

RESEARCH ARTICLE | APRIL 01 2024

Relative cooperativity in neutral and charged molecular clusters using QM/MM calculations

Jorge Nochebuena ; Shubin Liu ; G. Andrés Cisneros  



J. Chem. Phys. 160, 134301 (2024)

<https://doi.org/10.1063/5.0203020>



Articles You May Be Interested In

Polarizable MD and QM/MM investigation of acrylamide-based leads to target the main protease of SARS-CoV-2

J. Chem. Phys. (November 2022)

Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations

J. Chem. Phys. (August 2017)

Mapping the Drude polarizable force field onto a multipole and induced dipole model

J. Chem. Phys. (June 2017)



The Journal of Chemical Physics

Special Topics Open for Submissions

[Learn More](#)

Relative cooperativity in neutral and charged molecular clusters using QM/MM calculations

Cite as: J. Chem. Phys. 160, 134301 (2024); doi: 10.1063/5.0203020

Submitted: 8 February 2024 • Accepted: 13 March 2024 •

Published Online: 1 April 2024



Jorge Nochebuena,¹  Shubin Liu,^{2,3}  and G. Andrés Cisneros^{1,4,a)} 

AFFILIATIONS

¹ Department of Physics, University of Texas at Dallas, Richardson, Texas 75080, USA

² Research Computing Center, University of North Carolina, Chapel Hill, North Carolina 27599, USA

³ Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599, USA

⁴ Department of Chemistry and Biochemistry, University of Texas at Dallas, Richardson, Texas 75080, USA

^{a)} Author to whom correspondence should be addressed: andres@utdallas.edu

ABSTRACT

QM/MM methods have been used to study electronic structure properties and chemical reactivity in complex molecular systems where direct electronic structure calculations are not feasible. In our previous work, we showed that non-polarizable force fields, by design, describe intermolecular interactions through pairwise interactions, overlooking many-body interactions involving three or more particles. In contrast, polarizable force fields account partially for many-body effects through polarization, but still handle van der Waals and permanent electrostatic interactions pairwise. We showed that despite those limitations, polarizable and non-polarizable force fields can reproduce relative cooperativity achieved using density functional theory due to error compensation mechanisms. In this contribution, we assess the performance of QM/MM methods in reproducing these phenomena. Our study highlights the significance of the QM region size and force field choice in QM/MM calculations, emphasizing the importance of parameter validation to obtain accurate interaction energy predictions.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0203020>

I. INTRODUCTION

In the field of computational chemistry, hybrid quantum mechanics/molecular mechanics (QM/MM) methods have emerged as an essential tool for the study of complex molecular systems.^{1,2} QM/MM methods combine quantum mechanics (QM), which deals with the precise description of electrons, with the efficiency of molecular mechanics (MM). The choice of the QM region size plays a crucial role in QM/MM calculations and can have a significant impact on the accuracy of the results obtained because it determines how many electrons and atoms are included in the QM calculation. If the QM region is too small, critical interactions between electrons and atoms may be missed, leading to inaccurate results. Conversely, if the QM region is too large, the calculation can become computationally expensive and challenging to handle.

Selecting the appropriate size for the QM region in QM/MM calculations is an actively researched topic.³ For instance, some studies have shown that free energy reaction profiles converge rapidly as

the QM region size increases, indicating a minimal impact beyond a certain threshold.^{4–6} However, other research highlights significant shifts in electronic properties and free energy barriers when the QM region size changes, highlighting the critical role of size selection for achieving accurate simulation results.^{7–9} One of the most common techniques for the systematic construction of the QM region is based on the distance of the atoms to the active site.¹⁰ However, alternative criteria have also been proposed.^{6,11–18}

QM/MM calculations offer numerous advantages compared to purely QM or MM calculations for specific systems. For example, it provides an approach to study chemical reactions of large biomolecules, without sacrificing precision in the electronic description of the regions of interest and reducing the computational effort for large systems.^{1,19,20} Additionally, QM/MM calculations can be used to study the influence of solvation on chemical reactions.^{21–23} In some of these cases, cooperativity can have a significant influence, although a systematic investigation of cooperativity in QM/MM calculations has not been reported to our knowledge.

Cooperativity is a phenomenon of utmost importance in various biological processes, such as molecular recognition, homochirality, protein folding, and self-assembly.^{24–31} A notable example of cooperativity is observed in the binding of oxygen to hemoglobin. When an oxygen molecule binds to one of the four binding sites on hemoglobin, it triggers a structural change that facilitates the binding of oxygen to the remaining sites. The meaning of cooperativity varies across different fields, such as biochemistry or physics. However, this phenomenon occurs when the interactions between multiple components exceed their individual contributions, resulting in enhanced stability or activity. In this work, we are interested in intermolecular cooperativity, which involves the synergistic effects that occur when molecules interact with each other in a system.³² Cooperativity can be positive or negative depending on whether the interactions between system components benefit or hinder each other. Conversely, when the interactions within a system can be explained solely by the individual interactions of its components, cooperativity is not observed.

The intermolecular cooperativity phenomenon has been studied using density functional theory in neutral and charged molecular clusters consisting of up to 20 building blocks.^{33–36} We have recently studied the same systems with force fields and have shown that force fields are capable of qualitatively reproducing the cooperative effects observed with electronic structure calculations.³⁷ In the case of neutral systems, positive cooperative effects were identified, whereas charged systems exhibited negative cooperativity. This study aims to examine the performance of QM/MM simulations with regard to cooperativity to assess the effectiveness of this methodology in replicating the aforementioned phenomena. Hence, it is important to assess if QM/MM calculations can replicate the cooperativity phenomena in molecular clusters, as they are crucial for precise determination of properties such as solvation energy, diffusion, and residence times, which have been studied computationally using Born–Oppenheimer QM/MM MD simulations.^{21–23}

II. THEORETICAL FRAMEWORK

The purpose of this work was to investigate the performance of the combination of classical non-polarizable or polarizable potentials with DFT-based methods in a QM/MM context with respect to cooperativity effects. To this end, we used the $\text{NH}_3(\text{H}_2\text{O})_n$, $\text{Li}^+(\text{H}_2\text{O})_n$, and $\text{F}^-(\text{H}_2\text{O})_n$ systems as representatives of neutral, positively charged, and negatively charged systems, respectively. These systems can be described in the form AB_n , where A is the accessory component, B is the building block, and n is the number of building blocks. The accessory components are the NH_3 molecule and the Li^+ and F^- ions, while the building blocks are water molecules. We have recently conducted prior investigations on the previous systems by exclusively employing the AMOEBA polarizable force field, alongside the OPLS and AMBER non-polarizable force fields.³⁷

