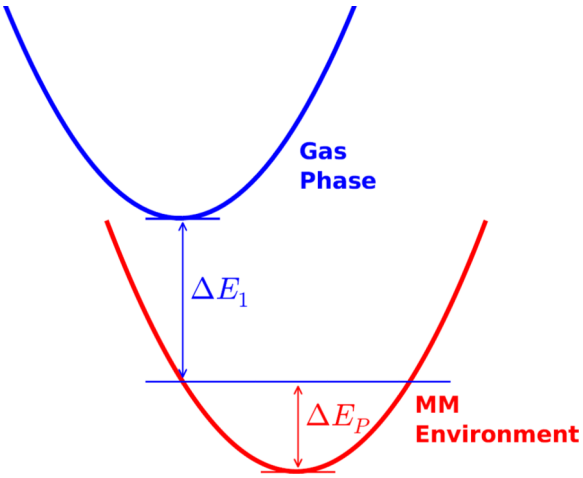
[a]Author to whom correspondence should be addressed; yihan.shao@ou.edu.

SUPPORTING INFORMATION

Drude force field parameters for aniline. Functional and basis dependence of QM/MM polarization energy, a comparison of QM/MM energy components, and the distance cutoff errors. Hydration free energies at a sampling frequency of 1ps. This information is available free of charge via the Internet at http://pubs.acs.org.

and PAD models to twelve small molecules, which are solvated by TIP3P water. The PAC model reproduces the QM/MM polarization energy with a R^2 value of 0.71 for aniline (in 10,000 TIP3P water configurations) and 0.87 or higher for other eleven solute molecules, while the PAD model has a much better performance with R^2 values of 0.98 or higher. The PAC model reproduces reference QM/MM hydration free energies for twelve solute molecules with a RMSD of 0.59 kcal/mol. The PAD model is even more accurate, with a much smaller RMSD of 0.12 kcal/mol, with respect to the reference. This suggests that polarization effects, including both local charge distortion and intramolecular charge transfer, can be well captured by induced dipole type models with proper parameterization.

Graphical abstract



I. INTRODUCTION

Over the last couple of decades, many polarizable molecular mechanics (polMM) force fields have been developed for molecular and macromolecular simulations[1–4]. In terms of the external perturbation, the polMM force field models fall into three general categories. The fluctuating charge (FQ) and related models belong to the first category, where a system responds to changes in the electrostatic *potential* at polarizable sites through charge migration within the molecule[5–22]. In the second category, the induced dipole (ID)[23–37] employ local induced dipoles to respond to changes in the electrostatic *field* at polarizable sites. In the third category, Drude oscillator (DO) models[38–45] attach Drude particles to polarizable atoms to describe instantaneous responses to changes in the environment.

In practice, molecular and macromolecular simulations using polMM force fields mostly employ ID and DO models. It remains not entirely clear how well the ID/DO models can describe the FQ effects, except that, in principle, midpoints of covalent bonds can be included as polarizable sites to account for through-bond charge fluctuation. Conversely, additional sites need to be added to FQ models to account for out-of-plane polarization. While there have been efforts from Stern *et al.* to combine FQ and ID/DO models[46], a common theoretical framework has yet to emerge, which unifies these different categories of polMM models. Such a theory would compare different polarization models on the same

footing, and provide insight into which model provides the most accurate and concise description of molecular polarizability in different environments.

In this work, such a theoretical framework is introduced in the context of estimating the polarization effect in hybrid quantum mechanical molecular mechanical (QM/MM) calculations. Specifically, by "probing" a quantum mechanical (QM) region with a large number of fixed-point-charge water molecules and modeling the electronic structure responses, we can learn how a molecular system gets polarized by an external electrostatic environment. This detailed knowledge about molecular polarizability then provides a new framework for further polMM force field development.

The first key component of our theoretical framework is a direct use of the *ab initio* response kernel, namely the energy second derivatives with respect to occupied-virtual orbital rotations, which was also utilized in our recent multiple environment single system QM/MM (MESS-QM/MM) models[47,48]. Similar response kernels were also used in Atom-Condensed Kohn-Sham Density Functional Theory approximated to second order (ACKS2) by Verstraelen, Vandenbrande, and Ayers[49], and in the fitting of atomic polarizabilities for polarizable force fields by Wang and Yang[50]. The second key component of our framework is to approximate the electrostatic potential of each occupied-virtual molecular orbital pair in two different ways: (a) as a potential of atomic charges (PAC); or (b) as a potential of atomic dipoles (PAD), where all charges or dipoles are located on QM atomic sites.

In this work, a combination of the above two components leads to two polMM-type models, PAC and PAD, for approximating the QM response with respect to multiple molecular mechanical (MM) electrostatic environments and thus for estimating the QM/MM polarization energy. The PAC model involves a response kernel for the perturbation of atomic charges by changes in the external electrostatic embedding potential at atomic sites.[51–53] The PAD model yields distributed polarizabilities, where the diagonal blocks are the generalized atomic polarizabilities and the off-diagonal blocks correspond to the coupling between induced atomic dipole moments. When the coupling between induced atomic dipole moments is neglected from the PAD model, it leads to an "inexpensive" potential of atomic dipoles (IPAD) model.

These three models (PAC, PAD, and IPAD) are described in detail in Section II. In section III, computational details are presented for their applications to twelve small molecules (water, methanol, ethanol, methanethiol, acetamide, tetrahydrofuran, benzene, phenol, aniline, ethane, n-hexane, and cyclohexane) as these molecules are being solvated in water (described with the TIP3P model[54]). Results and discussions are presented in Section IV, where polMM/MM polarization energies from these three models are compared against fully-converged QM/MM values and MESS-QM/MM values, and hydration free energy results are shown with corrections using each of these models. Concluding remarks are made in Section V.

The PAC, PAD, and IPAD models were developed as polMM extensions to our previous MESS-QM/MM modelss.[47,48] In hindsight, these new models are closely related to two previously reported theoretical methods: Misquitta and Stone's distributed polarizability

method using a constrained density-fitting algorithm[55] and Madjet, Abdurahman and Renger's fitting of the transition density as a set of atomic charges or dipoles in their computation of intermolecular Coulomb coupling.[56] The key difference of our PAC/PAD models to Misquitta and Stone's distributed polarizability method is that we represent each occupied-virtual molecular orbital pair as a set of ESP-fitted atomic charges/dipoles rather than as a linear combination of auxiliary basis functions. Of course, our PAD model is also related to many distributed polarizability models[49,50,57–66] and other models for describing the electron structure response to an external electrostatic potential.[51–53,67–72]

II. THEORY

Given an external MM environment of point charges, one can compute the QM/MM permanent electrostatics energy (also known as the vertical interaction energy[67,73,74]),

$$\Delta E_1 = \sum_{m,A} \frac{Z_A q_m}{|\vec{A} - \vec{R}_m|} + \int d\vec{r}\, \rho^{(0)}(\vec{r}) \sum_m \frac{q_m}{|\vec{r} - \vec{R}_m|} \quad (1)$$

where q_m and $\vec{R}_m$ are the charges and coordinates of each MM point charge. It includes the interaction of MM point charges with both the QM nuclei (A) and gas-phase electron density, $\rho^{(0)}(\vec{r})$.

As shown in Fig. 1, this MM electrostatic perturbation also polarizes the QM subsystem, which relaxes from its gas-phase electron structure. This leads to an energy stabilization (ΔE_P) — the QM/MM polarization energy. Within the Kohn-Sham density functional theory (KS-DFT), the electron-structure relaxation takes place through occupied-virtual orbital rotations.

One can represent the external electrostatic perturbation to the QM subsystem as

$$(\Delta h)_{ai} = \left\langle \psi_a \left| \sum_m \frac{q_m}{|\vec{r} - \vec{R}_m|} \right| \psi_i \right\rangle, \quad (2)$$

where ψ_a refer to virtual (unoccupied) molecular orbitals and ψ_i occupied molecular orbitals of the QM subsystem. The perturbation matrix in Eq. 2 contains only the occupied-virtual (OV) block, because the occupied-occupied (OO) and virtual-virtual (VV) blocks lead to unitary transformations within the occupied or virtual molecular orbital spaces that affect neither the electron density nor the system energy.

The second order approximation to the QM/MM polarization energy within the KS-DFT framework can be written as[47]

$$\Delta E_2 = -\frac{1}{2} \sum_{ai,bj} (\Delta h)_{ai} \left[H^{-1}\right]_{ai,bj} (\Delta h)_{bj}, \quad (3)$$

where H^{-1} is the inverse of the response kernel (second derivatives of energy with respect to occupied-virtual orbital rotations). When the inverse Hessian (H^{-1}) in Eq. 3 is computed within an incomplete subspace, this leads to the MESS-H approximation to the QM/MM polarization energy.[47] If one takes a diagonal approximation to the electronic Hessian, namely using only orbital energy differences, one obtains the GMIPp model for estimating the QM/MM polarization from Gao, Luque, and coworkers.[67,75,76] This is equivalent to an extrapolation of the Fock matrix from gas-phase to the MM environment, leading to the MESS-E approximation of the QM/MM polarization energy.[47] The MESS-E model was employed successfully in the computation of QM/MM-corrected hydration free energies.[48]

In this work, we demonstrate further approximations to the QM/MM polarization energy in Eq. 3, leading to a series of computational models shown in Fig. 2. This begins by rewriting the MM electrostatic perturbation in Eq. 2,

$$(\Delta h)_{ai}=\sum_m q_m \left\langle \psi_a \left| \frac{1}{|\vec{r}-\vec{R}_m|} \right| \psi_i \right\rangle = \sum_m q_m \phi_{ai}(\vec{R}_m), \quad (4)$$

which is a summation over the electrostatic potential from molecular orbital pairs at MM atomic sites

$$\phi_{ai}(\vec{R}_m)=\left\langle \psi_a \left| \frac{1}{|\vec{r}-\vec{R}_m|} \right| \psi_i \right\rangle \quad (5)$$

and can be approximated in the following three ways.

A. Potential of atomic charges (PAC) model

Here we use fitted atomic charges (Q_{ai}^A) to approximate the electrostatic potential of each pair of occupied-virtual molecular orbitals in Eq. 5:

$$\phi_{ai}(\vec{R}_m)\sim\phi_{ai}^{(PAC)}(\vec{R}_m)=\sum_A \frac{Q_{ai}^A}{|\vec{A}-\vec{R}_m|} \quad (6)$$

where A refer to QM atoms and $\vec{A}$ their coordinates. This allows us to approximate the MM electrostatic perturbation as

$$(\Delta h)_{ai}\sim\sum_m q_m \sum_A \frac{Q_{ai}^A}{|\vec{A}-\vec{R}_m|}=\sum_A Q_{ai}^A \phi_A \quad (7)$$

where ϕ_A's are simply the MM electrostatic embedding potential at QM atomic sites:

$$\phi_A = \sum_m \frac{q_m}{|\vec{A} - \vec{R}_m|} \quad (8)$$

Using a coupling matrix

$$S_{AB} = \sum_{ai,bj} Q^A_{ai} \left[H^{-1}\right]_{ai,bj} Q^B_{bj} \quad (9)$$

one obtains the PAC approximation to the second-order QM/MM polarization energy in Eq. 3,

$$\Delta E_2^{PAC} = -\frac{1}{2} \sum_{A,B} \phi_A S_{AB} \phi_B \quad (10)$$

To the extent that the PAC polarization energy depends on the MM electrostatic *potential*, it is related to various fluctuating-charge descriptions for the molecular polarization.[5–22] Specifically, with a change in the local electrostatic potential at one atomic site (ϕ_B), the "fluctuation" in the atomic charges in the molecule is given by $S_{AB}\phi_B$ values in Eq.10.

In other words, with the polarization energy definition in Eq. 10, the PAC coupling matrix can be written as

$$S_{AB} = \frac{\partial^2 E}{\partial \phi_A \partial \phi_B} = \frac{\partial q_A}{\partial \phi_B}, \quad (11)$$

which is closely related to the concept of local softness.[77–80] Due to the difference between the linear-response kernel and the local softness (see, for example, Eq. 21 in Reference[80]), we will call S_{AB} the pseudo-softness supermatrix.

B. Potential of atomic dipoles (PAD) model

Here the electrostatic potential in Eq. 5 is approximated using fitted atomic dipoles:

$$\phi_{ai}(\vec{R}_m) \sim \phi^{(PAD)}_{ai}(\vec{R}_m) = -\sum_A \frac{\vec{\mu}^A_{ai} \cdot (\vec{A} - \vec{R}_m)}{|\vec{A} - \vec{R}_m|^3} \quad (12)$$

which leads to the following estimation for the MM electrostatic perturbation

$$(\Delta h)_{ai} \sim -\sum_m q_m \sum_A \frac{\vec{\mu}^A_{ai} \cdot (\vec{A} - \vec{R}_m)}{|\vec{A} - \vec{R}_m|^3} = \sum_A \mu^A_{ai} \cdot \vec{F}_A \quad (13)$$

where

$$\vec{F}_A = -\sum_m \frac{q_m(\vec{A} - \vec{R}_m)}{|\vec{A} - \vec{R}_m|^3} \quad (14)$$

are the MM electrostatic embedding field at QM atomic sites. Using a coupling matrix defined as

$$T_{AB} = \sum_{ai,bj} \vec{\mu}^A_{ai} \left[H^{-1}\right]_{ai,bj} \vec{\mu}^B_{bj} \quad (15)$$

one obtains the PAD approximation to the QM/MM polarization energy

$$\Delta E_2^{PAD} = -\frac{1}{2}\sum_{A,B} \vec{F}_A \cdot T_{AB} \cdot \vec{F}_B \quad (16)$$

This is closely related to the induced dipole description of the polarization energy.[25–28,30–37] Clearly, the coupling matrix as defined in Eq. 15 correspond to distributed polarizabilities: its diagonal blocks are the (anisotropic) atomic polarizabilities, because $T_{BB} \cdot \vec{F}_B$ yields the induced dipole on atom B due to a small change in the electrostatic field at the same site ($\vec{F}_B$).

There are two ways to interpret the off-diagonal blocks (T_{AB}, $A \neq B$). On one hand, we can regard $T_{AB} \cdot \vec{F}_B$, ($A \neq B$) as *non-local* responses, namely induced dipoles at other atomic sites caused by a change in the local electrostatic field at site B. On the other hand, we can rewrite the PAD energy as

$$\Delta E_2^{PAD} = -\frac{1}{2}\sum_{A,B} \left[\vec{F}_A \cdot T_{AA}\right] \cdot \left[(T_{AA})^{-1} T_{AB} (T_{BB})^{-1}\right] \cdot \left[T_{BB} \cdot \vec{F}_B\right] \quad (17)$$

where $(T_{AA})^{-1} T_{AB} (T_{BB})^{-1}$ couples *local* induced dipoles at different atomic sites ($T_{AA} \cdot \vec{F}_A$ and $T_{BB} \cdot \vec{F}_B$).

C. Inexpensive potential of atomic dipoles (IPAD) scheme

Following the inexpensive AMOEBA model,[36] a further approximation can be made to the PAD model, where the off-diagonal terms (i.e. T_{AB}, $A \neq B$) in Eq. 16 are not included. The IPAD QM/MM polarization energy is

$$\Delta E_2^{IPAD} = -\frac{1}{2}\sum_{A} \vec{F}_A \cdot T_{AA} \cdot \vec{F}_A. \quad (18)$$

The IPAD model assumes that, upon a local electrostatic field change at atomic site A, an induced dipole is generated only at the same site, while the electron density distortion at all other sites can be ignored. Alternatively, according to Eq.17, one can say that the IPAD model neglects the coupling between induced dipoles at different atomic sites.

