Seamless integration of GEM, a density based-force field, for QM/MM simulations via LICHEM, Psi4 and Tinker-HP

Jorge Nochebuena,[1] Andrew C. Simmonett,[2] and G. Andrés Cisneros[1,3]

[1)] *Department of Physics. University of Texas at Dallas. Richardson, TX 75080, USA*

[2)] *Laboratory of Computational Biology, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, USA*

[3)] *Department of Chemistry and Biochemistry. University of Texas at Dallas. Richardson, TX 75080, USA*

(Dated: 27 January 2024)

Hybrid quantum mechanics/molecular mechanics (QM/MM) simulations have become an essential tool in computational chemistry, particularly for analyzing complex biological and condensed phase systems. Building on this foundation, our work presents a novel implementation of the Gaussian Electrostatic Model (GEM), a polarizable density-based force field, within the QM/MM framework. This advancement provides seamless integration, enabling efficient and optimized QM/GEM calculations in a single step using the LICHEM Code. We have successfully applied our implementation to water dimers and hexamers, demonstrating the ability to handle water systems with varying numbers of water molecules. Moreover, we have extended the application to describe the double proton transfer of the aspartic acid dimer in a box of water, which highlights the method's proficiency in investigating heterogeneous systems. Our implementation offers the flexibility to perform on-the-fly density fitting or to utilize pre-fitted coefficients to estimate exchange and Coulomb contributions. This flexibility enhances efficiency and accuracy in modeling molecular interactions, especially in systems where polarization effects are significant.

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

I. INTRODUCTION

Hybrid quantum mechanics/molecular mechanics (QM/MM) calculations represents an important tool in computational chemistry, especially for analyzing complex biological systems and condensed-phase systems.[1–3] These simulations combine quantum mechanics, which treats electrons explicitly, with molecular mechanics, which uses force fields to model larger-scale molecular interactions. This division is predicated on the assumption that not all atoms in a system are equally critical for the process under investigation. Therefore, the proper description of the MM region is often not given the attention it deserves.

In the traditional framework of QM/MM calculations, atom charges are often represented statically.[4–6] This approach, while might be useful in many contexts, overlooks the dynamic nature of atomic interactions, particularly in complex biological systems like enzymes. Therefore, it has become increasingly clear that considering the polarization effects is crucial in accurately modeling enzymatic reactions within the QM/MM approach.[7]

QM/MM calculations that incorporate polarizable force fields represent a significant step forward for achieving greater accuracy. The AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications) force field, in particular, stands out due to its advanced approach in modeling molecular interactions.[1,8,9] The AMOEBA force field differentiates itself from traditional force fields by including atomic multipole terms and explicit polarization effects.[10,11] Atomic multipole terms allow a more realistic representation of the electrostatic potential around atoms, going beyond the simple point charge model. This is crucial in capturing the anisotropic effects. Additionally, the explicit inclusion of polarization means that the AMOEBA force field can dynamically mimic the electron distribution in response to the local electrostatic environment. Despite these advancements, there are still challenges that need to be addressed. While adding explicit polarization improves the representation of the MM environment, other effects are treated with functions that approximate various terms, like Van der Waals forces, or not considered at all, such as charge transfer.[1] Furthermore, while the AMOEBA force field effectively models long-range electrostatics and polarization, it struggles with short-range interactions due to the limitations of the point multipolar approach. This approach gives rise to the charge penetration error, which arises from the lack of accounting of the overlap of electron clouds at close distances, and requires the consideration of their charge distributions for accurate modeling.[12–15]

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

For that reason, advanced potentials have been developed. A recent technique involves using many-body decomposition-based potentials, including MB-pol and MB-DFT, by providing a highly precise environmental representation through the explicit Coulomb and polarization embedding of the QM wavefunction and a many-body potential, eliminate the need for additional functions or approximations in adaptive QM/MM calculations.[16,17] Recently, advanced models of the embedding environment have been proposed, such as the charged quantum harmonic oscillators[18] and the frozen density embedding[19] (FDE). Another strategy is to use advanced potentials that explicitly include one or more of the missing components in the embedding environment. Examples include the effective fragment potential (EFP) that covers exchange and charge transfer,[20] the induced dipole model in the polarizable embedding model,[7] the fragment exchange potential,[21] the QMFF,[22] the exchange fragment potential,[23] QM:QM embedding approaches[24], among others.

In this work, we introduce a new implementation of the Gaussian Electrostatic Model (GEM) for QM/MM calculations. GEM is a polarizable force field that uses Gaussian auxiliary basis sets (ABSs) to reproduce molecular electronic densities and calculates individual contributions in a similar way to the symmetry-adapted perturbation theory (SAPT).[12,13,25–28] The GEM* version employs the same frozen molecular densities to calculate Coulomb and exchange-repulsion contributions. The key distinction between this new implementation and the previous one for QM/MM calculations,[29] lies in the ability to perform on-the-fly density fitting or to utilize pre-fitted expansion coefficients, the ability to study systems larger than dimers, and a complete integration of the full QM/GEM potential by combining a GEM implementation in TINKER-HP,[30] with a modified version of Psi4[31] via LICHEM,[32,33] resulting in increased efficiency and applicability.

The theoretical framework of GEM* and the QM/GEM* integration is detailed in the following section. Our implementation of GEM* in Psi4[31] enables the calculation of Coulomb and exchange-repulsion interactions between the QM and MM regions. Additionally, we have utilized Tinker-HP[30] to model the Coulomb and exchange interactions within the MM subsystem. Tinker-HP is also employed to calculate the polarization and Halgren's attractive term for the entire system. Our implementation has been applied to model systems, starting with water dimers and hexamers, as well as the aspartic acid dimer in a water box. Subsequent sections present the computational details and the results obtained from these applications. Our results include comparisons with QM/AMOEBA calculations, a QM/MM

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

method that has demonstrated significant efficacy in studying various systems.[8,9,34–37] The final section discusses our conclusions derived from these studies.

II. THEORETICAL FRAMEWORK

The total energy in QM/GEM* calculations can be expressed as follows:

$$E_{\text{tot}} = E^{\text{QM}} + E^{\text{GEM*}} + E^{\text{QM/GEM*}} \tag{1}$$

where the first term, E^{QM}, is the total potential energy of the QM region. The second term, $E^{\text{GEM*}}$, corresponds to the total potential energy of the MM subsystem evaluated with GEM and is defined as follows:

$$E^{\text{GEM*}} = E^{\text{GEM}}_{\text{Coul}} + E^{\text{GEM}}_{\text{exch-rep}} + E^{\text{GEM}}_{\text{pol}} + E^{\text{AMOEBA}}_{\text{bonded}} + E^{\text{modHalgren}}_{\text{disp+CT}} \tag{2}$$

where $E^{\text{GEM}}_{\text{Coul}}$ and $E^{\text{GEM}}_{\text{exch-rep}}$ are the Coulomb and exchange-repulsion contributions described with GEM. $E^{\text{AMOEBA}}_{\text{bonded}}$ is the bonded contribution and $E^{\text{GEM}}_{\text{pol}}$ is the polarization term, which can be calculated using electrostatic potential (ESP) derived charges,[38,39] or Mulliken[40] point charges to approximate the QM wavefunction. The last two terms employing the same functional forms as AMOEBA. However, one important difference in GEM* is that the permanent electric fields for the calculation of the induced dipoles are calculated with the distributed multipoles obtained from the fitted densities.[28] The charge transfer and dispersion are approximated by fitting them together to the modified Halgren potential:

$$E^{\text{modHalgren}}_{\text{disp+CT}} = \varepsilon_{AB} \left[\frac{1.07 R^0_{AB}}{R_{AB} + 0.07 R^0_{AB}} \right]^7 \tag{3}$$

where R_{AB} and R^0_{AB} are the distance and the equilibrium distance between atoms A and B and ε_{AB} is the depth of the potential well. GEM* uses only the attractive term of the Halgren potential because repulsive interactions are already considered by the exchange-repulsion term.

