Nuclear-Electronic Orbital QM/MM Approach: Geometry Optimizations and Molecular Dynamics

Mathew Chow[†], Eleftherios Lambros[‡], Xiaosong Li[‡], and Sharon Hammes-Schiffer*[, †]

[†] *Department of Chemistry, Yale University, New Haven, CT 06520, United States*

[‡] *Department of Chemistry, University of Washington, Seattle, WA 98195, United States*

E-mail: sharon.hammes-schiffer@yale.edu

Abstract

Hybrid quantum mechanical/molecular mechanical (QM/MM) methods allow simulations of chemical reactions in atomistic solvent and heterogeneous environments such as proteins. Herein the nuclear-electronic orbital (NEO) QM/MM approach is introduced to enable the quantization of specified nuclei, typically protons, in the QM region using a method such as NEO density functional theory (NEO-DFT). This approach includes proton delocalization, polarization, anharmonicity, and zero-point energy in geometry optimizations and dynamics. Expressions for the energies and analytical gradients associated with the NEO-QM/MM method, as well as the previously developed polarizable continuum model (NEO-PCM), are provided. Geometry optimizations of small organic molecules hydrogen bonded to water in either dielectric continuum solvent or explicit atomistic solvent illustrate that aqueous solvation can strengthen hydrogen-bonding interactions for the systems studied, as indicated by shorter intermolecular distances at the hydrogen bond interface. We then performed a real-time direct dynamics simulation of a phenol molecule in explicit water using the NEO-QM/MM method. These developments and initial examples provide the foundation for future studies of nuclear-electronic quantum dynamics in complex chemical and biological environments.

Introduction

The nuclear-electronic orbital (NEO) approach is a multicomponent quantum chemistry framework for studying coupled nuclear-electronic quantum effects.[1-3] Nuclear quantum effects and non-Born-Oppenheimer effects are incorporated through the quantum mechanical treatment of select nuclei, typically protons, at the same level as the electrons. NEO methods can therefore be useful in many biological and chemical applications for which nuclear delocalization, hydrogen tunneling, and nonadiabatic effects are important.[4-6] Inclusion of the solvent or protein environment in a computationally practical manner is important for describing such processes.

A variety of ground and excited state wave function[7-9] and density functional theory (DFT)[10-13] methods have been developed within the NEO framework. Compared to their purely electronic counterparts, NEO methods allow for the inclusion of important nuclear quantum effects, such as hydrogen tunneling, zero-point energy, and quantized vibrational states, directly into quantum chemistry calculations and molecular dynamics (MD) simulations. The exploration of coupled nuclear-electronic dynamics is an especially promising direction within the NEO framework, as exemplified by the NEO real-time time-dependent density functional theory (RT-TDDFT) method[14] and NEO-Ehrenfest dynamics approach.[15] Recently, implicit solvation effects have been included in NEO calculations[16, 17] through the polarizable continuum model (PCM).[18, 19] This NEO-PCM approach has been applied to a variety of chemical systems and has been used to simulate the dynamics of an excited state intramolecular proton transfer reaction in solution.[16]

In this paper, we build from these developments by implementing NEO-PCM analytical gradients, as well as introducing the NEO quantum mechanical/molecular mechanical (QM/MM)[20, 21] approach for including an explicit solvent or other type of chemical environment. Various implicit and explicit solvation models, as well as QM/MM approaches, have been

developed and applied to chemical and biological systems.[22-32] These approaches can be extended and generalized to the NEO framework to allow the description of quantized nuclei and non-Born-Oppenheimer effects in chemical and biological environments. After implementing these methods, we use both the NEO-PCM and NEO-QM/MM approaches to optimize geometries of small organic molecules hydrogen bonded to water. These calculations allow us to investigate the impact of proton quantization and implicit or explicit solvation on the geometries of the hydrogen-bonded interfaces. We also perform a NEO-QM/MM MD simulation of phenol in water. These initial applications pave the way for future simulations that describe nuclear-electronic quantum dynamics in complex environments.

Methods

NEO-PCM

Implicit solvation models such as PCM have been used extensively in quantum chemistry calculations and are frequently used to study spectroscopic phenomena.[33-35] In implicit solvation, a solute molecule is placed within a molecular-shaped cavity constructed from interlocking atom-centered spheres. The cavity surface serves as the boundary separating the solute and the continuum solvent regions, and the solvent polarization manifests as a charge density $\sigma(\mathbf{r})$ evaluated on the surface of the cavity. In PCM, the charge density is represented over a set of N_{tess} tesserae centered on grid points with spatial coordinates $\mathbf{s}_k$. The surface charge per unit area is given as

$$\sigma(\mathbf{r}) = \sum_{k}^{N_{\text{tess}}} \frac{q(\mathbf{s}_k)}{a_k} \phi_k\left(\mathbf{r}; \mathbf{s}_k, \zeta_k\right) \tag{1}$$

where $q(\mathbf{s}_k)$ is the apparent surface charge (ASC) and a_k is the surface area of the k^{th} tessera. Here $\phi_k\left(\mathbf{r};\mathbf{s}_k,\zeta_k\right)$ is a weighting function that is typically unity or a Gaussian function with exponent ζ_k,[27, 36] as discussed further below.

The charges $q(\mathbf{s}_k)$ are obtained by solving the matrix equation

$$\mathbf{Kq} = \mathbf{Rv} \tag{2}$$

at each self-consistent field step for Hartree-Fock (HF) or DFT calculations. Here, $\mathbf{q}$ and $\mathbf{v}$ are vectors containing the ASCs and the solute electrostatic potential, respectively, evaluated at tesserae grid points. The matrices $\mathbf{K}$ and $\mathbf{R}$ are defined according to the specific PCM formulation employed and are described in detail elsewhere.[27, 32] In this paper, we focus on the conductor-like PCM (C-PCM)[37] approach because of its computational efficiency, but the extension to various other PCM schemes is straighforward.

For NEO-HF (or NEO-DFT) calculations, the nuclear-electronic wave function (or reference system) is represented as the product of an electronic and protonic Slater determinant Φ^{e} and Φ^{p} composed of electronic and protonic orbitals, respectively:

$$\Psi_{\text{NEO}} = \Phi^{\text{e}}\Phi^{\text{p}} \tag{3}$$

For simplicity, throughout this paper we will refer to Ψ_{NEO} as a wave function, although it is rigorously the Kohn-Sham reference system for NEO-DFT. Solution of the mixed nuclear-electronic time-independent Schrödinger equation requires that the coupled electronic and protonic Hartree-Fock-Roothaan or Kohn-Sham equations be solved self-consistently to obtain the spatial orbitals and NEO energy:

$$\begin{aligned}\mathbf{F}^{\text{e}}\mathbf{C}^{\text{e}} &= \mathbf{S}^{\text{e}}\mathbf{C}^{\text{e}}\boldsymbol{\varepsilon}^{\text{e}}\\ \mathbf{F}^{\text{p}}\mathbf{C}^{\text{p}} &= \mathbf{S}^{\text{p}}\mathbf{C}^{\text{p}}\boldsymbol{\varepsilon}^{\text{p}}\end{aligned} \tag{4}$$

Here $\mathbf{F}$, $\mathbf{C}$, $\mathbf{S}$, and $\boldsymbol{\varepsilon}$ are the Fock or Kohn-Sham matrix, orbital coefficient matrix, overlap matrix, and orbital energy matrix, respectively, for electrons and protons with superscripts e and p, respectively.