III. COMPUTATIONAL DETAILS

In this work, twelve solute molecules in Fig. 3 were considered: water, methanol, ethanol, methanethiol, acetamide, tetrahydrofuran, benzene, phenol, aniline, ethane, *n*-hexane, and cyclohexane. This set of molecules, many of which are protein side-chain analogs, were previously studied in QM/MM-corrected hydration free energy calculations.[48] In all calculations, the solute molecules adopted a fixed gas-phase geometry optimized with the ωB97X-D functional[81] and 6-311++G** basis set,[82] using the Q-Chem software package.[83]

All computational models (fully-converged QM/MM, MESS-H, MESS-E, PAC, PAD, and IPAD) were analyzed using a large number of configurations generated from classical hydration free energy calculations using the CHARMM software package.[84] In these calculations, each solute molecule was constrained to its gas-phase geometry and was solvated in a truncated octahedral box of 1687 TIP3P water molecules.[54]

For the computation of hydration free energy at the classical level, we followed our earlier procedure[48] where 24 λ points were sampled with various scaling factors for the electrostatic and vdW interactions between the solute and solvent molecules. Each λ point consisted of a 1 ns NVT trajectory at 300K. From the first MM trajectory, configurations were collected at an interval of 100fs, leading to 10000 configurations for further QM/MM analyses. To study the effect of the correlation length, a longer sampling interval of 1ps was also employed, with the results presented in the supporting information.

For each solute molecule, the *ab initio* response kernel ($H_{ai,bj}$) was computed within a subspace at the B3LYP/6-31G* level of theory[85–87] using the stability analysis[88] module as implemented in Q-Chem. As before, the stability analysis was performed iteratively,[47] with the number of eigenvalues requested being 3.5 times the total number of electrons. Separately, the electrostatic potential of each occupied-virtual molecular pair was fitted as a potential of atomic charges (Eq. 6) or atomic dipoles (Eq. 12) following a procedure using single value decomposition for deriving Merz-Kollman charges.[89,90] From the fitted charges/dipoles, the gas-phase coupling matrix elements in the PAC, PAD, and IPAD models were computed for each solute molecule according to Eqs. 9 and 15.

For each of the 10000 configurations for each solute molecule, we computed fully-converged QM/MM energy, MESS-H-QM/MM energy, and MESS-E-QM/MM energy. For PAC, PAD and IPAD calculations, the MM electrostatic embedding potential (Eq. 8) and MM electrostatic embedding field (Eq. 14) were computed at all atomic sites in the solute molecule. When combined with the coupling matrices, this yielded the QM/MM polarization energy approximated with the PAC (Eq. 10), PAD (Eq. 16), and IPAD (Eq. 18) models. All these energy values were used in the computation of the QM/MM correction of the hydration free energies using the non-Boltzmann Bennett (NBB) method.[91,92]

For a comparison to a polarizable force field in wide use, we also performed Drude oscillator calculations[45] on the solute molecules using the CHARMM software package.[84] First, a gas-phase Drude oscillator optimization was performed for each solute molecule, and the Drude particle positions are recorded. Then, 1687 TIP3P water molecules were added to the calculation, and the Drude particle positions are reoptimized to minimize the energy. The polarization energy in this mixed Drude/TIP3P calculation was simply the energy difference caused by moving the Drude particles from their gas-phase positions to condensed-phase ones. Drude force field parameters are available for all solute molecules except for aniline, whose parameters were empirically transferred from those of phenol and cytosine, and provided in the supporting information. For water as a solute molecule, both the 4-site (SMW4) model[40] and the 6-site (SWM6) model[41] were employed and the results were very similar. So only the SWM4 results are reported in the next section.

IV. RESULTS AND DISCUSSIONS

A. Single-point energies

We begin by comparing approximate QM/MM polarization energies using five computational models (MESS-H, MESS-E, PAC, PAD, and IPAD) against the fully-converged QM/MM values. The QM region would be one solute molecule described at the B3LYP/6-31G* level of theory, and the MM region contains 1687 TIP3P water molecules. For each solute molecule, 10,000 configurations were collected from the classical hydration free energy calculations.

Scatter plots are shown in Figs. 4, 5, 6, 7, and 8, one for each of the five computational models. In those plots, each point indicates one of the 10,000 configurations, with the x-axis corresponding to the reference QM/MM polarization energy computed by fully converging Kohn-Sham SCF equations under each MM environment, and the y-axis values being the estimated QM/MM polarization energies using one of the five computational models. So with the best agreement, all points in a plot should lie perfectly along the diagonal. The R^2 values are shown in these plots, and the mean signed deviations (MSD) and root-mean square deviations (RMSD) are listed in Table I.

In our calculations, the reference QM/MM polarization energy (x-axis values) is the largest with acetamide ranging from −8 to 0 kcal/mol, depending on the solvent environment. This reflects that it can be strongly polarized by the solute water molecules. The energy ranges from −6 to 0 kcal/mol for water, methanol, ethanol, tetrahydrofuran, phenol and anline molecules, about −3 to 0 kcal/mol with methanethiol, benzene, hexane and cyclohexane. And it becomes even smaller (−1.5 to 0 kcal/mol) for ethane.

1. MESS-H and MESS-E models—As shown in Fig. 4, the MESS-H model well reproduces the reference QM/MM polarization energy. The R^2 value is found to be 1.0 for all twelve molecules. While points lie nearly perfectly along the diagonal for most molecules, it is noticeable that, for water, methanol, ethanol, and phenol molecules, the QM/MM polarization energies get slightly over-estimated (i.e. even more negative). Overall, as shown in, Table I its MSD ranges from −0.087 kcal/mol (water) to 0.006 kcal/mol (hexane and cyclohexane). The RSMD value ranges from 0.002 (ethane) to 0.1 kcal/mol

(water). This sub-0.1 kcal/mol agreement indicates that, for these neutral molecules, the QM/MM polarization is dominated by the second-order term in Eq. 3. Namely, hyperpolarizability and higher-order responses can be neglected. This is rather encouraging for the development of PAC, PAD and IPAD models, because they are further approximations to the MESS-H model.

Within the MESS-E model, an extrapolated Fock matrix is diagonalized to predict the orbital response. As we reported previously,[47,48] if the MM electrostatic potential in Eq. 2 is added to the gas-phase Fock matrix without any scaling, it resembles a steepest descent step and thus can lead to an overestimation in the response in the QM electron structure. To correct this systematic error, here we use the same scaling factors listed in Table S13 of our earlier publication.[48]

Since an approximate Hessian is employed, the MESS-E model is expected to display larger deviations than the MESS-H model. But, from Fig. 5, it can be seen that the MESS-E model still performs reasonably well with the smallest R^2 values being 0.96 (aniline). In Table I, its MSD value is found to be between −0.082 (anline) and 0.061 kcal/mol (benzene), and its RMSD value falls between 0.015 (ethane) to 0.181 kcal/mol (aniline). With a small deviation (< 0.2 kcal/mol) and very low computational costs (one matrix diagonalization and no SCF cycles for each MM configuration), the MESS-E model should be preferable for performing accelerated QM/MM simulations of other solvation processes, ligand-receptor binding, and enzymatic reactions.

2. PAC model—As a further approximation to the MESS-H model, PAC (and PAD or IPAD) is expected to yield larger deviations from the reference QM/MM polarization energies. As shown in Fig. 6, however, the PAC model yields reasonably good performance for most of the twelve molecules. The R^2 values exceed 0.94 for nine molecules: water, methanol, ethanol, acetamide, tetrahydrofuran, phenol, ethane, hexane, and cyclohexane. It performs less well for methanethiol (R^2=0.92) and benzene (R^2 = 0.87). The worst case is aniline with a R^2 value of 0.71.

In Table I, one observes a significant over-estimation of the polarization energy for water (MSD=−0.184 kcal/mol) and aniline (MSD=−0.433 kcal/mol), and a significant underestimation in terms of MSD values for acetamide (0.188 kcal/mol), benzene (0.219 kcal/mol), and phenol (0.175 kcal/mol), whereas the deviations are smaller than 0.15 kcal/mol for all other molecules. The RMSD values range from 0.059 kcal/mol (ethane) to 0.756 kcal/mol (aniline).

It is not surprising that the PAC model does not perform as well for these planar molecules (aniline, benzene, phenol, acetamide, and water) because, like other fluctuating charge models, it lacks a description for out-of-plane polarization.

3. PAD model—The PAD model yields noticeably better performance than the PAC model. As shown in Fig. 7, its R^2 value is found to be above 0.98 for all twelve molecules.

The good performance of the PAD model is also evident in Table I, where its MSD values range from −0.165 kcal/mol (water) to 0.054 kcal/mol (tetrahydrofuran). Its RMSD value is found to be below 0.15 kcal/mol for all molecules, except for water (0.216 kcal/mol).

The very encouraging performance of the PAD model suggests that the intramolecular charge transfer effect can be recovered, to a large extent, by induced-dipole-type models (such as PAD), even without the inclusion of polarizable mid-bond points.

4. IPAD model—The IPAD model, which neglects the off-diagonal blocks of the distributed polarizability tensor in Eq. 15, leads to slightly better results than PAD for the water molecule: the MSD value improves from −0.165 kcal/mol (PAD) to 0.025 kcal/mol (IPAD), and the RMSD value from 0.216 kcal/mol (PAD) to 0.089 kcal/mol (IPAD).

For all other molecules, the IPAD model performs less well. As shown in Fig. 8, the QM/MM polarization energy is over-estimated for ten molecules other than water and ethane. And the fitted slopes in Fig. 8 suggest that the energy is about twice too large for five molecules: tetrahydrofuran, benzene, phenol, aniline, and cyclohexane. This is not surprising, given the observation that the induced dipole moments from the iAMOEBA model typically get reduced by subsequent self-consistent field or perturbation theory treatments included to account for the interactions between induced dipoles in the AMOEBA model.[26,93]

On the other hand, the IPAD polarization energies do correlate reasonably well with the reference values: the smallest R^2 value is 0.86 (aniline). This compares less favorably than the PAD model (smallest R^2 is 0.98), but more favorably than the PAC model (smallest R^2 is 0.71). This suggests that a scaling of the atomic polarizabilities (T_{AA} in Eq. 15) can be employed to correct the systematic overestimation from the IPAD model. Attempts along this direction will be pursued in the near future.

5. Drude oscillator model—The QM/MM polarization energy as estimated using the Drude oscillator (DO) model are shown in Fig. 9. It is encouraging to see an overall good correlation between the DO and reference values: the best correlation occurs with tetrahydrofuran (R^2 = 0.95), and the smallest R^2 value is still relatively high at 0.59 (aniline). On the other hand, it is clear from Fig. 9 that the DO model significantly underestimates the QM/MM polarization energy for five molecules (methanethiol, benzene, phenol, hexane, and cyclohexane), with a less significant systematic deviation for other eight molecules.

Since most of the atomic polarizabilities employed in our DO calculations come from MP2 calculations using basis sets larger than 6-31G*, it is worthwhile to analyze the basis-set and method/functional dependence of the QM/MM polarization energies. So QM/MM polarization energies were computed for two solute molecules (water and ethanol) with four other density functionals (BLYP, PBE0, M06-2X, and ωB97X-D) and with four larger basis sets (6-311++G**, 6-311++G(3df,3pd), aug-cc-pVDZ, and aug-cc-pVTZ). As shown in the top panels in Figs. S1 and S2, the QM/MM polarization energy is rather insensitive to the choice of density functionals. However, as shown in the lower panels, the QM/MM

polarization energy *does* depend heavily on the size of the basis set. For both molecules, the QM/MM polarization energies become about 10–15% larger with the 6-311++G** basis, and about 50% larger with the 6-311++G(3df,3pd), aug-cc-pVDZ, and aug-cc-pVTZ. This is exactly what one would expect, because the computed molecular polarizability becomes larger with increased size of basis sets.

Going back to the DO model, it will uniformly underestimate the polarization energy for all twelve molecules, had the reference QM/MM values been computed with a larger basis. This systematic difference is also expected, because the atomic polarizabilities are usually scaled down in the DO model to damp the interactions associated with induced dipoles.

6. Permanent electrostatics versus polarization—As shown in Fig.S4, the QM/MM polarization energies (ΔE_P) for these molecules are several times smaller than the QM/MM permanent electrostatics energy in Eq.1.

According to Eq. 1, MM charges have a first-order contribution to the QM/MM permanent electrostatics energy. For TIP3P molecules, if we denote their distance from the solvent by r, then their contributions to ΔE_1 should decay asymptotically as $1/r^2$ if the QM molecule has a net charge, and $1/r^3$ for QM molecules with a net gas-phase dipole moment (such as eight of the solute molecules including acetamide and tetrahydrofuran), and $1/r^4$ for those with only quadrupole moments in the gas-phase (such as benzene, ethane, hexane, and cyclohexane).

Given that (a) the PAD model was found to well reproduce the QM/MM polarization energy, and (b) the electrostatic field from each TIP3P water decays as $1/r^3$, it might appear that the quadratic formula for the PAD polarization energy in Eq. 16 would suggest that faraway TIP3P water molecules would have a second-order contribution to the QM/MM polarization energy, which should scale as $1/r^6$ and might allow us to impose more aggressive distance cutoffs for the polarization effect than for the permanent electrostatics.

This is clearly not true as shown in Fig. S5, where the contributions to both ΔE_1 and ΔE_P from TIP3P molecules beyond a cutoff distance from the solute are presented for two randomly selected frames of solvated acetamide and tetrahydrofuran. There, as shown in the right panels, the contributions to the QM/MM polarization energy from those faraway TIP3P water molecules are clearly proportional to their contributions to the QM/MM electrostatics energy (which should scale as $1/r^3$, because both solute molecules have a permanent dipole moment). In other words, the ΔE_P contribution should be pseudo-first-order and also decay as $1/r^3$ for faraway water molecules, as a result of their interactions with the induced dipoles on QM atoms due to water molecules in the first few solvation shells.

A more detailed analysis for the decay behavior of QM/MM electrostatics and polarization energies will be presented in another publication in the context of boundary effects for QM/MM calculations.

B. Hydration free energies

The reference QM/MM polarization energy and its approximate values from using four computational models can be employed in the computation of the QM/MM-correction to

hydration free energy for the twelve solute molecules, and the results are collected in Table II and Fig. 10. Here the IPAD model is not applied, because proper scaling is needed to correct its tendency to systematically overestimate the QM/MM polarization energy.

As shown in the top panel of Fig. 10, all five models (full QM/MM, MESS-H, MESS-E, PAC, PAD) tend to well reproduce the hydration energy of seven solute molecules: ethanol, methanol, benzene, phenol, ethane, hexane and cyclohexane. On the other hand, all these models tend to overestimate the hydration free energy for five other molecules (water, methanethiol, acetamide, tetrahydrofuran, and anline). As shown in Table II, the over-all MSD is −0.63 kcal/mol with full QM/MM, −0.67 kcal/mol with MESS-H, −0.61 kcal/mol with MESS-E, −0.82 kcal/mol with PAC, and −0.69 kcal/mol with PAD. As indicated in our previous publication,[48] this systematic overestimation arises from the incompatibility of the QM/MM electrostatics (using the B3LYP/6-31G* density for the QM region) and the QM/MM vdW parameters in use. The compatibility of QM methods with water models was first reported by Shaw, Woods and Mulholland.[94]

While it would be desirable to improve the QM/MM methodology to overcome this systematic error, our goal of this work is to develop practical approximations to full QM/MM calculations. Therefore, it is more important to compare the computed hydration free energies using approximate energy values against full QM/MM calculations. As shown in the bottom panel of Fig. 10, the MESS-H, MESS-E and PAD models well reproduce the full QM/MM hydration free energies. In Table II, the MSD value with respect to full QM/MM is shown to be −0.04 kcal/mol for MESS-H, 0.02 kcal/mol for MESS-E, and −0.06 kcal/mol for PAD. The corresponding RMSD values are 0.07 kcal/mol for MESS-H, 0.05 kcal/mol for MESS-E, and 0.12 kcal/mol for PAD. With the PAD model, the largest deviation in the hydration free energy occurs with water (8.4 kcal/mol versus the 8.08 kcal/mol theoretical reference value).