In GEM* fitted molecular densities, $\tilde{\rho}(r)$, are expressed as an expansion of primitive Cartesian Hermite Gaussian functions:

$$\tilde{\rho}(r) = \sum_k c_k \Lambda(r) \tag{4}$$

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

where the expansion coefficients, c_k, for the approximate density may be obtained by minimizing the following equation using some metric:

$$E_{\text{self}} = \langle \rho(r) - \tilde{\rho}(r) | \tilde{O} | \rho(r) - \tilde{\rho}(r) \rangle \tag{5}$$

Several operators can be employed including the overlap operator, the Coulomb operator or the damped Coulomb operator. The minimization of equation (5) with respect to the expansion coefficients leads to a linear system of equations:

$$\frac{\partial E_{\text{self}}}{\partial c_l} = -\sum_{\mu,\nu} P_{\mu,\nu} \langle \mu\nu | \tilde{O} | l \rangle + \sum_k c_k \langle k | \tilde{O} | l \rangle \tag{6}$$

where $P_{\mu,\nu}$ is the density matrix. Equation (6) may be used for the determination of the coefficients by setting:

$$\mathbf{c} = \mathbf{G}^{-1}\mathbf{j} \tag{7}$$

where

$$\mathbf{G} = \langle k | \tilde{O} | l \rangle, \qquad \mathbf{j} = P_{\mu,\nu} \langle \mu\nu | \tilde{O} | l \rangle \tag{8}$$

In principle, $\mathbf{G}$ should be symmetric and positive definite. However, in practice, this matrix often exhibits near singularities. Consequently, the process of diagonalization to derive its inverse necessitates meticulous handling.[41] Then, the approximate density, $\tilde{\rho}(r)$, is used to compute the Coulomb and exchange-repulsion terms:

$$\begin{aligned} E_{\text{Coul}}^{\text{GEM}} = & \int \frac{\tilde{\rho}_A(r_A)\tilde{\rho}_B(r_B)}{r_{AB}} \mathrm{d}r + \frac{Z_A Z_B}{r_{AB}} \\ & + \int \frac{Z_A \tilde{\rho}_B(r_B)}{r_{AB}} \mathrm{d}r + \int \frac{Z_B \tilde{\rho}_A(r_A)}{r_{AB}} \mathrm{d}r \end{aligned} \tag{9}$$

$$E_{\text{exch-rep}}^{\text{GEM}} = k_{\text{exch}} \int \tilde{\rho}_A(r_A)\tilde{\rho}_B(r_B) \mathrm{d}r \tag{10}$$

where $Z_{A,B}$ represents the nuclei of atoms A and B and r_{AB} is the distance between atoms A and B and k_{exch} is a proportionality coefficient.

The third term of the equation (1) is expressed in a similar way to equation (2):

$$E^{\text{QM/GEM*}} = E_{\text{Coul}}^{\text{QM/GEM}} + E_{\text{exch-rep}}^{\text{QM/GEM}} + E_{\text{pol }(MM \to QM)}^{\text{QM/GEM}} + E_{\text{pol }(QM \to MM)}^{\text{QM/GEM}} + E_{\text{disp+CT}}^{\text{modHalgren}} \tag{11}$$

However, there are some important differences. First, the Coulomb and exchange repulsion terms are computed by evaluating the interaction between the electron density originating from the QM region and the fitted density derived from GEM:

$$E_{\text{Coul}}^{\text{GEM}} = \int \frac{\rho_A(r_A)\tilde{\rho}_B(r_B)}{r_{AB}} \mathrm{d}r \tag{12}$$

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

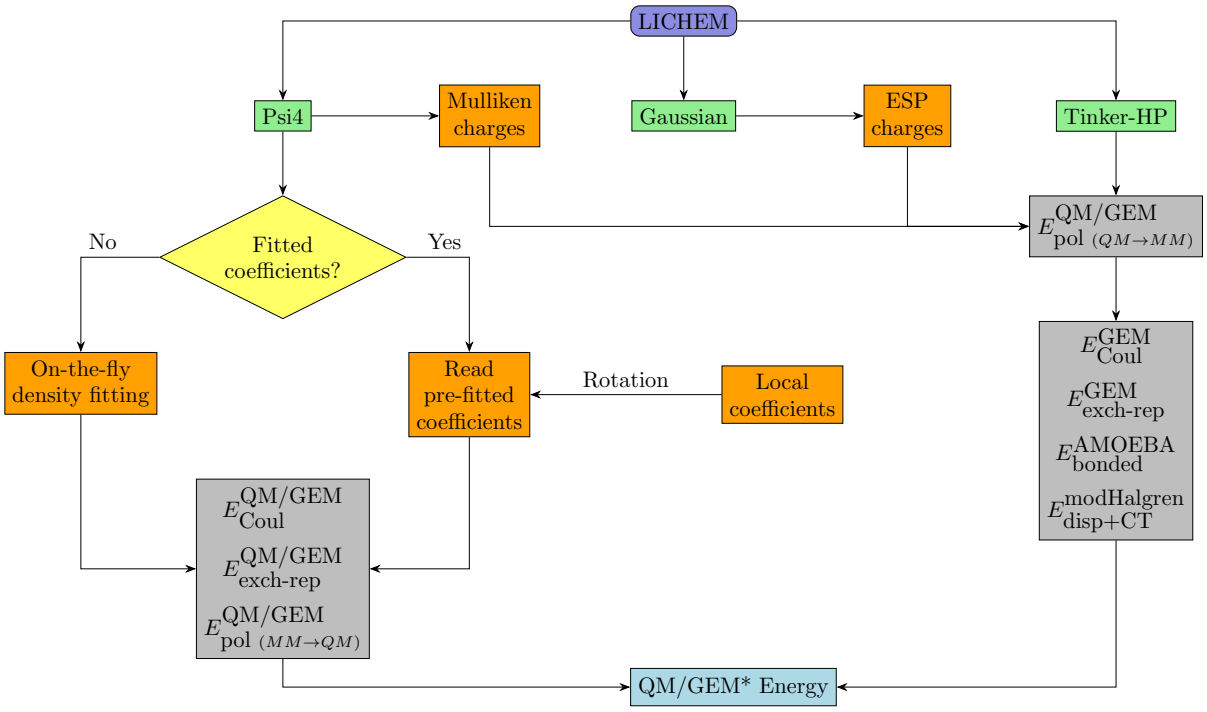

FIG. 1. General process to calculate the intermolecular QM/GEM* energy using the LICHEM code.

$$E^{\text{GEM}}_{\text{exch-rep}} = k_{\text{exch}} \int \rho_A(r_A)\tilde{\rho}_B(r_B)\mathrm{d}r \tag{13}$$

and second, the polarization of the QM region due to the presence of the MM region (third term in Eq. 13 is explicitly included in the effective Hamiltonian.

Equations (12) and (13) can also be included in the effective Hamiltonian:

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{core}} + \hat{V}_{\text{GEM}} \tag{14}$$

where:

$$\hat{V}_{\text{GEM}} = \sum_l x_l \sum_{\mu\nu} \langle \mu\nu || l \rangle + K_{\text{exch}} \sum_l x_l \sum_k \langle k|\tilde{O}|l\rangle \tag{15}$$

The fitting of the coeefficients x_l can be carried out using either Cholesky decomposition or Tikhonov regularization. A Lagrange multiplier is employed to guarantee that the density correctly integrates to the desired number of electrons.

