Chemical Physics Reviews

RESEARCH ARTICLE | OCTOBER 09 2024

Constrained nuclear-electronic orbital QM/MM approach for simulating complex systems with quantum nuclear delocalization effects incorporated

FREE

Xianyuan Zhao ; Zehua Chen ; Yang Yang

Chem. Phys. Rev. 5, 041401 (2024)
https://doi.org/10.1063/5.0226271

Chemical Physics Reviews ARTICLE pubs.aip.org/aip/cpr

Constrained nuclear-electronic orbital QM/MM approach for simulating complex systems with quantum nuclear delocalization effects incorporated

Cite as: Chem. Phys. Rev. 5, 041401 (2024); doi: 10.1063/5.0226271
Submitted: 30 June 2024 · Accepted: 10 September 2024 ·
Published Online: 9 October 2024

View Online Export Citation CrossMark

Xianyuan Zhao, Zehua Chen, and Yang Yang[a)]

AFFILIATIONS

Theoretical Chemistry Institute and Department of Chemistry, University of Wisconsin-Madison, 1101 University Avenue, Madison, Wisconsin 53706, USA

[a)]Author to whom correspondence should be addressed: yyang222@wisc.edu

ABSTRACT

The hybrid quantum mechanics/molecular mechanics (QM/MM) approach, which combines the accuracy of QM methods with the efficiency of MM methods, is widely used in the study of complex systems. However, past QM/MM implementations often neglect or face challenges in addressing nuclear quantum effects, despite their crucial role in many key chemical and biological processes. Recently, our group developed the constrained nuclear-electronic orbital (CNEO) theory, a cost-efficient approach that accurately addresses nuclear quantum effects, especially quantum nuclear delocalization effects. In this work, we integrate CNEO with the QM/MM approach through the electrostatic embedding scheme and apply the resulting CNEO QM/MM to two hydrogen-bonded complexes. We find that both solvation effects and nuclear quantum effects significantly impact hydrogen bond structures and dynamics. Notably, in the glutamic acid–glutamate complex, which mimics a common low barrier hydrogen bond in biological systems, CNEO QM/MM accurately predicts nearly equal proton sharing between the two residues. With an accurate description of both quantum nuclear delocalization effects and environmental effects, CNEO QM/MM is a promising new approach for simulating complex chemical and biological systems.

I. INTRODUCTION

Hybrid quantum mechanics/molecular mechanics (QM/MM) is a powerful tool in computational chemistry.[1–4] It enables the investigation of intricate chemical properties and processes within complex systems and has been widely used in the study of biological problems, including enzymatic processes[4–6] and drug design.[7–9] Additionally, it has been used in various other fields such as heterogeneous catalysis[10,11] and nanochemistry.[12,13]

QM/MM is unique in its multiscale nature, where higher-level accurate QM methods are applied to regions of primary interest while lower-level cost-effective MM methods are used for the surrounding environment, thereby minimizing the total computational expense. Although QM/MM is less accurate than pure QM methods and presents challenges such as the proper partitioning of QM and MM regions,[14,15] managing QM/MM boundary treatments,[16–20] and addressing overpolarization issues of QM electron densities near the MM region,[21–24] it remains a preferred method for studying complex systems due to its balanced accuracy and computational efficiency.

Despite the remarkable achievements of QM/MM in practical applications,[12,25–28] the majority of current approaches still treat nuclei in the key QM region classically, resulting in the neglect of nuclear quantum effects. This neglect is particularly problematic in systems where hydrogen, the lightest element, plays a significant role, as seen in many enzymatic reactions involving proton transfer, hydrogen atom transfer, and/or hydride transfer processes.[29–33]

To address this challenge, several theories have been developed to conduct QM/MM calculations with nuclear quantum effects included. One such successful approach is path-integral-based methods, which represent quantum systems using ensembles of replicas connected by harmonic springs.[34–40] Path-integral-based QM/MM methods have been successfully applied to study proton transfer,[41] hydride transfer,[42] and RNA cleavage reactions.[43] Although a major limitation of path-integral based methods is the high computational cost, especially if *ab* *initio* potential energy surfaces for the QM part are to be used, recent developments have been able to accelerate the calculations and reduce their computational costs.[44–48]

Another promising approach is the nuclear-electronic orbital (NEO) method, which employs multicomponent wave functions to simultaneously describe the quantum behavior of both nuclei and electrons.[49–57] When integrated with QM/MM, NEO has provided insights into the impact of nuclear quantum effects on molecular geometries in condensed phases.[58–60] Although static NEO calculations are limited by the assumption of instantaneous quantum nuclear response to the motion of classical nuclei,[58,61,62] the recently developed real-time NEO QM/MM can address this limitation and incorporate nonadiabaticity between nuclei and electrons, thus offering insights into short-time vibronic dynamics.[59,63,64]

Additionally, semi-classical trajectory methods have been integrated with QM/MM calculations with nuclear quantum effects incorporated.[65] They have been utilized to simulate the vibrational spectra of small biological molecules in condense phases.[65]

Recently, our group developed the constrained nuclear-electronic orbital (CNEO)[66,67] theory to incorporate nuclear quantum effects, particularly quantum nuclear delocalization effects, into classical molecular simulations through a quantum-corrected effective potential energy surface.[68,69] CNEO shows great potential to be a widely used method for its simple physical picture, high computational efficiency, and accuracy for describing quantum nuclear delocalization effects.[70–76] Due to its similarity to conventional electronic structure methods, with the addition of a more physically accurate quantum delocalized nuclear picture, CNEO is naturally capable of being integrated with the QM/MM framework.

In this work, we develop such an integration using the QM/MM electrostatic embedding scheme. By studying two bimolecular complex systems, one of which is of strong biological relevance, we show that CNEO QM/MM outperforms conventional QM/MM in describing hydrogen bonds and hydrogen-bond dynamics in the condensed phase, aligning well with experimental evidence.

II. METHODS

A. Conventional electrostatic QM/MM embedding scheme

In QM/MM calculations, a key aspect is the effective description of the interactions between the QM and MM regions. Two major embedding schemes are commonly used: mechanical embedding and electrostatic embedding.[1–4,77] In general, the electrostatic embedding scheme offers greater accuracy and is more widely used in computations.[1,2,78,79] In this scheme, the QM system is influenced by the electrostatic potential provided by the MM environment, and the MM portion interacts with the charge obtained from the quantum mechanical calculations of the QM system. Additionally, building upon the electrostatic scheme, polarization of the MM system may be considered through polarized embedding using polarizable force fields.[1,80,81]

In the conventional QM/MM electrostatic embedding scheme, the total energy of the whole system can be decomposed into three parts,

$$E_{\text{system}} = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM–MM}}, \tag{1}$$

where E_{QM} and E_{MM} are the energies of the QM and MM regions, respectively, and $E_{\text{QM–MM}}$ represents the QM–MM interaction energy. When QM and MM atoms interact only through non-bonded interactions, $E_{\text{QM–MM}}$ mainly includes two terms: the electrostatic interactions $E^{\text{electrostatic}}_{\text{QM–MM}}$ and the van der Waals interactions $E^{\text{vdW}}_{\text{QM–MM}}$. However, when there are covalent bonds connecting QM and MM atoms, special considerations on the boundary are needed.[4,16–18,21,82–85] In this development, we will focus our discussion on cases where there are no such covalent bonds.

The van der Waals term $E^{\text{vdW}}_{\text{QM–MM}}$ is easier to deal with. It describes both the short-range repulsion and dispersion interactions between QM and MM atoms, and it is often modeled with the Lennard–Jones (L-J) potential

$$E^{\text{vdW}}_{\text{QM–MM}} = \sum_{A}^{N_{\text{QM}}} \sum_{M}^{N_{\text{MM}}} \varepsilon_{AM} \left[\left(\frac{\sigma_{AM}}{|\mathbf{R}_A - \mathbf{R}_M|} \right)^{12} - \left(\frac{\sigma_{AM}}{|\mathbf{R}_A - \mathbf{R}_M|} \right)^{6} \right]. \tag{2}$$

Here, ε and σ are pairwise L-J parameters, $\mathbf{R}$ is the position of nuclei, N_{QM} is the number of classical nuclei in the QM region, and N_{MM} is the number of atoms in the MM region.