The PAC model displays larger deviations, especially for the aniline molecule where the full QM/MM value is −7.42 kcal/mol and the PAC value is −9.31 kcal/mol. This is a direct result of the systematic over-estimation of the QM/MM polarization energy shown in Fig. 6 from applying the PAC model to this planar molecule. For the entire set of twelve molecules, the MSD value of the PAC model is −0.19 kcal/mol, and the RMSD value is 0.59 kcal/mol.

Overall, it is very encouraging to see the PAD model produces hydration free energies within a RMSD of 0.12 kcal/mol from the reference values. Given the simplicity and apparent deficiency of the PAC model, it is still encouraging to see that it predicts the hydration free energies within a RMSD of 0.59 kcal/mol.

All above hydration free energy results were obtained with a sampling frequency of 100fs. Since the water configurations might be still correlated with this sampling interval, the NBB correction to the hydration free energy was also computed using a longer sampling interval of 1ps. As shown in Fig. S6, the error bars become noticeably larger because fewer data points are employed in the sampling. On the other hand, there are no substantial changes in the computed hydration free energy values.

C. Analysis of coupling matrices

1. PAC model—The PAC coupling matrix in Eqs. 9 and 11, which we call pseudo-softness, describes the linear response of a molecule to an external electrostatic potential. Its diagonal elements, S_{AA}, measure the tendency for an atom to lose electrons (to its neighboring atoms) with an increase in the local chemical potential at the same atomic site. Their values are shown in Fig. 11 for each molecule and also in Fig. S3, where atoms from the twelve molecules are grouped into several types.

The pseudo-softness values for hydroxyl oxygen and hydrogen atoms vary slightly among molecules in this study. At the current level of theory (B3LYP/6-31G*), the oxygen atom in the water molecule has a pseudo-softness of about 4.04, while the hydrogen atoms have a softness of about 1.8. In methanol, the pseudo-softness increases slightly to 4.26 for the oxygen atom, while the value remains around 1.8 for the hydroxyl hydrogen. In ethanol, the pseudo-softness further increases to 4.77 for the oxygen atom, while the value also changes to 1.94 for the hydroxyl hydrogen. In phenol, the pseudo-softness value is 5.50 for the oxygen atom and 2.03 for the hydroxyl hydrogen.

As shown in Fig. S3, the pseudo-softness values vary much more among the carbon atoms. The highest value occurs with methylene (CH_2) carbon atoms. In *n*-hexane, its value ranges from 39.1 to 42.1. The value is similar for cyclohexane (40.8), but lower for tetrahydrofuran (31.7 and 35.8), and even lower for ethanol (24.6). This, in general, correlates with the amount of partial charges on these atoms. For example, the ethanol methylene carbon atom, which is the least polarizable among this group, has the least negative Mulliken charge of −0.02.

In general, methyl carbon atoms are less polarizable than methylene carbons. The pseudosoftness value for methyl carbons is 26.5 for ethane, 25.5 for hexane, and 24.7 for acetamide. A lower value is found for methanol (19.0), and a higher value is found for ethanol (32.1) and methanethiol (34.5). This also roughly correlates with atomic charges: the carbon atom in methanol has the least negative Mulliken charge of −0.20, and it is also the least polarizable. On the other end, the Mulliken charge is the highest for the carbon atom in methanethiol, making it the most polarizable.

Most carbon atoms on the benzene ring have a lower pseudo-softness value: 18.9 for benzene carbon atoms, and 18.6–19.8 for carbon atoms in phenol. In aniline, the value goes up slightly to 21.2–24.8 for five carbon atoms not directly bonded to the amine (NH_2) group, and it increases dramatically to 42.8 for the carbon atom next to NH_2.

The off-diagonal elements in the pseudo-softness matrix, $S_{AB}(A \neq B)$, correspond to the energy associated with intramolecular charge transfer. For two atoms connected through chemical bonds, its value should be negative. Namely, if the potential is raised on one atom, its bonded neighbors are expected to gain some electronic charges. This is indeed what we found for all twelve molecules.

The off-diagonal elements are expected to decay with the distance between the atoms. Since *n*-hexane is the longest molecule in this study, we divided the coupling between three

categories: carbon-carbon, carbon-hydrogen, and hydrogen-hydrogen. In Fig.12, the values of the pseudo-softness matrix elements are plotted against the atom-atom distances for each of these three categories. A general trend of exponential decay can be observed.

If the potential is raised by the same amount for all atoms in a molecule, it should not cause any intramolecular charge transfer. In terms of the pseudo-softness matrices, this means that the sum of each row (or column), $\sum_A S_{AB}$, should be zero. This is confirmed in this study, where the sums are mostly below 0.01, and the largest deviation is 0.055 for the methanethiol molecule.

2. PAD model—The atomic polarizabilities from the PAD model are the diagonal blocks (T_{AA}) of the distributed polarizability matrix in Eq. 15. They correspond to the dipolar response of each atom with a change in the local electrostatic embedding field. In Fig. 13, the three eigenvectors of the atomic polarizability matrix were shown as arrows from each non-hydrogen atom, with the length of each arrow proportional to the eigenvalues (principal polarizabilities). The scalar polarizability (average of the three principal polarizabilities) is shown next to each non-hydrogen atom.

From Fig. 13, it can easily be seen that atomic polarizability can be highly non-isotropic for some atoms. For example, the carbon atom in methanethiol is most polarizable along the C-S bond, and less polarizable in two other directions. The carbon atom in benzene is most polarizable along the ring, and least polarizable perpendicular to the ring.

In terms of scalar polarizability, the oxygen atom has relatively small values: 3.59 (water); 4.26 (acetamide); 4.33 (methanol), 4.67 (ethanol) and 5.06 (phenol), all in atomic units. This is comparable to the atomic polarizability for sp^3 oxygen in the range of 2.88 to 5.21 in Table 1 in the recent paper from Wang and Yang.[50]

The scalar polarizability of methyl carbon is 7.69 for methanol, 8.24 for ethane, 9.22 for *n*-hexane, 9.53 for ethanol, and 10.8 for acetamide. But its value jumped to 17.1 for methanethiol (which also displayed the largest value for the pseudo-softness). Compared to methyl carbon atoms, the methylene carbon atoms usually have a higher scalar polarizability: 11.8 for ethanol, 14.2 and 14.9 for *n*-hexane, 16.9 for cyclohexane, 14.9 and 19.1 for tetrahydrofuran. While the carbon atom in benzene has a similar scalar polarizability of 15.2, the value goes up in phenol (16.0, 16.5, 16.7, 23.6) and aniline (15.2, 15.3, 17.4, 25.8), displaying a clear trend of larger scalar polarizability for carbon atoms closer to the hydroxyl or amine groups. Overall, these atomic polarizability values are slightly higher than the values of 5.00-9.13 for sp^3 carbon atoms and 4.51–13.40 for sp^2 carbon atoms as reported by Wang and Yang.[50]

The off-diagonal blocks (T_{AB}, $A \neq B$) tell us how induced dipoles at different atomic sites get coupled together. And the coupling is expected to decay with the distance. To confirm that, we plot the 2-norm of each off-diagonal block for *n*-hexane against the atom-atom distances. As shown in Fig. 14, an exponential decay is displayed for all three cases: C-C, C-H, and H-H.

V. CONCLUSIONS

In this work, we have developed two new polarizable-force-field type models to describe the response in the electron structure of twelve solute molecules to external electrostatic perturbations from water solvent molecules. They are named the potential of atomic charges (PAC) model and the potential of atomic dipoles (PAD) model.

Following common polarizable force field models, the electrostatic embedding potential from the water molecules (as represented by the TIP3P model) are approximated by a local Taylor expansion. The PAC model employs a zeroth-order Taylor expansion that only uses the electrostatic embedding *potential* at solute atom positions. The PAD model utilizes only the first-order term in the local Taylor expansion, namely the electrostatic embedding *field* at solute atom positions.

To obtain response kernels to the embedding potential or field, we fit each pair of occupied-unoccupied orbitals as a set of atomic charges (in the PAC model) or a set of atomic dipoles (in the PAD model), and then combine the fitting results with the DFT electron Hessian (also called the orbital Hessian) to obtain the response kernels for the two models: pseudo-softness matrix in the PAC model and distributed polarizability matrix in the PAD model. An inexpensive variation of the PAD model, termed the IPAD model, was also analyzed where the off-diagonal blocks of the distributed polarizability matrix are neglected. For a comparison, Drude/TIP3P calculations were also performed with the solute molecules described with the Drude oscillator model and the solvent molecules retaining a TIP3P description.

Our main observations are:

- The PAC model reproduces the polarization energy (for 10000 different water configurations) reasonably well, with a R^2 value of 0.71 (aniline), 0.87 (benzene), 0.92 (methanethiol), 0.94 and above (other molecules). Compared to theoretical reference values, the PAC model well reproduces the hydration free energy of most molecules, but over-estimates the hydration free energy of aniline by 1.89 kcal/mol. For the entire set of twelve molecules, the PAC model has a MSD of −0.19 kcal/mol and a RMSD of 0.59 kcal/mol.
- The PAD model reproduces the polarization energy for individual water configurations even better, with a R^2 value of 0.98 and above for all molecules. It well reproduces the reference theoretical values for the hydration energies of the set of twelve solute molecules, with a MSD value of −0.06 kcal/mol and a RMSD value of 0.12 kcal/mol. The largest deviation occurs with water, where PAD overestimates the hydration energy by 0.32 kcal/mol.
- The IPAD model, which neglects the off-diagonal blocks of the distributed polarizability matrix, tends to significantly over-estimate the polarization energy for all molecules except for water and ethane. But its results still correlate well with reference theoretical values with a R^2 value of 0.86 and above, suggesting that the IPAD model can potentially be used with proper scaling.

- The Drude oscillator model produces polarization energies that correlate reasonably well with theoretical reference values, with R^2 values of 0.59 – 0.96.

In short, we have shown that response kernels in our polarizable-force-field-type models can be derived directly from *ab initio* electron structure — in particular the density functional theory orbital Hessian. Given the simplicity of the PAD model, it is very encouraging to observe its good performance. This suggests that *polarization effects, including both local charge distortion and intramolecular charge transfer, can be well captured by induced dipole type models with proper parameterization*, which is consistent with the findings by Ren and Ponder in 2002.[95]

On the other hand, this work has several limitations:

- It might be difficult to improve the performance of the PAC model for planar molecules like aniline, which displays non-negligible out-of-plane polarizability.
- The fitting of each pair of occupied-unoccupied MOs as a set of atomic charge or dipoles is performed through a least-square fitting of the electrostatic potential.[89,90] This procedure is numerically undesirable for large molecules with many buried atoms.[96–99]
- In principle, one can represent each pair of occupied-unoccupied MOs as a linear combination of atomic charges *and* atomic dipoles, for which the ESP fitting might resemble the work by Stern *et al*.[46] This will allow us to: (a) combine a fluctuating-charge-type model and an induced-dipole-type model, which can in principle help improve the PAD model; and (b) build a continuous connection between PAC and PAD models. This has yet to be explored.
- The water solvent molecules are treated using a fixed-charge TIP3P model. We have yet to extend the models to average molecular electrostatic potential (MEP) fitted charges,[100] multipole-preserving charges,[101–103] or multipolar/polarizable representations[35,37,104–107] for the environment atoms.
- A rather small 6-31G* basis is used in this work. In practical calculations, a larger basis set with diffuse functions might be needed to fully describe the polarization effect of the system. While the current procedure for formulating the PAC and PAD response kernels can be applied to larger basis sets without any adaption, its performance there needs to be checked.
- The pseudo-softness and distributed polarizability matrices should also be compared to existing fluctuating charge, induced dipole, and Drude oscillator models. PAC should also be compared against (a) other models for generating the pseudo-softness matrix from Morita and Kato[51,52] and from Yang and coworkers[53] and (b) the chemical potential equalization model from Field and coworkers[11,13].
- All twelve molecules in this work are neutral, and the environment contains only water molecules. While our earlier publication[47] included encouraging results for the luciferin-luciferase complex, which is an anion-protein system, we have yet to apply the new PAC and PAD models to ionic systems, other solvents, and

macromolecular environments, and to compute ligand-receptor binding energies and enzymatic reaction energies.

- The response kernels were formulated for fixed-geometry solute molecules, and therefore only electron degrees of freedom are accounted for. In order to allow some flexibility in the solute geometry, nuclear degrees of freedom need to be added and coupled to the electron degrees of freedom. While there are several options for handling this, it has yet to be explored.
- The QM/MM polarization energy in this work is strictly a second-order function of the external perturbation. In the PAD model, for example, this means that the "induced dipoles" only interact among themselves, and do not interact with "permanent multipoles" on the solute atoms. It is unclear how this might change with flexible-geometry solute molecules.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

YS acknowledges financial support by NIH grant GM096678-02, DOE grant No. DESC0011297, and the University of Oklahoma startup fund. ADM acknowledges partial support from NIH grants GM072558 and GM051501. JH, GK, ACS, FCP and BRB are supported by the Intramural Research Program of the NIH, NHLBI. Computational resources and services used in this work were provided by the LoBoS cluster of the National Institutes of Health, the OU Supercomputing Center for Education and Research, and the Center of Functional Nanomaterials at the Brookhaven National Laboratory supported by DOE OFFce of Science under Contract No. DE-SC0012704. YS thanks Drs. Qiang Cui, Shervin Fatehi, Mark Hybertsen, Toshiko Ichiye, Michael Jones, Daniel Lambrecht, Juyong Lee, Yuezhi Mao, Kwangho Nam, Richard Pastor, Edina Rosta, Lyudmila Slipchenko, Peng Tao, Florentina Tofoleanu, Richard Venable, Feng Wang, and Lee Woodcock for helpful discussions. The authors thank the editor (J. Gao) and reviewers for very helpful comments.

REFERENCES

1. Halgren TA, Damm W. Polarizable force fields. Curr. Op. Struct. Bio. 2001; 11:236–242.
2. Ponder JW, Case DA. Force Fields for Protein Simulations. Adv. Prot. Chem. 2003; 66:27–85.
3. Mackerell AD. Empirical force fields for biological macromolecules: overview and issues. J. Comput. Chem. 2004; 25:1584–1604. [PubMed: 15264253]
4. Demerdash O, Yap E-H, Head-Gordon T. Advanced Potential Energy Surfaces for Condensed Phase Simulation. Ann. Rev. Phys. Chem. 2014; 65:149–174. [PubMed: 24328448]
5. Mortier WJ, Van Genechten K, Gasteiger J. Electronegativity equalization: application and parametrization. J. Am. Chem. Soc. 1985; 107:829–835.
6. Mortier W, Ghosh SK, Shankar S. Electronegativity Equalization Method for the Calculation of Atomic Charges in Molecules. J. Am. Chem. Soc. 1986; 108:4315–4320.
7. Rappe AA, Goddard WA III. Charge equilibration for Molecular Dynamics Simulations. J. Phys. Chem. 1991; 95:3358–3363.
8. Rick SW, Stuart SJ, Berne BJ. Dynamical fluctuating charge force fields: Application to liquid water. J. Chem. Phys. 1994; 101:6141.
9. Rick SW, Berne BJ. Dynamical Fluctuating Charge Force Fields: The Aqueous Solvation of Amides. J. Am. Chem. Soc. 1996; 118:672–679.
10. York DM, Yang W. A chemical potential equalization method for molecular simulations. J. Chem. Phys. 1996; 104:159.
11. Field MJ. Hybrid quantum mechanical/molecular mechanical fluctuating charge models for condensed phase simulations. Mol. Phys. 1997; 91:835–846.