Equations (3), (9) and (10) have been implemented in Tinker-HP, while equation (14) has been implemented in Psi4. The general process for calculating QM/GEM energies is shown in Figure 1. The process begins with LICHEM, which is the central interface capable

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

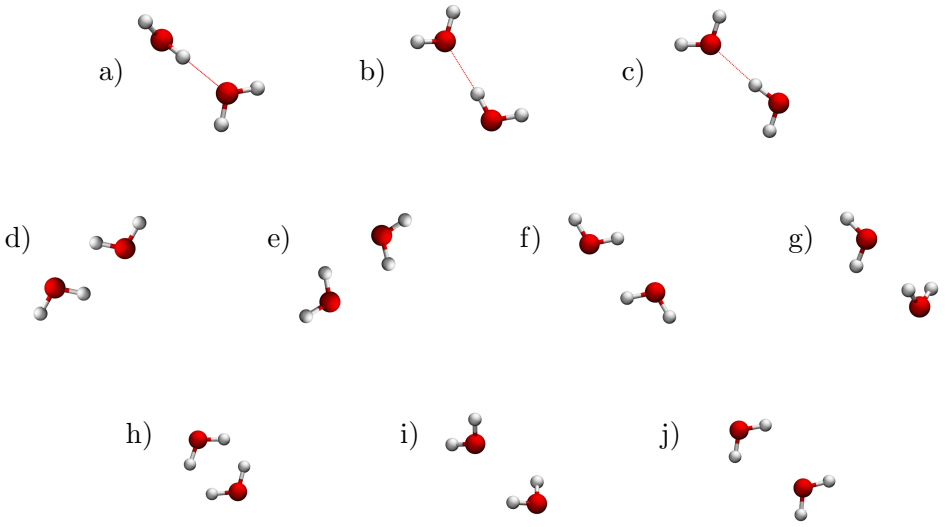

FIG. 2. a) to j) Smith water dimers.

of coordinating the QM/MM calculations. The modified versions of LICHEM, Psi4 and Tinker-HP are available at https://github.com/CisnerosResearch. The Psi4 code is used to calculate the Coulomb and exchange-repulsion contributions between the QM and MM regions, as well as to calculate the Mulliken charges. Alternatively, Gaussian[42] can be used to calculate ESP charges. The polarization term of the QM region due to the presence of the MM region is implicitly evaluated when calculating the Coulomb interaction. We can decide if we want to perform on-the-fly density fitting or to utilize pre-fitted coefficients, which speeds up the process. The pre-fitted coefficients are obtained by rotating local coefficients, which were calculated previously from the average molecular densities of 500 individual water molecules.[25]

On the other side, Tinker-HP is used to calculate the Halgren attractive term, as well as polarization of the MM region due to QM region employing the same functional forms as AMOEBA but using GEM multipoles. To calculate the polarization energy, the Mulliken or the ESP charges calculated are used to represent the QM region in the MM calculations. Tinker-HP also calculates the Coulomb and Exchange interactions of the MM subsystem. Finally, all these elements are combined to obtain the total intermolecular QM/GEM* energy and force of the system.

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

III. COMPUTATIONAL DETAILS

In this work, we studied various systems to assess the implementation of GEM for conducting QM/MM calculations. Initially, we calculated the interaction energies of the ten Smith water dimers[43] (see Figure 2). These dimers are representative of systems with attractive and repulsive interactions. Each water molecule in these dimers follows the AMOEBA internal geometry, characterized by an O-H distance of 0.9572 Å and an H-O-H angle of 108.5° (see Figure 3).

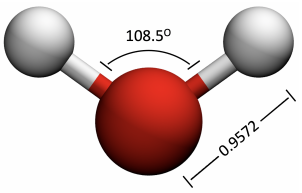

FIG. 3. AMOEBA water.

We utilized the SAPT(DFT) method, with the SAPT expansion truncated at the same level of SAPT0, and DFT calculations as reference values. Both approaches were calculated in Psi4 employing the ωB97X-D[44] exchange-correlation functional and the aug-cc-pVDZ basis set. The interaction energies calculated with ωB97X-D/aug-cc-pVDZ[45,46] were corrected for the basis set superposition error (BSSE) using the counterpoise method.[47] We determined the QM/MM interaction energies, treating one water molecule in the QM region and the other in the MM region. We examined two possible scenarios for differentiating between the QM and MM regions. That is, A-B and B-A dimers were taken into account. To benchmark our implementation, we included QM/MM interaction energies computed using the AMOEBA force field for the MM region. Consequently, the QM/MM calculations were run using the ωB97X-D/aug-cc-pVDZ//AMOEBA and ωB97X-D/aug-cc-pVDZ//GEM methods. In both scenarios, we considered two distinct charge schemes to represent the QM region in the MM calculations, which is crucial for calculating the polarization of each dimer. The first scheme employed Mulliken charges, calculated in Psi4, while the second scheme utilized ESP charges, calculated in Gaussian 16.

Additionally, we evaluated six water hexamers (see Figure 4), which also have the AMOEBA internal geometry (see Figure 3). To assess the total intermolecular QM/MM energies, we explored all feasible combinations, ranging from 1 to 5 water molecules in

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

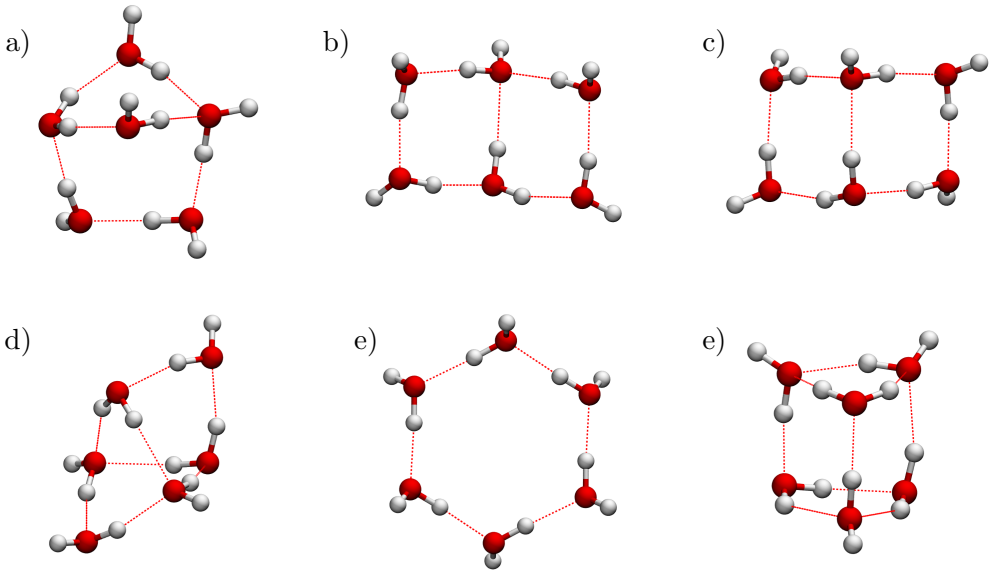

FIG. 4. Hexamers. a) bag, b) book1, c) book2, d) cage, e) chair, f) prism

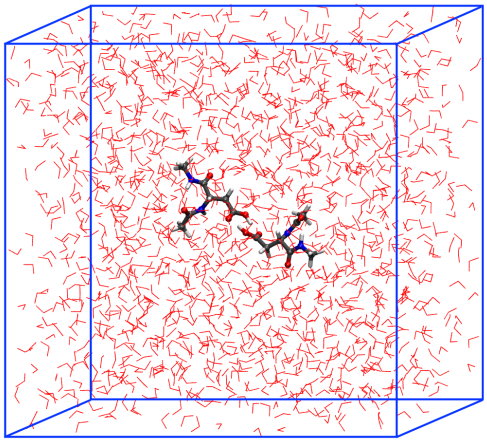

FIG. 5. di-aspartic acid dimer in a water box. Red lines represent water molecules.

the QM region, with the remaining molecules described in the MM region. We considered the total intermolecular energy obtained with ωB97X-D/aug-cc-pVDZ, BSSE corrected, as our reference values. Similar to the previous cases, two atomic charge schemes were contemplated for representing the QM region in these systems.