In this framework, implicit solvation effects can be incorporated through the inclusion of an additional term $\hat{V}^{\mathrm{R}}$ into the Hamiltonian of the system:

$$\hat{H}_{\mathrm{implicit}} = \hat{H}_{\mathrm{NEO}} + \hat{V}^{\mathrm{R}} \tag{5}$$

Here $\hat{H}_{\mathrm{NEO}}$ is the NEO gas-phase Hamiltonian, which includes the kinetic energies of the electrons and quantum nuclei (assumed to be protons for simplicity) and all Coulomb interactions among the electrons, quantum nuclei, and classical nuclei. $\hat{V}^{\mathrm{R}}$ is the potential term that represents the density-dependent electrostatic interactions between the solute molecule and the implicit solvent. It is defined in terms of a summation over ASCs and an operator $\hat{V}_k$ that generates the solute electrostatic potential on the surface of the molecular cavity:

$$\hat{V}^{\mathrm{R}} = \sum_{k}^{N_{\mathrm{tess}}} q_k \hat{V}_k \tag{6}$$

The solute electrostatic potential, which depends on the electronic and protonic densities, as well as the positions and charges of the classical nuclei, is evaluated at the surface tesserae positions. The ASCs obtained by solving Eq. (2) enter as additional one-electron (one-proton) solvation contributions to the electronic (protonic) NEO Fock or Kohn-Sham matrix elements:

$$\begin{aligned} F_{\mu\nu}^{\mathrm{e,CPCM}} &= -\sum_{k}^{N_{\mathrm{tess}}} q_k L_{\mu\nu,k}^{\mathrm{e}} \\ F_{\mu'\nu'}^{\mathrm{p,CPCM}} &= \sum_{k}^{N_{\mathrm{tess}}} q_k L_{\mu'\nu',k}^{\mathrm{p}} \end{aligned} \tag{7}$$

where $L^{\mathrm{e}}_{\mu\nu,k}$ and $L^{\mathrm{p}}_{\mu'\nu',k}$ are the interaction integrals over basis functions and are defined below.[38] These additional terms augment the NEO gas-phase Fock matrices as follows:

$$\begin{aligned} F^{\mathrm{e,pol}}_{\mu\nu} &= F^{\mathrm{e}}_{\mu\nu} + F^{\mathrm{e,CPCM}}_{\mu\nu} \\ F^{\mathrm{p,pol}}_{\mu'\nu'} &= F^{\mathrm{p}}_{\mu'\nu'} + F^{\mathrm{p,CPCM}}_{\mu'\nu'} \end{aligned} \tag{8}$$

Here the unprimed and primed indices are used to distinguish between electronic and protonic basis functions $\varphi^{\mathrm{e}}_{\mu}$ and $\varphi^{\mathrm{p}}_{\mu'}$, respectively. The NEO-PCM problem is solved self-consistently using the solvent polarized $\mathbf{F}^{\mathrm{e,pol}}$ and $\mathbf{F}^{\mathrm{p,pol}}$ matrices in Eq. (4) while using Eq. (2) to obtain the ASCs.

Since the PCM integral equations are solved numerically over a finite set of grid points, the quality of the results and the density of grid points are directly linked. Furthermore, the choice of the type of surface discretization scheme is especially important for molecular geometry optimizations with PCM. As the molecular coordinates are updated after each step of the geometry optimization procedure, the cavity surface must frequently be regenerated, altering the number and positions of the surface grid points. This process can lead to a discontinuous potential energy surface as well as artifacts such as Coulomb singularities that occur when adjacent surface elements are allowed to get too close to each other. For NEO-PCM, we adopt the switching/Gaussian (SwiG) cavity discretization formalism to avoid some of these issues.

The SwiG formalism involves spatially smeared surface charges, where each point charge is weighted by a spherical Gaussian function. Specifically, $\phi_k\left(\mathbf{r};\mathbf{s}_k,\zeta_k\right)$ in Eq. (1) is a Gaussian function of the form

$$\phi_k\left(\mathbf{r};\mathbf{s}_k,\zeta_k\right) = \left(\frac{\zeta_k^2}{\pi}\right)^{3/2} \exp\left(-\zeta_k^2\left|\mathbf{r}-\mathbf{s}_k\right|^2\right) \tag{9}$$

where ζ_k is an optimally chosen parameter for the k^{th} tessera,[39] and $L^{\mathrm{e}}_{\mu\nu,\mathrm{k}}$ and $L^{\mathrm{p}}_{\mu'\nu',\mathrm{k}}$ in Eq. (7) are defined as

$$L^{\mathrm{e}}_{\mu\nu,k} = \int d\mathbf{r}^{\mathrm{e}}_1 \varphi^{\mathrm{e}*}_{\mu}\left(\mathbf{r}^{\mathrm{e}}_1\right)\left[\frac{\mathrm{erf}\left(\zeta_k\left|\mathbf{r}^{\mathrm{e}}_1 - \mathbf{s}_k\right|\right)}{\left|\mathbf{r}^{\mathrm{e}}_1 - \mathbf{s}_k\right|}\right]\varphi^{\mathrm{e}}_{\nu}\left(\mathbf{r}^{\mathrm{e}}_1\right)$$
$$L^{\mathrm{p}}_{\mu'\nu',k} = \int d\mathbf{r}^{\mathrm{p}}_1 \varphi^{\mathrm{p}*}_{\mu'}\left(\mathbf{r}^{\mathrm{p}}_1\right)\left[\frac{\mathrm{erf}\left(\zeta_k\left|\mathbf{r}^{\mathrm{p}}_1 - \mathbf{s}_k\right|\right)}{\left|\mathbf{r}^{\mathrm{p}}_1 - \mathbf{s}_k\right|}\right]\varphi^{\mathrm{p}}_{\nu'}\left(\mathbf{r}^{\mathrm{p}}_1\right) \quad (10)$$

For the simpler point charge description, $\phi_k\left(\mathbf{r};\mathbf{s}_k,\zeta_k\right)$ is unity, and the quantity in square brackets is replaced by $1/\left|\mathbf{r}-\mathbf{s}_k\right|$. The SwiG approach was originally developed by York and Karplus[39] and further refined by Lange and Herbert, who subsequently showed that this approach overcomes spurious oscillations during molecular geometry optimizations due to discontinuous gradients of the solute-solvent electrostatic interaction energy.[32, 40-42] This approach avoids problems associated with Coulomb singularities and discontinuities in the potential energy landscape.

We found that the SwiG approach yields smoother convergence for NEO-PCM geometry optimizations, as illustrated by our application to a water pentamer with all ten hydrogen nuclei treated quantum mechanically. These PCM calculations were performed with NEO-DFT using the 6-31G electronic basis set[43], an *sp* protonic basis set that consists of one *s*-type and one *p*-type basis function with exponent 4, the B3LYP[44, 45] electron exchange-correlation functional, and the epc17-2[11, 12] electron-proton correlation functional. Figure 1 illustrates that the geometry optimization is more efficient for the SwiG scheme than for the variable tesserae number (VTN)[46] discretization scheme, which utilizes a point charge representation approach. Note that the spikes that appear for the VTN scheme in this example do not appear in all cases and often can be mitigated by altering the number of tesserae or starting at a different initial geometry.

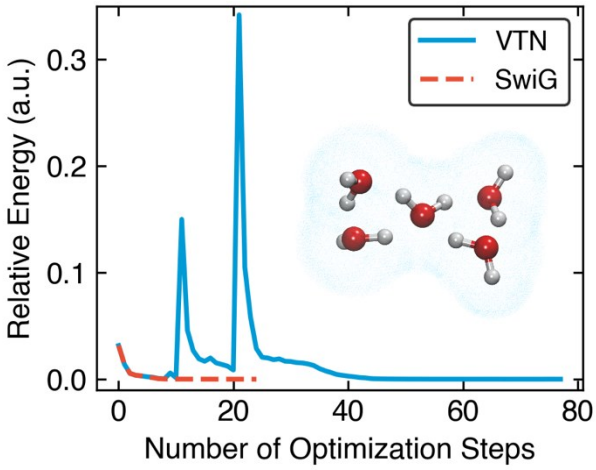

Figure 1. Geometry optimization of a water pentamer using the NEO-PCM method with a dielectric constant corresponding to water, showing the difference in energy at each optimization step relative to the energy of the minimized structure. All ten hydrogen nuclei were treated quantum mechanically using NEO-DFT.