12. Liu Y-P, Kim K, Berne BJ, Friesner RA, Rick SW. Constructing ab initio force fields for molecular dynamics simulations. J. Chem. Phys. 1998; 108:4739.
13. Bret C, Field MJ, Hemmingsen L. A chemical potential equalization model for treating polarization in molecular mechanical force fields. Mol. Phys. 2000; 98:751–763.
14. Chelli R, Procacci P. A transferable polarizable electrostatic force field for molecular mechanics based on the chemical potential equalization principle. J. Chem. Phys. 2002; 117:9175–9189.
15. Patel S, Brooks CL. CHARMM Fluctuating Charge Force Field for Proteins : I Parameterization and Application to Bulk Organic Liquid Simulations. J. Comput. Chem. 2004; 25:1–15. [PubMed: 14634989]
16. Patel S, Mackerell AD, Brooks CL. CHARMM Fluctuating Charge Force Field for Proteins : II Protein/Solvent Properties from Molecular Dynamics Simulations Using a Nonadditive Electrostatic Model. J. Comput. Chem. 2004; 25:1504–1514. [PubMed: 15224394]
17. Chen J, Martínez TJ. QTPIE: Charge transfer with polarization current equalization. A fluctuating charge model with correct asymptotics. Chem. Phys. Lett. 2007; 438:315–320.
18. Chen J, Hundertmark D, Martínez TJ. A unified theoretical framework for fluctuating-charge models in atom-space and in bond-space. J. Chem. Phys. 2008; 129:214113. [PubMed: 19063550]
19. Chen J, Martínez TJ. Charge conservation in electronegativity equalization and its implications for the electrostatic properties of fluctuating-charge models. J. Chem. Phys. 2009; 131:044114. [PubMed: 19655844]
20. Zhong Y, Patel S. Nonadditive empirical force fields for short-chain linear alcohols: Methanol to butanol. Hydration free energetics and kirkwood-buff analysis using charge equilibration models. J. Phys. Chem. B. 2010; 114:11076–11092. [PubMed: 20687517]
21. Lee AJ, Rick SW. The effects of charge transfer on the properties of liquid water. J. Chem. Phys. 2011; 134:184507. [PubMed: 21568521]
22. Soniat M, Rick SW. The effects of charge transfer on the aqueous solvation of ions. J. Chem. Phys. 2012; 137:044511. [PubMed: 22852635]
23. Gao J, Habibollazadeh D, Shao L. A Polarizable Intermolecular Potential Function for Simulation of Liquid Alcohols. J. Phys. Chem. 1995; 99:16460–16467.
24. Gao J, Pavelites JJ, Habibollazadeh D. Simulation of Liquid Amides Using a Polarizable Intermolecular Potential Function. J. Phys. Chem. 1996; 100:2689–2697.
25. Day PN, Jensen JH, Gordon MS, Webb SP, Stevens WJ, Krauss M, Garmer D, Basch H, Cohen D. An effective Fragment Method for Modeling Solvent Effects in Quantum Mechanical Calculations. J. Chem. Phys. 1996; 105:1968–1986.
26. Ren P, Ponder JW. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. J. Phys. Chem. B. 2003; 107:5933–5947.
27. Li H, Netzloff HM, Gordon MS. Gradients of the Polarization Energy in the Effective Fragment Potential Method. J. Chem. Phys. 2006; 125:194103. [PubMed: 17129085]
28. Defusco A, Schofield DP, Jordan KD. Comparison of models with distributed polarizable sites for describing water clusters. Mol. Phys. 2007; 105:2681–2696.
29. Xie W, Pu J, MacKerell AD Jr, Gao J. Development of a Polarizable Intermolecular Potential (PIPF) for Liquid Amides and Alkanes. J. Chem. Theory Comput. 2007; 3:1878–1889. [PubMed: 18958290]
30. Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, DiStasio RA, Head-Gordon M, Clark GNI, Johnson ME, Head-Gordon T. Current Status of the AMOEBA Polarizable Force Field. J. Phys. Chem. B. 2010; 114:2549–2564. [PubMed: 20136072]
31. Kumar R, Wang FF, Jenness GR, Jordan KD. A second generation distributed point polarizable water model. J. Chem. Phys. 2010; 132:014309. [PubMed: 20078163]
32. Ren P, Wu C, Ponder JW. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. J. Chem. Theory Comput. 2011; 7:3143–3161. [PubMed: 22022236]
33. Gordon MS, Fedorov DG, Pruitt SR, Slipchenko LV. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. Chem. Rev. 2012; 112:632–672. [PubMed: 21866983]

34. Kaliman IA, Slipchenko LV. LIBEFP: A New Parallel Implementation of the Effective Fragment Potential Method as a Portable Software Library. J. Comput. Chem. 2013; 34:2284–2292. [PubMed: 24159627]
35. Ghosh D, Kosenkov D, Vanovschi V, Flick J, Kaliman I, Shao Y, Gilbert ATB, Krylov AI, Slipchenko LV. Effective Fragment Potential Method in Q-CHEM: A Guide for Users and Developers. J. Comput. Chem. 2013; 34:1060–1070. [PubMed: 23319180]
36. Wang L-P, Head-Gordon T, Ponder JW, Ren P, Chodera JD, Eastman PK, Martinez TJ, Pande VS. Systematic improvement of a classical molecular model of water. J. Phys. Chem. B. 2013; 117:9956–9972. [PubMed: 23750713]
37. Albaugh A, Boateng HA, Bradshaw RT, Demerdash ON, Dziedzic J, Mao Y, Margul DT, Swails JM, Zeng Q, Case DA, Kenneth P, Essex JW, Head-Gordon M, Pande VS, Ponder JW, Shao Y, Skylaris C-K, Todorov IT, Tuckerman ME, Head-Gordon T. Advanced Potential Energy Surfaces for Molecular Simulation. J. Phys. Chem. B. 2016; 120:9811–9832. [PubMed: 27513316]
38. Straatsma TP, McCammon JA. Molecular Dynamics Simulations with Interaction Potentials Including Polarization Development of a Noniterative Method and Application to Water. Mol. Sim. 1990; 5:181–192.
39. Lamoureux G, MacKerell Jr AD, Roux B. A simple polarizable model of water based on classical Drude oscillators. J. Chem. Phys. 2003; 119:5185–5197.
40. Lamoureux G, Harder E, Vorobyov IV, Roux B, Mackerell AD. A polarizable model of water for molecular dynamics simulations of biomolecules. Chem. Phys. Lett. 2006; 418:245–249.
41. Yu W, Lopes PEM, Roux B, MacKerell AD Jr. Six-site polarizable model of water based on the classical Drude oscillator. J. Chem. Phys. 2013; 138:034508. [PubMed: 23343286]
42. Szklarczyk OM, Bachmann SJ, Van Gunsteren WF. A polarizable empirical force field for molecular dynamics simulation of liquid hydrocarbons. J. Comput. Chem. 2014; 35:789–801. [PubMed: 26248885]
43. Savelyev A, Mackerell AD. All-Atom Polarizable Force Field for DNA Based on the Classical Drude Oscillator Model. J. Comput. Chem. 2014; 35:1219–1239. [PubMed: 24752978]
44. Savelyev A, Mackerell AD. Balancing the Interactions of Ions, Water, and DNA in the Drude Polarizable Force Field. J. Phys. Chem. B. 2015; 118:6742.
45. Lemkul JA, Huang J, Roux B, Mackerell AD. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. Chem. Rev. 2016; 116:4983–5013. [PubMed: 26815602]
46. Stern HA, Kaminski GA, Banks JL, Zhou R, Berne BJ, Friesner RA. Fluctuating Charge, Polarizable Dipole, and Combined Models: Parameterization from ab Initio Quantum Chemistry. J. Phys. Chem. B. 1999; 103:4730–4737.
47. Sodt AJ, Mei Y, König G, Tao P, Steele RP, Brooks BR, Shao Y. Multiple Environment Single System Quantum Mechanical / Molecular Mechanical (MESS-QM/MM) Calculations. 1. Estimation of Polarization Energies. J. Phys. Chem. A. 2015; 119:1511–1523. [PubMed: 25321186]
48. König G, Mei Y, Pickard FC, Simmonett AC, Miller BT, Herbert JM, Woodcock HL, Brooks BR, Shao Y. Computation of Hydration Free Energies Using the Multiple Environment Single System Quantum Mechanical/Molecular Mechanical (MESS-QM/MM) Method. J. Chem. Theory Comput. 2016; 12:332–344. [PubMed: 26613419]
49. Verstraelen T, Vandenbrande S, Ayers PW. Direct computation of parameters for accurate polarizable force fields. J. Chem. Phys. 2014; 141:194114. [PubMed: 25416881]
50. Wang H, Yang W, Wang H, Yang W. Determining Polarizable Force Fields With Electrostatic Potentials from Quantum Mechanical Linear Response Theory. J. Chem. Phys. 2016; 144:224107. [PubMed: 27305996]
51. Morita A, Kato S. Ab Initio Molecular Orbital Theory on Intramolecular Charge Polarization: Effect of Hydrogen Abstraction on the Charge Sensitivity of Aromatic and Nonaromatic Species. J. Am. Chem. Soc. 1997; 119:4021–4032.
52. Morita A, Kato S. The Charge Response Kernel with Modified Electrostatic Potential Charge Model. J. Phys. Chem. A. 2002; 106:3909–3916.

53. Lu Z, Yang W. Reaction path potential for complex systems derived from combined ab initio quantum mechanical and molecular mechanical calculations. J. Chem. Phys. 2004; 121:89–100. [PubMed: 15260525]
54. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 1983; 79:926–936.
55. Misquitta AJ, Stone AJ. Distributed polarizabilities obtained using a constrained density-fitting algorithm. J. Chem. Phys. 2006; 124:024111. [PubMed: 16422575]
56. Madjet ME, Abdurahman A, Renger T. Intermolecular coulomb couplings from ab initio electrostatic potentials: application to optical transitions of strongly coupled pigments in photosynthetic antennae and reaction centers. J. Phys. Chem. B. 2006; 110:17268–17281. [PubMed: 16928026]
57. Le Sueur CR, Stone AJ. Practical schemes for distributed polarizabilities. Mol. Phys. 1993; 78:1267–1291.
58. Angyan JG, Jansen G, Loos M, Hattig C, Hess BA. Distributed polarizabilities using the topological theory of atoms in molecules. Chem. Phys. Lett. 1994; 219:267–273.
59. Williams GJ, Stone AJ. Distributed dispersion: A new approach. J. Chem. Phys. 2003; 119:4620–4628.
60. Gagliardi L, Lindh R, Karlström G. Local properties of quantum chemical systems: The LoProp approach. J. Chem. Phys. 2004; 121:4494–4500. [PubMed: 15332879]
61. Wheatley R, Lillestolen T. Local polarizabilities and dispersion energy coefficients. Mol. Phys. 2008; 106:1545–1556.
62. Soderhjelm P, Krogh JW, Karlström G, Lindh R. Accuracy of Distributed Multipoles and Polarizabilities: Comparison Between the LoProp and MpProp Models. J. Comput. Chem. 2010; 28:1083–1090.
63. Bauer BA, Lucas TR, Krishtal A, Van Alsenoy C, Patel S. Variation of ion polarizability from vacuum to hydration: Insights from Hirshfeld partitioning. J. Phys. Chem. A. 2010; 114:8984–8992. [PubMed: 20684565]
64. Rob F, Szalewicz K. Distributed molecular polarisabilities and asymptotic intermolecular interaction energies. Mol. Phys. 2013; 111:1430–1455.
65. Marenich AV, Cramer CJ, Truhlar DG. Reduced and quenched polarizabilities of interior atoms in molecules. Chem. Sci. 2013; 4:2349.
66. Mei Y, Simmonett AC, Pickard FC, Distasio RA, Brooks BR, Shao Y. Numerical Study on the Partitioning of the Molecular Polarizability into Fluctuating Charge and Induced Atomic Dipole Contributions. J. Phys. Chem. A. 2015; 119:5865–5882. [PubMed: 25945749]
67. Cubero E, Luque FJ, Orozco M, Gao J. Perturbation Approach to Combined QM/MM Simulation of Solute-Solvent Interactions in Solution. J. Phys. Chem. B. 2003; 107:1664–1671.
68. Hu H, Lu Z, Parks JM, Burger SK, Yang W. Quantum mechanics/molecular mechanics minimum free-energy path for accurate reaction energetics in solution and enzymes: sequential sampling and optimization on the potential of mean force surface. J. Chem. Phys. 2008; 128:034105. [PubMed: 18205486]
69. Hu H, Yang W. Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. Ann. Rev. Phys. Chem. 2008; 59:573–601. [PubMed: 18393679]
70. Pulay P, Janowski T. Efficient Calculation of the Energy of a Molecule in an Arbitrary Electric Field. Int J. Quantum Chem. 2009; 109:2113–2120.
71. Janowski T, Wolinski K, Pulay P. Ultrafast Quantum Mechanics/Molecular Mechanics Monte Carlo simulations using generalized multipole polarizabilities. Chem. Phys. Lett. 2012; 530:1–9.
72. Janowski T, Wolinski K, Pulay P. Efficient calculation of the density response function from generalized polarizabilities. Theor. Chem. Acc. 2016; 135:1–6.
73. Gao J, Xia X. A Priori Evaluation of Aqueous Polarization Effects Through Monte- Carlo QM-MM Simulations. Science. 1992; 258:631–635. [PubMed: 1411573]
74. Gao J. Energy components of aqueous solution: insight from hybrid QM/MM simulations using a polarizable solvent model. J. Comput. Chem. 1997; 18:1061–1071.