The electrostatic interaction $E^{\text{electrostatic}}_{\text{QM–MM}}$ describes the Coulombic interactions between the QM system and MM charges. Specifically, it usually includes the Coulombic interactions of electron density and classical nuclear point charges in the QM region with MM charges, denoted by $E^{\text{electrostatic}}_{\text{e–MM}}$ and $E^{\text{electrostatic}}_{\text{nuc–MM}}$, respectively,

$$\begin{aligned} E^{\text{electrostatic}}_{\text{QM–MM}} &= E^{\text{electrostatic}}_{\text{e–MM}} + E^{\text{electrostatic}}_{\text{nuc–MM}} \\ &= -\int d\mathbf{r} V^{\text{ext}}_{\text{MM}}(\mathbf{r}) \rho^{\text{e}}(\mathbf{r}) + \sum_{A}^{N_{\text{QM}}} V^{\text{ext}}_{\text{MM}}(\mathbf{R}_A) Z_A. \end{aligned} \tag{3}$$

Here, $\rho^{\text{e}}(\mathbf{r})$ is the electron density, Z_A is the nuclear charge of the Ath nucleus, and $V^{\text{ext}}_{\text{MM}}(\mathbf{r})$ is the external potential produced by MM charges. Usually, the MM charges are represented by point charges and the MM potential can be expressed as

$$V^{\text{ext}}_{\text{MM}}(\mathbf{r}) = \sum_{M}^{N_{\text{MM}}} \frac{q_M}{|\mathbf{r} - \mathbf{R}_M|}, \tag{4}$$

in which q_M is the effective charge of the Mth MM atom. Note that instead of point charges, Gaussian charges can also be used, which have been shown to be able to avoid overpolarization of the QM electron density.[17,21,22,84]

Because the QM/MM electrostatic interaction energy depends on the electron density, the solution to the electron density must come from the variational optimization of $E_{\text{QM}} + E^{\text{electrostatic}}_{\text{QM–MM}}$. In practical calculations, if Kohn–Sham density functional theory (DFT) is used, the Kohn–Sham equation will incorporate the MM electrostatic potential $V^{\text{ext}}_{\text{MM}}(\mathbf{r})$ in addition to the external potential generated by classical nuclei in the QM region

$$\begin{aligned} &\left[-\frac{1}{2}\nabla^2 + \int d\mathbf{r}' \frac{\rho^{\text{e}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - V^{\text{ext}}_{\text{MM}}(\mathbf{r}) - \sum_{A}^{N_c} \frac{Z_A}{|\mathbf{r} - \mathbf{r}_N|} + V^{\text{e}}_{\text{xc}}(\mathbf{r}) \right] \psi^{\text{e}}_i(\mathbf{r}) \\ &= \varepsilon^{\text{e}}_i \psi^{\text{e}}_i(\mathbf{r}). \end{aligned} \tag{5}$$

Upon convergence of self-consistent field (SCF) calculations for $E_{\text{QM}} + E^{\text{electrostatic}}_{\text{QM–MM}}$, the total energy E_{system} can be calculated by adding the MM energy, as well as the van der Waals term $E^{\text{vdW}}_{\text{QM–MM}}$. Afterward, the forces on QM and MM atoms can be calculated through analytic gradient expressions of the total energy with respect to QM and MM coordinates.

B. Past development of CNEO-DFT

In the past few years, our group developed the CNEO framework to incorporate nuclear quantum effects, especially quantum nuclear delocalization effects, into quantum chemical calculations and molecular dynamics (MD) simulations.[66,67,70–73] This is achieved within the multicomponent quantum chemistry framework by imposing positional constraints on quantum nuclei,

$$\langle \psi_I^{\rm n} | \hat{\mathbf{r}}_I | \psi_I^{\rm n} \rangle = \mathbf{R}_I, \tag{6}$$

where $\psi_I^{\rm n}$ is the nuclear orbital of the Ith quantum nucleus, and $\hat{\mathbf{r}}_I$ and $\mathbf{R}_I$, respectively, are its associated quantum position operator and classical position specified by the molecular geometry. In this way, CNEO treats nuclei quantum mechanically but also retains the intuitive classical molecular picture.

With the introduction of constraints on the expectation value of quantum nuclear position operators, the multicomponent electronic Kohn–Sham equation remains the same as conventional NEO theory,

$$\left[-\frac{1}{2}\nabla^2 + \int d\mathbf{r}' \frac{\rho^{\rm e}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \sum_A^{N_{\rm c}} \frac{Z_A}{|\mathbf{r}-\mathbf{r}_N|} - \sum_I^{N_{\rm q}} Z_I \int d\mathbf{r}' \frac{\rho_I^{\rm n}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V_{\rm xc}^{\rm e}(\mathbf{r}) \right] \psi_i^{\rm e}(\mathbf{r}) = \varepsilon_i^{\rm e} \psi_i^{\rm e}(\mathbf{r}), \tag{7}$$

where $\rho_I^{\rm n}$ denotes the density of the Ith quantum nucleus, and Z_I denotes its charge. $N_{\rm c}$ and $N_{\rm q}$ are the total numbers of classical and quantum nuclei, respectively. The terms in the bracket represent in order the electronic kinetic energy term, Hartree potential, external potential due to classical nuclei, external potential due to quantum nuclei, and exchange-correlation potential for electrons.

In contrast, the nuclear Kohn–Sham equation is modified with an extra term $(\mathbf{f}_I \cdot \mathbf{r})$ associated with the constraint on the expectation position[66]

$$\left[-\frac{1}{2m_I}\nabla^2 - Z_I \int d\mathbf{r}' \frac{\rho^{\rm e}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + Z_I \sum_A^{N_{\rm c}} \frac{Z_A}{|\mathbf{r}-\mathbf{r}_A|} + Z_I \sum_{J\neq I}^{N_{\rm q}} Z_J \int d\mathbf{r}' \frac{\rho_J^{\rm n}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V_{{\rm c},I}^{\rm n}(\mathbf{r}) + \mathbf{f}_I \cdot \mathbf{r} \right] \psi_I^{\rm n}(\mathbf{r}) = \varepsilon_I^{\rm n} \psi_I^{\rm n}(\mathbf{r}). \tag{8}$$

Here, the first four terms corresponds to those in Eq. (7) but are now for nuclei. $V_{{\rm c},I}^{\rm n}(\mathbf{r})$ is the correlation potential for the Ith quantum nucleus. Note that there is no nuclear exchange within CNEO because of the distinguishable nucleus assumption and nuclear self-Coulomb is explicitly excluded.[66] The Lagrange multiplier $\mathbf{f}_I$ needs to be solved iteratively together with electronic and nuclear orbitals, subject to the geometric constraints on quantum nuclear expectation positions via Eq. (6). The converged orbitals can be subsequently used to evaluate the multicomponent energies as a function of both classical and quantum nuclear positions, leading to quantum-corrected effective potential energy surfaces.

Analytic gradients[67] and Hessians[70] with respect to both classical and quantum nuclear positions have also been developed.

C. Development of CNEO-DFT QM/MM

For the current CNEO-DFT QM/MM development, because some or all nuclei in the QM region are now described quantum mechanically, additional terms in the QM–MM interaction energy $E_{\rm QM-MM}$ will arise due to the interactions between the quantum nuclei and the MM environment. Specifically, these terms include both the electrostatic interactions and the van der Waals interactions between quantum nuclei and MM atoms.

For the simpler van der Waals interactions, the additional quantum nuclei-MM (q-MM) term can be calculated with

$$E_{\rm q-MM}^{\rm vdW} = \sum_I^{N_{\rm q}} \sum_M^{N_{\rm MM}} \varepsilon_{IM} \left[\left(\frac{\sigma_{IM}}{|\mathbf{R}_I - \mathbf{R}_M|} \right)^{12} - \left(\frac{\sigma_{IM}}{|\mathbf{R}_I - \mathbf{R}_M|} \right)^{6} \right]. \tag{9}$$

Then, the total QM–MM van der Waals interaction energy becomes

$$E_{\rm QM-MM}^{\rm vdW} = E_{\rm c-MM}^{\rm vdW} + E_{\rm q-MM}^{\rm vdW}. \tag{10}$$

where we now use c-MM to denote the interactions between the classical nuclei in the QM region and MM atoms.

For electrostatic interactions, the additional term can be calculated with

$$E_{\rm q-MM}^{\rm eletrostatic} = \sum_I^{N_{\rm q}} Z_I \int d\mathbf{r} V_{\rm MM}^{\rm ext}(\mathbf{r}) \rho_I^{\rm n}(\mathbf{r}), \tag{11}$$

and the total electrostatic QM–MM interaction energy becomes

$$E_{\rm QM-MM}^{\rm electrostatic} = E_{\rm e-MM}^{\rm electrostatic} + E_{\rm q-MM}^{\rm electrostatic} + E_{\rm c-MM}^{\rm electrostatic}. \tag{12}$$

Similar to the conventional QM/MM approach, the variational energy minimization of $E_{\rm QM} + E_{\rm QM-MM}^{\rm electrostatic}$ with respect to QM densities leads to Kohn–Sham equations for quantum particles in the QM region. The resulting electronic and nuclear Kohn–Sham equations are highly similar to those in CNEO-DFT, except that the MM potential now enters both the electronic equation [Eq. (7)] and the nuclear equation [Eq. (8)] as an additional external potential term.

In CNEO-DFT QM/MM, analytic gradients with respect to the displacement of classical nuclei in the QM region, the displacement of the expectation positions of quantum nuclei in the QM region, and the displacement of MM atom positions can be derived in a similar way to what has been done for conventional DFT QM/MM. These details are provided in Sec. 3 of the supplementary material.