Iterative, self-consistent convergence of Eqs. (2) and (4) leads to the solvated nuclear-electronic wave function and corresponding total energy of the solvated molecular system. The total energy $E_{\mathrm{PCM}}^{\mathrm{total}}$ is the expectation value of the total Hamiltonian given in Eq. (5) with respect to the solvent polarized nuclear-electronic wave function Ψ_{NEO}, subtracting the work required to create the set of polarization charges:

$$\begin{aligned} E_{\mathrm{PCM}}^{\mathrm{total}} &= \left\langle \Psi_{\mathrm{NEO}} \middle| \hat{H}_{\mathrm{implicit}} \middle| \Psi_{\mathrm{NEO}} \right\rangle - \frac{1}{2} \left\langle \Psi_{\mathrm{NEO}} \middle| \hat{V}^{\mathrm{R}} \middle| \Psi_{\mathrm{NEO}} \right\rangle \\ &= \left\langle \Psi_{\mathrm{NEO}} \middle| \hat{H}_{\mathrm{NEO}} \middle| \Psi_{\mathrm{NEO}} \right\rangle + \frac{1}{2} \mathbf{v}^{\mathrm{T}} \mathbf{q} \end{aligned} \tag{11}$$

In the final expression, the first term is the NEO electronic energy computed as the expectation value of $\hat{H}_{\mathrm{NEO}}$ with respect to the solvent polarized nuclear-electronic wave function, and $\frac{1}{2}\mathbf{v}^{\mathrm{T}}\mathbf{q}$ is the additional solvent polarization term corresponding to the solute-solvent electrostatic interaction energy.

We implemented the PCM approach with NEO-HF and NEO-DFT in Q-Chem[47] 6.0.2. The energies for both the NEO-HF-PCM and NEO-DFT-PCM approaches were validated by comparison to an implementation in Chronus Quantum.[48] We also derived the analytical gradients and validated our implementation in Q-Chem by comparison to numerical gradients, as given in the Supporting Information (SI).

In this work, we assume that each quantum proton is represented by a single basis function center for its electronic and protonic basis functions. In this case, the form of the gradients for the classical nuclei and the basis function centers is the same. As shown in the SI, the NEO-HF and NEO-DFT PCM energy gradients are analogous to the conventional electronic counterparts with additional terms associated with the quantum protons. In particular, the gradient of the first term in Eq. (11) includes additional terms associated with the protonic Fock (or Kohn-Sham) matrix, and the gradient of the second term, which is the solute-solvent electrostatic energy, includes an additional term accounting for the interaction between the induced surface charges and the quantum proton densities.

NEO-QM/MM

In addition, we developed a NEO-QM/MM approach to provide an atomistic description of the solvent or other chemical environment. In QM/MM approaches, the system is partitioned into QM and MM regions. Often the reactive part is treated quantum mechanically with a method such as DFT, and the environment is described by an MM force field. A variety of methods for treating the interactions between the QM and MM regions have been developed.[49] Here we describe the coupling between the QM and MM regions with the electrostatic embedding scheme, where inclusion of the MM point charges into the NEO Hamiltonian enables the solvent to polarize

the charge density of the QM solute. The field of MM point charges acts as an external potential that is included during optimization of the NEO wave function:

$$\hat{H}_{\text{explicit}} = \hat{H}_{\text{NEO}} + \hat{H}_{\text{QM/MM}}^{\text{electrostatic}} \tag{12}$$

For the NEO-QM/MM approach, $\hat{H}_{\text{QM/MM}}^{\text{electrostatic}}$ is analogous to the operator used for QM/MM methods with conventional electronic structure but includes an additional term for the electrostatic interaction between the MM point charges (q_m) and the QM protons:

$$\hat{H}_{\text{QM/MM}}^{\text{electrostatic}} = -\sum_{i}^{N_{\text{e}}}\sum_{m}^{N_{\text{MM}}} \frac{q_m}{\left|\mathbf{r}_i^{\text{e}} - \mathbf{r}_m^{\text{MM}}\right|} + \sum_{i'}^{N_{\text{p}}}\sum_{m}^{N_{\text{MM}}} \frac{q_m}{\left|\mathbf{r}_{i'}^{\text{p}} - \mathbf{r}_m^{\text{MM}}\right|} + \sum_{A}^{N_{\text{c}}}\sum_{m}^{N_{\text{MM}}} \frac{Z_A q_m}{\left|\mathbf{r}_A^{\text{c}} - \mathbf{r}_m^{\text{MM}}\right|} \tag{13}$$

Here N_{MM} is the number of MM point charges, and N_{e}, N_{p}, and N_{c} are the number of electrons, quantum protons, and classical nuclei, respectively, in the QM region. The spatial coordinates, **r**, are labeled with their respective superscripts. The MM point charges enter as additional one-electron (one-proton) contributions to the electronic (protonic) NEO core Hamiltonian matrix elements:

$$\begin{aligned} H_{\mu\nu}^{\text{e,eff}} &= H_{\mu\nu}^{\text{e}} + H_{\mu\nu}^{\text{e,electrostatic}} \\ H_{\mu'\nu'}^{\text{p,eff}} &= H_{\mu'\nu'}^{\text{p}} + H_{\mu'\nu'}^{\text{p,electrostatic}} \end{aligned} \tag{14}$$

$$\begin{aligned} H_{\mu\nu}^{\text{e,electrostatic}} &= \int d\mathbf{r}_1^{\text{e}} \varphi_\mu^{\text{e}*}\left(\mathbf{r}_1^{\text{e}}\right)\left[-\sum_{m}^{N_{\text{MM}}} \frac{q_m}{\left|\mathbf{r}_1^{\text{e}} - \mathbf{r}_m^{\text{MM}}\right|}\right]\varphi_\nu^{\text{e}}\left(\mathbf{r}_1^{\text{e}}\right) \\ H_{\mu'\nu'}^{\text{p,electrostatic}} &= \int d\mathbf{r}_1^{\text{p}} \varphi_{\mu'}^{\text{p}*}\left(\mathbf{r}_1^{\text{p}}\right)\left[\sum_{m}^{N_{\text{MM}}} \frac{q_m}{\left|\mathbf{r}_1^{\text{p}} - \mathbf{r}_m^{\text{MM}}\right|}\right]\varphi_{\nu'}^{\text{p}}\left(\mathbf{r}_1^{\text{p}}\right) \end{aligned} \tag{15}$$

The iterative self-consistent convergence of Eq. (4) using $\mathbf{H}^{\text{e,eff}}$ and $\mathbf{H}^{\text{p,eff}}$ leads to the environmentally polarized nuclear-electronic wave function. Employing the additive QM/MM energy approach, the total NEO-QM/MM energy $E_{\text{QM/MM}}^{\text{total}}$ can be expressed as

$$E_{\mathrm{QM/MM}}^{\mathrm{total}} = \left\langle \Psi_{\mathrm{NEO}} \left| \hat{H}_{\mathrm{explicit}} \right| \Psi_{\mathrm{NEO}} \right\rangle + E_{\mathrm{QM/MM}}^{\mathrm{vdw}} + E_{\mathrm{QM/MM}}^{\mathrm{bonded}} + E_{\mathrm{MM}} \tag{16}$$

The first term is the expectation value of the Hamiltonian in Eq. (12) with respect to the environmentally polarized nuclear-electronic wave function. $E_{\mathrm{QM/MM}}^{\mathrm{vdw}}$ denotes the van der Waals interactions between the QM and MM atoms,[50] and E_{MM} denotes the interactions within the MM region. $E_{\mathrm{QM/MM}}^{\mathrm{bonded}}$ denotes the interactions that arise in some QM/MM formulations if QM and MM atoms are directly bonded to each other, which can be treated in a variety of ways.[51-53]

In addition to the QM/MM energy, we also implemented the analytical gradient of $E_{\mathrm{QM/MM}}^{\mathrm{total}}$ with respect to both the QM and MM nuclei. These gradients are obtained by a straightforward extension of the conventional electronic structure QM/MM counterpart to yield a set of forces. The forces acting on the QM nuclei are obtained from the gradients of the first three terms of Eq. (16) . The resulting NEO-QM/MM forces are analogous to the conventional electronic counterparts with additional terms associated with the protonic Fock (or Kohn-Sham) matrix. The forces on the MM nuclei involve gradients of the last three terms of Eq. (16), as well as the following term corresponding to the electrostatic interaction between the QM and MM regions:

$$\mathbf{F}_m^{\mathrm{MM}} = q_m \mathbf{E}(\mathbf{r}_m^{\mathrm{MM}}) \tag{17}$$

Here $\mathbf{F}_m^{\mathrm{MM}}$ is the force on the m^{th} MM nucleus, and $\mathbf{E}(\mathbf{r}_m^{\mathrm{MM}})$ is the electric field at the m^{th} MM atom due to the electronic and protonic densities, as well as the classical nuclei, in the QM region. The expression for the electric field is analogous to the conventional electronic counterpart with additional terms associated with the protonic densities. We implemented the NEO-QM/MM method in a developer version of Q-Chem 6.1.0 interfaced with the GROMACS[54] 4.6.5 software using a locally modified version of the INAQS[55] interface. The analytical gradients for the NEO-

HF and NEO-DFT QM/MM approach have been verified by comparison to numerical gradients, as given in the SI.