75. Luque FJ, Orozco M. Polarization effects in generalized molecular interaction potential: New Hamiltonian for reactivity studies and mixed QM/MM calculations. J. Comput. Chem. 1998; 19:866–881.
76. Chipot C, Luque F. Fast evaluation of induction energies: a second-order perturbation theory approach. Chem. Phys. Lett. 2000; 332:190–198.
77. Yang W, Parr RG. Hardness, softness, and the fukui function in the electronic theory of metals and catalysis. Proc. Nat. Acad. Sci. USA. 1985; 82:6723–6726. [PubMed: 3863123]
78. Berkowitz M, Parr RG. Molecular hardness and softness, local hardness and softness, hardness and softness kernels, and relations among these quantities. J. Chem. Phys. 1988; 88:2554–2557.
79. Ayers PW. Strategies for computing chemical reactivity indices. Theor. Chem. Acc. 2001; 106:271–279.
80. Cuevas-Saavedra R, Rabi N, Ayers PW. The unconstrained local hardness: an intriguing quantity, beset by problems. Phys. Chem. Chem. Phys. 2011; 13:19594–19600. [PubMed: 21984043]
81. Chai J-D, Head-Gordon M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. Phys. Chem. Chem. Phys. 2008; 10:6615–6620. [PubMed: 18989472]
82. Krishnan R, Binkley JS, Seeger R, Pople JA. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. J. Chem. Phys. 1980; 72:650–654.
83. Shao Y, Gan Z, Epifanovsky E, Gilbert AT, Wormit M, Kussmann J, Lange AW, Behn A, Deng J, Feng X, Ghosh D, Goldey M, Horn PR, Jacobson LD, Kaliman I, Khaliullin RZ, Kuś T, Landau A, Liu J, Proynov EI, Rhee YM, Richard RM, Rohrdanz MA, Steele RP, Sundstrom EJ, Woodcock HL, Zimmerman PM, Zuev D, Albrecht B, Alguire E, Austin B, Beran GJO, Bernard YA, Berquist E, Brandhorst K, Bravaya KB, Brown ST, Casanova D, Chang C-M, Chen Y, Chien SH, Closser KD, Crittenden DL, Diedenhofen M, DiStasio Ra, Do H, Dutoi AD, Edgar RG, Fatehi S, Fusti-Molnar L, Ghysels A, Golubeva-Zadorozhnaya A, Gomes J, Hanson-Heine MW, Harbach PH, Hauser AW, Hohenstein EG, Holden ZC, Jagau T-C, Ji H, Kaduk B, Khistyaev K, Kim J, Kim J, King RA, Klunzinger P, Kosenkov D, Kowalczyk T, Krauter CM, Lao KU, Laurent A, Lawler KV, Levchenko SV, Lin CY, Liu F, Livshits E, Lochan RC, Luenser A, Manohar P, Manzer SF, Mao S-P, Mardirossian N, Marenich AV, Maurer SA, Mayhall NJ, Neuscamman E, Oana CM, Olivares-Amaya R, O'Neill DP, Parkhill JA, Perrine TM, Peverati R, Prociuk A, Rehn DR, Rosta E, Russ NJ, Sharada SM, Sharma S, Small DW, Sodt A, Stein T, Stuck D, Su Y-C, Thom AJ, Tsuchimochi T, Vanovschi V, Vogt L, Vydrov O, Wang T, Watson MA, Wenzel J, White A, Williams CF, Yang J, Yeganeh S, Yost SR, You Z-Q, Zhang IY, Zhang X, Zhao Y, Brooks BR, Chan GK, Chipman DM, Cramer CJ, Goddard WA, Gordon MS, Hehre WJ, Klamt A, Schaefer HF, Schmidt MW, Sherrill CD, Truhlar DG, Warshel A, Xu X, Aspuru-Guzik A, Baer R, Bell AT, Besley NA, Chai J-D, Dreuw A, Dunietz BD, Furlani TR, Gwaltney SR, Hsu C-P, Jung Y, Kong J, Lambrecht DS, Liang W, Ochsenfeld C, Rassolov VA, Slipchenko LV, Subotnik JE, Van Voorhis T, Herbert JM, Krylov AI, Gill PM, Head-Gordon M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. Mol. Phys. 2015; 113:184–215.
84. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. CHARMM : The Biomolecular Simulation Program. J. Comput. Chem. 2009; 30:1564–1614.
85. Becke AD. Density-functional exchange-energy approximation with correct asymptotic behavior. Phys. Rev. A. 1988; 38:3098–3100.
86. Becke AD. A new mixing of Hartree-Fock and local density functional theories. J. Chem. Phys. 1993; 98:1372.
87. Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B. 1988; 37:785–789.
88. Seeger R, Pople JA. Self-consistent molecular orbital methods. XVIII. Constraints and stability in Hartree-Fock theory. J. Chem. Phys. 1977; 66:3045–3050.
89. Singh UC, Kollman PA. An approach to computing electrostatic charges for molecules. J. Comput. Chem. 1984; 5:129–145.
90. Besler BH, Merz KM, Kollman PA. Atomic charges derived from semiempirical methods. J. Comput. Chem. 1990; 11:431–439.

91. König G, Boresch S. Non-Boltzmann Sampling and Bennett's Acceptance Ratio Method: How to Profit from Bending the Rules. J. Comput. Chem. 2011; 32:1082–1090. [PubMed: 21387335]
92. König G, Hudson PS, Boresch S, Woodcock HL. Multiscale Free Energy Simulations: An Efficient Method for Connecting Classical MD Simulations to QM or QM/MM Free Energies Using Non-Boltzmann Bennett Reweighting Schemes. J. Chem. Theory Comput. 2014; 10:1406–1419. [PubMed: 24803863]
93. Simmonett AC, Pickard FC, Shao Y, Cheatham TE, Brooks BR. Efficient treatment of induced dipoles. J. Chem. Phys. 2015; 143:074115. [PubMed: 26298123]
94. Shaw KE, Woods CJ, Mulholland AJ. Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. J. Phys. Chem. Lett. 2010; 1:219–223.
95. Ren P, Ponder JW. Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations. J. Comput. Chem. 2002; 23:1497–1506. [PubMed: 12395419]
96. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. J. Phys. Chem. 1993; 97:10269–10280.
97. Stouch TR, Williams DE. Conformational dependence of electrostatic potential-derived charges: Studies of the fitting procedure. J. Comput. Chem. 1993; 14:858–866.
98. Hu H, Lu Z, Yang W. Fitting molecular electrostatic potentials from quantum mechanical calculations. J. Chem. Theory Comput. 2007; 3:1004–1013. [PubMed: 26627419]
99. Zeng J, Duan L, Zhang JZH, Mei Y. A numerically stable restrained electrostatic potential charge fitting method. J. Comput. Chem. 2013; 34:847–853. [PubMed: 23281029]
100. Gao J, Luque FJ, Orozco M. Induced dipole moment and atomic charges based on average electrostatic potentials in aqueous solution. J. Chem. Phys. 1993; 98:2975–2982.
101. Thole BT, van Duijnen PT. A general population analysis preserving the dipole moment. Theor. Chim. Acta. 1983; 63:209–221.
102. Swart M, Van Duijnen PT, Snijders JG. A Charge Analysis Derived from an Atomic Multipole Expansion. J. Comput. Chem. 2001; 22:79–88.
103. Zhang P, Bao P, Gao J. Dipole Preserving and Polarization Consistent Charges. J. Comput. Chem. 2011; 32:2127–2139. [PubMed: 21541954]
104. Cisneros GA, Karttunen M, Ren P, Sagui C. Classical Electrostatics for Biomolecular Simulations. Chem. Rev. 2014; 114:779. [PubMed: 23981057]
105. Loco D, Polack E, Caprasecca S, Lagardere L, Lipparini F, Piquemal JP, Mennucci B. A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations. J. Chem. Theory Comput. 2016; 12:3654–3661. [PubMed: 27340904]
106. Kratz EG, Walker AR, Lagardère L, Lipparini F, Piquemal JP, Cisneros AG. LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields. J. Comput. Chem. 2016; 37:1019–1029. [PubMed: 26781073]
107. Dziedzic J, Mao Y, Shao Y, Ponder J, Head-Gordon T, Head-Gordon M, Skylaris C-K. TINKTEP: A fully self-consistent, mutually polarizable QM/MM approach based on the AMOEBA force field. J. Chem. Phys. 2016; 145:124106. [PubMed: 27782640]

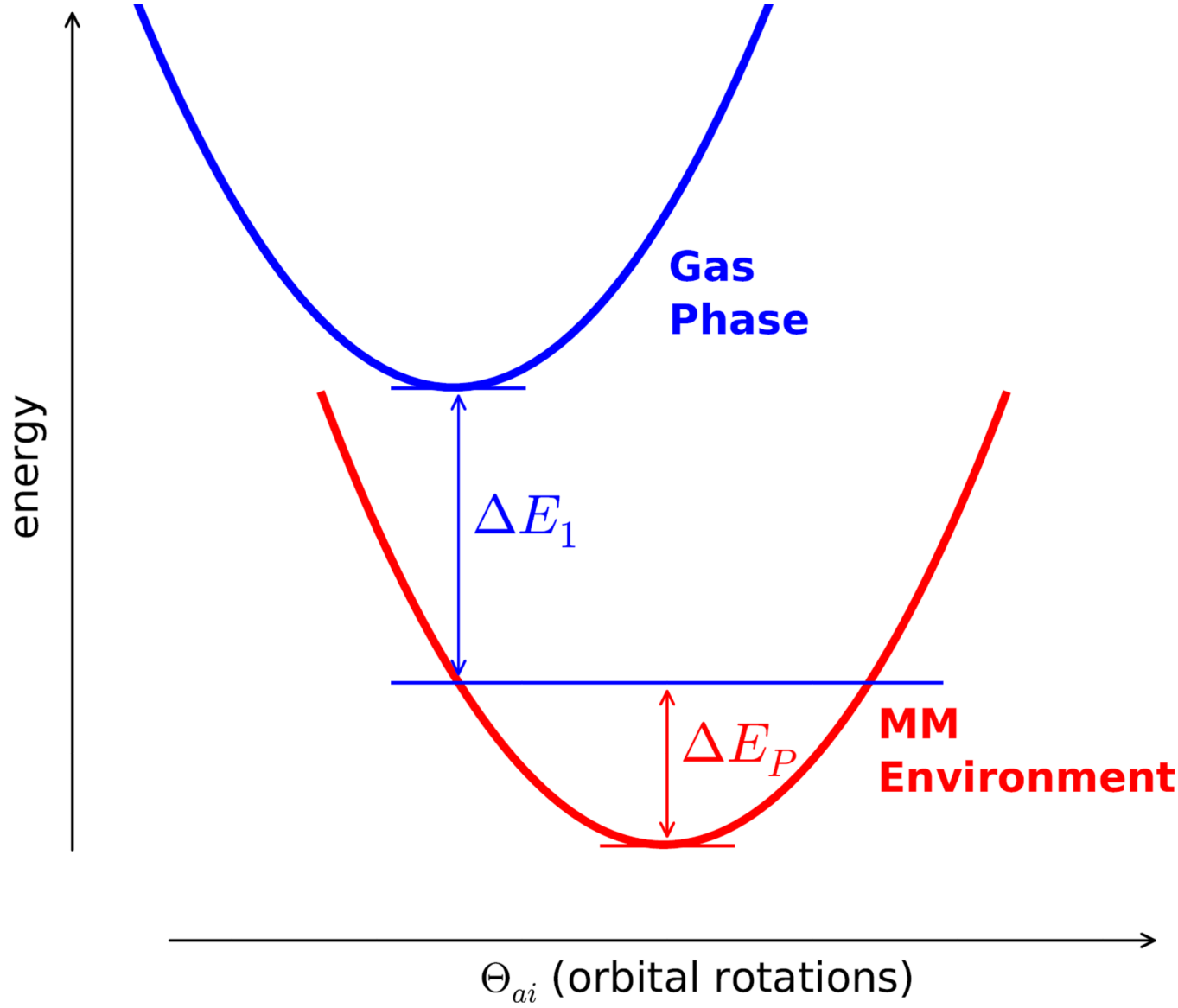


FIG. 1.

A schematic illustration for two QM/MM energy components. ΔE_1, the QM/MM permanent electrostatics energy, can be either negative or positive (see Fig.S4). ΔE_P, the polarization energy, is always negative and arises from relaxation of the occupied molecular orbitals of the QM subsystem from the gas-phase to the MM environment.

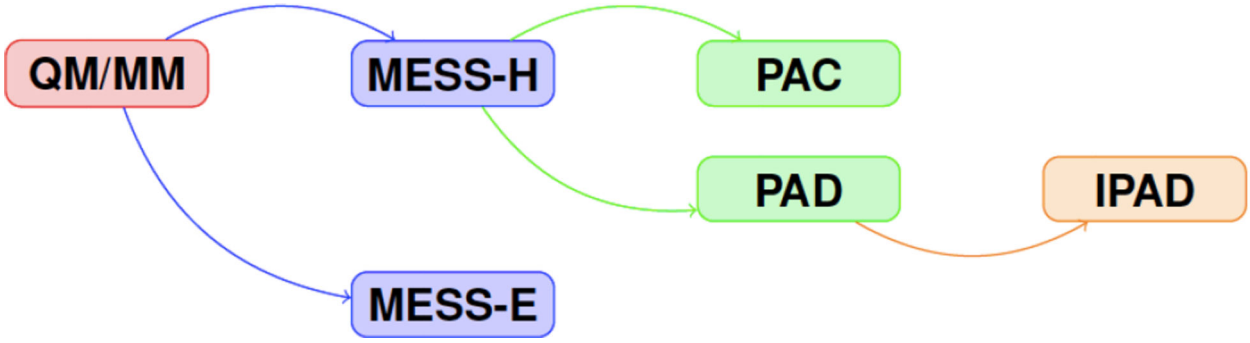


FIG. 2.
Various computational models for computing the QM/MM polarization energy

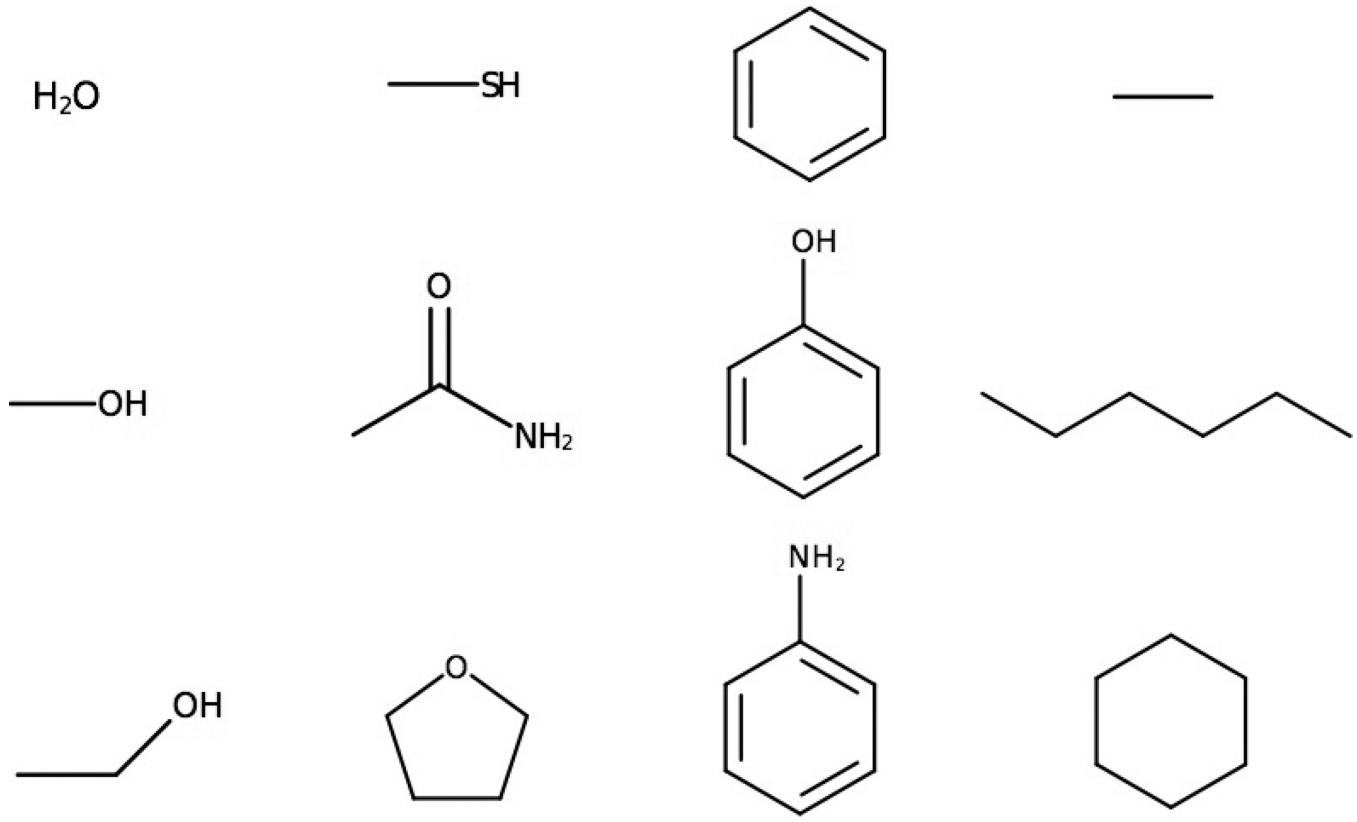


FIG. 3.
The twelve solute molecules considered in this work.