We determined the relative cooperativity by utilizing the adiabatic interaction energy per building block, which is derived from the total adiabatic interaction energy. While we have previously introduced the mathematical expression in our prior research, a slight modification was necessary to make it suitable for

integration into QM/MM calculations. The adiabatic interaction energy is defined as

$$E_{\text{int}}^{\text{adiab}} = E(\text{AB}_n) - nE_{\text{B}}(\text{R}_{0\text{B}}) - E_{\text{A}}(\text{R}_{0\text{A}}), \quad (1)$$

where $E(\text{AB}_n)$ is the total energy of the cluster, E_{B} is the total energy of a building block, and E_{A} is the total energy of the accessory component. $\text{R}_{0\text{A}}$ and $\text{R}_{0\text{B}}$ indicate that the total energies of the accessory component and the building blocks, respectively, are calculated from optimized DFT molecular geometries. The total QM/MM energy, $E_{\text{QMMM}}^{\text{total}}$, of the cluster involves a combination of the QM and MM regions,

$$E(\text{AB}_n) = E_{\text{QMMM}}^{\text{total}} = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QMMM}}, \quad (2)$$

where E_{QM} is the total energy of the QM region, E_{MM} is the total energy of the MM region, and E_{QMMM} is the interaction energy between the QM and MM regions. Therefore,

$$E_{\text{int}}^{\text{adiab}} = E_{\text{QMMM}}^{\text{total}} - (n - m)E_{\text{B}}^{\text{MM}}(\text{R}_{0\text{B}}) - mE_{\text{B}}^{\text{QM}}(\text{R}_{0\text{B}}) - E_{\text{A}}^{\text{QM}}(\text{R}_{0\text{A}}), \quad (3)$$

where $1 \leq m \leq 4$ indicates the number of water molecules in the QM region. When $m = 0$, it simplifies to a scenario where the system is exclusively characterized within the MM region, like our previous work. When $m = n$, it signifies a situation where the system is exclusively characterized within the QM region, like in the previous work of Liu.³⁶

The interaction energy per building block, denoted as E_n , is defined as the adiabatic interaction energy divided by the number of building blocks, n , that is,

$$E_n = E_{\text{int}}^{\text{adiab}} / n. \quad (4)$$

The cooperativity index, represented as k , is defined as the negative change in adiabatic interaction energy per change in the number of building blocks, that is,

$$k = -(\partial E_n / \partial n). \quad (5)$$

The parameter k can have three distinct scenarios. When k assumes a positive value, it implies positive cooperativity, signifying that the introduction of an extra building block enhances interactions. Conversely, if k takes on a negative value, it denotes negative cooperativity, indicating that the addition of a building block diminishes interactions. Finally, when k equals zero, there is no cooperativity, signifying that the inclusion of an extra building block has no effect on interactions.

III. COMPUTATIONAL DETAILS

We performed single point QM/MM calculations on the previously optimized DFT molecular geometries taken from the work of Liu and Rong.³⁶ Each cluster is the global minimum for each number of building blocks. Additionally, we optimized a water molecule and an ammonia molecule using density functional theory implemented in Gaussian16.³⁸ For all calculations, the exchange–correlation functionals and basis sets for the QM regions are the same as the ones employed by Liu and Rong,³⁶ specifically $\omega\text{B97X-D/ aug-cc-pVDZ}$ for $\text{NH}_3(\text{H}_2\text{O})_n$,^{39–42} $\text{M06-2X/6-311+G(d,p)}$ for $\text{Li}^+(\text{H}_2\text{O})_n$,^{43–47} and $\text{B3LYP/ aug-cc-pVDZ}$ for

$F^-(H_2O)_n$.^{41,48,49} To describe the MM region, we employed the AMOEBA model for all polarizable force field calculations. For non-polarizable calculations, we used the OPLS force field for the ammonia and fluoride systems and the AMBER force field for the lithium systems.^{43–47,50–55} Water molecules were described using the flexible TIP3P implemented in Tinker for the non-polarizable force fields.^{52,56} The O–H bond stretching force constants are 553.0 and 600 kcal/mol/Å² and the H–O–H angle bending force constant are 100.0 and 75.0 kcal/mol/rad² for the Amber and OPLS force fields, respectively.⁵⁷

To separate each cluster into distinct QM and MM regions, we explored various options by adjusting the QM region's size (see Fig. 1). In the initial scenario, exclusively the accessory component occupied the QM region, with all water molecules situated in the MM region. In subsequent configurations, we included one to four water molecules into the QM region. The choice of which water molecules to include in the QM region was determined based on the shortest distances between the center of mass of each water molecule and the accessory component. We utilized the LICHEM software to conduct QM/MM calculations.^{58,59} LICHEM serves as a versatile interface, enabling seamless integration between electronic structure codes and molecular mechanics codes. We employed Gaussian code to calculate the QM energies, while we utilized the Tinker7 code to acquire the MM energies. The electrostatic and polarization (for polarizable QM/MM simulations with AMOEBA) embedding procedures have been described in detail in Refs. 58 and 59. In addition, we performed SAPT(DFT) calculations, with the SAPT expansion

truncated at the same level of SAPT0, using the Psi4 code to calculate the interaction energy errors between QM and MM regions.⁶⁰

IV. RESULTS AND DISCUSSION

It has been previously shown that both non-polarizable and polarizable classical force fields can provide a qualitative description (up to a constant factor) of cooperativity.³⁷ For the former, this is due to a combination of the fact that two-body interactions can comprise over 80% of the total interaction energy^{63–65} and the use of bulk properties to parameterize non-bonded interactions. For polarizable potentials, such as AMOEBA, in addition to the above-mentioned factors, these force fields also partially account for many-body contributions through the explicit polarization term.

In the present contribution, we evaluated the performance of QM/MM methods in reproducing both the interaction energy and the cooperativity phenomenon in three representative neutral and charged systems: $NH_3(H_2O)_n$, $Li^+(H_2O)_n$, and $F^-(H_2O)_n$. The objective of this study was to determine the impact of the integration of QM and MM regions for the description of these systems. To achieve this, we systematically expand the size of the QM region, starting from a single molecule and extending it to encompass up to five molecules, that is, the accessory component and four water molecules. We refrain from further increasing the number of molecules because with each additional molecule, the available number of clusters for study diminishes.

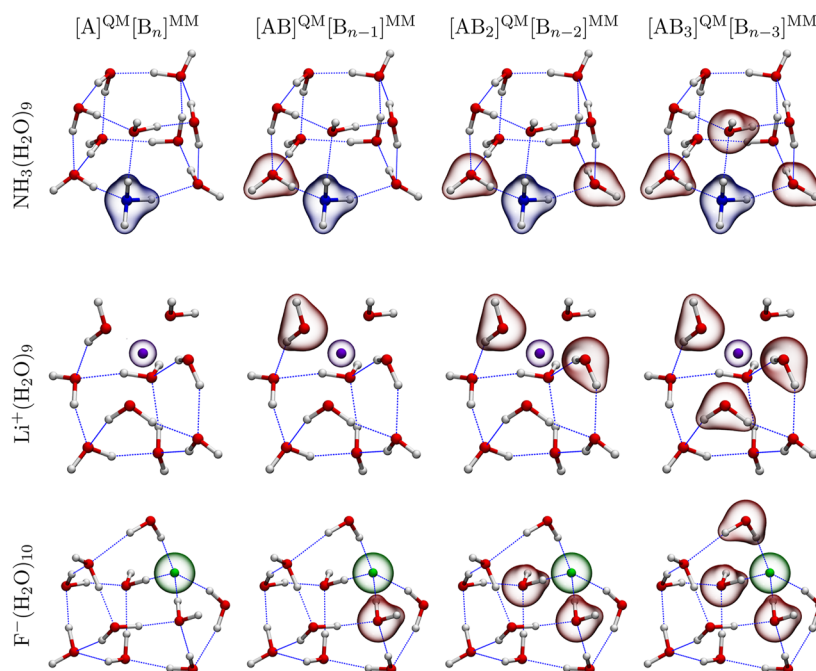


FIG. 1. Diagram illustrating the systematic expansion of the QM region within a cluster of specified size. The QM region is highlighted using blown glass isosurfaces. Molecules were visualized using VMD.^{61,62}

The results of the QM/MM calculations are shown in Figs. 2–4. Figures 2–4 show the total adiabatic interaction energies, Eq. (3), and the interaction energies per building block, Eq. (4). The cooperativity index, Eq. (5), is determined by the shape of the interaction energy per building block. If the interaction energy per building block increases (decreases) as the number of building blocks increases (decreases), it indicates that the system exhibits positive (negative) cooperativity. The blue line corresponds to systems where the QM region only includes the accessory component, which can be NH_3 , Li^+ , or F^- . The orange, green, red, and purple lines correspond to systems where the QM region contains the accessory component along with one, two, three, or four water molecules, respectively. We have included the results previously reported using only force fields as in Ref. 37. Total interaction energy errors and interaction energy errors per building block with respect to the electron structure calculations are shown in Figs. S1–S3.