Computational Details

We performed geometry optimizations using both conventional DFT and NEO-DFT in the gas phase, implicit solvent, and explicit MM solvent. We also performed geometry optimizations and MD simulations for fully MM, conventional QM/MM, and NEO-QM/MM potentials. For the NEO-DFT calculations, the 6-31G(d,p) electronic basis set[56] and the PB4-F2 protonic basis set,[57] as well as the B3LYP electronic exchange-correlation functional and the epc-17-2 electron-proton correlation functional, were used. The same electronic basis set and functional were used for the conventional DFT calculations. An energy convergence threshold of 1.0×10^{-8} a.u. and a maximum gradient component criterion of 3.0×10^{-4} Hartree/Bohr were utilized for all calculations performed entirely within Q-Chem, namely the gas phase and PCM calculations. For the PCM calculations, the cavity was constructed using interlocking spheres centered at the positions of each of the classical nuclei and, for the NEO calculations, the quantum proton basis function center positions. These cavity spheres were sized according to Bondi's set of radii[58] and scaled by a factor of 1.2. The molecular cavity was discretized according to the SwiG prescription utilizing a grid density of 302 surface elements per cavity sphere.

For the QM/MM and fully MM energy calculations, the TIP3P[59] water model and the OPLS-AA force field[60-62] were used. This force field does not include van der Waals interactions for polar hydrogens, therefore allowing a consistent comparison between conventional QM/MM and NEO-QM/MM calculations. In future calculations utilizing force fields that include Lennard-Jones parameters for hydrogen, the van der Waals interactions could be neglected for the quantized protons or could be included based on the expectation value of the quantized proton position

operator. Although TIP3P was parameterized as a rigid water model, we allowed the water to be flexible in order to use sufficiently tight convergence criteria for geometry optimizations. For the flexible water calculations, the corresponding OPLS-AA terms were used for the intramolecular angle bending and bond stretching water terms. The MD simulations were performed with both rigid and flexible water models, producing qualitatively similar results. For the rigid water calculations, the bond angles and lengths were constrained using the SETTLE[63] algorithm. For all geometry optimizations involving molecular mechanics, namely the QM/MM and fully MM energy minimizations, the limited-memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newtonian minimizer with a maximum force criterion of 15.0 kJ mol^{-1} nm^{-1} was used with GROMACS configured in double precision mode. Molecular structures visualized in this paper were created using VMD,[64] and utilities provided through the MDAnalysis[65] library were used to analyze GROMACS trajectory information.

Results and Discussion

To showcase the effects of aqueous solvation on equilibrium structures obtained within the NEO framework, we performed a series of molecular geometry optimizations for three hydrogen-bonded complexes, each composed of a small organic molecule and its neighboring water molecule (Fig. 2). The organic molecule and hydrogen-bonded water were treated quantum mechanically with conventional DFT or NEO-DFT in a gas phase, implicit solvent, or explicit MM solvent environment (Fig. 3). For the NEO calculations, the hydrogen-bonding proton and the two protons on the water were treated quantum mechanically. The analogous calculations with only the hydrogen-bonding proton treated quantum mechanically are provided in the SI.

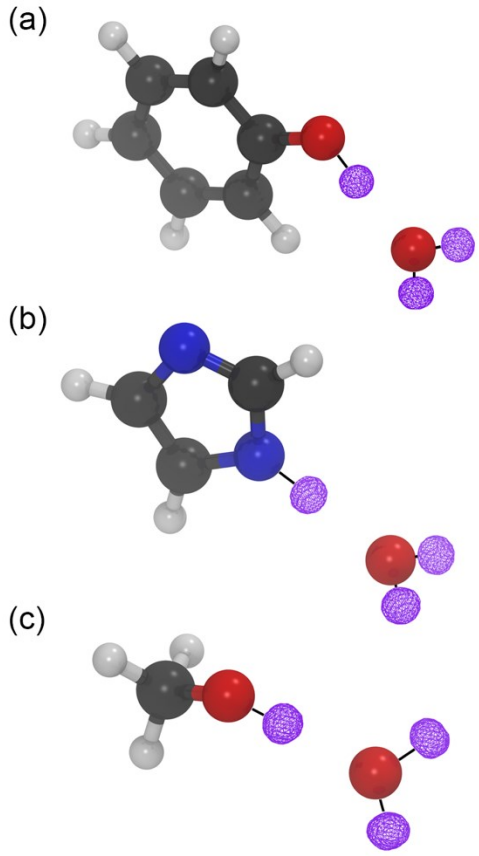

Figure 2. Three organic molecules acting as hydrogen-bond donors to an adjacent water molecule: (a) phenol, (b) imidazole, and (c) methanol. Three protons are treated quantum mechanically with NEO-DFT, as indicated by the three protonic densities depicted with purple mesh.

For the QM/MM calculations, each of the organic molecules previously optimized in the gas phase using the MM force field was placed in its own simulation box that was then filled with water. For the phenol system, 4756 water molecules were added to a 629.54 nm^3 volume cubic box. The other two systems were prepared similarly. Water was then equilibrated in the NVT and NPT ensembles (temperature of 300 K and pressure of 1 bar) for 2 ns each using the MM force field while keeping the organic molecule fixed and utilizing periodic boundary conditions. Following these steps, a droplet was created by retaining only the water molecules within 10 Å of any atom in the hydrogen-bonded complex. The resulting system composed of the organic

molecule in a water droplet was then energy minimized with the conventional DFT or NEO-DFT QM/MM approach. These energy minimizations were performed without periodic boundary conditions and including all non-bonded interactions (i.e., there were no cutoffs for the non-bonded interactions).

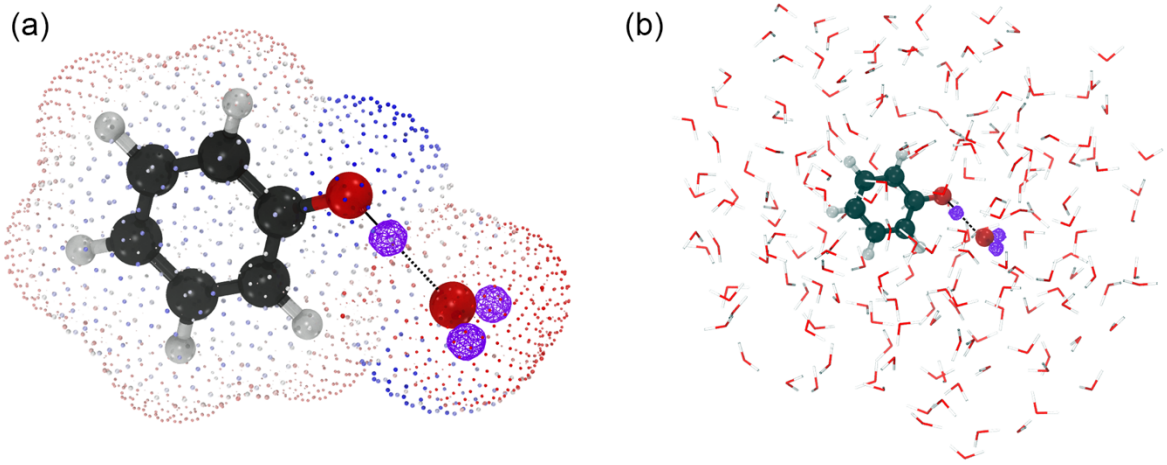

Figure 3. Optimized geometries of the phenol molecule hydrogen bonded to water, where the hydrogen-bonded complex is treated at the NEO-DFT level and is embedded in (a) C-PCM water or (b) a droplet of explicit TIP3P water molecules. The proton densities of the three quantized protons are depicted with purple mesh. The ASCs in (a) are represented on a color scale where red indicates a negative charge and blue indicates a positive charge. The geometries of these systems were optimized with NEO-DFT-PCM and NEO-DFT-QM/MM for (a) and (b), respectively.