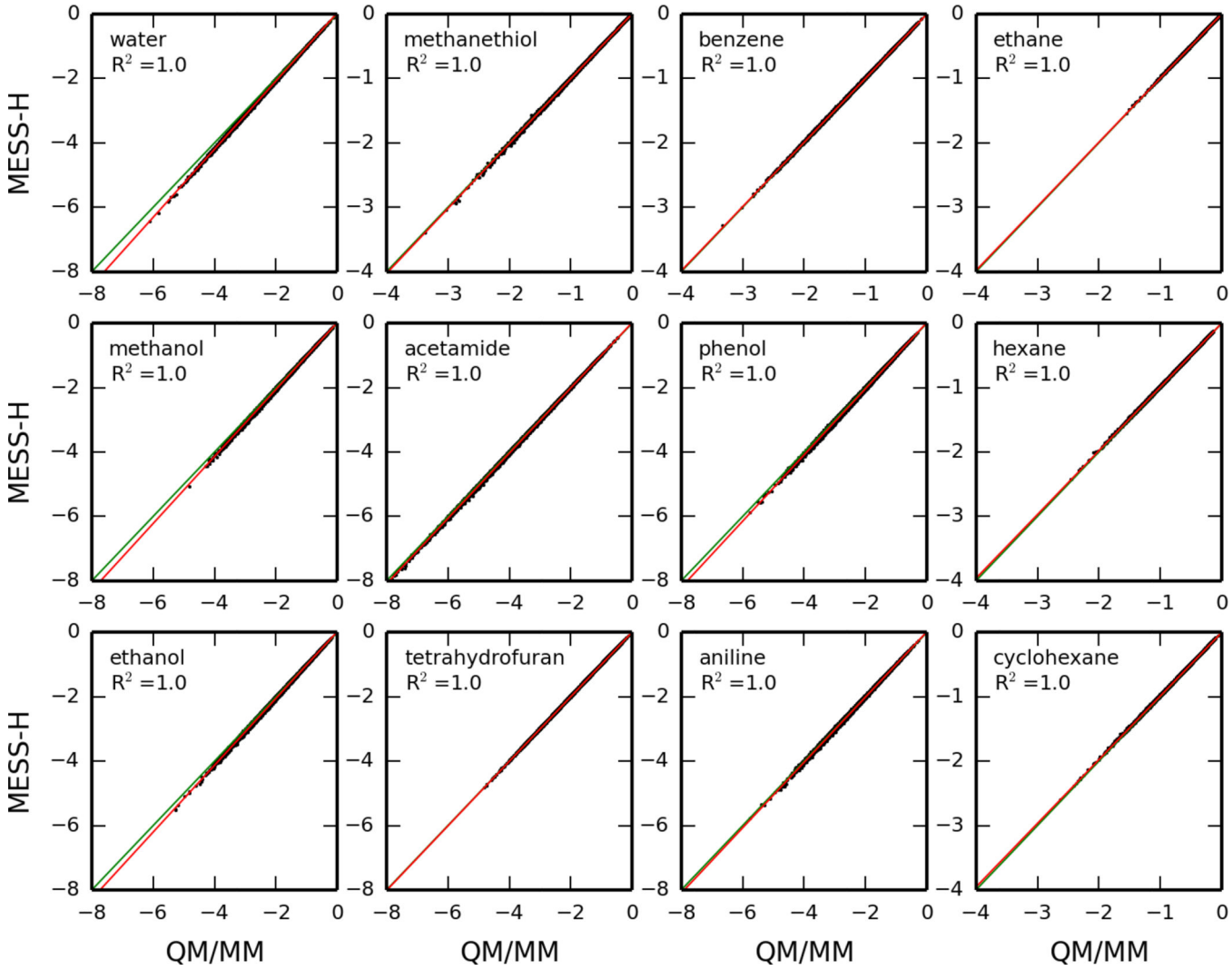


FIG. 4.

Estimated QM/MM polarization energies (in kcal/mol) using the MESS-H model for 10000 configurations of twelve small molecules solvated in a box of water molecules, as compared against fully-converged QM/MM values.

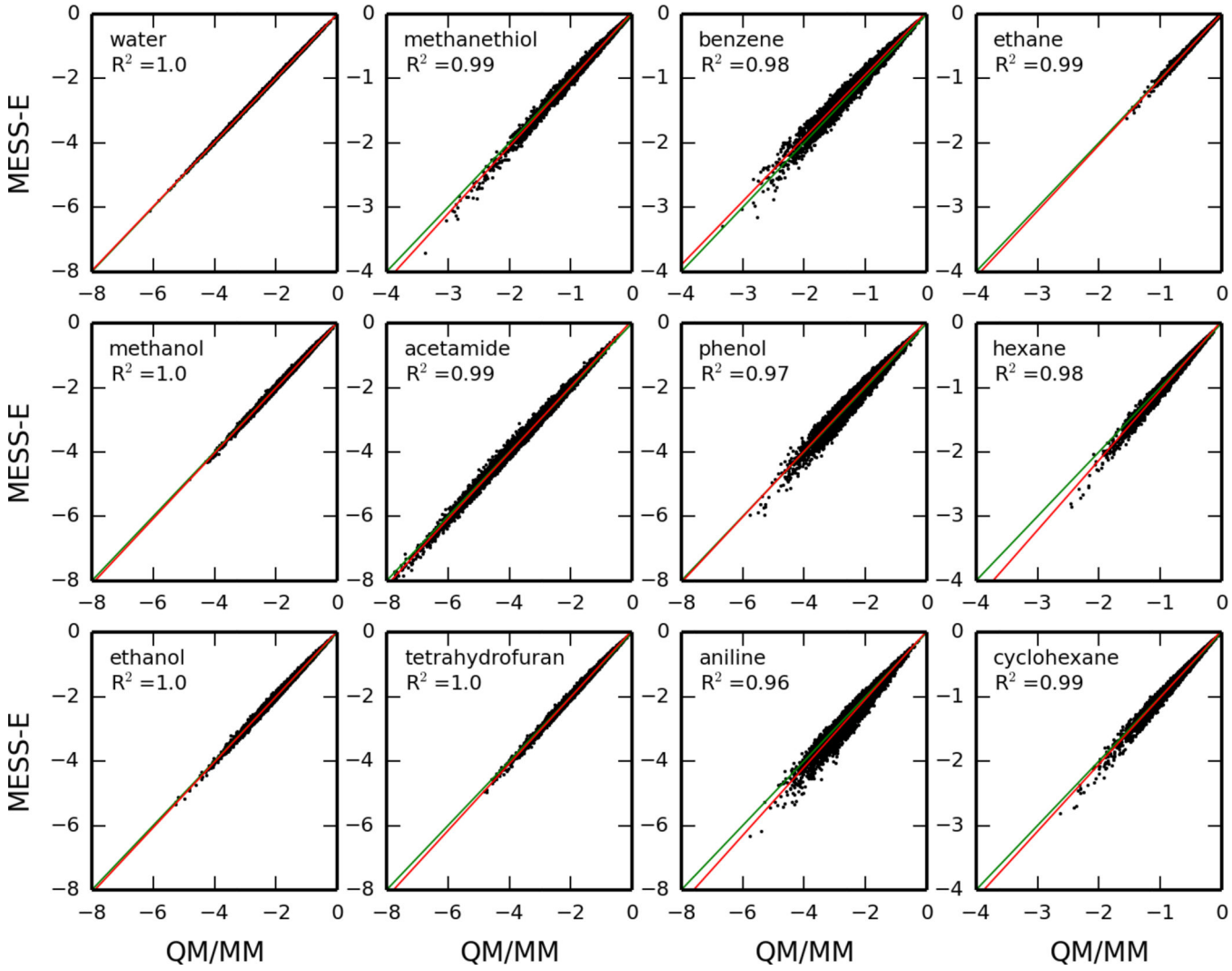


FIG. 5.

Estimated QM/MM polarization energies (in kcal/mol) using the MESS-E model for 10000 configurations of twelve small molecules solvated in a box of water molecules, as compared against fully-converged QM/MM values.

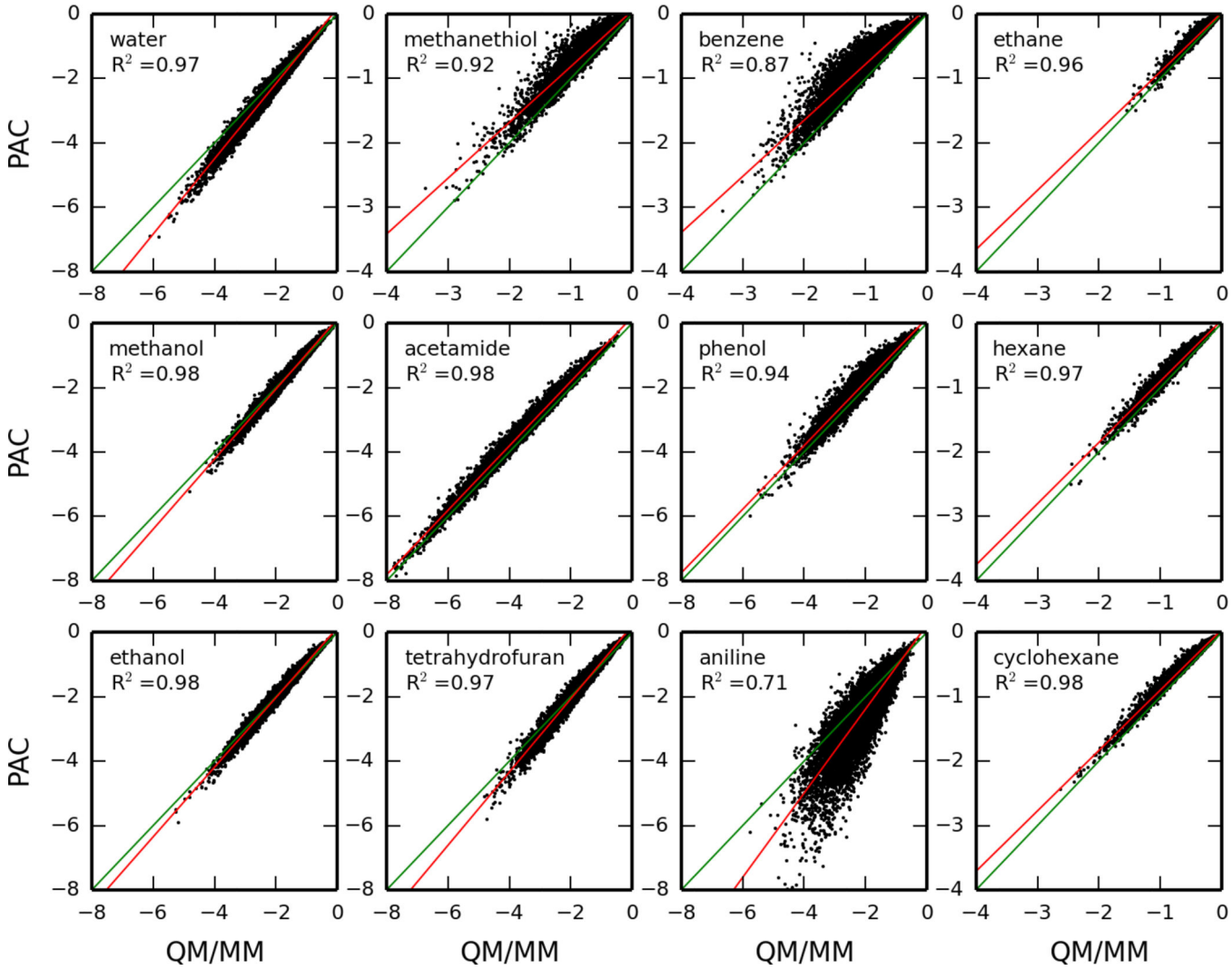


FIG. 6.

Estimated QM/MM polarization energies (in kcal/mol) using the PAC model for 10000 configurations of twelve small molecules solvated in a box of water molecules, as compared against fully-converged QM/MM values.

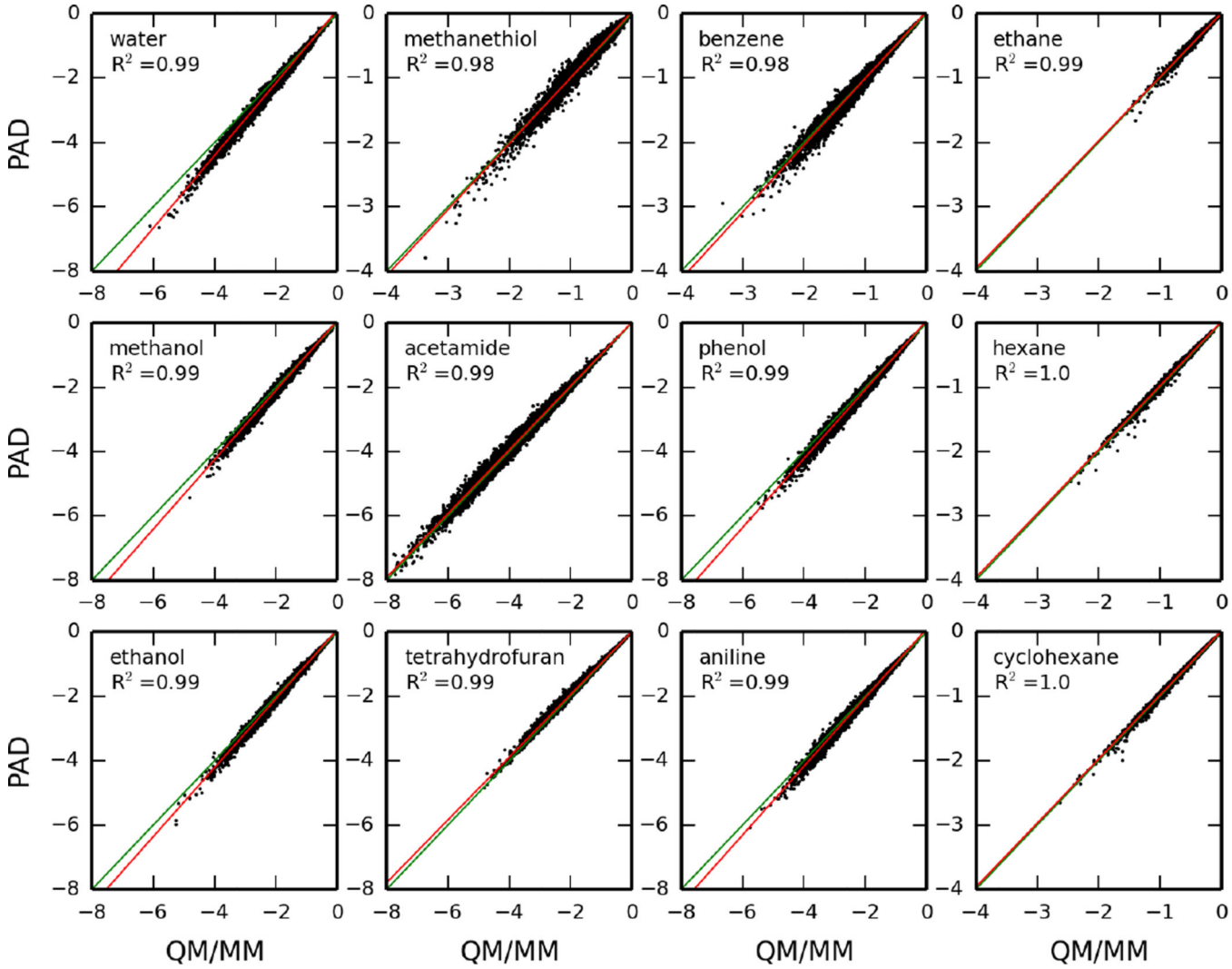


FIG. 7.

Estimated QM/MM polarization energies (in kcal/mol) using the PAD model for 10000 configurations of twelve small molecules solvated in a box of water molecules, as compared against fully-converged QM/MM values.