For the ammonia system (Fig. 2 and Fig. S1), it is observed that the errors in the total adiabatic interaction energies exhibit a nearly linear increase as the cluster size expands, regardless of the force field chosen within the MM region. Surprisingly, when the AMOEBA force field is employed, none of the combinations of QM and MM regions offer an improvement compared to

previous results relying solely on force fields, especially for the larger groups. In fact, the most substantial errors arise when the QM region spans one ammonia molecule and four water molecules, as indicated by the purple line. However, when examining the interaction energies per building block, the same combination demonstrates the fewest errors for the $\text{NH}_3(\text{H}_2\text{O})_5$ cluster, where only one water molecule is described within the classical region. In contrast, when the OPLS force field is used to describe the MM region, an opposite trend emerges. All combinations of QM/MM regions present smaller errors than those reported previously. In fact, as the size of the QM region increases, the errors in the energies per building block decrease, and the most favorable results are obtained by combining one ammonia molecule with four water molecules.

Similar results are observed for the Li^+ system (Fig. 3 and Fig. S2) when only the accessory component is included in the QM region (indicated by the blue line). However, as the size of the QM region increases, the errors decrease. This behavior suggests that as the system description is improved by increasing the size of the QM region, the errors are reduced. This effect is consistent with the expectation of improved accuracy with increasing QM subsystem size, although this is not observed for the neutral system. We can observe that the M06-2X/AMBER combination is more sensitive

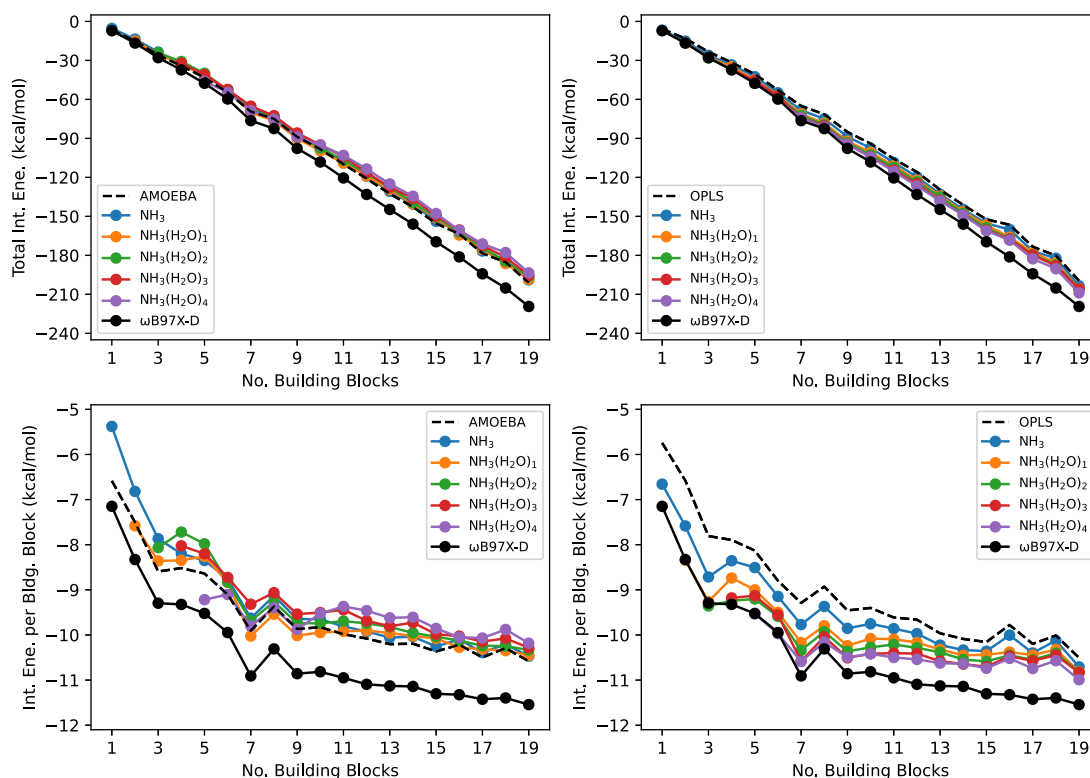


FIG. 2. Total interaction energies (top) and interaction energy per building block (bottom) for the $\text{NH}_3(\text{H}_2\text{O})_n$ system. Left and right sides correspond to AMOEBA and OPLS force fields as descriptors for MM regions, respectively. The black solid line represents the reference values calculated using $\omega\text{B97X-D/aug-cc-pVDZ}$. The dashed line represents the interaction energies reported previously using only the force field. The remaining lines indicate the molecules included in the QM region, with all other molecules being calculated using their respective force fields.