Finally, our implementation was tested to describe the reaction energy of a double proton transfer in a model aspartic acid dimer in water solution (see Figure 5). This model system had been previously studied.[32,48,49] For this system, we positioned the aspartic acid dimer at the center of a $36 \times 36 \times 36$ Å^3 box, filled with 1,500 water molecules. Then the system

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

was minimized using the AMOEBA force field and Tinker. To heat the system from 0 K, we performed MD simulations increasing the temperature by 50K every 500 ps until reaching 300 K. This was followed by a 3 ns molecular dynamics simulation with Tinker-HP and the AMOEBA force field to facilitate system equilibration.

With the system equilibrated, we selected a representative structure and optimized the molecules within a 15 Å radius centered around the COOH groups of each aspartic acid molecule. This structure was designated as the reactant (see Figure 8a). The aspartic dimer is describe in the QM region and the water molecules in the MM region. To form the product, we modified the optimized reactant by transferring protons between groups and then re-optimized the structure (see Figure 8c). We performed a simple interpolation between the reactant and product and optimized the path using the quadratic string method (QSM) implemented in LICHEM.[50] QM convergence criteria were: RMS deviation of 0.001 Å, RMS force of 0.025 Hartree/Bohr and maximum force of 0.05 Hartree/Bohr. MM convergence criteria were: RMS deviation of 0.1 Å and RMS force of 0.1 kcal/mol Å. For the QM region, we employed the ωB97X-D functional with the 6-311G(d,p) basis set,[51] utilizing the Gaussian16 code. The MM region was characterized using the AMOEBA force field and Tinker 7 code.[11,52,53] For the optimized path, single-point calculations were conducted in each bead using our GEM implementation. In both cases, the polarization was calculated using ESP charges to represent the QM region.

IV. RESULTS AND DISCUSSION

Water dimers. In this work, we compared the computational methods ωB97X-D//AMOEBA and ωB97X-D//GEM for estimating interaction energies in the ten Smith water dimers, in order to assess their accuracy against SAPT(DFT) and ωB97X-D reference values. We calculated QM/GEM energies from electronic densities generated using pre-fitted coefficients and calculated on-the-fly. Two different schemes, denoted as Mulliken charges and ESP charges, were used to represent the molecule in the QM region within the MM region.

The results in Figure 6a and 6b show QM/GEM values calculated by using pre-fitted densities, however, comparable results were obtained by using on-the-fly calculated densities (see Figures S1-S2 and Tables S1-S4). In Figure 6a, the ωB97X-D/GEM (ESP) method

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

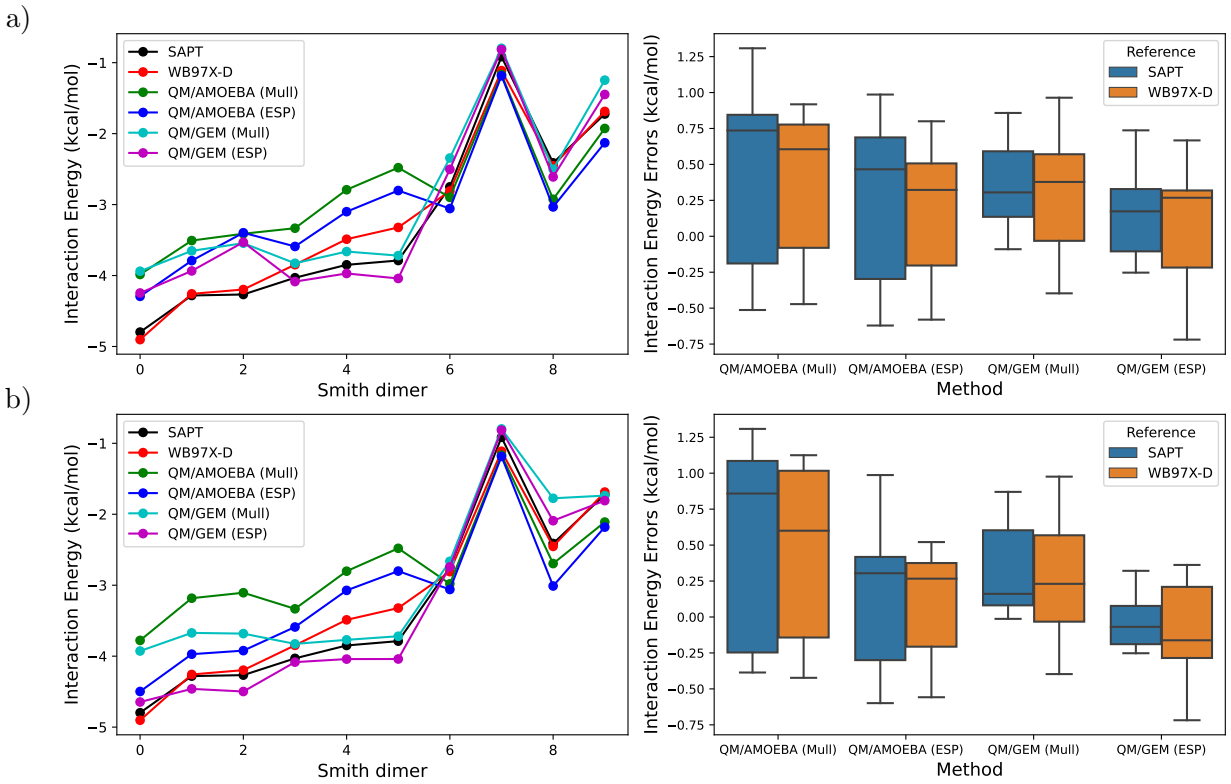

FIG. 6. Interaction energies (left) and deviations (right) for the ten Smith water dimers. a) and b) show the two different options for the selection of QM region. Mull and ESP represent Mulliken and electrostatic potential derived charges, respectively.

had the lowest mean absolute error (MAE) of 0.29, closely followed by ωB97X-D/GEM (Mulliken) with an MAE of 0.38 compared to the SAPT reference values. The ωB97X-D/AMOEBA methods showed a higher MAE, being 0.66 and 0.56 when using Mulliken charges and ESP charges, respectively. These results suggest that the ωB97X-D/GEM method is more accurate and consistent in replicating the reference energies provided by SAPT.

When comparing the same methods with the ωB97X-D reference value set, similar results were observed in terms of accuracy. The QM/GEM method maintained the smallest difference compared to the reference values, with MAE of 0.41 for the Mulliken and ESP charge schemes. On the other hand, the ωB97X-D/AMOEBA methods had MAEs of 0.54 and 0.44 for the same schemes.

Figure 6b shows the results when the molecule selection in the QM region is reversed. In comparison with the SAPT reference values, the QM/GEM (ESP) method demonstrated

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

the highest accuracy, with an MAE of 0.16. This result was considerably lower than those obtained by the other methods (0.33, 0.48 and 0.75 for ωB97X-D/GEM (Mulliken), ωB97X-D/AMOEBA (ESP) and ωB97X-D/AMOEBA (Mulliken), respectively), suggesting that QM/GEM (ESP) might be the most reliable for estimating interaction energies in water dimers under the studied conditions.