D. Computational details

We implemented the CNEO-DFT QM/MM in our locally modified version of PySCF,[86–88] which is available through our group GitHub page.[89] Molecular dynamics simulations were carried out with GROMACS.[90–92] In all the following calculations, the aug-cc-pVDZ electronic basis set[93] was used for both CNEO-DFT and conventional DFT. For the glutamic acid–glutamate complex, density fitting with the aug-cc-pVDZ-RI auxiliary basis set was used for electronic integrals.[94,95] With CNEO-DFT, all hydrogen atoms in the QM region were treated as quantum nuclei with the PB4D protonic basis set.[96]

The B3LYP[97–99] electronic exchange-correlation functional was used and no electron–proton correlation (epc) functional was used. Note that our preliminary tests found that the currently developed epc functionals[100–103] tend to make negligible difference to molecular geometries and vibrational frequencies in the CNEO framework. Therefore, we chose to present results without epc functionals here and leave the detailed investigation of epc effects for future work. For MM-related calculations, we utilized a modified OPLS all-atom force field for the hydrogen-bonded complexes[104–106] and a modified TIP3P water.[107] The polar hydrogens (hydrogen in N–H and O–H) of the hydrogen-bonded molecules were assigned Lennard–Jones coefficients from Ref. 108. Additional computational details can be found in Sec. 1 of the supplementary material. Regarding computational cost, we observed that incorporating nuclear quantum effects in CNEO-DFT QM/MM results in only a minor increase in computational cost (typically 10% to 20% longer wall time) compared to conventional DFT-based QM/MM. This comparison is summarized in Sec. 2 of the supplementary material.

III. RESULTS AND DISCUSSION

We first considered a phenol–water complex (Fig. 1), which is a hydrogen bonded system that has been studied in the past during the development NEO-DFT QM/MM.[58] In this system, the hydrogen bond to be investigated is the one between the hydrogen atom in the phenol hydroxyl group and the oxygen atom in the water. Therefore, the QM region constitutes the phenol molecule and the hydrogen-bonded water molecule.

A. Phenol–water complex

Following the procedure in Ref. 58, we investigated four key geometric properties for the optimized geometries of the complex in both gas phase and in a water droplet: the OH bond length of the phenol hydroxyl group (O–H), the hydrogen bond distance (O…H), the distance between the two oxygen atoms (O…O), and the ∠OHO bond angle. The results from the methods based on pure MM, DFT, and CNEO-DFT are presented in Table I.

For all three methods, the hydroxyl O–H bond length is consistently very close to 1 Å. Additionally, switching from gas phase to aqueous phase has little impact on the distance. These results indicate that all three methods can describe the equilibrium bond length well. In contrast, the hydrogen bond O…H distance varies dramatically with the underlying method. Specifically, pure MM predicts the largest distance with 1.93 Å in the gas phase and 1.91 Å in the aqueous phase, whereas CNEO predicts the smallest distance with 1.82 Å in the gas phase and 1.59 Å in the aqueous phase. Interestingly, switching from the gas phase to the aqueous phase barely changes the pure MM results ($\Delta = -0.02$ Å), but it leads to a large distance decrease for both DFT QM/MM ($\Delta = -0.15$ Å) and CNEO-DFT QM/MM ($\Delta = -0.23$ Å). As to the O…O distance, because the ∠OHO bond angle is almost linear and thus the O…O distance is roughly the sum of O–H and O…H distances, the behavior of the O…O distance is similar to that of the O…H distance.

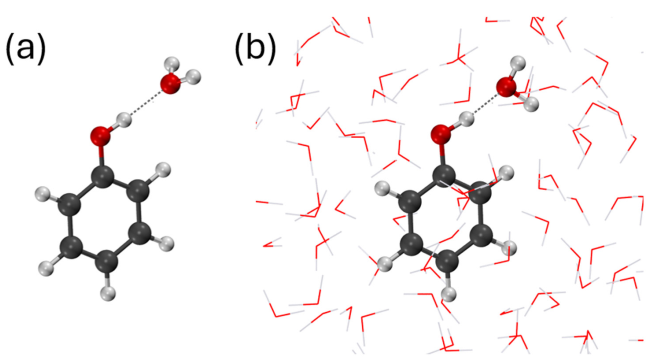

FIG. 1. Phenol–water complex in the (a) gas phase and (b) aqueous phase.

The long hydrogen bond distance by pure MM and its insensitivity to environmental change is a manifestation of its failure in describing hydrogen bonds. This is because using only Coulombic and van der Waals interaction terms in pure MM tends to inadequately capture the intricate nature of hydrogen bonds.[109,110] In contrast, both DFT QM/MM and CNEO-DFT QM/MM can qualitatively describe the significant environmental effect, although CNEO-DFT QM/MM predicts shorter O…H and O…O distances than DFT QM/MM by 0.14 and 0.10 Å, respectively. Additionally, in CNEO-DFT QM/MM, the hydrogen atom is located closer to the center of the two oxygen atoms. These results are qualitatively consistent with previous computational studies that also found neutral hydrogen bond complexes contract when solvated by water.[111]

One notable point is that for this type of static QM/MM geometry optimization, theoretically, CNEO QM/MM and NEO QM/MM should yield the same equilibrium geometry results. This can be confirmed by comparing with the data in Ref. 58, which shows that our CNEO QM/MM results match well the optimized geometric parameters obtained from NEO QM/MM, with negligible differences attributed to different basis sets and MM water environment. This consistency in results is a strong indicator that both developments have correctly implemented their respective theories.

Next, we performed MD simulations on the phenol–water complex system starting from optimized geometries obtained from the respective methods. Within the *NVT* ensemble at 270 K, all MD simulations are performed for 10 ps, with the first 2 ps used for equilibration and the remaining 8 ps for data collection. The geometric properties as well as their standard deviation during the latter 8 ps MD simulations are also shown in Table I. Note that we only performed the simulation in the aqueous phase because in the gas phase, the phenol–water easily dissociates at the picosecond timescale.

Compared to the geometry optimization results, MD simulations barely change the hydroxyl O–H bond distance on the phenol group and the bond distance standard deviations remain small. This is because the simulation temperature is low compared to the bond strength and thus show negligible thermal fluctuation effects. In contrast, the hydrogen bonded O…H distance becomes larger with MD simulations as a result of the significant thermal fluctuation, which also makes its standard deviation significantly larger than that of the O–H bond distance. Note that with pure MM, the average O…H distance increased by about 1.4 Å and the average ∠OHO reduces to about 125°. This large change is due to the inadequately weak hydrogen bond being broken and re-formed many times during the 8 ps sampling time. Note that all *NVT* simulations for this complex were conducted at a relatively low temperature of 270 K. We anticipate that at room temperature, hydrogen bonds would break and form more frequently, potentially altering the water molecule to which phenol is

TABLE I. Geometric properties of phenol–water complex.

Environment	Method	Distance (Å)			∠OHO (°)
		O–H	O…H	O…O	
Geometry optimization (gas phase)	MM	0.96	1.93	2.88	177
	DFT	0.97	1.88	2.85	173
	CNEO-DFT	1.01	1.82	2.82	172
Geometry optimization (aqueous phase)	Full MM	0.96	1.91	2.85	166
	DFT QM/MM	0.99	1.73	2.71	169
	CNEO-DFT QM/MM	1.03	1.59	2.61	171
Molecular dynamics (aqueous phase)	Full MM	0.95 ± 0.04	3.33 ± 0.59	3.92 ± 0.56	124 ± 17
	DFT QM/MM	0.98 ± 0.02	1.87 ± 0.20	2.81 ± 0.18	161 ± 9
	CNEO-DFT QM/MM	1.03 ± 0.03	1.67 ± 0.15	2.70 ± 0.15	166 ± 7

hydrogen-bonded. In contrast, the hydrogen bond predicted by both DFT QM/MM and CNEO-DFT QM/MM remains unbroken and is much less affected by thermal fluctuation. Specifically, for the O…H distance, an increase of 0.14 and 0.08 Å is observed for DFT QM/MM and CNEO-DFT QM/MM, respectively, and for the bond angle, DFT QM/MM observes a change of about 8°, which is slightly larger than the 5° change observed by CNEO-DFT QM/MM. Compared to conventional DFT results, the reduced susceptibility to thermal fluctuations in CNEO-DFT suggests a stronger hydrogen bond, attributed to its quantum treatment of hydrogen nuclei. Given the great performance of CNEO-DFT in describing hydrogen-bonded systems from the past studies,[70,72,73,76] this stronger hydrogen bond is likely to be physically correct, although the absence of experimental data on the position of the hydrogen atom makes it challenging to reach a definitive conclusion.