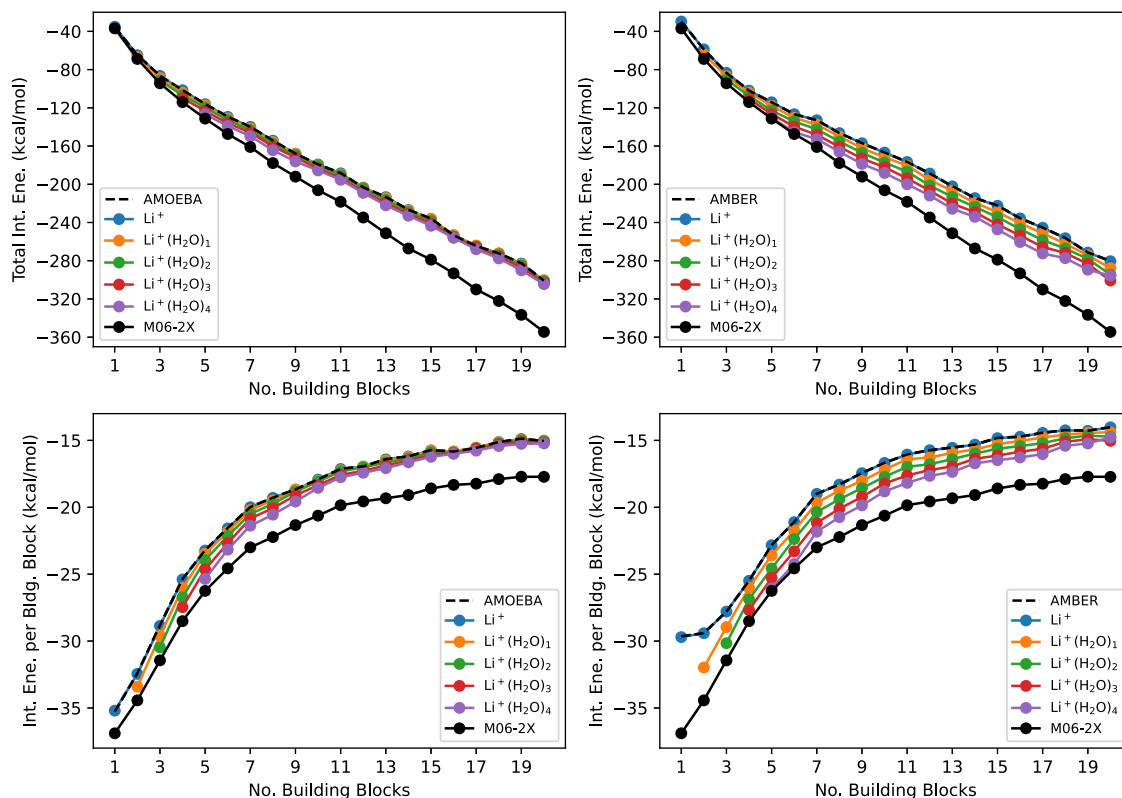


FIG. 3. Total interaction energies (top) and interaction energy per building block (bottom) for the $\text{Li}^+(\text{H}_2\text{O})_n$ system. Left and right sides correspond to AMOEBA and AMBER force fields as descriptors for MM regions, respectively. The black solid line represents the reference values calculated using M06-2X/6-311+G(d,p). The dashed line represents the interaction energies reported previously using only the force field. The remaining lines indicate the molecules included in the QM region, with all other molecules being calculated using their respective force fields.

to the increase in size of the QM region compared to the M06-2X/AMOEBA combination. However, better results were obtained with the AMOEBA force field for both smaller and larger groups.

Figure 4 and Fig. S3 show the results for the $\text{F}^-(\text{H}_2\text{O})_n$ system. It is evident that the errors are more pronounced in smaller clusters, whether we consider the polarizable or non-polarizable force field. Specifically, the AMOEBA force field exhibits errors of 13 kcal/mol for the smallest cluster, while the OPLS force field demonstrates errors exceeding 20 kcal/mol for the same cluster. Remarkably, none of the combinations of QM and MM appear to enhance the description when compared to the previously published results for the AMOEBA force field. Nevertheless, an improvement in results becomes increasingly evident as we expand the size of the QM region. For larger clusters, the error diminishes considerably, reaching values of less than 1 kcal/mol per construction block. Notably, the polarizable force field exhibits a reduction in errors as we increase the size of the QM region.

The previous results indicate that, in most of the cases, including more molecules in the QM region generally enhances the description compared to systems where only the accessory component is described in the QM region for some systems. The size dependence of the QM region in QM/MM calculations has been

explored in several publications.^{6,10,11,15,16,18,66–69} Our findings also reveal that, for the $\text{NH}_3(\text{H}_2\text{O})_n$ (QM/AMOEBA) system and, particularly, for the $\text{F}^-(\text{H}_2\text{O})_n$ system, the errors in QM/MM energies are generally greater than those obtained previously using only force fields. This may seem counter-intuitive since one would expect the interaction energy errors to decrease as the QM region size increases. In other words, we would anticipate that all three systems would exhibit behavior similar to that observed in the $\text{Li}^+(\text{H}_2\text{O})_n$ system.

From Figs. S1–S3, it is evident that a significant portion of errors come from the MM region description. To confirm this, we conducted linear regressions on these errors to evaluate the fitting coefficients. By applying linear regression to these errors, we aimed to quantify and understand their systematic nature through the evaluation of fitting coefficients. This approach not only can be used to confirm the source of inaccuracies but also could serve as a foundational step toward refining QM/MM methodologies for enhanced accuracy in computational simulations. Figure S4 reveals that as the size of the QM region increases, the slope becomes more pronounced for the $\text{NH}_3(\text{H}_2\text{O})_n$ system analyzed using $\omega\text{B97X-D/aug-cc-pVDZ//AMOEBA}$. This suggests that increasing the QM region size results in more significant errors in larger systems. Conversely,

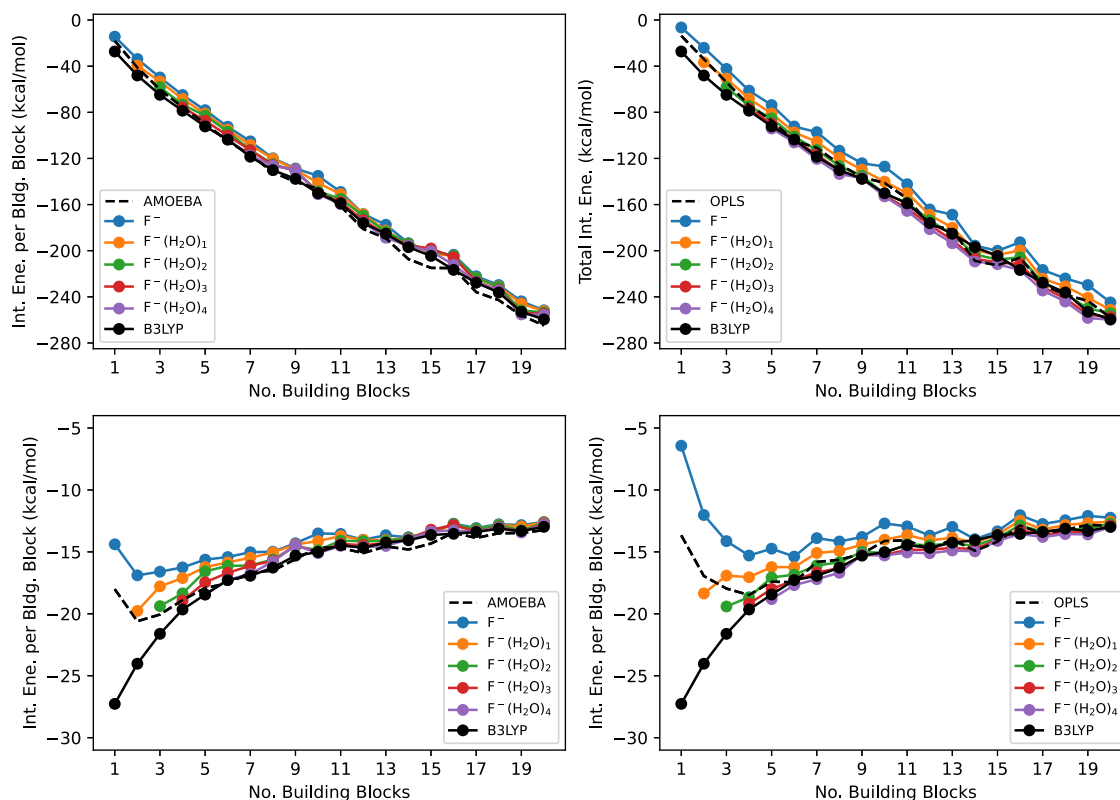


FIG. 4. Total interaction energies (top) and interaction energy per building block (bottom) for the $F^-(H_2O)_n$ system. Left and right sides correspond to AMOEBA and OPLS force fields as descriptors for MM regions, respectively. The black solid line represents the reference values calculated using B3LYP/aug-cc-pVDZ. The dashed line represents the interaction energies reported previously using only the force field. The remaining lines indicate the molecules included in the QM region, with all other molecules being calculated using their respective force fields.

for the same system calculated with ω B97X-D/aug-cc-pVDZ//OPLS, the errors decrease as the system size grows (see Fig. S5).

A similar trend is observed for the $Li^+(H_2O)_n$ system. The slope increases with an expanding QM region by using the M06-2X/6-311+G(d,p)//AMOEBA method (Fig. S6). In contrast, the slope diminishes when analyzed with the M06-2X/6-311+G(d,p)//AMBER method (Fig. S7). Interestingly, for the $F^-(H_2O)_n$ system, the slopes are negative, suggesting that the most substantial errors are associated with the smallest clusters (Figs. S8 and S9).