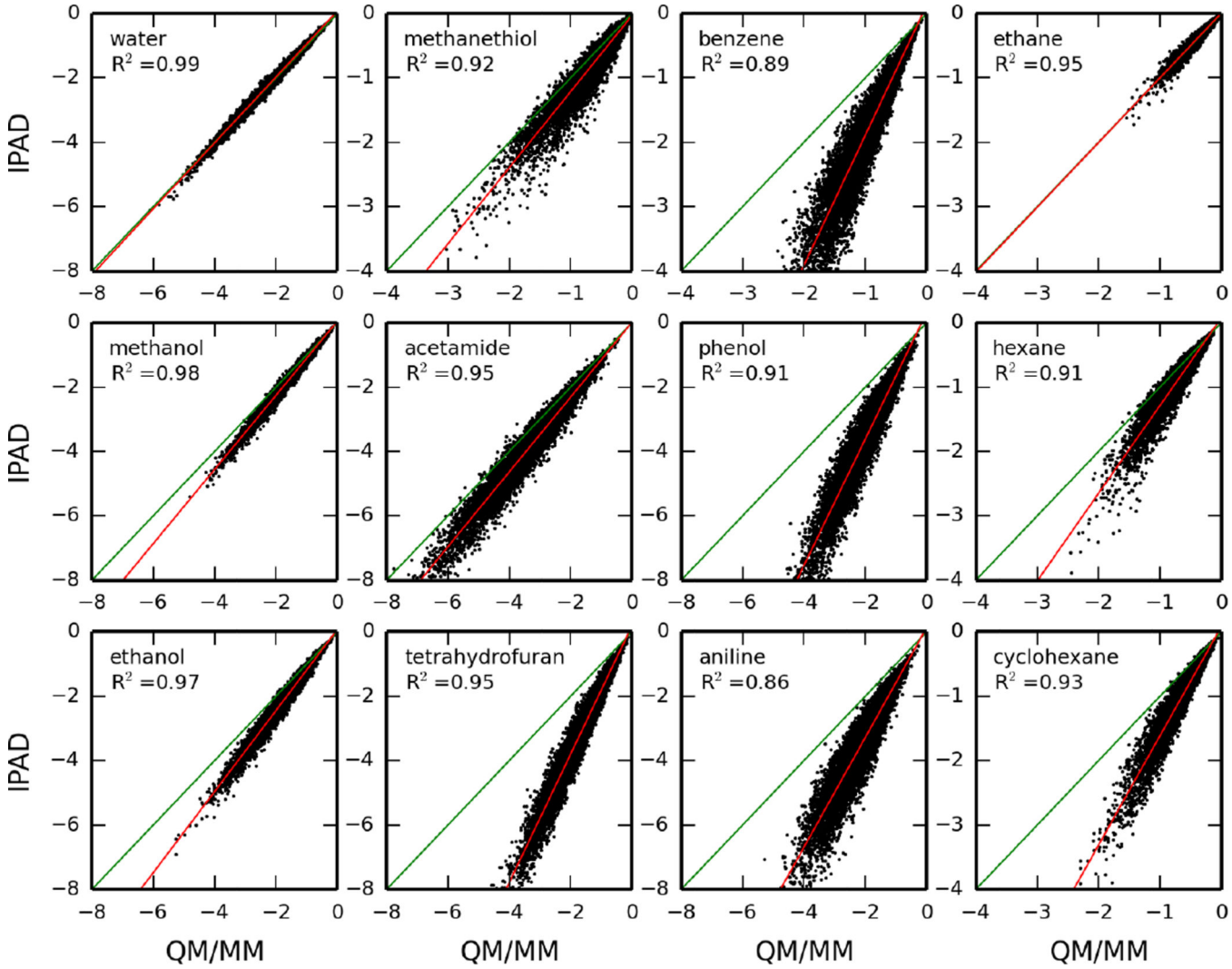


FIG. 8.

Estimated QM/MM polarization energies (in kcal/mol) using the IPAD model for 10000 configurations of twelve small molecules solvated in a box of water molecules, as compared against fully-converged QM/MM values.

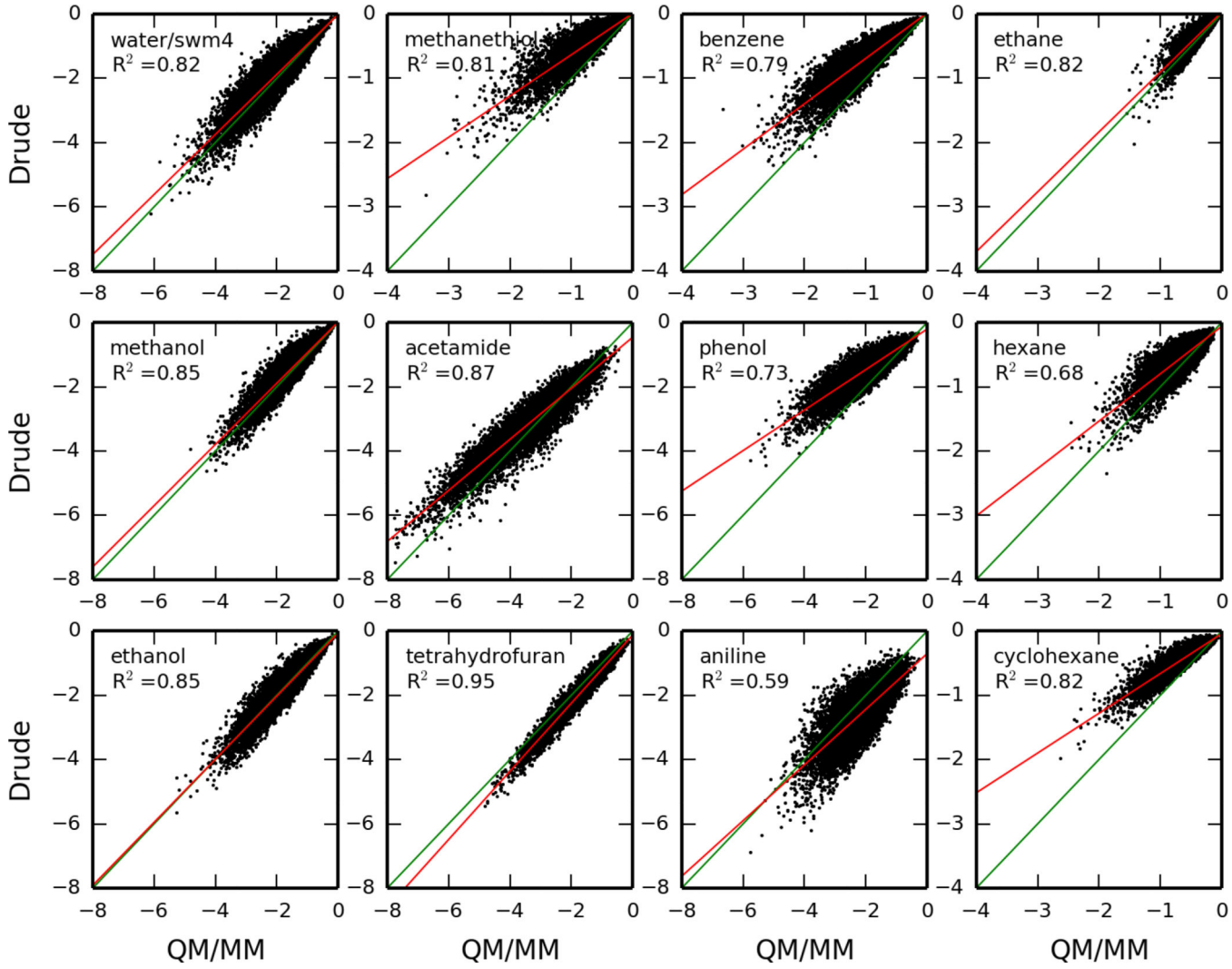


FIG. 9.

Solute polarization energies (in kcal/mol) using the DRUDE model for 10000 configurations of twelve small molecules solvated in a box of water molecules, as compared against fully-converged QM/MM values.

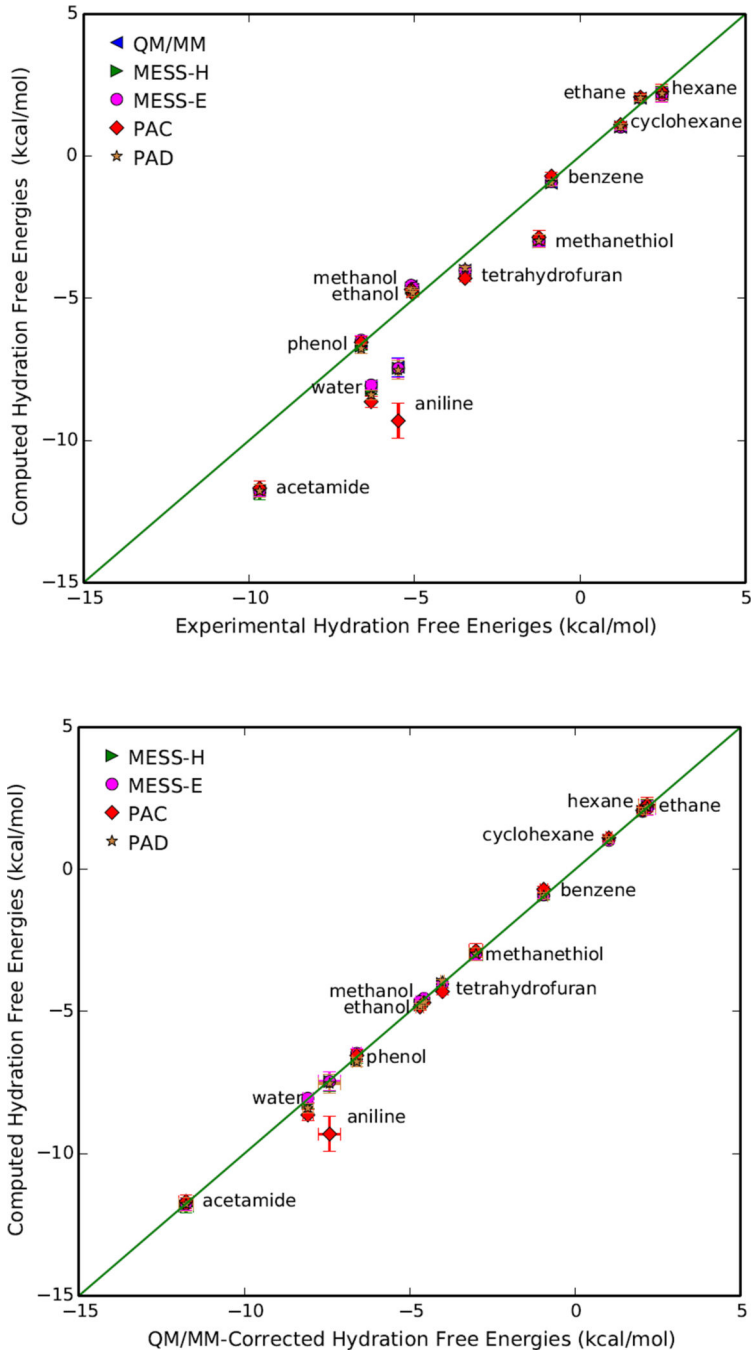


FIG. 10.

Experimental versus computed hydration free energies of the twelve solutes, using various methods for the computation of the polarization energy. Error bars represent standard deviations.

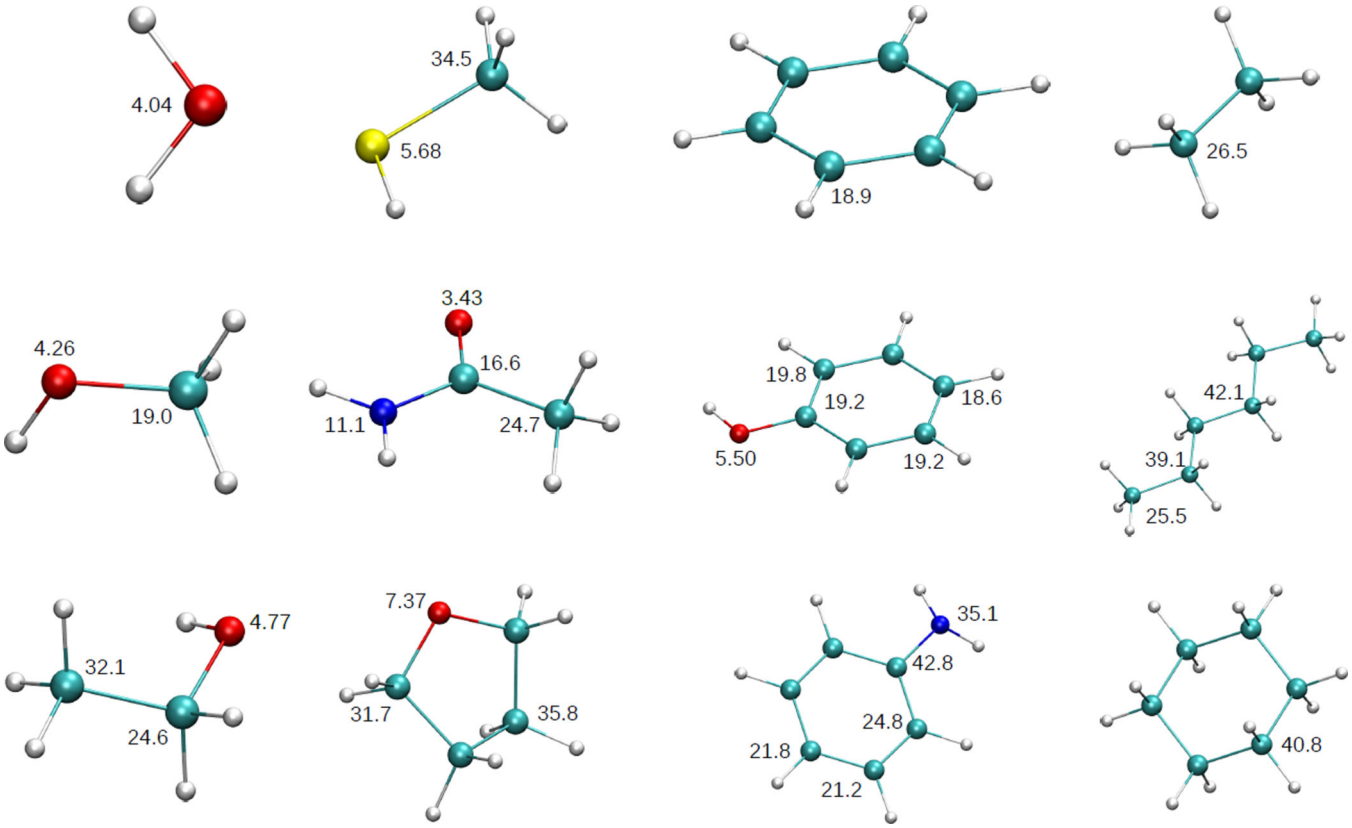


FIG. 11.

Computed diagonal elements of the pseudo-softness matrix (in a.u.) for twelve small molecules.