In the same evaluation using ωB97X-D calculations as reference values, a similar trend was observed. The QM/GEM (ESP) method maintained superiority in terms of accuracy, with an MAE of 0.31, although with a reduced difference compared to the other methods (0.40, 0.35 and 0.62 for ωB97X-D/GEM (Mulliken), ωB97X-D/AMOEBA (ESP) and ωB97X-D/AMOEBA (Mulliken), respectively) compared to the previous analysis. It is noteworthy that the MAE for the ωB97X-D/AMOEBA (Mulliken) and ωB97X-D/AMOEBA (ESP) methods decreased compared to the SAPT reference, indicating a relative improvement in accuracy when using ωB97X-D as a reference.

Combining the results in Figures 6a and 6b, the ωB97X-D/GEM (ESP) method again presented the lowest MAE with respect to SAPT, with a value of 0.22, indicating a high concordance with the reference values. This suggests that the ωB97X-D/GEM (ESP) method is the most accurate of the evaluated methods for predicting the interaction energies of water dimers compared to the values obtained by SAPT. Similarly, when compared with the ωB97X-D reference values, the ωB97X-D/GEM (ESP) method showed the lowest MAE, with a value of 0.36, suggesting good accuracy, although slightly inferior to that observed when compared with SAPT.

Water hexamers.

The results related to the total interaction energies for the hexamers using the ωB97X-D/AMOEBA and ωB97X-D/GEM methods for the different QM/MM combinations are shown in Figure 7. The mean absolute errors (MAE) and the standard deviations compared to pure QM calculations are shown in Tables S3 and S4. These values are complemented by a graphical representation in Figures S3 to S8, providing an easier visualization of the variations and trends observed in the data.

In the bag hexamer, the ωB97X-D/GEM (ESP) method shows the best accuracy, especially in the combination with the majority of water molecules in the QM region (denoted as 5/1), where the MAE is 1.14 kcal/mol. On the other hand, the ωB97X-D/AMOEBA (Mull) method presents the highest MAE values in several combinations, indicating lower accu-

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

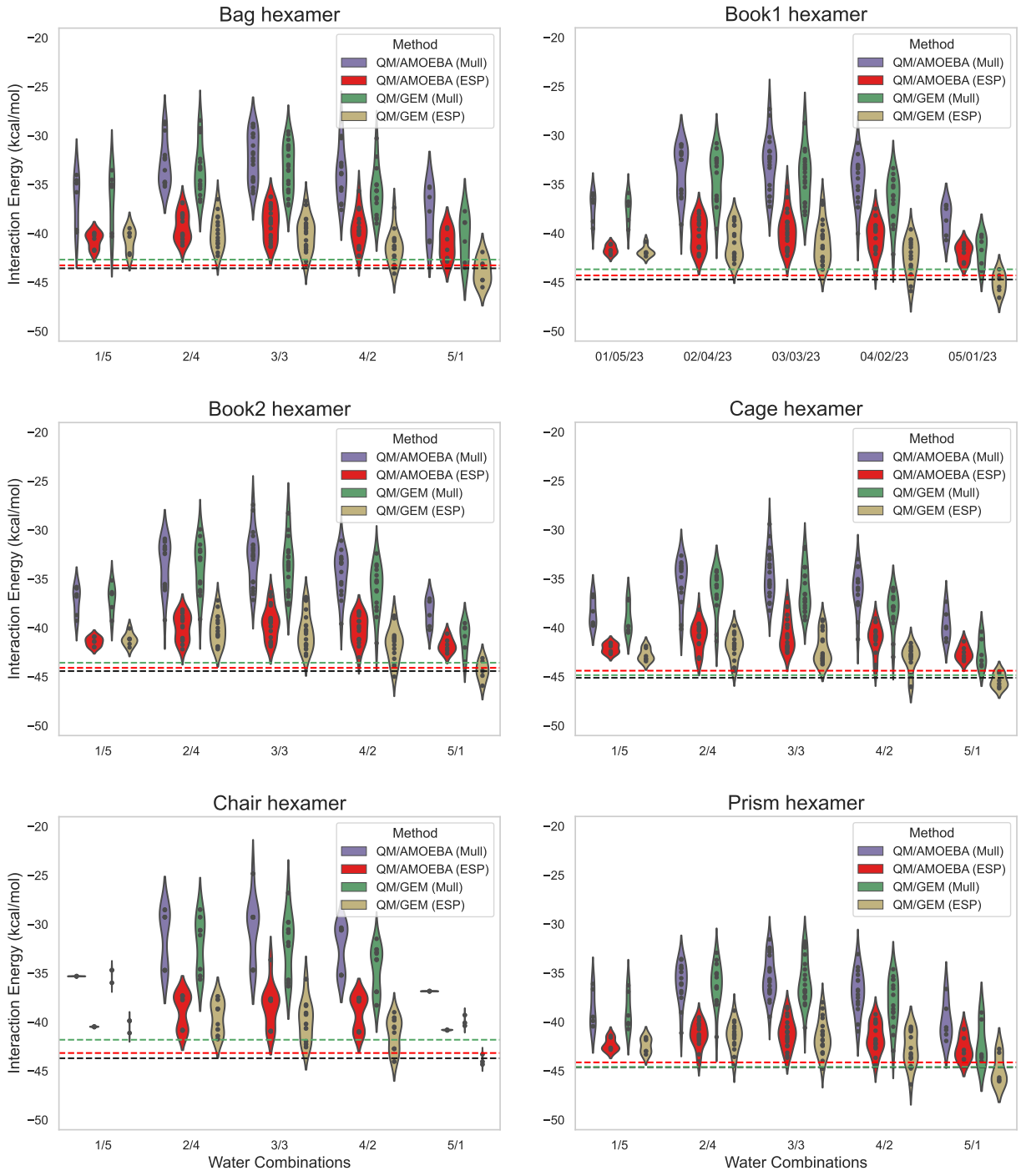

FIG. 7. Violin plots comparing ωB97X-D//AMOEBA and ωB97X-D//GEM calculations for the different QM/MM combinations of the six water hexamers. The dashed horizontal lines in black, red, and green represent ωB97X-D-pure, AMOEBA-pure, and GEM-pure calculations, respectively.

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

racy compared to other methods. Regarding standard deviations, the ωB97X-D/AMOEBA (ESP) and ωB97X-D/GEM (ESP) methods show less variability in most combinations, suggesting greater consistency in their predictions. Conversely, the ωB97X-D/GEM (Mull) and ωB97X-D/AMOEBA (Mull) methods present the highest standard deviations in several combinations, indicating greater dispersion in the results and, therefore, less reliability.

For the book1 hexamer, the ωB97X-D/GEM (ESP) method shows the lowest MAEs for most combinations, suggesting greater accuracy in estimating interaction energies compared to other methods. It is particularly outstanding in the 5/1 combination, where the MAE is 0.87 kcal/mol. On the other hand, the ωB97X-D/AMOEBA (Mull) and ωB97X-D/GEM (Mull) methods tend to have higher MAEs, indicating less accuracy. This is particularly evident in the 3/3 and 2/4 combinations, where the MAE values are the highest among all methods and combinations. The ωB97X-D/AMOEBA (ESP) and ωB97X-D/GEM (ESP) methods show the lowest standard deviations in the 1/5 and 5/1 combinations, indicative of greater consistency in these configurations. In contrast, the ωB97X-D/AMOEBA (Mull) and ωB97X-D/GEM (Mull) methods show the highest standard deviations, especially in the 3/3 combination, suggesting greater variability in the results obtained with this method.