Figures S4–S9 show that linear regressions can be used to estimate interaction energy errors. Thus, we employed these fitting coefficients to correct the total interaction energies and subsequently the interaction energies per building block. The results are represented in Fig. 5. It is noteworthy that post-correction interaction energies in QM/MM calculations produce very accurate values. However, it is important to note that these corrections vary depending on the type of system, the size of the QM region, and the level of theory applied.

Our results suggest that for larger clusters, errors can be associated with the errors originating from the force fields. However, it is clear that among the charged systems, the $F^-(H_2O)_n$ system exhibits

the largest errors for the smaller clusters. This suggests that the error likely arises from the QM–MM interaction. Therefore, we calculated ESP and Hirshfeld charges on the Li^+ and F^- ions in their respective systems to perform a comparison. From the results shown in Tables S1–S8, we observed that Hirshfeld charges predict significant changes as the QM region size increases, especially in the case of the Li^+ ion, where the values range from 1.0 to 0.27 electrons. On the other hand, ESP charges predict a minimal change for this ion. However, both methods agree that the charge of the F^- ion is distributed among the water molecules as the QM region size increases. This issue has been previously observed, and flexible boundary schemes have been proposed to facilitate partial charge transfer between QM and MM subsystems.^{70–73} These studies have suggested the importance of an open boundary for improving accuracy and efficiency in QM/MM calculations. Our calculations do not show significant differences depending on the force field utilized. Similar results were obtained for polarizable and non-polarizable force fields.

In order to understand the source of the errors for the description of cooperativity within a QM/MM context, it is necessary to consider the individual contributions from Eq. (2). The errors within the QM region are minima since both the reference and test systems are described using the same methodology. For the

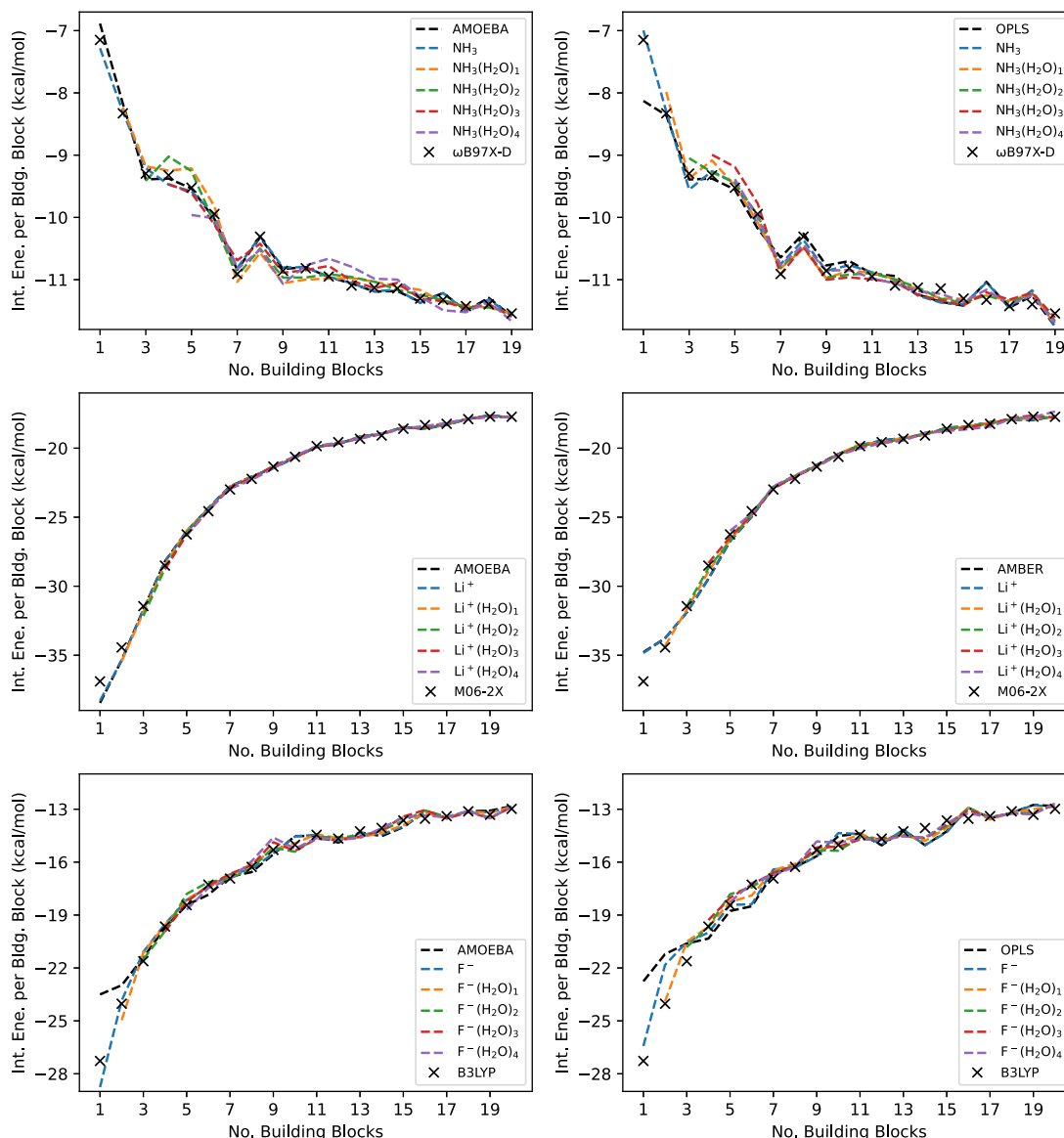


FIG. 5. Corrected interaction energies per building block of $\text{NH}_3(\text{H}_2\text{O})_n$ (top), $\text{Li}^+(\text{H}_2\text{O})_n$ (middle), and $\text{F}^-(\text{H}_2\text{O})_n$ (bottom) systems. On the left side, the results are for the AMOEBA force field, while on the right side, the results are for both OPLS and AMBER force fields. Force field names in the legend indicate FF-only results, while XC functional names denote DFT-only results. The remaining names signify the QM size.

MM regions, the errors in the interaction energies are the same as those reported previously in Ref. 37. Therefore, an additional source of error lies in the final term of Eq. (2), namely, the interaction between the QM and MM regions. To investigate the source of the errors arising from the interaction between the two subsystems, we have carried out an energy decomposition analysis based on SAPT(DFT).

For these SAPT(DFT) calculations, the systems were set up such that one fragment corresponds to the molecules in the QM subsystem, and the other fragment corresponds to all the molecules

in the MM subsystem. We computed the errors for the first three combinations of QM/MM regions, that is, only with the solute molecule in the QM subsystem and with the solute and one or two water molecules in the QM subsystem. We considered only the first ten clusters because the major deviations in the interaction energy per building block in the $\text{F}^-(\text{H}_2\text{O})_n$ system occur in this range. As shown in Fig. 6 (top row), increasing the size of the QM region results in an increase in error for the total inter-molecular interaction energy compared to SAPT(DFT) energies. The $\text{NH}_3(\text{H}_2\text{O})_n$ system exhibited the smallest differences, whereas the charged