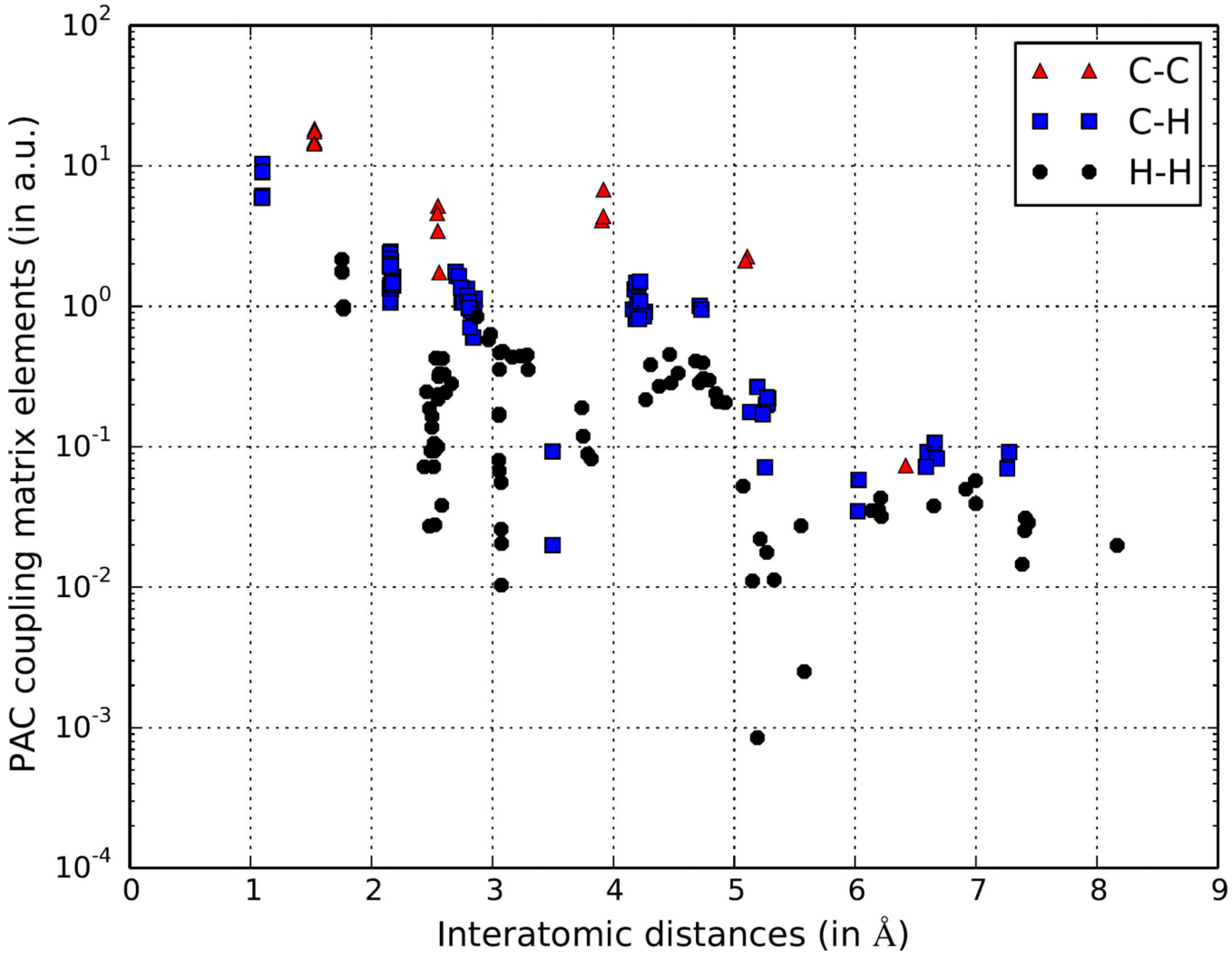


FIG. 12.

Decay of the off-diagonal pseudo-softness matrix elements with distance for the *n*-hexane molecule.

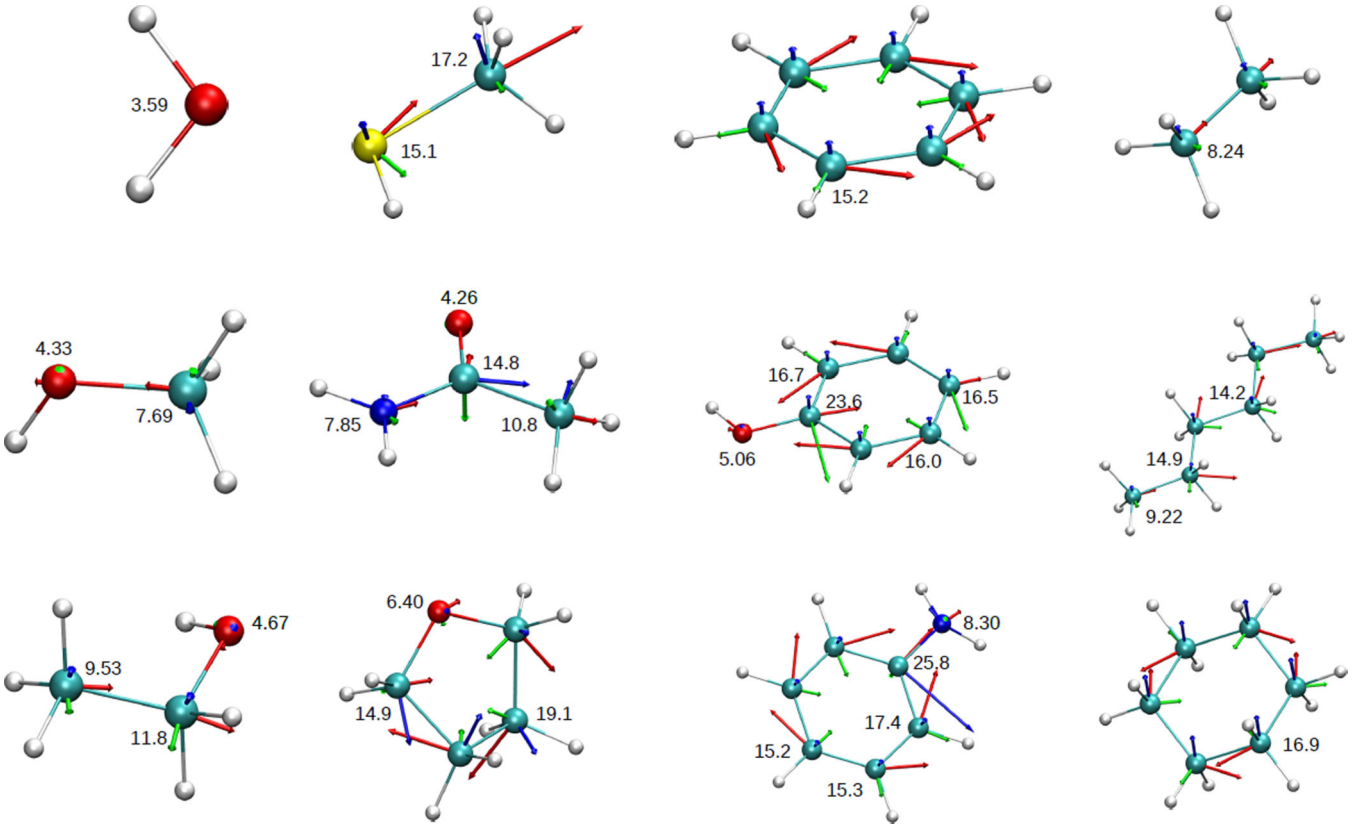


FIG. 13.

Computed atomic polarizabilities (the diagonal blocks of the distributed polarizability matrix, in a.u.) for twelve small molecules. The three arrows on each atom represent the direction of three eigenvectors, and the length of each vector is proportional to the eigenvalues. The scalar value (the average of the three eigenvalues) are shown next to each non-hydrogen atom.

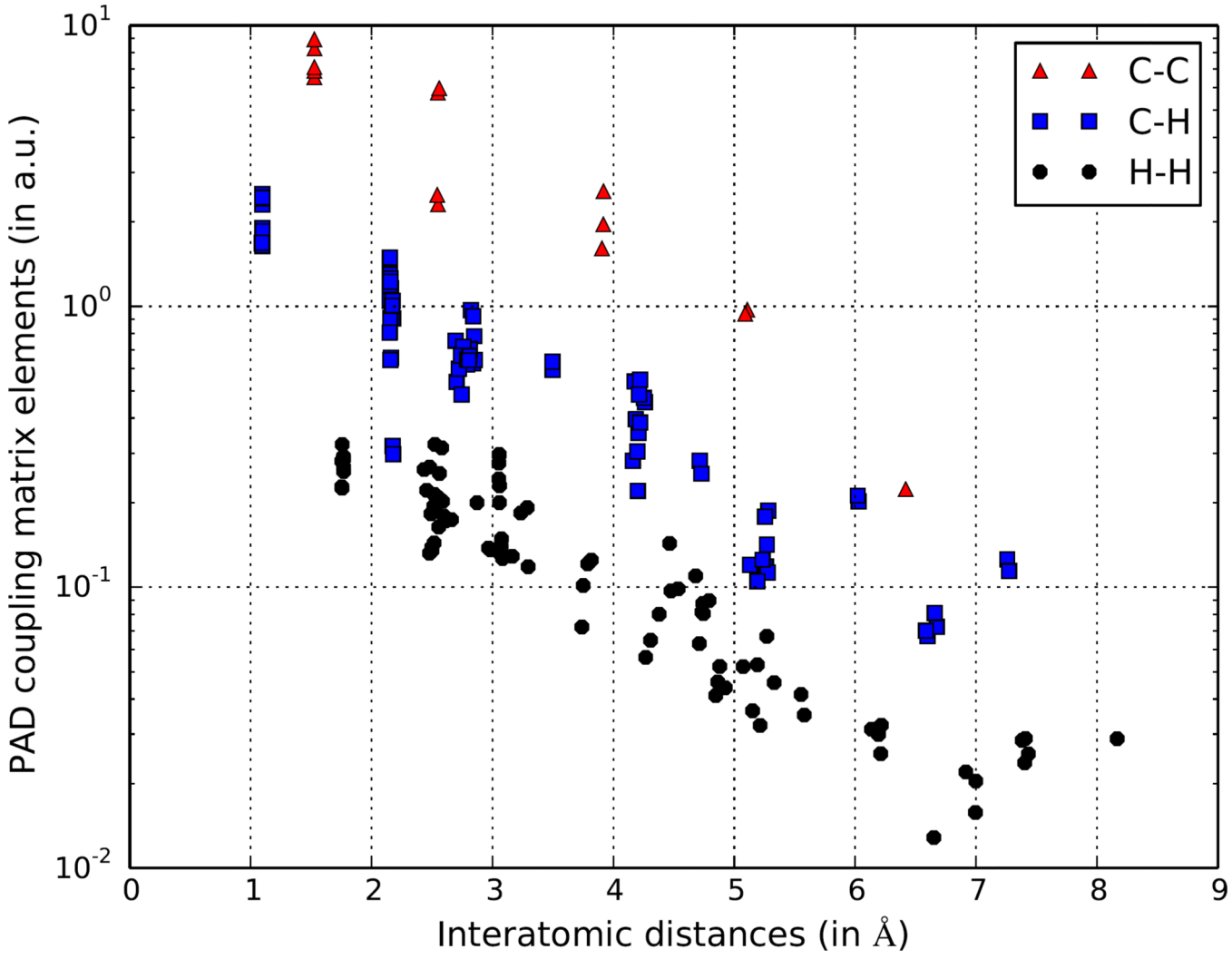


FIG. 14.

Decay of the 2-norm of the off-diagonal blocks in the distributed polarizability matrix with distance for the *n*-hexane molecule.

TABLE I

Mean signed deviations (MSD) and root-mean square deviations (RMSD) of the approximate QM/MM polarization energy using five computational models against fully-converged QM/MM values. All deviations are shown in kcal/mol.

molecule		MESS-H	MESS-E	PAC	PAD	IPAD	Drude
water	MSD	−0.087	0.014	−0.184	−0.165	0.025	0.131
	RMSD	0.1	0.022	0.275	0.216	0.089	0.373
methanol	MSD	−0.041	−0.004	−0.011	−0.058	−0.196	0.092
	RMSD	0.053	0.038	0.122	0.109	0.24	0.28
ethanol	MSD	−0.042	0.007	−0.006	−0.055	−0.387	−0.031
	RMSD	0.054	0.049	0.123	0.104	0.451	0.282
methanethiol	MSD	−0.002	−0.0	0.137	0.015	−0.149	0.24
	RMSD	0.008	0.041	0.184	0.062	0.224	0.311
acetamide	MSD	−0.029	−0.007	0.188	0.05	−0.533	0.233
	RMSD	0.039	0.101	0.237	0.128	0.64	0.483
tetrahydrofuran	MSD	0.001	−0.001	−0.029	0.054	−1.527	−0.23
	RMSD	0.007	0.055	0.177	0.076	1.717	0.29
benzene	MSD	0.003	0.061	0.219	−0.003	−0.875	0.305
	RMSD	0.006	0.088	0.268	0.061	1.025	0.365
phenol	MSD	−0.036	0.058	0.175	−0.077	−1.666	0.556
	RMSD	0.049	0.151	0.258	0.13	1.871	0.681
aniline	MSD	−0.012	−0.082	−0.433	−0.049	−1.395	−0.413
	RMSD	0.027	0.181	0.756	0.109	1.554	0.658
ethane	MSD	0.001	0.002	0.045	0.009	0.005	0.047
	RMSD	0.002	0.015	0.059	0.017	0.04	0.092
hexane	MSD	0.006	−0.02	0.083	0.021	−0.174	0.122
	RMSD	0.008	0.052	0.098	0.028	0.243	0.21
cyclohexane	MSD	0.006	0.006	0.079	0.018	−0.348	0.194
	RMSD	0.008	0.039	0.092	0.024	0.432	0.242

TABLE II

NBB-corrected hydration free energy (in kcal/mol) of twelve small solute molecules computed using different QM/MM models and using the PAC and PAD models.

molecule	Expt	QM/MM	MESS-H	MESS-E	PAC	PAD
water	−6.31	−8.08 ± 0.15	−8.24 ± 0.16	−8.05 ± 0.15	−8.64 ± 0.19	−8.4 ± 0.17
methanol	−5.1	−4.57 ± 0.07	−4.64 ± 0.07	−4.54 ± 0.07	−4.69 ± 0.07	−4.69 ± 0.07
ethanol	−5.05	−4.69 ± 0.11	−4.77 ± 0.11	−4.65 ± 0.11	−4.84 ± 0.12	−4.81 ± 0.11
methanethiol	−1.24	−3.0 ± 0.2	−3.01 ± 0.21	−2.98 ± 0.21	−2.86 ± 0.24	−2.98 ± 0.23
acetamide	−9.68	−11.76 ± 0.22	−11.85 ± 0.23	−11.77 ± 0.21	−11.68 ± 0.24	−11.78 ± 0.23
tetrahydrofuran	−3.47	−4.01 ± 0.06	−4.02 ± 0.06	−4.02 ± 0.06	−4.3 ± 0.12	−3.94 ± 0.06
benzene	−0.86	−0.95 ± 0.12	−0.94 ± 0.12	−0.9 ± 0.12	−0.7 ± 0.12	−0.93 ± 0.12
phenol	−6.61	−6.6 ± 0.14	−6.68 ± 0.15	−6.47 ± 0.15	−6.55 ± 0.2	−6.77 ± 0.16
aniline	−5.49	−7.42 ± 0.33	−7.46 ± 0.34	−7.46 ± 0.33	−9.31 ± 0.61	−7.53 ± 0.33
ethane	1.83	2.04 ± 0.12	2.04 ± 0.12	2.04 ± 0.12	2.08 ± 0.12	2.04 ± 0.12
hexane	2.48	2.18 ± 0.26	2.19 ± 0.26	2.15 ± 0.26	2.27 ± 0.26	2.2 ± 0.26
cyclohexane	1.23	1.02 ± 0.1	1.03 ± 0.1	1.02 ± 0.1	1.11 ± 0.1	1.04 ± 0.1
MSD (wrt Expt)		−0.63	−0.67	−0.61	−0.82	−0.69
RMSD (wrt Expt)		1.12	1.16	1.13	1.52	1.18
R^2 (wrt Expt)		0.96	0.96	0.96	0.93	0.96
MSD (wrt QM/MM)			−0.04	0.02	−0.19	−0.06
RMSD (wrt QM/MM)			0.07	0.05	0.59	0.12
R^2 (wrt QM/MM)			1.0	1.0	0.99	1.0