We analyzed the hydrogen-bonding distances for the DFT and NEO-DFT calculations in the gas phase, implicit solvent, and explicit solvent environments. Table 1 provides the $O\cdots O$ (or $O\cdots N$) distance between the oxygen (or nitrogen) atoms of the hydrogen-bonded molecule and water, the $O\cdots H$ distance between the oxygen atom of the water and the hydrogen atom of the molecule, and the $H-O$ (or $H-N$) distance within the molecule. In the NEO-DFT calculations, the distances involving the quantum proton were computed from the expectation value of the proton position operator.

Table 1. Equilibrium Distances and Angles for Molecule and Hydrogen-Bonded Water Treated Quantum Mechanically in Various Environments[a]

	Distance or Angle	DFT	NEO-DFT	DFT PCM	NEO-DFT PCM	DFT MM	NEO-DFT MM
Phenol	H−O	0.98	1.00	0.99	1.02	0.99	1.02
	O···H	1.83	1.77	1.73	1.65	1.71	1.61
	O···O	2.80	2.77	2.72	2.67	2.68	2.61
	∠OHO	174.5	174.2	177.6	177.4	167.7	165.9
Methanol	H−O	0.97	0.99	0.98	1.00	0.99	1.01
	O···H	1.91	1.85	1.82	1.75	1.77	1.69
	O···O	2.86	2.83	2.80	2.76	2.72	2.68
	∠OHO	165.7	165.6	175.6	176.0	162.5	164.3
Imidazole	H−N	1.02	1.04	1.03	1.06	1.04	1.08
	O···H	1.91	1.85	1.83	1.74	1.71	1.61
	O···N	2.93	2.89	2.86	2.80	2.71	2.67
	∠OHN	176.1	176.0	178.4	179.9	160.0	166.0

[a]The molecule and hydrogen-bonded water are treated with DFT or NEO-DFT in the gas phase, C-PCM water, or MM water. Distances are given in Å for the hydrogen bond between the specified molecule and water in the molecule-water system, where O···O or (O···N) indicates the distance between the water oxygen and molecule oxygen (or nitrogen) atoms, O···H indicates the distance between the molecule hydrogen and water oxygen atoms, and H−O or (H−N) indicates the distance between the molecule hydrogen and its bonded oxygen (or nitrogen) atoms. The angle measured along the hydrogen bond is given in degrees. For the NEO calculations, the expectation value of the quantum proton coordinate was used to compute distances and angles.

In all cases, introduction of solvent results in a decrease in the O···O (or O···N) and O···H distances, with a slightly greater decrease obtained with explicit solvation compared to

implicit solvation. These trends manifest for both conventional DFT and NEO-DFT approaches. Moreover, the O···O (or O···N) and O···H distances are observed to be slightly shorter when the hydrogen-bonded proton is quantized with the NEO approach, which we make note of in discussions below. The distances between the hydrogen-bonded proton and its bonded oxygen (or nitrogen), H−O (or H−N), are slightly elongated due to solvation, consistent with previous findings.[17] For the systems studied herein, the distance changes indicate that solvation can strengthen the hydrogen bond. Furthermore, we found that the hydrogen-bond angle is similar for both conventional DFT and NEO-DFT calculations. In particular, the angles obtained with explicit solvation are slightly less linear than those obtained with implicit solvation. A possible explanation for these differences is that the angle is modulated by the surrounding MM water molecules, which can form hydrogen bonds with the QM complex through electrostatic and van der Waals interactions. These interactions may also influence the quantum proton densities in the NEO-QM/MM calculations.

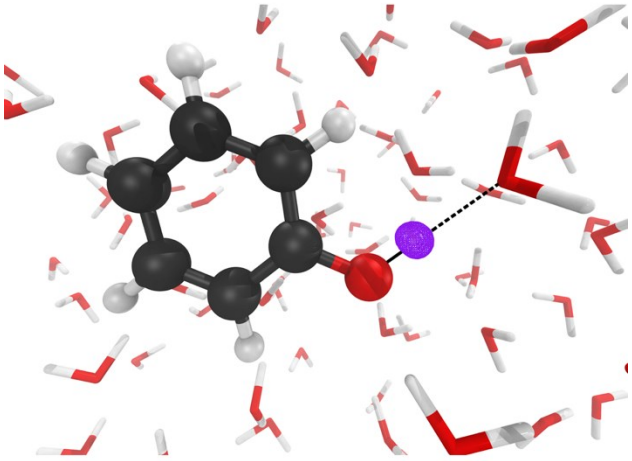

Figure 4. Representative configuration obtained from the NEO-QM/MM MD trajectory. The purple mesh depicts the quantized proton density. The hydrogen-bonded water molecule did not exchange on the timescale of the MD trajectories propagated.

To illustrate real-time dynamics, we also performed a short NEO-QM/MM MD simulation on the ground state potential energy surface of the phenol-water complex (Fig. 4). Starting from the optimized conventional DFT QM/MM geometry of the phenol-water droplet, initial velocities for the classical nuclei (i.e., all nuclei except the quantum proton) were assigned from a Maxwell-Boltzmann distribution at 300 K. A conventional QM/MM MD trajectory under the microcanonical ensemble was propagated for 2 ps prior to beginning the 100 fs production trajectory using the fully MM, conventional QM/MM, and NEO-QM/MM approaches. These MD simulations were performed without periodic boundary conditions and including all non-bonded interactions. The rigid TIP3P water model was used for these MD simulations. To be consistent with the geometry optimizations, analogous MD simulations with flexible TIP3P water molecules were also performed, and the results are provided in the SI. The velocity Verlet algorithm was utilized to propagate the classical nuclear coordinates using a time step of 0.5 fs. At each time step during the NEO-QM/MM MD trajectory, the quantum proton basis function center position was variationally optimized using NEO-DFT.