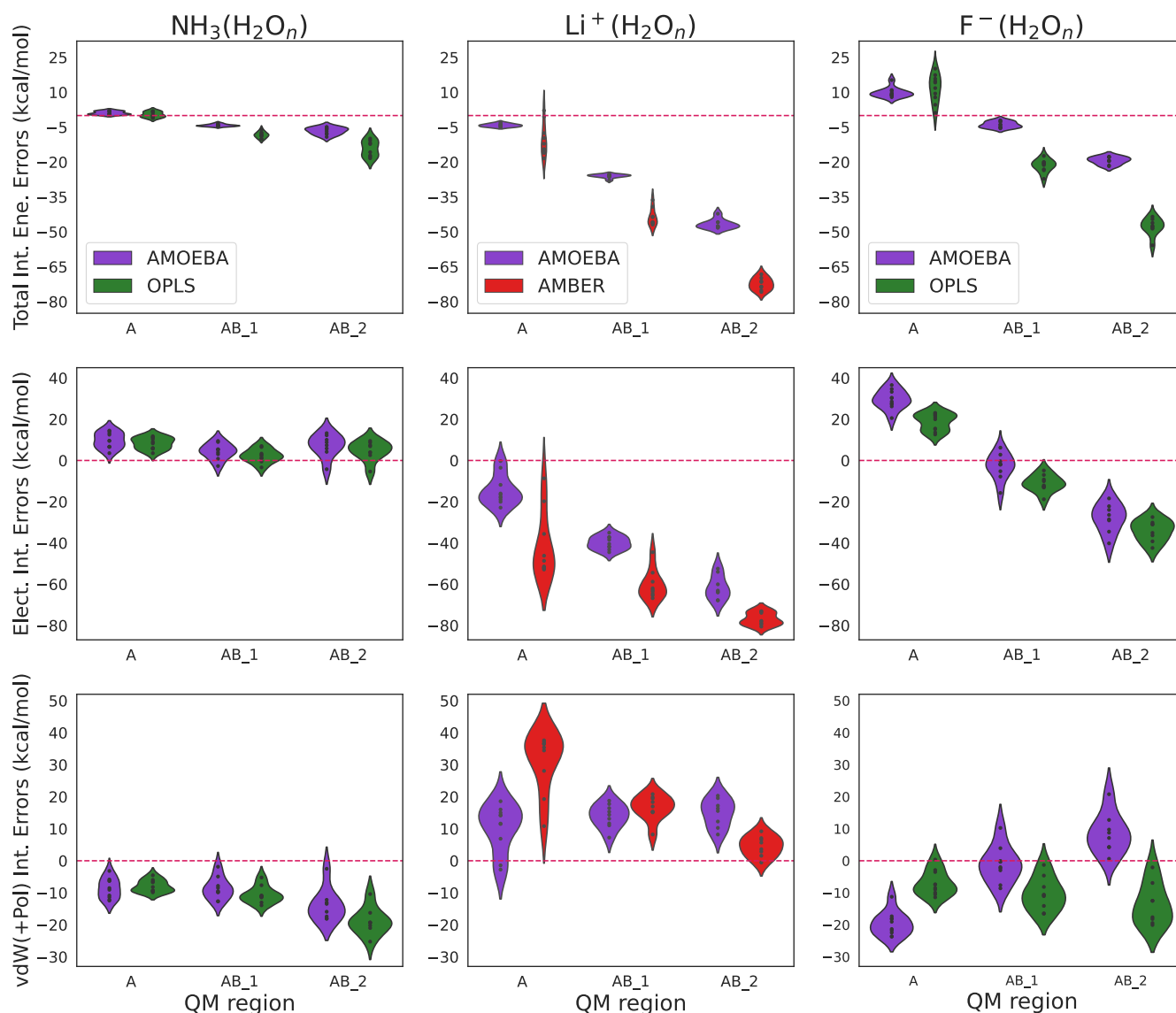


FIG. 6. Violin plots for QM–MM interaction errors for $\text{NH}_3(\text{H}_2\text{O})_n$ (left), (b) $\text{Li}^+(\text{H}_2\text{O})_n$ (middle), and (c) $\text{F}^-(\text{H}_2\text{O})_n$ (right) systems with respect to SAPT(DFT) interaction energies. The first row displays errors in total interaction energies. The second row displays errors in electrostatic interaction energies. The third row shows errors in vdW (AMBER and OPLS) or vdW + polarization (AMOEBA) interaction energies. Purple, green, and red correspond to AMOEBA, OPLS, and AMBER results, respectively. A, AB_1, and AB_2, indicate the size of the QM region.

systems, particularly the $\text{Li}^+(\text{H}_2\text{O})_n$ system, showed significantly more substantial changes. Our findings indicate that the expansion of the QM region has a greater impact on the results of QM/OPLS or QM/AMBER calculations, making them more sensitive to changes in the QM size. In contrast, the errors in interaction energies calculated with QM/AMOEBA are more consistent across all analyzed systems.

The interaction energies in SAPT(DFT) analysis can be decomposed into four distinct components: electrostatic, exchange, induction, and dispersion forces. Then, the next step was to separate the

interaction energy contributions in QM/MM calculations to compare them with their counterparts in SAPT(DFT) calculations. First, we conducted a comparison between the electrostatic contribution, which emerges from the interactions between charges in the MM region and electron density, and its SAPT(DFT) counterpart. When employing OPLS and AMBER force fields, we then proceeded to contrast van der Waals interactions with the combined values of exchange, induction, and dispersion contributions. In the case of the AMOEBA force field, we considered the sum of polarization and van der Waals interactions for the comparison.

The electrostatic interaction energy errors (Fig. 6, middle row) indicate that the $\text{NH}_3(\text{H}_2\text{O})_n$ system shows minimal changes when increasing the size of the QM region. However, errors in electrostatic energy increase for charged systems. In particular, the $\text{Li}^+(\text{H}_2\text{O})_n$ system exhibits the most significant deviations. Similar to the previous results, larger errors are observed with non-polarizable force fields as the QM size increases.

Regarding errors in van der Waals or van der Waals + polarization interaction energies, the results reveal better consistency as the size of the QM region increases for all three systems (see Fig. 6, bottom row). Therefore, these results suggest that the deviations observed in total interaction energies predominantly originate from electrostatic contributions. We observed an error compensation between the electrostatic interaction energies and the van der Waals or vdW + polarization interactions. As illustrated in Fig. 6, one tends to overestimate while the other underestimates. However, this error compensation is not perfect and gives rise to the errors in the total interaction energies discussed previously.

V. CONCLUSIONS

In this work, we evaluated the performance of QM/MM methods to reproduce the relative cooperativity previously calculated with density functional theory and force fields. This study showed that the size of the QM region in QM/MM calculations plays a critical role in the accuracy of interaction energy calculations. In all three systems studied, increasing the size of the QM region generally led to improved results, indicating that a larger QM region enhances the description of the systems under investigation. The choice of a force field within the MM region significantly influenced the accuracy of interaction energy predictions. Notably, the performance of the AMOEBA force field varied from system to system. Conversely, the non-polarizable force fields consistently provided improved results when used in combination in a QM/MM approach, particularly for larger clusters. However, the results also indicate that even when the size of the QM region is increased, improved accuracy for cooperativity is not guaranteed compared with force field-only results. This is because, while the description of the molecules in the QM region improves, the errors in the interaction energy between the QM and MM regions also increase simultaneously. The results of this study have important implications for molecular simulations involving solvation systems and related chemical processes. They emphasize the need for rigorous validation and optimization of QM/MM calculation parameters, including the QM region size and the choice of a force field, to achieve accurate predictions of interaction energies. These insights could guide researchers in making informed decisions to improve the accuracy of simulations and pave the way for more reliable predictions of chemical and physical properties in complex molecular environments.

SUPPLEMENTARY MATERIAL

The supplementary material is available online in a PDF format, including interaction energy errors, linear regressions, and K^+ and F^- ion charges.