For the book2 hexamer, the ωB97X-D/GEM (ESP) and ωB97X-D/AMOEBA (ESP) methods generally present lower MAE values compared to their Mulliken counterparts. Specifically, the ωB97X-D/GEM (ESP) method in the 5/1 combination stands out for having the lowest MAE of 0.72 kcal/mol, suggesting high accuracy in this configuration. In contrast, the highest MAE values are observed in the QM/AMOEBA (Mull) method, particularly in the 3/3 and 2/4 combinations. Regarding Standard Deviations, the ωB97X-D/AMOEBA (ESP) and ωB97X-D/GEM (ESP) methods show generally lower values than the Mulliken counterparts. This indicates greater uniformity in the results obtained using these methods, especially in combinations such as 5/1 and 1/5. It is interesting to note how combinations with a higher proportion of molecules in the MM region, such as 5/1 and 1/5, tend to present not only lower errors, but also lower variability compared to more balanced combinations such as 3/3. and 2/4.

For the cage hexamer, the ωB97X-D/GEM (ESP) method stands out for having the lowest errors in all combinations, especially in the 5/1 combination, where the MAE of 0.64 kcal/mol is notably lower than in other methods. This suggests high accuracy and reliability in the predictions of ωB97X-D/GEM (ESP), particularly when most of the water molecules

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

are located in the QM region. Conversely, the ωB97X-D/AMOEBA (Mull) method consistently shows the highest MAEs, indicating lower accuracy in its predictions. The 2/4 and 3/3 combinations are particularly notable in this regard, with MAEs exceeding 11 kcal/mol. Regarding standard deviations, we find that methods using Mulliken charges show greater variability in their results compared to those using ESP charges. This is evident in combinations like 2/4 and 3/3, where the standard deviations for the ωB97X-D/AMOEBA (Mull) and ωB97X-D/GEM (ESP) methods exceed 2 kcal/mol. In contrast, the ωB97X-D/GEM (ESP) and ωB97X-D/AMOEBA (ESP) methods maintain lower standard deviations, indicating greater consistency in their measurements.

With respect to the chair hexamer, we observe that the ωB97X-D/AMOEBA (Mull) method shows the highest MAEs in all combinations. On the other hand, the ωB97X-D/GEM (ESP) method tends to have lower MAEs, especially in the combination with more water molecules in the QM region (5/1), where the MAE is only 0.46 kcal/mol, suggesting greater accuracy in this configuration. Regarding standard deviations, the results also vary significantly between methods and combinations. The ωB97X-D/AMOEBA (Mull) and ωB97X-D/GEM (Mull) methods tend to show higher standard deviations in combinations with a more balanced distribution of molecules between the QM and MM regions (2/4, 3/3, and 4/2), indicating greater variability in their predictions. In contrast, ωB97X-D/AMOEBA (Mull) and ωB97X-D/AMOEBA (ESP) show pretty lover standard deviations for combinations with one molecule in the MM region (1/5) and one molecula in the QM region (5/1) regions, indicating greater consistency.

In the prism hexamer, the ωB97X-D/AMOEBA (ESP) and ωB97X-D/GEM (ESP) methods showed lower MAEs compared to the ωB97X-D/AMOEBA (Mull) and ωB97X-D/GEM (Mull) methods, suggesting greater accuracy in their predictions. Specifically, the combination with the majority of molecules in the QM region (5/1) calculated with ωB97X-D/GEM (ESP) show an MAE of 1.40 kcal/mol. On the other hand, when observing standard deviations, the ωB97X-D/AMOEBA (ESP) and ωB97X-D/GEM (ESP) methods also showed lower values compared to the Mulliken-based methods (ωB97X-D/AMOEBA (Mull) and ωB97X-D/GEM (Mull)).

Generally, the results for the Mean Absolute Errors (MAE) and standard deviations in the analysis of ωB97X-D/MM interaction energies in the six water hexamers reveal significant differences in the performance of the four evaluated methods. We observe that

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

methods including Mulliken charges exhibit larger errors, while those including ESP charges show smaller errors. The ωB97X-D/GEM (ESP) method consistently proves to be more accurate in all studied hexamers (bag, book1, book2, cage, chair, prism). This is evidenced by the lower MAEs, particularly in configurations where the majority of water molecules are located in the QM region. This higher accuracy indicates that ωB97X-D/GEM (ESP) is more reliable for predicting interactions and energies in these complex systems. In contrast, the ωB97X-D/AMOEBA (Mull) method shows significantly lower accuracy, reflected in the highest MAEs in various combinations. The ωB97X-D/AMOEBA (ESP) and ωB97X-D/GEM (ESP) methods not only surpass in accuracy but also in consistency. They present lower standard deviations compared to the Mulliken-based methods, indicating that their results are more reproducible and reliable across different molecular configurations. Nevertheless, the spread of the errors is still significant, this spread is due to the approximation of the polarization calculation in LICHEM, which involves a partial self-consistent polarization approach. The spread of the errors is expected to be reduced with a fully self-consistent polarization approach.[54]

Computation times for the evaluation of Coulomb and exchange interactions for the hexamer systems calculated with Psi4, using pre-fitted densities and on-the-fly density fitting on a desktop computer using a single core of an 11th Gen Intel(R) Core(TM) i7-11700 @ 2.50GHz and 70 GB of RAM are shown in Figure S3. Generally, both methods exhibit an increase in average time as the size of the QM region expands. However, the results reveal a significant difference in average times between the two methods. The on-the-fly density fitting method demonstrates higher average times compared to the method using pre-fitted densities across all combinations. The most pronounced difference is observed in the 1/5 combination, where the difference is approximately an order of magnitude. These values indicate that using pre-fitted densities substantially enhances performance, particularly when the number of molecules in the MM region is larger.

Double Proton Transfer in di-Asp. Figure 8 shows the results for a double proton transfer in a model aspartic acid dimer. We have employed ωB97X-D/AMOEBA and ωB97X-D/GEM, to calculate the relative energies along the reaction coordinate. It is observed that both methods show a generally similar trend in the evolution of these relative energies in relation to the reaction coordinate.

From the perspective of the energy of the reactants, we identified that the energy of the

https://doi.org/10.26434/chemrxiv-2024-8k45x ORCID: https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. License: CC BY-NC-ND 4.0

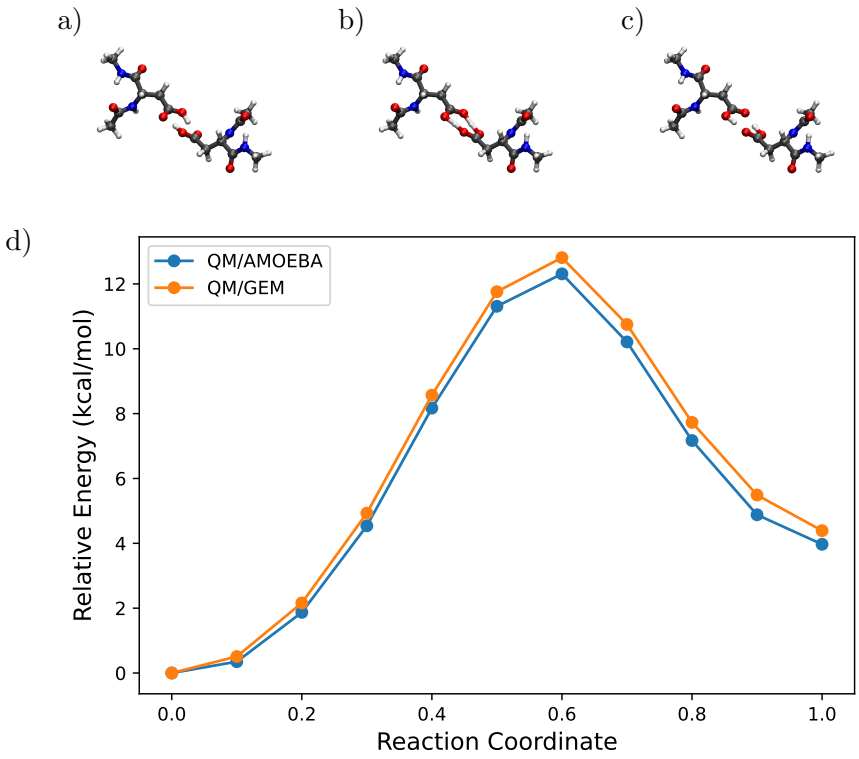

FIG. 8. Double proton transfer in a model aspartic acid dimer. a) to c) Reactant, Transition state, and Product structures, respectively. d) Energy profile from the reaction coordinate.

transition state is 12.31 kcal/mol and 12.81 kcal/mol for ωB97X-D/AMOEBA and ωB97X-D/GEM, respectively. On the other hand, the energy associated with the product was calculated to be 3.97 kcal/mol and 4.39 kcal/mol for each method. These results indicate that, at a quantitative level, there is an agreement between both methods in terms of the description of the proton transfer process.