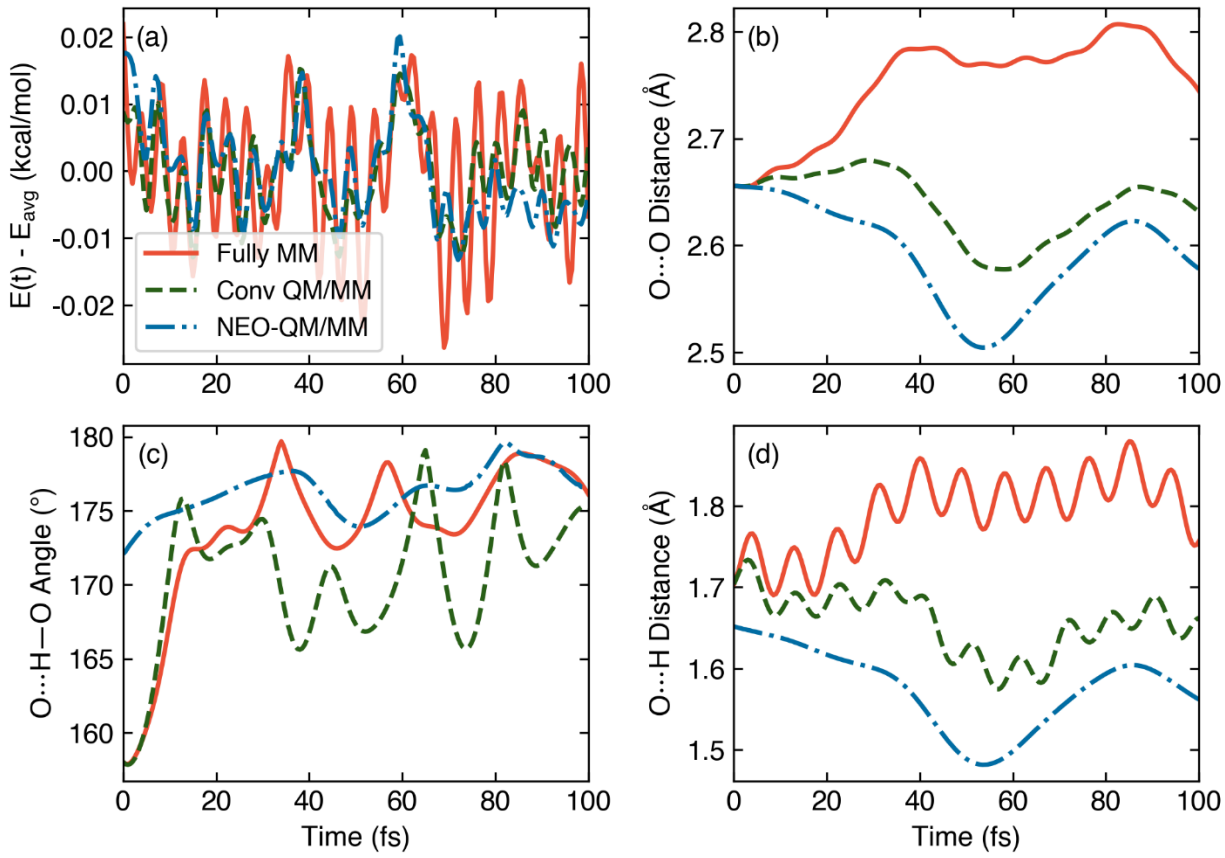

Figure 5. Analysis of energy fluctuations and hydrogen-bonding distances and angles between a phenol molecule and the oxygen of the closest water molecule. Time evolution of (a) energy relative to average energy, (b) O···O distance between the phenol and hydrogen-bonded water oxygen atoms, (c) O···H—O hydrogen-bond angle, and (d) O···H distance between the phenol hydrogen and water oxygen atoms. The phenol molecule was treated with an MM force field, conventional DFT, or NEO-DFT and was surrounded by 240 rigid TIP3P water molecules. Each trajectory started from the same initial coordinates and velocities to allow a direct comparison. For the NEO calculations, the proton basis function center was optimized at each time step, and the distances were computed using the expectation value of the quantum proton coordinate. The analogous figure with flexible TIP3P water is provided in Figure S1 of the SI.

Analyses of the MD trajectories propagated with the fully MM, conventional QM/MM, and NEO-QM/MM approaches are given in Figure 5. We chose to start each trajectory from the same initial conditions to allow us to systematically compare the differences among the various levels of theory. As mentioned above, the system was equilibrated for only 2 ps with the conventional QM/MM approach. As a result, these trajectories are starting out of equilibrium. For the fully MM approach, the distances change significantly because the MM force field differs

substantially from the QM/MM potential. Figure 5a shows that energy is conserved well for all three trajectories.

The trends in distances and angles shown in Figure 5 are consistent with those obtained from corresponding geometry optimizations. For the optimized geometries reported in Table 2, all water molecules were treated at the MM level using the flexible TIP3P water model, and the organic molecule was described at the MM level or at the QM level with conventional DFT or NEO-DFT. Consistent with our results in Table 1, which treated both the organic molecule and the hydrogen-bonded water molecule quantum mechanically, the intermolecular hydrogen-bonding distances are shorter with NEO-DFT compared to conventional DFT. These results suggest a stronger hydrogen-bonding interaction for these systems when protons are quantized with NEO-DFT/epc17-2. However, such subtle differences in distances may depend on various factors, such as the electron-proton correlation functional and the proton basis set, and have not been validated experimentally or with other levels of theory for this system. As shown in previous studies using path integral molecular dynamics methods, the impact of proton quantization on hydrogen-bonding interactions is system dependent.[66-68] Thus, this subtle effect of proton quantization on the hydrogen-bonding interaction in these systems should not be viewed as definitive.

An important observation to highlight is the absence of oscillations in the $O\cdots H$ distances for the NEO-QM/MM approach (blue line in Figure 5d). This behavior emerges because for the Born-Oppenheimer MD approach within the NEO framework, the quantum proton responds instantaneously to the classical nuclear positions, analogous to the instantaneous response of the electrons. For every time step in the NEO-QM/MM trajectory, the electronic and protonic densities are relaxed to the ground vibronic state by solving the Kohn-Sham equations (Eq. (4)), where the

Fock matrix depends on the external potential created by the classical nuclei in the QM region and the MM point charges. In order to capture the proton vibrations, the motion of the classical subsystem must be coupled with a dynamical treatment of the nuclear-electronic quantum subsystem. This can be accomplished using the real-time NEO time-dependent DFT approach with Ehrenfest dynamics.[15, 69] Efforts in this direction are currently underway.

Table 2. Equilibrium Distances and Angles for Molecule and Hydrogen-Bonded MM Water Obtained from Geometry Optimizations[a]

	Distance and Angle	Fully MM	DFT MM	NEO-DFT MM
Phenol	O···H	1.80	1.63	1.56
	O···O	2.74	2.60	2.57
	∠OHO	165.4	168.2	169.1
Methanol	O···H	1.88	1.72	1.64
	O···O	2.77	2.66	2.62
	∠OHO	152.9	160.4	162.8
Imidazole	O···H	1.63	1.57	1.48
	O···N	2.68	2.61	2.56
	∠OHN	174.0	171.1	172.8

[a]The molecule is treated with MM, DFT, or NEO-DFT in an MM water environment. Distances are given in Å for the hydrogen bond between the specified molecule and water in the molecule-water system, where O···O (or O···N) indicates the distance between the water oxygen and molecule oxygen (or nitrogen) atoms, and O···H indicates the distance between the molecule hydrogen and water oxygen atoms. The angle measured along the hydrogen bond is given in degrees. For the NEO calculations, the expectation value of the quantum proton coordinate was used to compute distances and angles.

Conclusion

This manuscript presents the development and implementation of the NEO-QM/MM approach for quantizing specified nuclei in the QM region. In this approach, the QM region is treated with a NEO method such as NEO-HF or NEO-DFT, and the remainder of the system is treated with an MM force field. We compare this approach to an implementation of the NEO-PCM method, where the solvent environment is represented as a polarizable dielectric continuum. The expressions for the NEO-QM/MM and NEO-PCM energies and analytical nuclear gradients are provided. These analytical gradients enable geometry optimizations and MD simulations for molecules in solvent and, in the case of NEO-QM/MM, more complex heterogeneous environments. Our applications to hydrogen-bonded systems illustrate that aqueous solvation can strengthen hydrogen-bonding interactions, as indicated by the shortening of the intermolecular hydrogen bond distances for the systems studied. However, given the many competing factors influencing hydrogen-bonding interactions, this effect may be system dependent.[70] The solvated equilibrium structures obtained with NEO-PCM or NEO-QM/MM can be used to compute pK_a values[22, 71-74] and solvation free energies[75, 76] in a manner that includes proton delocalization and anharmonicity.

We also show that the NEO-QM/MM method can be used to perform real-time direct dynamics simulations in complex environments. For these types of Born-Oppenheimer MD simulations, the quantum proton(s) as well as the electrons respond instantaneously to the classical nuclear motions. The proton quantum dynamics can be described more accurately by removing the Born-Oppenheimer separation between the quantum protons and the other nuclei. For this purpose, we can propagate the proton density as well as the classical nuclei on the electronic ground state using the Born-Oppenheimer real-time NEO-TDDFT Ehrenfest approach.[77] This approach can also be used to describe processes occurring on excited electronic states by

propagating both the electronic and protonic densities with the original real-time NEO-TDDFT Ehrenfest approaches.[14, 15] The developments presented herein serve as a stepping stone for simulating nuclear-electronic quantum dynamics in complex chemical and biological environments within the NEO framework, providing the infrastructure for studying a wide range of nonadiabatic condensed phase processes.