ACKNOWLEDGMENTS

The authors acknowledge the University of North Texas CAS-CaM CRUNTCh3 high-performance cluster supported by NSF, Grant Nos. CHE-1531468 and OAC-2117247, for the computational time provided. This work was supported by NIH under Grant No. R01GM108583.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jorge Nochebuena: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal). **Shubin Liu:** Conceptualization (equal); Supervision (equal); Writing – original draft (equal). **G. Andrés Cisneros:** Conceptualization (equal); Funding acquisition (equal); Resources (equal); Supervision (equal); Writing – original draft (equal).

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 Comput. Mol. Sci.* **11**, e1515 (2021).
- 2 K.-S. Csizi and M. Reiher, “Universal QM/MM approaches for general nanoscale applications,” *WIREs Comput. Mol. Sci.* **13**, e1656 (2023).
- 3 J.-N. Wang, W. Liu, P. Li, Y. Mo, W. Hu, J. Zheng, X. Pan, Y. Shao, and Y. Mei, “Accelerated computation of free energy profile at *ab initio* quantum mechanical/molecular mechanics accuracy via a semiempirical reference potential. 4. Adaptive QM/MM,” *J. Chem. Theory Comput.* **17**, 1318–1325 (2021).
- 4 D. Demapan, J. Kussmann, C. Ochsenfeld, and Q. Cui, “Factors that determine the variation of equilibrium and kinetic properties of QM/MM enzyme simulations: QM region, conformation, and boundary condition,” *J. Chem. Theory Comput.* **18**, 2530–2542 (2022).
- 5 S. Das, K. Nam, and D. T. Major, “Rapid convergence of energy and free energy profiles with quantum mechanical size in quantum mechanical–molecular mechanical simulations of proton transfer in DNA,” *J. Chem. Theory Comput.* **14**, 1695–1705 (2018).
- 6 G. Jindal and A. Warshel, “Exploring the dependence of QM/MM calculations of enzyme catalysis on the size of the QM region,” *J. Phys. Chem. B* **120**, 9913–9921 (2016).
- 7 R. Mehmood and H. J. Kulik, “Both configuration and QM region size matter: Zinc stability in QM/MM models of DNA methyltransferase,” *J. Chem. Theory Comput.* **16**, 3121–3134 (2020).
- 8 J. Chen, J. Kato, J. B. Harper, Y. Shao, and J. Ho, “On the accuracy of QM/MM models: A systematic study of intramolecular proton transfer reactions of amino acids in water,” *J. Phys. Chem. B* **125**, 9304–9316 (2021).
- 9 H. J. Kulik, “Large-scale QM/MM free energy simulations of enzyme catalysis reveal the influence of charge transfer,” *Phys. Chem. Chem. Phys.* **20**, 20650–20660 (2018).
- 10 Á. Pérez-Barcia, G. Cárdenas, J. J. Nogueira, and M. Mandado, “Effect of the QM size, basis set, and polarization on QM/MM interaction energy decomposition analysis,” *J. Chem. Inf. Model.* **63**, 882–897 (2023).

- ¹¹M. A. Hix, E. M. Leddin, and G. A. Cisneros, "Combining evolutionary conservation and quantum topological analyses to determine quantum mechanics sub-systems for biomolecular quantum mechanics/molecular mechanics simulations," *J. Chem. Theory Comput.* **17**, 4524–4537 (2021).
- ¹²F. Brandt and C. R. Jacob, "Systematic QM region construction in QM/MM calculations based on uncertainty quantification," *J. Chem. Theory Comput.* **18**, 2584–2596 (2022).
- ¹³F. Brandt and C. R. Jacob, "Protein network centralities as descriptor for QM region construction in QM/MM simulations of enzymes," *Phys. Chem. Chem. Phys.* **25**, 20183–20188 (2023).
- ¹⁴H. W. Qi, M. Karelina, M. Karelina, and H. J. Kulik, "Quantifying electronic effects in QM and QM/MM biomolecular modeling with the Fukui function," *Acta Phys.-Chim. Sin.* **34**, 81 (2018).
- ¹⁵M. Karelina and H. J. Kulik, "Systematic quantum mechanical region determination in qm/mm simulation," *J. Chem. Theory Comput.* **13**, 563–576 (2017).
- ¹⁶H. J. Kulik, J. Zhang, J. P. Klinman, and T. J. Martinez, "How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase," *J. Phys. Chem. B* **120**, 11381–11394 (2016).
- ¹⁷R.-Z. Liao and W. Thiel, "Convergence in the QM-only and QM/MM modeling of enzymatic reactions: A case study for acetylene hydratase," *J. Comput. Chem.* **34**, 2389–2397 (2013).
- ¹⁸D. Flaig, M. Beer, and C. Ochsenfeld, "Convergence of electronic structure with the size of the QM region: Example of QM/MM NMR shieldings," *J. Chem. Theory Comput.* **8**, 2260–2271 (2012).
- ¹⁹C. E. Tzeliou, M. A. Mermigki, and D. Tzeli, "Review on the QM/MM methodologies and their application to metalloproteins," *Molecules* **27**, 2660 (2022).
- ²⁰M. W. van der Kamp and A. J. Mulholland, "Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology," *Biochemistry* **52**, 2708–2728 (2013).
- ²¹T. S. Hofer, H. T. Tran, C. F. Schwenk, and B. M. Rode, "Characterization of dynamics and reactivities of solvated ions by *ab initio* simulations," *J. Comput. Chem.* **25**, 211–217 (2004).
- ²²A. Tongraar and B. M. Rode, "Ab initio QM/MM dynamics of anion–water hydrogen bonds in aqueous solution," *Chem. Phys. Lett.* **403**, 314–319 (2005).
- ²³A. Tongraar, K. R. Liedl, and B. M. Rode, "The hydration shell structure of Li⁺ investigated by Born–Oppenheimer *ab initio* QM/MM dynamics," *Chem. Phys. Lett.* **286**, 56–64 (1998).
- ²⁴T. Haino, "Cooperativity in molecular recognition of feet-to-feet-connected biscavitands," *Pure Appl. Chem.* **95**, 343–352 (2023).
- ²⁵E. Gatto, C. Toniolo, and M. Venanzi, "Peptide self-assembled nanostructures: From models to therapeutic peptides," *Nanomaterials* **12**, 466–521 (2022).
- ²⁶S. Parida, S. K. Patra, and S. Mishra, "Self-assembling behaviour of perylene, perylene diimide, and thionated perylene diimide deciphered through non-covalent interactions," *ChemPhysChem* **23**, e202200361 (2022).
- ²⁷K. Rudharachari Maiyavaganan, M. Prakash, and M. Kumar Ravva, "Simultaneous interaction of graphene nanoflakes with cations and anions: A cooperativity study," *Comput. Theor. Chem.* **1209**, 113601 (2022).
- ²⁸S. Liu, "Principle of chirality hierarchy in three-blade propeller systems," *J. Phys. Chem. Lett.* **12**, 8720–8725 (2021).
- ²⁹S. Liu, "Homochirality originates from the handedness of helices," *J. Phys. Chem. Lett.* **11**, 8690–8696 (2020).
- ³⁰X. Yang and S. Liu, "Cationic cyclophanes-in-cucurbit[10]uril: Host-in-host complexes showing cooperative recognition towards neutral phenol guests," *Supramol. Chem.* **33**, 693–700 (2021).
- ³¹X. Hu, M. Liao, H. Gong, L. Zhang, H. Cox, T. A. Waigh, and J. R. Lu, "Recent advances in short peptide self-assembly: From rational design to novel applications," *Curr. Opin. Colloid Interface Sci.* **45**, 1–13 (2020).
- ³²A. S. Mahadevi and G. N. Sastry, "Cooperativity in noncovalent interactions," *Chem. Rev.* **116**, 2775–2825 (2016).
- ³³C. Rong, D. Zhao, D. Yu, and S. Liu, "Quantification and origin of cooperativity: Insights from density functional reactivity theory," *Phys. Chem. Chem. Phys.* **20**, 17990–17998 (2018).
- ³⁴T. Zhou, S. Liu, D. Yu, D. Zhao, C. Rong, and S. Liu, "On the negative cooperativity of argon clusters containing one lithium cation or fluorine anion," *Chem. Phys. Lett.* **716**, 192–198 (2019).
- ³⁵C. Rong, D. Zhao, T. Zhou, S. Liu, D. Yu, and S. Liu, "Homogeneous molecular systems are positively cooperative, but charged molecular systems are negatively cooperative," *J. Phys. Chem. Lett.* **10**, 1716–1721 (2019).
- ³⁶S. Liu and C. Rong, "Quantifying frustrations for molecular complexes with noncovalent interactions," *J. Phys. Chem. A* **125**, 4910–4917 (2021).
- ³⁷J. Nochebuena, J. P. Piquemal, S. Liu, and G. A. Cisneros, "Cooperativity and frustration effects (or lack thereof) in polarizable and non-polarizable force fields," *J. Chem. Theory Comput.* **19**, 7715 (2023).
- ³⁸M. 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.
- ³⁹J.-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).
- ⁴⁰B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson, "Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg," *Theor. Chem. Acc.* **128**, 69–82 (2011).
- ⁴¹T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," *J. Chem. Phys.* **90**, 1007–1023 (1989).
- ⁴²R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," *J. Chem. Phys.* **96**, 6796–6806 (1992).
- ⁴³Y. Zhao and D. G. Truhlar, "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals," *Theor. Chem. Acc.* **120**, 215–241 (2008).
- ⁴⁴T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. V. R. Schleyer, "Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F," *J. Comput. Chem.* **4**, 294–301 (1983).
- ⁴⁵R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, "Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions," *J. Chem. Phys.* **72**, 650–654 (1980).
- ⁴⁶Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder, and P. Ren, "Polarizable atomic multipole-based AMOEBA force field for proteins," *J. Chem. Theory Comput.* **9**, 4046–4063 (2013).
- ⁴⁷C. 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," *J. Chem. Theory Comput.* **14**, 2084–2108 (2018).
- ⁴⁸A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *J. Chem. Phys.* **98**, 5648–5652 (1993).
- ⁴⁹C. 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**, 785–789 (1988).
- ⁵⁰P. Ren, C. Wu, and J. W. Ponder, "Polarizable atomic multipole-based molecular mechanics for organic molecules," *J. Chem. Theory Comput.* **7**, 3143–3161 (2011).
- ⁵¹J. W. Ponder and D. A. Case, "Force fields for protein simulations," in *Protein Simulations, Advances in Protein Chemistry* (Academic Press, 2003), Vol. 66, pp. 27–85.