We note that for MD simulations, CNEO-DFT QM/MM and NEO-DFT QM/MM will be significantly different. CNEO-DFT QM/MM does not assume the instantaneous response of the quantum nuclei to the motion of classical nuclei.[66,70] Therefore, it carries a more physically correct picture and can accurately describe the O–H and O…H vibrational pictures.[71–73] Nonetheless, the recently developed real-time NEO QM/MM dynamics work can mitigate the problem of NEO-DFT QM/MM with the desired capability of describing electron-nuclear nonadiabaticity, although the computational cost will be much higher.[59]

B. Glutamic acid–glutamate complex

To further demonstrate the power of CNEO-DFT QM/MM and its potential in biological studies, we next investigated a glutamic acid–glutamate complex (Fig. 2) in both gas phase and aqueous phase. This complex is a typical low-barrier hydrogen bond system[112–115] in which two glutamate anions share a proton, and it is known to play a vital role in some biological systyems.[114,116–118] For example, in human transketolase, this complex (between E366' and E160) is believed to participate in a proton transfer wire, which is the structural origin of the enzyme's cooperativity,[114] and in bacteriorhodopsin, a proton pump that uses photon energy to establish transcellular proton gradient, this complex (between E194 and E204) is directly involved in the key pump process by releasing the shared proton to the extracellular environment.[116–118]

Due to the high significance in biological systems, the location of the shared proton in this complex and its real-time dynamics is of particular interest. For human transketolase, high-resolution x-ray crystallography concludes that the proton is almost equally shared by the two carboxylic oxygens of E366' and E160 residues, and the O…O distance is 2.55 Å,[114] which is much shorter than that of a normal hydrogen bond (2.7–3.1 Å). Note that these high-resolution x-ray crystallography studies can obtain hydrogen positions, which are different from the conventional impression that hydrogen positions cannot be obtained by x-ray crystallography. For bacteriorhodopsin, although the location of the proton is not exactly known, the determined O…O distance is 2.48 Å,[116] which is even shorter and may imply a more equally shared proton.

To mimic the real biological environment, we should ideally embed the two amino acids in relatively rigid protein backbones. However, for this proof-of-concept study, we simplified the problem by applying distance constraints to two pairs of carbon atoms according to the experimental x-ray structure of human transketolase.[119] Specifically, the distance between the two α-carbons is constrained to 7.84 Å, and the distance between the two carboxylic carbons is constrained to 9.01 Å (Fig. 2). These constraints serve a role similar to the

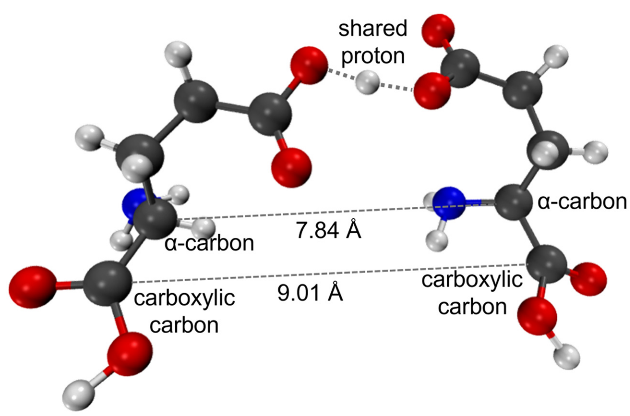

FIG. 2. Structure of the glutamic acid–glutamate complex and the applied distance constraints between two α-carbons and two carboxylic carbons to mimic the real structure in the enzyme environment.

protein backbones, which preserve the intercarboxylic hydrogen bond and prevent strong conformational changes that would be unnatural in a protein environment. Additionally, we acknowledge that in a real biological setting, the system would not be fully immersed in an aqueous environment, which is another difference between our current treatment and actual biological conditions.

As with the phenol–water complex, we first optimized the geometries of the glutamic acid–glutamate complex in both gas phase and aqueous phase. As shown in Table II, classical MM again yields the longest O…O distances (around 2.65 Å) and the largest difference between O–H and O…H distances (around 0.7 Å), thus incorrectly predicting the proton to be owned by one residue. With DFT, the O…O distance is predicted to be significantly shorter (2.46–2.49 Å) and the length difference between O…H and O–H becomes smaller, which is 0.29 Å in the gas phase and 0.40 Å in solution. For CNEO-DFT, the O…O distance is similarly short, and the O…H and O–H distance difference further reduces to about 0.12 Å in both phases, which aligns well with the experimental result that the H is equally shared by the two residues,[114] although as acknowledged above, there are some differences between the current setup and real biological conditions.

There is a major difference between the current glutamic acid–glutamate complex and the previous phenol–water complex: for the phenol–water complex, the hydrogen bond distance and the O…O distance shorten in aqueous phase, whereas for the glutamic acid–glutamate complex, these distances barely change with CNEO-DFT or even becomes slightly longer with DFT. This phenomenon has been observed in the past and it was attributed to the differences between neutral complexes and negatively charged complexes.[111] Heuristically speaking, the attractive interactions between the solvent and solute compress the solute and thereby shorten the hydrogen bonds, as observed in neutral complexes. However, in charged complexes, the dipole–ion interactions are much stronger than the dipole–dipole interactions present in neutral complexes. Consequently, the hydrogen bonds in charged complexes are already much shorter in the gas phase, and placing the charged complex in a polar solvent makes little difference to the hydrogen bond length. However, more rigorously, the bond length changes reflect an interplay between electronic effects, nuclear quantum effects, and solvation effects. The collective impact of these effects leads CNEO-DFT QM/MM to predict that the hydrogen bond barely changes upon solvation for the glutamic acid–glutamate complex, whereas DFT QM/MM predicts slightly elongated O…H and O…O differences (by ~0.05 Å).

Next, molecular dynamics simulations were carried out with each method from the corresponding optimized geometries, and the geometric properties are also reported in Table II. With pure MM, the hydrogen bond becomes considerably longer (~0.15 Å) and inadequately weaker, again indicating the failure of the force field in describing hydrogen bonds. In contrast, the hydrogen bonds treated by both DFT QM/MM and CNEO-DFT QM/MM are stronger with the increase in the hydrogen bond length always within 0.1 Å.

Since the bond lengths statistics within the *NVT* ensemble do not provide dynamical information in this low-barrier hydrogen bonded system, in order to investigate the important proton transfer dynamics, we performed 4 ps *NVE* simulations with each method and plotted the O_1–H and O_2–H (O_1 and O_2 are the two carboxylic oxygens sharing the proton) distances as a function of time. The results are shown in Fig. 3.

Due to the inability to describe bond formation and bond dissociation, proton transfer never occurs in classical MM. The bonded O–H always vibrates around its local minimum with a small length variance whereas the hydrogen bonded O…H distance can fluctuate significantly. This fluctuation becomes smaller in the aqueous phase owing to confinement from environmental molecules. In contrast, proton transfer can be observed with both DFT and CNEO-DFT *ab initio* molecular dynamics with the relative O_1–H and O_2–H distances swapped frequently during the gas phase simulation. It is noticeable that proton transfer is more frequent in CNEO-DFT than in DFT. Interestingly, in the aqueous phase, proton transfer is now nearly prohibited in DFT QM/MM simulations but can still occasionally take place with CNEO-DFT QM/MM. Although a possible reason for this difference between DFT and CNEO-DFT is the slightly smaller O…O

TABLE II. Geometric properties of glutamic acid–glutamate complex.

Environment	Method	Distance (Å)			∠OHO (°)
		O–H	O…H	O…O	
Geometry optimization (gas phase)	MM	0.97	1.69	2.65	170
	DFT	1.09	1.38	2.46	171
	CNEO-DFT	1.17	1.30	2.46	170
Geometry optimization (aqueous phase)	Full MM	0.97	1.68	2.64	170
	DFT QM/MM	1.05	1.45	2.49	170
	CNEO-DFT QM/MM	1.17	1.29	2.45	170
Molecular dynamics (gas phase)	MM	0.97 ± 0.03	1.83 ± 0.21	2.74 ± 0.19	152 ± 27
	DFT	1.11 ± 0.06	1.38 ± 0.11	2.47 ± 0.07	168 ± 6
	CNEO-DFT	1.16 ± 0.05	1.34 ± 0.09	2.48 ± 0.07	168 ± 5
Molecular dynamics (aqueous phase)	Full MM	0.96 ± 0.02	1.83 ± 0.17	2.75 ± 0.16	153 ± 26
	DFT QM/MM	1.05 ± 0.04	1.49 ± 0.13	2.53 ± 0.09	168 ± 7
	CNEO-DFT QM/MM	1.13 ± 0.06	1.38 ± 0.12	2.50 ± 0.08	170 ± 5