Supporting Information

Analytical gradient expressions for NEO-PCM and NEO-QM/MM; validation of analytical gradients; single quantum proton geometry optimization results; flexible water MD trajectories; tests for proton basis function center optimization; sample user inputs for Q-Chem and INAQS.

Acknowledgements

The development of fundamental NEO theory is supported by the National Science Foundation Grant No. CHE-1954348 (S.H.-S.). The development of solvated NEO methods in the Chronus Quantum computational software is supported by the Department of Energy in the Computational Chemical Sciences program Grant No. DE-SC0023284. The authors thank Dr. Jonathan Fetherolf, Dr. Tao Li, Dr. Christopher Malbon, Dr. Eno Paenurk, Dr. Alexander Soudackov, and Joseph Dickinson for useful discussions. We also thank Dr. Vale Cofe-Shabica for his help and correspondence related to using the INAQS interface. Figures were generated using the *columnplots* wrapper, which is available on Github (https://github.com/TaoELi/columnplots).

References

(1) Webb, S. P.; Iordanov, T.; Hammes-Schiffer, S., Multiconfigurational nuclear-electronic orbital approach: Incorporation of nuclear quantum effects in electronic structure calculations. *J. Chem. Phys.* **2002,** *117*, 4106-4118.
(2) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Quantum Chemistry: Integrating Electronic and Nuclear Quantum Effects via the Nuclear–Electronic Orbital Method. *Chem. Rev.* **2020,** *120*, 4222-4253.
(3) Hammes-Schiffer, S., Nuclear–electronic orbital methods: Foundations and prospects. *J. Chem. Phys.* **2021,** *155*, 030901.
(4) Huynh, M. H. V.; Meyer, T. J., Proton-Coupled Electron Transfer. *Chem. Rev.* **2007,** *107*, 5004-5064.
(5) Layfield, J. P.; Hammes-Schiffer, S., Hydrogen Tunneling in Enzymes and Biomimetic Models. *Chem. Rev.* **2014,** *114*, 3466-3494.
(6) Wang, L.; Fried, S. D.; Boxer, S. G.; Markland, T. E., Quantum delocalization of protons in the hydrogen-bond network of an enzyme active site. *Proc. Natl. Acad. Sci.* **2014,** *111*, 18454-18459.
(7) Pavošević, F.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Coupled Cluster Singles and Doubles Theory within the Nuclear-Electronic Orbital Framework. *J. Chem. Theory Comput.* **2019,** *15*, 338-347.
(8) Pavošević, F.; Hammes-Schiffer, S., Multicomponent equation-of-motion coupled cluster singles and doubles: Theory and calculation of excitation energies for positronium hydride. *J. Chem. Phys.* **2019,** *150*, 161102.
(9) Pavošević, F.; Rousseau, B. J. G.; Hammes-Schiffer, S., Multicomponent Orbital-Optimized Perturbation Theory Methods: Approaching Coupled Cluster Accuracy at Lower Cost. *J. Phys. Chem. Lett.* **2020,** *11*, 1578-1583.
(10) Pak, M. V.; Chakraborty, A.; Hammes-Schiffer, S., Density Functional Theory Treatment of Electron Correlation in the Nuclear−Electronic Orbital Approach. *J. Phys. Chem. A* **2007,** *111*, 4522-4526.
(11) Yang, Y.; Brorsen, K. R.; Culpitt, T.; Pak, M. V.; Hammes-Schiffer, S., Development of a practical multicomponent density functional for electron-proton correlation to produce accurate proton densities. *J. Chem. Phys.* **2017,** *147*, 114113.
(12) Brorsen, K. R.; Yang, Y.; Hammes-Schiffer, S., Multicomponent Density Functional Theory: Impact of Nuclear Quantum Effects on Proton Affinities and Geometries. *J. Phys. Chem. Lett.* **2017,** *8*, 3488-3493.
(13) Yang, Y.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Time-Dependent Density Functional Theory: Proton and Electron Excitation Energies. *J. Phys. Chem. Lett.* **2018,** *9*, 1765-1770.
(14) Zhao, L.; Tao, Z.; Pavošević, F.; Wildman, A.; Hammes-Schiffer, S.; Li, X., Real-Time Time-Dependent Nuclear−Electronic Orbital Approach: Dynamics beyond the Born–Oppenheimer Approximation. *J. Phys. Chem. Lett.* **2020,** *11*, 4052-4058.
(15) Zhao, L.; Wildman, A.; Tao, Z.; Schneider, P.; Hammes-Schiffer, S.; Li, X., Nuclear–electronic orbital Ehrenfest dynamics. *J. Chem. Phys.* **2020,** *153*, 224111.
(16) Wildman, A.; Tao, Z.; Zhao, L.; Hammes-Schiffer, S.; Li, X., Solvated Nuclear–Electronic Orbital Structure and Dynamics. *J. Chem. Theory Comput.* **2022,** *18*, 1340-1346.
(17) Lambros, E.; Link, B.; Chow, M.; Hammes-Schiffer, S.; Li, X., Solvent Induced Proton Polarization within the Nuclear−Electronic Orbital Framework. *J. Phys. Chem. Lett.* **2023,** *14*, 2990-2995.

(18) Miertuš, S.; Scrocco, E.; Tomasi, J., Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981,** *55*, 117-129.
(19) Tomasi, J.; Persico, M., Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994,** *94*, 2027-2094.
(20) Field, M. J.; Bash, P. A.; Karplus, M., A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J. Comput. Chem.* **1990,** *11*, 700-733.
(21) Åqvist, J.; Warshel, A., Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem. Rev.* **1993,** *93*, 2523-2544.
(22) Simonson, T.; Carlsson, J.; Case, D. A., Proton Binding to Proteins: pKa Calculations with Explicit and Implicit Solvent Models. *J. Am. Chem. Soc.* **2004,** *126*, 4167-4180.
(23) Tomasi, J.; Mennucci, B.; Cammi, R., Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005,** *105*, 2999-3094.
(24) Senthilkumar, K.; Mujika, J. I.; Ranaghan, K. E.; Manby, F. R.; Mulholland, A. J.; Harvey, J. N., Analysis of polarization in QM/MM modelling of biologically relevant hydrogen bonds. *J R Soc Interface* **2008,** *5 Suppl 3*, S207-S216.
(25) Senn, H. M.; Thiel, W., QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009,** *48*, 1198-1229.
(26) Mennucci, B., Continuum Solvation Models: What Else Can We Learn from Them? *J. Phys. Chem. Lett.* **2010,** *1*, 1666-1674.
(27) Mennucci, B., Polarizable continuum model. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012,** *2*, 386-404.
(28) Liu, M.; Wang, Y.; Chen, Y.; Field, M. J.; Gao, J., QM/MM through the 1990s: The First Twenty Years of Method Development and Applications. *Israel Journal of Chemistry* **2014,** *54*, 1250-1263.
(29) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K., The ONIOM Method and Its Applications. *Chem. Rev.* **2015,** *115*, 5678-5796.
(30) Raghavachari, K.; Saha, A., Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015,** *115*, 5643-5677.
(31) Zhou, Y.; Wang, S.; Li, Y.; Zhang, Y., Chapter Five - Born–Oppenheimer Ab Initio QM/MM Molecular Dynamics Simulations of Enzyme Reactions. In *Methods in Enzymology*, Voth, G. A., Ed. Academic Press: 2016; Vol. 577, pp 105-118.
(32) Herbert, J. M., Dielectric continuum methods for quantum chemistry. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021,** *11*, e1519.
(33) Barone, V.; Bloino, J.; Biczysko, M.; Santoro, F., Fully Integrated Approach to Compute Vibrationally Resolved Optical Spectra: From Small Molecules to Macrosystems. *J. Chem. Theory Comput.* **2009,** *5*, 540-554.
(34) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Guido, C. A.; Mennucci, B.; Scalmani, G.; Frisch, M. J., Practical computation of electronic excitation in solution: vertical excitation model. *Chem. Sci.* **2011,** *2*, 2143-2161.
(35) Ding, F.; Lingerfelt, D. B.; Mennucci, B.; Li, X., Time-dependent non-equilibrium dielectric response in QM/continuum approaches. *J. Chem. Phys.* **2015,** *142*, 034120.
(36) Scalmani, G.; Frisch, M. J., Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **2010,** *132*, 114110.