Furthermore, it was noted that the differences in relative energies between both methods vary slightly along the reaction coordinate, with a maximum difference of 0.61 kcal/mol. This discrepancy is considered reasonable when compared to the absolute values of the relative energies. This pattern suggests that, although there are some numerical differences between the methods, together, they offer a coherent perspective on the relative energy in the studied system. Therefore, these results demonstrate the feasibility of applying ωB97X-D/GEM calculations in more complex systems, extending their use beyond dimers and water hexamers, and indicate the potential of these methods for exploring similar phenomena in more complex systems.

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

V. CONCLUSIONS

In the present work we have successfully implemented the Gaussian Electrostatic Model (GEM) to carry out QM/MM calculations in a seamless manner. This new implementation improves accessibility for the study of complex molecular systems. A key feature of this new implementation is the flexibility to calculate on-the-fly density fitting or use pre-fitted coefficients. This allows for an optimal balance between efficiency and accuracy. Our results show greater accuracy when using electrostatic potential (ESP) derived charges instead of Mulliken charges to represent the QM region in MM calculations. The ωB97X-D/GEM (ESP) method demonstrated superior performance compared to the ωB97X-D/AMOEBA (ESP) method in estimating the interaction energy in water dimers, when compared to the reference values of SAPT(DFT) and ωB97X-D. Furthermore, our results with water hexamers suggest that the ωB97X-D/GEM (ESP) method achieves a favorable balance between accuracy and consistency in various QM/MM combinations. The successful application of the method to the aspartic acid dimer exemplifies its effectiveness in the analysis of heterogeneous systems. This versatility will allow it to expand its applicability to describe complex biological and chemical interactions. Our work not only advances the field of computational chemistry but also lays the foundation for future developments.

SUPPLEMENTARY MATERIAL

Supplementary material is available online in PDF format, including details about individual energy contributions for water dimers, CPU times, mean absolute errors, and standard deviations for water hexamers. Input files for sample systems are also provided in a ZIP file.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. R01GM108583. The authors thank to University of North Texas CASCaM CRUNTCh3 high-performance cluster supported by NSF Grants No. CHE-1531468 and OAC-2117247 for the computational time provided.

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]J. Nochebuena, S. Naseem-Khan, and G. A. Cisneros, "Development and application of quantum mechanics/molecular mechanics methods with advanced polarizable potentials," WIREs Computational Molecular Science **11**, e1515 (2021).

[2]C. R. Vosmeer, A. S. Rustenburg, J. E. Rice, H. W. Horn, W. C. Swope, and D. P. Geerke, "Qm/mm-based fitting of atomic polarizabilities for use in condensed-phase biomolecular simulation," Journal of Chemical Theory and Computation **8**, 3839–3853 (2012).

[3]C. Curutchet, A. Muñoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, and B. Mennucci, "Electronic energy transfer in condensed phase studied by a polarizable qm/mm model," Journal of Chemical Theory and Computation **5**, 1838–1848 (2009).

[4]S. Bonfrate, N. Ferré, and M. Huix-Rotllant, "An efficient electrostatic embedding QM/MM method using periodic boundary conditions based on particle-mesh Ewald sums and electrostatic potential fitted charge operators," The Journal of Chemical Physics **158**, 021101 (2023).

[5]J. P. Pederson and J. G. McDaniel, "DFT-based QM/MM with particle-mesh Ewald for direct, long-range electrostatic embedding," The Journal of Chemical Physics **156**, 174105 (2022).

[6]A. O. Dohn, "Multiscale electrostatic embedding simulations for modeling structure and dynamics of molecules in solution: A tutorial review," International Journal of Quantum Chemistry **120**, e26343 (2020).

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

[7]M. Bondanza, M. Nottoli, L. Cupellini, F. Lipparini, and B. Mennucci, "Polarizable embedding qm/mm: the future gold standard for complex (bio)systems?" Phys. Chem. Chem. Phys. **22**, 14433–14448 (2020).

[8]D. Loco, É. Polack, S. Caprasecca, L. Lagardère, F. Lipparini, J.-P. Piquemal, and B. Mennucci, "A QM/MM Approach Using the AMOEBA Polarizable Embedding: From Ground State Energies to Electronic Excitations," Journal of Chemical Theory and Computation **12**, 3654–3661 (2016).

[9]D. Loco, L. Lagardère, S. Caprasecca, F. Lipparini, B. Mennucci, and J.-P. Piquemal, "Hybrid QM/MM Molecular Dynamics with AMOEBA Polarizable Embedding," Journal of Chemical Theory and Computation **13**, 4025–4033 (2017).

[10]C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder, and P. Ren, "AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids," Journal of Chemical Theory and Computation **14**, 2084–2108 (2018).

[11]Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, "Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins," Journal of Chemical Theory and Computation **9**, 4046–4063 (2013).

[12]G. A. Cisneros, J.-P. Piquemal, and T. A. Darden, "Intermolecular electrostatic energies using density fitting," The Journal of Chemical Physics **123**, 044109 (2005).

[13]G. A. Cisneros, D. Elking, J.-P. Piquemal, and T. A. Darden, "Numerical fitting of molecular properties to hermite gaussians," The Journal of Physical Chemistry A **111**, 12049–12056 (2007).

[14]J. A. Rackers, R. R. Silva, Z. Wang, and J. W. Ponder, "Polarizable water potential derived from a model electron density," Journal of Chemical Theory and Computation **17**, 7056–7084 (2021).

[15]M. L. Laury, L.-P. Wang, V. S. Pande, T. Head-Gordon, and J. W. Ponder, "Revised parameters for the amoeba polarizable atomic multipole water model," The Journal of Physical Chemistry B **119**, 9423–9437 (2015).

[16]A. Duster, C.-H. Wang, and H. Lin, "Adaptive QM/MM for Molecular Dynamics Simulations: 5. On the Energy-Conserved Permuted Adaptive-Partitioning Schemes," Molecules **23**, 2170 (2018).

[17]H. C. Watanabe and Q. Cui, "Quantitative Analysis of QM/MM Boundary Artifacts and Correction in Adaptive QM/MM Simulations," Journal of Chemical Theory and Compu-

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

tation **15**, 3917–3928 (2019).

[18]M. Ditte, M. Barborini, L. Medrano Sandonas, and A. Tkatchenko, "Molecules in environments: Toward systematic quantum embedding of electrons and drude oscillators," Phys. Rev. Lett. **131**, 228001 (2023).

[19]P. Lafiosca, F. Rossi, F. Egidi, T. Giovannini, and C. Cappelli, "Multiscale frozen density embedding/molecular mechanics approach for simulating magnetic response properties of solvated systems," Journal of Chemical Theory and Computation **20**, 266–279 (2024).