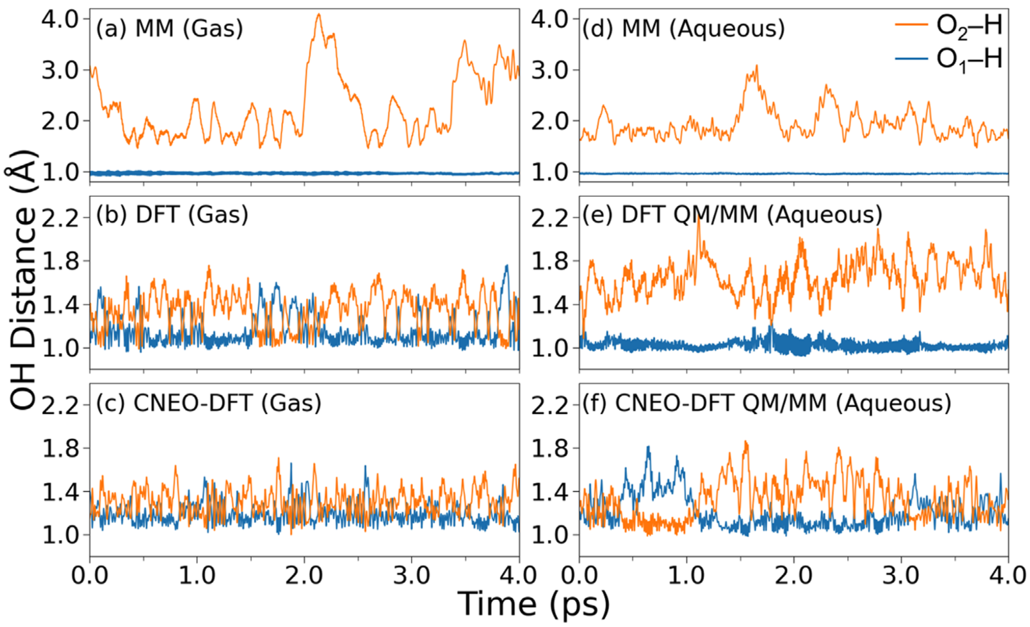

FIG. 3. Distances between the shared proton and its two adjacent oxygen atoms during *NVE* simulations of glutamic acid–glutamate complex in the gas and aqueous phases by classical molecular dynamics, DFT-based molecular dynamics, and CNEO-DFT-based molecular dynamics.

distance that facilitates hydrogen sharing and hydrogen transfer by CNEO-DFT, the major reason is that the incorporation of quantum nuclear delocalization effects in the CNEO effective surface can lower the proton-transfer barrier[72,73,75] and accelerate the proton transfer dynamics. The slower or prohibited proton transfers in aqueous phase compared to the gas phase may be explained by considering solvent fluctuations and their influence on the proton potential. In the gas phase, the absence of solvent leads to a relatively symmetric double well with similar depths on both sides. In contrast, in the aqueous phase, solvent fluctuations tend to create tilted double wells, and consequently, the proton is more likely to remain on one side of the well until significant changes in the solvent configuration cause the double well to tilt toward the other side, thereby facilitating proton transfer.[120]

With this dynamic information, we can now reinvestigate the bond length distributions in the prior *NVT* simulations and better interpret the proton location in the glutamic acid–glutamate complex. Because of the occurrence of proton transfer that weakens the identification of hydrogen bond donor and acceptor, we combined the bond length data of O–H and O…H and plotted them in an overall distribution in Fig. 4. In DFT simulations, because there is little to no proton transfer, the hydrogen bond tends to be asymmetric with the proton being more possessed by one oxygen, leading to two overlapping peaks in gas phase and more distinctly separated peaks in the aqueous phase. In contrast, with CNEO-DFT simulations, the overall distribution becomes a single peak in both gas phase and aqueous phase due to the much more frequent proton transfer. The peak positions are both at around 1.2 Å, and the average OH distances are 1.24 and 1.25 Å for gas phase and aqueous phase, respectively. This prediction is in great agreement with the experiment x-ray results on human transketolase,[119] in which the proton is shown to be equally shared by the two glutamate residue with the OH distance being 1.28 Å.

We further investigated the correlation between the OH distances and the O…O distance by plotting their joint probability in Fig. 5. The lower branch in each panel represents the bonded O–H distance while the higher branch represents the hydrogen bonded O…H distance. It can be observed that the smaller the O…O distance, the more likely that the two branches merge together, facilitating the proton transfer. This observation is consistent with the conventional understanding of proton transfer processes. However, in general, DFT and DFT QM/MM give larger O…O distances and more distinguishable O–H and

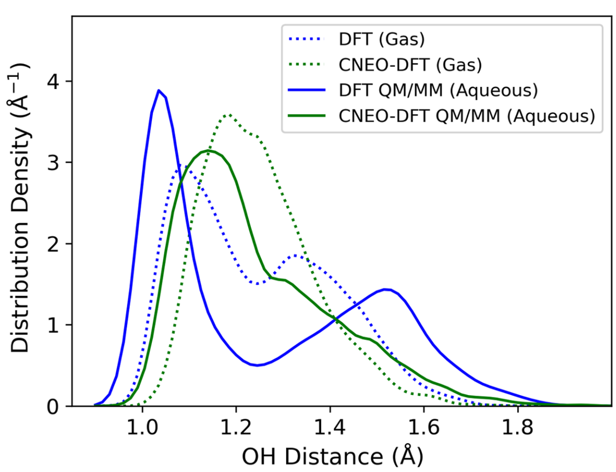

FIG. 4. Distance distributions between the shared proton and adjacent oxygen atoms of the glutamic acid–glutamate complex in the gas (dotted line) and aqueous (solid line) phases from DFT-based (blue) and CNEO-DFT-based (green) *NVT* simulations.

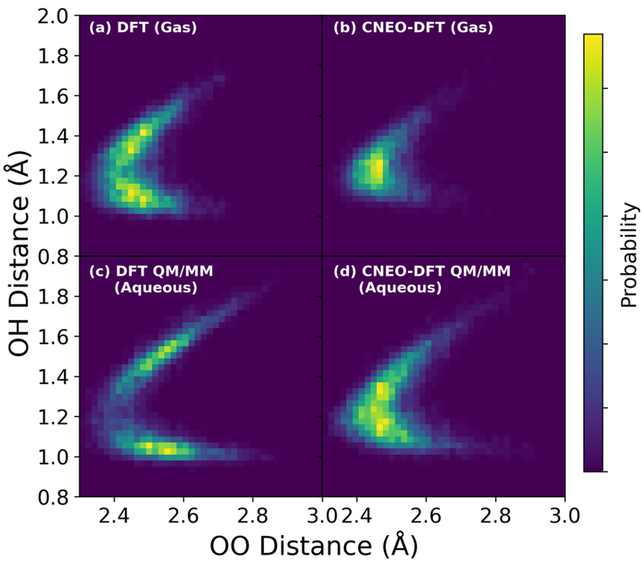

FIG. 5. Correlation between oxygen–proton (OH) and oxygen–oxygen (OO) distances in the glutamic acid–glutamate complex in the gas phase and aqueous phases from DFT-based and CNEO-DFT-based *NVT* simulations.

O…H distributions, whereas CNEO-DFT and CNEO-DFT QM/MM yield smaller O…O distances and more overlapped O–H and O…H branches that allow more proton transfers.

IV. CONCLUSIONS

In conclusion, we integrated CNEO with the QM/MM electrostatic embedding scheme and achieved the accurate and efficient incorporation of nuclear quantum effects, particularly quantum delocalization effects, in the QM region of QM/MM simulations. We applied the resulting CNEO QM/MM theory to the calculation of a phenol–water complex and a glutamic acid–glutamate complex in both gas phase and aqueous phase. We investigated the impact of nuclear quantum effects on both optimized geometries and molecular dynamics. For the neutral phenol–water complex, solvation reduces the hydrogen bond distance and the incorporation of nuclear quantum effects through CNEO-DFT leads to a shorter hydrogen bond than that of DFT simulations. In contrast, for the negatively charged glutamic acid–glutamate complex, solvation leads to a slight increase for the hydrogen bond distance as predicted by DFT and DFT QM/MM but a negligible change as predicted by CNEO-DFT and CNEO-DFT QM/MM. Through dynamics simulations, we observed much more frequent proton transfer in CNEO-DFT QM/MM simulations than in DFT QM/MM simulations for the glutamic acid–glutamate complex in both gas phase and aqueous phase due to the incorporation of quantum nuclear delocalization effects. Additionally, the location of the shared proton predicted by CNEO QM/MM is in great agreement with the experimental observations. All of these results demonstrate the significant impact of the solvation environment and quantum nuclear delocalization effects, both of which are key features of the CNEO QM/MM approach. As an accurate and efficient method, CNEO QM/MM holds great promise for future investigation of hydrogen-related processes in complex chemical and biological environments.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional computational details, benchmarks on computational cost, and detailed derivation of CNEO-DFT QM/MM energy, Kohn–Sham equations, and analytic gradients with both point charge and Gaussian charge models.