(37) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V., Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **2003,** *24*, 669-681.
(38) Liu, F.; Sanchez, D. M.; Kulik, H. J.; Martínez, T. J., Exploiting graphical processing units to enable quantum chemistry calculation of large solvated molecules with conductor-like polarizable continuum models. *Int. J. Quantum Chem.* **2019,** *119*, e25760.
(39) York, D. M.; Karplus, M., A Smooth Solvation Potential Based on the Conductor-Like Screening Model. *J. Phys. Chem. A* **1999,** *103*, 11060-11079.
(40) Lange, A. W.; Herbert, J. M., Polarizable Continuum Reaction-Field Solvation Models Affording Smooth Potential Energy Surfaces. *J. Phys. Chem. Lett.* **2010,** *1*, 556-561.
(41) Lange, A. W.; Herbert, J. M., A smooth, nonsingular, and faithful discretization scheme for polarizable continuum models: The switching/Gaussian approach. *J. Chem. Phys.* **2010,** *133*, 244111.
(42) Lange, A. W.; Herbert, J. M.; Albrecht, B. J.; You, Z.-Q., Intrinsically smooth discretisation of Connolly's solvent-excluded molecular surface. *Mol. Phys.* **2020,** *118*, e1644384.
(43) Hehre, W. J.; Ditchfield, R.; Pople, J. A., Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972,** *56*, 2257-2261.
(44) Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988,** *38*, 3098-3100.
(45) Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988,** *37*, 785-789.
(46) Li, H.; Jensen, J. H., Improving the efficiency and convergence of geometry optimization with the polarizable continuum model: New energy gradients and molecular surface tessellation. *J. Comput. Chem.* **2004,** *25*, 1449-1462.
(47) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; Wergifosse, M. d.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge, Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Vera, M. H.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, Á.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Landau, A.; Lawler, K. V.; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Liebenthal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V.; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, Á.;

Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sundstrom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Coriani, S.; Cramer, C. J.; Cserey, G.; DePrinceIII, A. E.; DiStasioJr., R. A.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; GoddardIII, W. A.; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lambrecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsenfeld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V.; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truhlar, D. G.; Voorhis, T. V.; Wesolowski, T. A.; Whaley, K. B.; WoodcockIII, H. L.; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I., Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021,** *155*, 084801.

(48) Williams-Young, D. B.; Petrone, A.; Sun, S.; Stetina, T. F.; Lestrange, P.; Hoyer, C. E.; Nascimento, D. R.; Koulias, L.; Wildman, A.; Kasper, J.; Goings, J. J.; Ding, F.; DePrince III, A. E.; Valeev, E. F.; Li, X., The Chronus Quantum software package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2020,** *10*, e1436.

(49) Cao, L.; Ryde, U., On the Difference Between Additive and Subtractive QM/MM Calculations. *Frontiers in Chemistry* **2018,** *6*.

(50) Freindorf, M.; Shao, Y.; Furlani, T. R.; Kong, J., Lennard–Jones parameters for the combined QM/MM method using the B3LYP/6-31G*/AMBER potential. *J. Comput. Chem.* **2005,** *26*, 1270-1278.

(51) Gao, J., Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials. In *Reviews in Computational Chemistry*, 1996; pp 119-185.

(52) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M., Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000,** *104*, 1720-1735.

(53) König, P. H.; Hoffmann, M.; Frauenheim, T.; Cui, Q., A Critical Evaluation of Different QM/MM Frontier Treatments with SCC-DFTB as the QM Method. *J. Phys. Chem. B* **2005,** *109*, 9082-9095.

(54) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013,** *29*, 845-854.

(55) Cofer-Shabica, D. V.; Menger, M. F. S. J.; Ou, Q.; Shao, Y.; Subotnik, J. E.; Faraji, S., INAQS, a Generic Interface for Nonadiabatic QM/MM Dynamics: Design, Implementation, and Validation for GROMACS/Q-CHEM simulations. *J. Chem. Theory Comput.* **2022,** *18*, 4601-4614.

(56) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A., Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982,** *77*, 3654-3665.

(57) Yu, Q.; Pavošević, F.; Hammes-Schiffer, S., Development of nuclear basis sets for multicomponent quantum chemistry methods. *J. Chem. Phys.* **2020,** *152*, 244123.

(58) Bondi, A., van der Waals Volumes and Radii. *J. Phys. Chem.* **1964,** *68*, 441-451.

(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.
(60) Jorgensen, W. L.; Tirado-Rives, J., Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl. Acad. Sci.* **2005,** *102*, 6665-6670.
(61) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L., LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017,** *45*, W331-W336.
(62) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L., 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *J. Phys. Chem. B* **2017,** *121*, 3864-3870.
(63) Miyamoto, S.; Kollman, P. A., Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992,** *13*, 952-962.
(64) Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996,** *14*, 33-38.
(65) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O., MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011,** *32*, 2319-2327.
(66) Swalina, C.; Wang, Q.; Chakraborty, A.; Hammes-Schiffer, S., Analysis of Nuclear Quantum Effects on Hydrogen Bonding. *J. Phys. Chem. A* **2007,** *111*, 2206-2212.
(67) Li, X.-Z.; Walker, B.; Michaelides, A., Quantum nature of the hydrogen bond. *Proc. Natl. Acad. Sci.* **2011,** *108*, 6369-6373.
(68) Ceriotti, M.; Fang, W.; Kusalik, P. G.; McKenzie, R. H.; Michaelides, A.; Morales, M. A.; Markland, T. E., Nuclear Quantum Effects in Water and Aqueous Systems: Experiment, Theory, and Current Challenges. *Chem. Rev.* **2016,** *116*, 7529-7550.
(69) Li, X.; Tully, J. C.; Schlegel, H. B.; Frisch, M. J., Ab initio Ehrenfest dynamics. *J. Chem. Phys.* **2005,** *123*, 084106.
(70) Meredith, N. Y.; Borsley, S.; Smolyar, I. V.; Nichol, G. S.; Baker, C. M.; Ling, K. B.; Cockroft, S. L., Dissecting Solvent Effects on Hydrogen Bonding. *Angew. Chem. Int. Ed.* **2022,** *61*, e202206604.
(71) Li, G.; Cui, Q., pKa Calculations with QM/MM Free Energy Perturbations. *J. Phys. Chem. B* **2003,** *107*, 14521-14528.
(72) Riccardi, D.; Schaefer, P.; Cui, Q., pKa Calculations in Solution and Proteins with QM/MM Free Energy Perturbation Simulations: A Quantitative Test of QM/MM Protocols. *J. Phys. Chem. B* **2005,** *109*, 17715-17733.
(73) Ho, J., Are thermodynamic cycles necessary for continuum solvent calculation of pKas and reduction potentials? *Phys. Chem. Chem. Phys.* **2015,** *17*, 2859-2868.
(74) Dutra, F. R.; Silva, C. d. S.; Custodio, R., On the Accuracy of the Direct Method to Calculate pKa from Electronic Structure Calculations. *J. Phys. Chem. A* **2021,** *125*, 65-73.
(75) Takano, Y.; Houk, K. N., Benchmarking the Conductor-like Polarizable Continuum Model (CPCM) for Aqueous Solvation Free Energies of Neutral and Ionic Organic Molecules. *J. Chem. Theory Comput.* **2005,** *1*, 70-77.
(76) Ho, J.; Ertem, M. Z., Calculating Free Energy Changes in Continuum Solvation Models. *J. Phys. Chem. B* **2016,** *120*, 1319-1329.

(77) Li, T. E.; Hammes-Schiffer, S., Electronic Born–Oppenheimer approximation in nuclear-electronic orbital dynamics. *J. Chem. Phys.* **2023,** *158*, 114118.

For Table of Contents Only