[20]C. I. Viquez Rojas and L. V. Slipchenko, "Exchange-repulsion in QM/EFP excitation energies – beyond polarizable embedding," Journal of Chemical Theory and Computation (2020), 10.1021/acs.jctc.9b01156.

[21]X. Chen, Z. Qu, B. Suo, and J. Gao, "A self-consistent coulomb bath model using density fitting," Journal of Computational Chemistry **41**, 1698–1708 (2020).

[22]T. J. Giese and D. M. York, "Quantum mechanical force fields for condensed phase molecular simulations." Journal of physics. Condensed matter : an Institute of Physics journal **29**, 383002 (2017).

[23]X. Chen and J. Gao, "Fragment Exchange Potential for Realizing Pauli Deformation of Interfragment Interactions," The Journal of Physical Chemistry Letters **11**, 4008–4016 (2020).

[24]L. O. Jones, M. A. Mosquera, G. C. Schatz, and M. A. Ratner, "Embedding Methods for Quantum Chemistry: Applications from Materials to Life Sciences," Journal of the American Chemical Society **142**, 3281–3295 (2020).

[25]R. E. Duke and G. A. Cisneros, "Ewald-based methods for Gaussian integral evaluation: application to a new parameterization of GEM," Journal of molecular modeling **25**, 307 (2019).

[26]J.-P. Piquemal and G. A. Cisneros, "Status of the Gaussian Electrostatic Model, a Density-Based Polarizable Force Field," in *Many-body effects and electrostatics*, multi-scale computations of Biomolecules No. 28, edited by P. R. Qiang Cui and M. Meuwly (Pan Standford Publishing, 2016) pp. 269–299.

[27]R. E. Duke, O. N. Starovoytov, J.-P. Piquemal, and G. A. Cisneros, "Gem*: A molecular electronic density-based force field for molecular dynamics simulations," Journal of Chemical Theory and Computation **10**, 1361–1365 (2014).

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

[28]G. A. Cisneros, "Application of Gaussian Electrostatic Model (GEM) Distributed Multipoles in the AMOEBA Force Field," Journal of Chemical Theory and Computation **8**, 5072–5080 (2012).

[29]H. Gökcan, E. Kratz, T. A. Darden, J.-P. Piquemal, and G. A. Cisneros, "QM/MM simulations with the Gaussian Electrostatic Model: A density-based polarizable potential," The Journal of Physical Chemistry Letters **9**, 3062–3067 (2018).

[30]L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M. J. Schnieders, N. Gresh, Y. Maday, P. Y. Ren, J. W. Ponder, and J.-P. Piquemal, "Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields," Chem. Sci. **9**, 956–972 (2018).

[31]D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O'Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, I. Schaefer, Henry F., A. Y. Sokolov, K. Patkowski, I. DePrince, A. Eugene, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill, "PSI4 1.4: Open-source software for high-throughput quantum chemistry," The Journal of Chemical Physics **152**, 184108 (2020).

[32]E. G. Kratz, A. R. Walker, L. Lagardère, F. Lipparini, J.-P. Piquemal, and G. A. Cisneros, "LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields," Journal of Computational Chemistry **37**, 1019–1029 (2016).

[33]H. Gökcan, E. A. Vázquez-Montelongo, and G. A. Cisneros, "LICHEM 1.1: Recent Improvements and New Capabilities," Journal of Chemical Theory and Computation **15**, 3056–3065 (2019).

[34]M. Nottoli, M. Bondanza, P. Mazzeo, L. Cupellini, C. Curutchet, D. Loco, L. Lagardère, J.-P. Piquemal, B. Mennucci, and F. Lipparini, "Qm/amoeba description of properties and dynamics of embedded molecules," WIREs Computational Molecular Science **13**, e1674 (2023).

[35]M. Nottoli, M. Bondanza, F. Lipparini, and B. Mennucci, "An enhanced sampling QM/AMOEBA approach: The case of the excited state intramolecular proton transfer in solvated 3-hydroxyflavone," The Journal of

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

Chemical Physics **154**, 184107 (2021), https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0046844/15588132/184107_1_online.pdf.

[36]J. Nochebuena and G. A. Cisneros, "Polarizable MD and QM/MM investigation of acrylamide-based leads to target the main protease of SARS-CoV-2," The Journal of Chemical Physics **157**, 185101 (2022).

[37]J. Nochebuena, L. Quintanar, A. Vela, and G. A. Cisneros, "Structural and electronic analysis of the octarepeat region of prion protein with four Cu2+ by polarizable MD and QM/MM simulations," Phys. Chem. Chem. Phys. **23**, 21568–21578 (2021).

[38]U. C. Singh and P. A. Kollman, "An approach to computing electrostatic charges for molecules," Journal of Computational Chemistry **5**, 129–145 (1984).

[39]B. H. Besler, K. M. Merz Jr., and P. A. Kollman, "Atomic charges derived from semiempirical methods," Journal of Computational Chemistry **11**, 431–439 (1990).

[40]R. S. Mulliken, "Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I," The Journal of Chemical Physics **23**, 1833–1840 (1955).

[41]G. A. Cisneros, J.-P. Piquemal, and T. A. Darden, "Quantum mechanics/molecular mechanics electrostatic embedding with continuous and discrete functions," The Journal of Physical Chemistry B **110**, 13682–13684 (2006).

[42]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, "Gaussian˜16 Revision C.01," (2016).

[43]B. J. Smith, D. J. Swanton, J. A. Pople, I. Schaefer, Henry F., and L. Radom, "Transition structures for the interchange of hydrogen atoms within the water dimer," The Journal of Chemical Physics **92**, 1240–1247 (1990).

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

[44]J.-D. Chai and M. Head-Gordon, "Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections," Phys. Chem. Chem. Phys. **10**, 6615–6620 (2008).

[45]D. E. Woon and J. Dunning, Thom H., "Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties," The Journal of Chemical Physics **100**, 2975–2988 (1994), https://pubs.aip.org/aip/jcp/article-pdf/100/4/2975/10771441/2975_1_online.pdf.

[46]R. A. Kendall, J. Dunning, Thom H., and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," The Journal of Chemical Physics **96**, 6796–6806 (1992).

[47]S. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors," Molecular Physics **19**, 553–566 (1970).

[48]E. G. Kratz, R. E. Duke, and G. A. Cisneros, "Long-range electrostatic corrections in multipolar/polarizable QM/MM simulations," Theoretical Chemistry Accounts **135**, 166 (2016).

[49]D. Fang, R. E. Duke, and G. A. Cisneros, "A new smoothing function to introduce long-range electrostatic effects in QM/MM calculations," The Journal of Chemical Physics **143**, 44103 (2015).

[50]S. K. Burger and W. Yang, "Quadratic string method for determining the minimum-energy path based on multiobjective optimization," The Journal of Chemical Physics **124**, 054109 (2006).

[51]R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, "Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions," The Journal of Chemical Physics **72**, 650–654 (2008).

[52]P. Ren, C. Wu, and J. W. Ponder, "Polarizable Atomic Multipole-based Molecular Mechanics for Organic Molecules." Journal of Chemical Theory and Computation **7**, 3143–3161 (2011).

[53]J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. J. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson, and T. Head-Gordon, "Current status of the AMOEBA polarizable force field," The Journal of Physical Chemistry B **114**, 2549–2564 (2010).

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0

[54]E. Lambros, F. Lipparini, G. A. Cisneros, and F. Paesani, "A many-body, fully polarizable approach to QM/MM simulations," Journal of Chemical Theory and Computation **16**, 7462–7472 (2020).

https://doi.org/10.26434/chemrxiv-2024-8k45x **ORCID:** https://orcid.org/0000-0003-0707-1066 Content not peer-reviewed by ChemRxiv. **License:** CC BY-NC-ND 4.0