ACKNOWLEDGMENTS

The authors thank Yuzhe Zhang, Tanner Culpitt, and Lin Mei for helpful discussions. The authors are grateful for the funding support from the National Science Foundation under Grant No. 2238473. The computational resource support from Center for High Throughput Computing at the University of Wisconsin-Madison is acknowledged.[121]

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Xianyuan Zhao: Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (equal); Software (equal); Visualization (lead); Writing – original draft (lead); Writing – review & editing (supporting). **Zehua Chen:** Conceptualization (supporting); Formal analysis (supporting); Investigation (supporting); Methodology (equal); Software (equal); Supervision (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). **Yang Yang:** Conceptualization (lead); Formal analysis (supporting); Funding acquisition (lead); Investigation (supporting); Methodology (supporting); Supervision (lead); Writing – review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material and available from the corresponding author upon reasonable request.

REFERENCES

[1]A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme," J. Mol. Biol. **103**(2), 227–249 (1976).
[2]M. J. Field, P. A. Bash, and M. Karplus, "A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations," J. Comput. Chem. **11**(6), 700–733 (1990).
[3]H. Lin and D. G. Truhlar, "QM/MM: What have we learned, where are we, and where do we go from here?," Theor. Chem. Acc. **117**(2), 185–199 (2007).
[4]H. M. Senn and W. Thiel, "QM/MM methods for biomolecular systems," Angew. Chem., Int. Ed. **48**(7), 1198–1229 (2009).
[5]H. M. Senn and W. Thiel, "QM/MM studies of enzymes," Curr. Opin. Chem. Biol. **11**(2), 182–187 (2007).
[6]M. J. Ramos and P. A. Fernandes, "Computational enzymatic catalysis," Acc. Chem. Res. **41**(6), 689–698 (2008).
[7]A. Lodola and M. De Vivo, "The increasing role of QM/MM in drug discovery," in *Advances in Protein Chemistry and Structural Biology*, Structural and Mechanistic Enzymology Vol. 87, edited by C. Christov and T. Karabencheva-Christova (Academic Press, 2012), Chap. 11, pp 337–362.
[8]M. Taguchi, R. Oyama, M. Kaneso, and S. Hayashi, "Hybrid QM/MM free-energy evaluation of drug-resistant mutational effect on the binding of an

inhibitor indinavir to HIV-1 protease," J. Chem. Inf. Model. **62**(5), 1328–1344 (2022).
[9]B. Raghavan, M. Paulikat, K. Ahmad, L. Callea, A. Rizzi, E. Ippoliti, D. Mandelli, L. Bonati, M. De Vivo, and P. Carloni, "Drug design in the exascale era: A perspective from massively parallel QM/MM simulations," J. Chem. Inf. Model. **63**(12), 3647–3658 (2023).
[10]G. Adrian Bramley, O. Tomos Beynon, P. Viktorovich Stishenko, and A. James Logsdail, "The application of QM/MM simulations in heterogeneous catalysis," Phys. Chem. Chem. Phys. **25**(9), 6562–6585 (2023).
[11]N. Abidi and S. N. Steinmann, "An electrostatically embedded QM/MM scheme for electrified interfaces," ACS Appl. Mater. Interfaces **15**(20), 25009–25017 (2023).
[12]K.-S. Csizi and M. Reiher, "Universal QM/MM approaches for general nanoscale applications," WIREs Comput. Mol. Sci. **13**(4), e1656 (2023).
[13]K. Cui and J. R. Schmidt, "Enabling efficient and accurate computational studies of MOF reactivity via QM/MM and QM/QM methods," J. Phys. Chem. C **124**(19), 10550–10560 (2020).
[14]H. J. Kulik, J. Zhang, J. P. Klinman, and T. J. Martínez, "How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase," J. Phys. Chem. B **120**(44), 11381–11394 (2016).
[15]M. Karelina and H. J. Kulik, "Systematic quantum mechanical region determination in QM/MM simulation," J. Chem. Theory Comput. **13**(2), 563–576 (2017).
[16]Y. Zhang, T.-S. Lee, and W. Yang, "A pseudobond approach to combining quantum mechanical and molecular mechanical methods," J. Chem. Phys. **110**(1), 46–54 (1999).
[17]M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, "A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields," J. Chem. Phys. **110**(21), 10452–10467 (1999).
[18]P. Sherwood, "Hybrid quantum mechanics/molecular mechanics approaches," in *Modern Methods and Algorithms of Quantum Chemistry*, NIC series (John von Neumann Institute for Computing, Jülich, 2000), pp. 285–305.
[19]J. Pu, J. Gao, and D. G. Truhlar, "Generalized hybrid orbital (GHO) method for combining ab initio Hartree–Fock wave functions with molecular mechanics," J. Phys. Chem. A **108**(4), 632–650 (2004).
[20]J. Pu, J. Gao, and D. G. Truhlar, "Generalized hybrid-orbital method for combining density functional theory with molecular mechanicals," ChemPhysChem **6**(9), 1853–1865 (2005).
[21]D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodošček, and B. R. Brooks, "Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method," J. Chem. Phys. **117**(23), 10534–10547 (2002).
[22]T. Laino, F. Mohamed, A. Laio, and M. Parrinello, "An efficient real space multigrid QM/MM electrostatic coupling," J. Chem. Theory Comput. **1**(6), 1176–1184 (2005).
[23]C. Várnai, N. Bernstein, L. Mones, and G. Csányi, "Tests of an adaptive QM/MM calculation on free energy profiles of chemical reactions in solution," J. Phys. Chem. B **117**(40), 12202–12211 (2013).
[24]L. Mones, A. Jones, A. W. Götz, T. Laino, R. C. Walker, B. Leimkuhler, G. Csányi, and N. Bernstein, "The adaptive buffered force QM/MM method in the CP2K and AMBER software packages," J. Comput. Chem. **36**(9), 633–648 (2015).
[25]A. Shurki and A. Warshel, "Structure/function correlations of proteins using MM, QM/MM, and related approaches: Methods, concepts, pitfalls, and current progress," in *Advances in Protein Chemistry*, Protein Simulations Vol. 66 (Academic Press, 2003), pp 249–313.
[26]S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, "Progress in ab initio QM/MM free-energy simulations of electrostatic energies in proteins: Accelerated QM/MM studies of pKa, redox reactions and solvation free energies," J. Phys. Chem. B **113**(5), 1253–1272 (2009).
[27]X. Lu, D. Fang, S. Ito, Y. Okamoto, V. Ovchinnikov, and Q. Cui, "QM/MM free energy simulations: Recent progress and challenges," Mol. Simul. **42**(13), 1056–1078 (2016).
[28]Q. Cui, T. Pal, and L. Xie, "Biomolecular QM/MM simulations: What are some of the 'burning issues?,'" J. Phys. Chem. B **125**(3), 689–702 (2021).
[29]V. R. I. Kaila, M. I. Verkhovsky, and M. Wikström, "Proton-coupled electron transfer in cytochrome oxidase," Chem. Rev. **110**(12), 7062–7081 (2010).
[30]Z. D. Nagel and J. P. Klinman, "Update 1 of: Tunneling and dynamics in enzymatic hydride transfer," Chem. Rev. **110**(12), PR41–PR67 (2010).
[31]J. P. Klinman and K. Amnon, "Hydrogen tunneling links protein dynamics to enzyme catalysis," Annu. Rev. Biochem. **82**, 471–496 (2013).
[32]L. Wang, S. D. Fried, S. G. Boxer, and T. E. Markland, "Quantum delocalization of protons in the hydrogen-bond network of an enzyme active site," Proc. Natl. Acad. Sci. U. S. A. **111**(52), 18454–18459 (2014).
[33]S. Hammes-Schiffer, "Proton-coupled electron transfer: Moving together and charging forward," J. Am. Chem. Soc. **137**(28), 8860–8871 (2015).
[34]R. P. Feynman, A. R. Hibbs, and D. F. Styer, *Quantum Mechanics and Path Integrals* (Courier Corp., 2010).
[35]B. J. Berne, G. Ciccotti, and D. F. Coker, "Classical and quantum dynamics in condensed phase simulations," in *Proceedings of the International School of Physics* (World Scientific, Lerici, Villa Marigola, 1998).
[36]"Path integration via molecular dynamics," in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, NIC series, edited by M. E. Tuckerman (John von Neumann Institute for Computing, Jülich, 2002).
[37]J. Cao and G. A. Voth, "The formulation of quantum statistical mechanics based on the Feynman path centroid density. II. Dynamical properties," J. Chem. Phys. **100**(7), 5106–5117 (1994).
[38]S. Jang and G. A. Voth, "A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables," J. Chem. Phys. **111**(6), 2371–2384 (1999).
[39]I. R. Craig and D. E. Manolopoulos, "Quantum statistics and classical mechanics: Real time correlation functions from ring polymer molecular dynamics," J. Chem. Phys. **121**(8), 3368–3373 (2004).
[40]T. E. Markland and M. Ceriotti, "Nuclear quantum effects enter the mainstream," Nat. Rev. Chem. **2**(3), 0109 (2018).
[41]J. Gao, K.-Y. Wong, and D. T. Major, "Combined QM/MM and path integral simulations of kinetic isotope effects in the proton transfer reaction between nitroethane and acetate ion in water," J. Comput. Chem. **29**(4), 514–522 (2008).
[42]N. Boekelheide, R. Salomón-Ferrer, and T. F. Miller, "Dynamics and dissipation in enzyme catalysis," Proc. Natl. Acad. Sci. U. S. A. **108**(39), 16159–16163 (2011).
[43]T. J. Giese, J. Zeng, Ş. Ekesan, and D. M. York, "Combined QM/MM, machine learning path integral approach to compute free energy profiles and kinetic isotope effects in RNA cleavage reactions," J. Chem. Theory Comput. **18**(7), 4304–4317 (2022).
[44]V. Kapil, J. VandeVondele, and M. Ceriotti, "Accurate molecular dynamics and nuclear quantum effects at low cost by multiple steps in real and imaginary time: Using density functional theory to accelerate wavefunction methods," J. Chem. Phys. **144**(5), 054111 (2016).
[45]O. Marsalek and T. E. Markland, "*Ab initio* molecular dynamics with nuclear quantum effects at classical cost: Ring polymer contraction for density functional theory," J. Chem. Phys. **144**(5), 054112 (2016).
[46]F. Uhl, D. Marx, and M. Ceriotti, "Accelerated path integral methods for atomistic simulations at ultra-low temperatures," J. Chem. Phys. **145**(5), 054101 (2016).
[47]Y. Xue, J.-N. Wang, W. Hu, J. Zheng, Y. Li, X. Pan, Y. Mo, Y. Shao, L. Wang, and Y. Mei, "Affordable *ab initio* path integral for thermodynamic properties via molecular dynamics simulations using semiempirical reference potential," J. Phys. Chem. A **125**(50), 10677–10685 (2021).
[48]X. Gui, W. Fan, J. Sun, and Y. Li, "New stable and fast ring-polymer molecular dynamics for calculating bimolecular rate coefficients with an example of OH + CH_4," J. Chem. Theory Comput. **18**(9), 5203–5212 (2022).
[49]I. L. Thomas, "Protonic structure of molecules. I. Ammonia molecules," Phys. Rev. **185**(1), 90–94 (1969).
[50]J. F. Capitani, R. F. Nalewajski, and R. G. Parr, "Non-Born–Oppenheimer density functional theory of molecular systems," J. Chem. Phys. **76**(1), 568–573 (1982).
[51]M. Tachikawa, K. Mori, H. Nakai, and K. Iguchi, "An extension of ab initio molecular orbital theory to nuclear motion," Chem. Phys. Lett. **290**(4), 437–442 (1998).