- ⁵²P. Ren and J. W. Ponder, "Polarizable atomic multipole water model for molecular mechanics simulation," *J. Phys. Chem. B* **107**, 5933–5947 (2003).
- ⁵³P. Ren and J. W. Ponder, "Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations," *J. Comput. Chem.* **23**, 1497–1506 (2002).
- ⁵⁴G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, "Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides," *J. Phys. Chem. B* **105**, 6474–6487 (2001).
- ⁵⁵J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, "ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB," *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
- ⁵⁶W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.* **79**, 926–935 (1983).
- ⁵⁷L. X. Dang and B. M. Pettitt, "Simple intramolecular model potentials for water," *J. Phys. Chem.* **91**, 3349–3354 (1987).
- ⁵⁸E. G. Kratz, A. R. Walker, L. Lagardère, F. Lipparini, J.-P. Piquemal, and G. Andrés Cisneros, "LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields," *J. Comput. Chem.* **37**, 1019–1029 (2016).
- ⁵⁹H. Gökcan, E. A. Vázquez-Montelongo, and G. A. Cisneros, "LICHEM 1.1: Recent improvements and new capabilities," *J. Chem. Theory Comput.* **15**, 3056–3065 (2019).
- ⁶⁰J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, "Psi4: An open-source *ab initio* electronic structure program," *WIREs Comput. Mol. Sci.* **2**, 556–565 (2012).
- ⁶¹W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," *J. Mol. Graphics* **14**, 33–38 (1996).
- ⁶²J. Stone, "An efficient library for parallel ray tracing and animation," M.S. thesis, Computer Science Department, University of Missouri-Rolla, 1998.
- ⁶³G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, J. Lu, Y. Xu, H. Torabifard, A. P. Bartók, G. Csányi, V. Molinero, and F. Paesani, "Modeling molecular interactions in water: From pairwise to many-body potential energy functions," *Chem. Rev.* **116**, 7501–7528 (2016).
- ⁶⁴J. Cui, H. Liu, and K. D. Jordan, "Theoretical characterization of the (H₂O)₂₁ cluster: Application of an *n*-body decomposition procedure," *J. Phys. Chem. B* **110**, 18872–18878 (2006).
- ⁶⁵M. P. Hodges, A. J. Stone, and S. S. Xantheas, "Contribution of many-body terms to the energy for small water clusters: A comparison of *ab initio* calculations and accurate model potentials," *J. Phys. Chem. A* **101**, 9163–9168 (1997).
- ⁶⁶V. Vennelakanti, A. Nazemi, R. Mehmood, A. H. Steeves, and H. J. Kulik, "Harder, better, faster, stronger: Large-scale QM and QM/MM for predictive modeling in enzymes and proteins," *Curr. Opin. Struct. Biol.* **72**, 9–17 (2022).
- ⁶⁷L. Hu, P. Söderhjelm, and U. Ryde, "On the convergence of QM/MM energies," *J. Chem. Theory Comput.* **7**, 761–777 (2011).
- ⁶⁸I. Solt, P. Kulhánek, I. Simon, S. Winfield, M. C. Payne, G. Csányi, and M. Fuxreiter, "Evaluating boundary dependent errors in QM/MM simulations," *J. Phys. Chem. B* **113**, 5728–5735 (2009).
- ⁶⁹C. V. Sumowski and C. Ochsenfeld, "A convergence study of QM/MM isomerization energies with the selected size of the QM region for peptidic systems," *J. Phys. Chem. A* **113**, 11734–11741 (2009).
- ⁷⁰Y. Zhang and H. Lin, "Flexible-boundary quantum-mechanical/molecular-mechanical calculations: Partial charge transfer between the quantum-mechanical and molecular-mechanical subsystems," *J. Chem. Theory Comput.* **4**, 414–425 (2008).
- ⁷¹Y. Zhang and H. Lin, "Flexible-boundary QM/MM calculations: II. Partial charge transfer across the QM/MM boundary that passes through a covalent bond," *Theor. Chem. Acc.* **126**, 315–322 (2010).
- ⁷²S. Pezeshki and H. Lin, "Molecular dynamics simulations of ion solvation by flexible-boundary QM/MM: On-the-fly partial charge transfer between QM and MM subsystems," *J. Comput. Chem.* **35**, 1778–1788 (2014).
- ⁷³S. Pezeshki and H. Lin, "Recent developments in QM/MM methods towards open-boundary multi-scale simulations," *Mol. Simul.* **41**, 168–189 (2015).