[52] T. Kreibich and E. K. U. Gross, "Multicomponent density-functional theory for electrons and nuclei," Phys. Rev. Lett. **86**(14), 2984–2987 (2001).
[53] S. P. Webb, T. Iordanov, and S. Hammes-Schiffer, "Multiconfigurational nuclear-electronic orbital approach: Incorporation of nuclear quantum effects in electronic structure calculations," J. Chem. Phys. **117**(9), 4106–4118 (2002).
[54] H. Nakai, "Nuclear orbital plus molecular orbital theory: Simultaneous determination of nuclear and electronic wave functions without Born–Oppenheimer approximation," Int. J. Quantum Chem. **107**(14), 2849–2869 (2007).
[55] T. Ishimoto, M. Tachikawa, and U. Nagashima, "Review of multicomponent molecular orbital method for direct treatment of nuclear quantum effect," Int. J. Quantum Chem. **109**(12), 2677–2694 (2009).
[56] F. Pavošević, T. Culpitt, and S. Hammes-Schiffer, "Multicomponent quantum chemistry: Integrating electronic and nuclear quantum effects via the nuclear–electronic orbital method," Chem. Rev. **120**(9), 4222–4253 (2020).
[57] S. Hammes-Schiffer, "Nuclear–electronic orbital methods: Foundations and prospects," J. Chem. Phys. **155**(3), 030901 (2021).
[58] M. Chow, E. Lambros, X. Li, and S. Hammes-Schiffer, "Nuclear–electronic orbital QM/MM approach: Geometry optimizations and molecular dynamics," J. Chem. Theory Comput. **19**(13), 3839–3848 (2023).
[59] M. Chow, T. E. Li, and S. Hammes-Schiffer, "Nuclear–electronic orbital quantum mechanical/molecular mechanical real-time dynamics," J. Phys. Chem. Lett. **14**(43), 9556–9562 (2023).
[60] E. Lambros, B. Link, M. Chow, F. Lipparini, S. Hammes-Schiffer, and X. Li, "Assessing implicit and explicit polarizable solvation models for nuclear–electronic orbital systems: Quantum proton polarization and solvation energetics," J. Phys. Chem. A **127**(44), 9322–9333 (2023).
[61] T. Iordanov and S. Hammes-Schiffer, "Vibrational analysis for the nuclear–electronic orbital method," J. Chem. Phys. **118**(21), 9489–9496 (2003).
[62] Y. Yang, P. E. Schneider, T. Culpitt, F. Pavošević, and S. Hammes-Schiffer, "Molecular vibrational frequencies within the nuclear–electronic orbital framework," J. Phys. Chem. Lett. **10**(6), 1167–1172 (2019).
[63] L. Zhao, Z. Tao, F. Pavošević, A. Wildman, S. Hammes-Schiffer, and X. Li, "Real-time time-dependent nuclear–electronic orbital approach: Dynamics beyond the Born–Oppenheimer approximation," J. Phys. Chem. Lett. **11**(10), 4052–4058 (2020).
[64] L. Zhao, A. Wildman, Z. Tao, P. Schneider, S. Hammes-Schiffer, and X. Li, "Nuclear–electronic orbital Ehrenfest dynamics," J. Chem. Phys. **153**(22), 224111 (2020).
[65] D. Moscato, G. Mandelli, M. Bondanza, F. Lipparini, R. Conte, B. Mennucci, and M. Ceotto, "Unraveling water solvation effects with quantum mechanics/molecular mechanics semiclassical vibrational spectroscopy: The case of thymidine," J. Am. Chem. Soc. **146**(12), 8179–8188 (2024).
[66] X. Xu and Y. Yang, "Constrained nuclear-electronic orbital density functional theory: Energy surfaces with nuclear quantum effects," J. Chem. Phys. **152**(8), 084107 (2020).
[67] X. Xu and Y. Yang, "Full-quantum descriptions of molecular systems from constrained nuclear–electronic orbital density functional theory," J. Chem. Phys. **153**(7), 074106 (2020).
[68] Z. Chen and Y. Yang, "Incorporating nuclear quantum effects in molecular dynamics with a constrained minimized energy surface," J. Phys. Chem. Lett. **14**(1), 279–286 (2023).
[69] Y. Wang, Z. Chen, and Y. Yang, "Calculating vibrational excited state absorptions with excited state constrained minimized energy surfaces," J. Phys. Chem. A **127**(25), 5491–5501 (2023).
[70] X. Xu and Y. Yang, "Molecular vibrational frequencies from analytic Hessian of constrained nuclear–electronic orbital density functional theory," J. Chem. Phys. **154**(24), 244110 (2021).
[71] X. Xu, Z. Chen, and Y. Yang, "Molecular dynamics with constrained nuclear electronic orbital density functional theory: Accurate vibrational spectra from efficient incorporation of nuclear quantum effects," J. Am. Chem. Soc. **144**(9), 4039–4046 (2022).
[72] Y. Zhang, X. Xu, N. Yang, Z. Chen, and Y. Yang, "Describing proton transfer modes in shared proton systems with constrained nuclear–electronic orbital methods," J. Chem. Phys. **158**(23), 231101 (2023).
[73] Y. Zhang, Y. Wang, X. Xu, Z. Chen, and Y. Yang, "Vibrational spectra of highly anharmonic water clusters: Molecular dynamics and harmonic analysis revisited with constrained nuclear-electronic orbital methods," J. Chem. Theory Comput. **19**(24), 9358–9368 (2023).
[74] X. Xu, "Constrained nuclear-electronic orbital density functional theory with a dielectric continuum solvent model," J. Phys. Chem. A **127**(30), 6329–6334 (2023).
[75] Z. Chen and Y. Yang, "Transition state theory within constrained nuclear-electronic orbital framework: Accurate hydrogen atom transfer reaction rates with direct incorporation of quantum nuclear delocalization effects," ChemRxiv, 2024.
[76] J. Langford, Y. Zhang, Y. Yang, and Z. Chen, "Where is the hidden intramolecular H-bonding vibrational signal in the proline matrix IR spectrum?," J. Chem. Phys. **161**, 134302 (2024).
[77] L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding, and K. Morokuma, "The ONIOM method and its applications," Chem. Rev. **115**(12), 5678–5796 (2015).
[78] T. Vreven, K. Morokuma, Ö. Farkas, H. B. Schlegel, and M. J. Frisch, "Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints," J. Comput. Chem. **24**(6), 760–769 (2003).
[79] A. O. Dohn, "Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: A tutorial review," Int. J. Quantum Chem. **120**(21), e26343 (2020).
[80] A. Warshel, M. Kato, and A. V. Pisliakov, "Polarizable force fields: History, test cases, and prospects," J. Chem. Theory Comput. **3**(6), 2034–2045 (2007).
[81] Z. Jing, C. Liu, S. Y. Cheng, R. Qi, B. D. Walker, J.-P. Piquemal, and P. Ren, "Polarizable force fields for biomolecular simulations: Recent advances and applications," Annu. Rev. Biophys. **48**, 371–394 (2019).
[82] I. Antes and W. Thiel, "Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods," J. Phys. Chem. A **103**(46), 9290–9295 (1999).
[83] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, "Frontier bonds in QM/MM methods: A comparison of different approaches," J. Phys. Chem. A **104**(8), 1720–1735 (2000).
[84] P. Amara and M. J. Field, "Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method," Theor. Chem. Acc. **109**(1), 43–52 (2003).
[85] H. Lin and D. G. Truhlar, "Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations," J. Phys. Chem. A **109**(17), 3991–4004 (2005).
[86] Q. Sun, "Libcint: An efficient general integral library for Gaussian basis functions," J. Comput. Chem. **36**(22), 1664–1671 (2015).
[87] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, "PySCF: The Python-based simulations of chemistry framework," WIREs Comput. Mol. Sci. **8**(1), e1340 (2018).
[88] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, "Recent developments in the PySCF program package," J. Chem. Phys. **153**(2), 024109 (2020).
[89] See our group GitHub page https://github.com/theorychemyang for information about the PySCF version with CNEO functionalities.
[90] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen, "GROMACS: A message-passing parallel molecular dynamics implementation," Comput. Phys. Commun. **91**(1), 43–56 (1995).
[91] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, "GROMACS: Fast, flexible, and free," J. Comput. Chem. **26**(16), 1701–1718 (2005).
[92] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," SoftwareX **1–2**, 19–25 (2015).

[93]T. H. Dunning, Jr., "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**(2), 1007–1023 (1989).
[94]B. I. Dunlap, "Robust and variational fitting," Phys. Chem. Chem. Phys. **2**(10), 2113–2116 (2000).
[95]F. Weigend, A. Köhn, and C. Hättig, "Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations," J. Chem. Phys. **116**(8), 3175–3183 (2002).
[96]Q. Yu, F. Pavošević, and S. Hammes-Schiffer, "Development of nuclear basis sets for multicomponent quantum chemistry methods," J. Chem. Phys. **152**(24), 244123 (2020).
[97]C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," Phys. Rev. B **37**(2), 785–789 (1988).
[98]A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," J. Chem. Phys. **98**(7), 5648–5652 (1993).
[99]P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields," J. Phys. Chem. **98**(45), 11623–11627 (1994).
[100]Y. Yang, K. R. Brorsen, T. Culpitt, M. V. Pak, and S. Hammes-Schiffer, "Development of a practical multicomponent density functional for electron-proton correlation to produce accurate proton densities," J. Chem. Phys. **147**(11), 114113 (2017).
[101]K. R. Brorsen, Y. Yang, and S. Hammes-Schiffer, "Multicomponent density functional theory: Impact of nuclear quantum effects on proton affinities and geometries," J. Phys. Chem. Lett. **8**(15), 3488–3493 (2017).
[102]K. R. Brorsen, P. E. Schneider, and S. Hammes-Schiffer, "Alternative forms and transferability of electron-proton correlation functionals in nuclear-electronic orbital density functional theory," J. Chem. Phys. **149**(4), 044110 (2018).
[103]Z. Tao, Y. Yang, and S. Hammes-Schiffer, "Multicomponent density functional theory: Including the density gradient in the electron-proton correlation functional for hydrogen and deuterium," J. Chem. Phys. **151**(12), 124102 (2019).
[104]W. L. Jorgensen and J. Tirado-Rives, "Potential energy functions for atomic-level simulations of water and organic and biomolecular systems," Proc. Natl. Acad. Sci. U. S. A. **102**(19), 6665–6670 (2005).
[105]L. S. Dodda, J. Z. Vilseck, J. Tirado-Rives, and W. L. Jorgensen, "1.14*CM1A-LBCC: Localized bond-charge corrected CM1A charges for condensed-phase simulations," J. Phys. Chem. B **121**(15), 3864–3870 (2017).
[106]L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives, and W. L. Jorgensen, "LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands," Nucl. Acids Res. **45**(W1), W331–W336 (2017).
[107]D. J. Price and C. L. Brooks, "A modified TIP3P water potential for simulation with Ewald summation," J. Chem. Phys **121**(20), 10096–10103 (2004).
[108]M. Freindorf, Y. Shao, T. R. Furlani, and J. Kong, "Lennard–Jones parameters for the combined QM/MM method using the B3LYP/6-31G*/AMBER potential," J. Comput. Chem. **26**(12), 1270–1278 (2005).
[109]A. G. Donchev, N. G. Galkin, A. A. Illarionov, O. V. Khoruzhii, M. A. Olevanov, V. D. Ozrin, M. V. Subbotin, and V. I. Tarasov, "Water properties from first principles: Simulations by a general-purpose quantum mechanical polarizable force field," Proc. Natl. Acad. Sci. U. S. A. **103**(23), 8613–8617 (2006).
[110]R. S. Paton and J. M. Goodman, "Hydrogen bonding and π-stacking: How reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions," J. Chem. Inf. Model. **49**(4), 944–955 (2009).
[111]A. J. A. Aquino, D. Tunega, G. Haberhauer, M. H. Gerzabek, and H. Lischka, "Solvent effects on hydrogen bonds: A theoretical study," J. Phys. Chem. A **106**(9), 1862–1871 (2002).
[112]W. W. Cleland and M. M. Kreevoy, "Low-barrier hydrogen bonds and enzymic catalysis," Science **264**(5167), 1887–1890 (1994).
[113]A. Warshel, A. Papazyan, and P. A. Kollman, "On low-barrier hydrogen bonds and enzyme catalysis," Science **269**(5220), 102–106 (1995).
[114]S. Dai, L.-M. Funk, F. R. von Pappenheim, V. Sautner, M. Paulikat, B. Schröder, J. Uranga, R. A. Mata, and K. Tittmann, "Low-barrier hydrogen bonds in enzyme cooperativity," Nature **573**(7775), 609–613 (2019).
[115]M. T. Kemp, E. M. Lewandowski, and Y. Chen, "Low barrier hydrogen bonds in protein structure and function," Biochim. Biophys. Acta, Proteins Proteomics **1869**(1), 140557 (2021).
[116]E. Nango, A. Royant, M. Kubo, T. Nakane, C. Wickstrand, T. Kimura, T. Tanaka, K. Tono, C. Song, R. Tanaka, T. Arima, A. Yamashita, J. Kobayashi, T. Hosaka, E. Mizohata, P. Nogly, M. Sugahara, D. Nam, T. Nomura, T. Shimamura, D. Im, T. Fujiwara, Y. Yamanaka, B. Jeon, T. Nishizawa, K. Oda, M. Fukuda, R. Andersson, P. Båth, R. Dods, J. Davidsson, S. Matsuoka, S. Kawatake, M. Murata, O. Nureki, S. Owada, T. Kameshima, T. Hatsui, Y. Joti, G. Schertler, M. Yabashi, A.-N. Bondar, J. Standfuss, R. Neutze, and S. Iwata, "A three-dimensional movie of structural changes in bacteriorhodopsin," Science **354**(6319), 1552–1557 (2016).
[117]P. Phatak, N. Ghosh, H. Yu, Q. Cui, and M. Elstner, "Amino acids with an intermolecular proton bond as proton storage site in bacteriorhodopsin," Proc. Natl. Acad. Sci. U. S. A. **105**(50), 19672–19677 (2008).
[118]D. Maag, T. Mast, M. Elstner, Q. Cui, and T. Kubař, "O to bR transition in bacteriorhodopsin occurs through a proton hole mechanism," Proc. Natl. Acad. Sci. U. S. A. **118**(39), e2024803118 (2021).
[119]S. Lüdtke, P. Neumann, K. M. Erixon, F. Leeper, R. Kluger, R. Ficner, and K. Tittmann, "Sub-Ångström-resolution crystallography reveals physical distortions that enhance reactivity of a covalent enzymatic intermediate," Nat. Chem. **5**(9), 762–767 (2013).
[120]P. M. Kiefer and J. T. Hynes, "Theoretical aspects of tunneling proton transfer reactions in a polar environment," J. Phys. Org. Chem. **23**(7), 632–646 (2010).
[121]See https://doi.org/10.21231/GNT1-HW21 for "Center for high throughput computing" (2006).