AIP | The Journal of Chemical Physics

# Charge-dependent many-body exchange and dispersion interactions in combined QM/MM simulations

Erich R. Kuechler, Timothy J. Giese, and Darrin M. York

Citation: The Journal of Chemical Physics **143**, 234111 (2015); doi: 10.1063/1.4937166
View online: http://dx.doi.org/10.1063/1.4937166
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/143/23?ver=pdfcov
Published by the AIP Publishing



# Charge-dependent many-body exchange and dispersion interactions in combined QM/MM simulations

Erich R. Kuechler,[1,2] Timothy J. Giese,[1] and Darrin M. York[1]
[1]*BioMaPS Institute and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854-8087, USA*
[2]*Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431, USA*

(Received 4 August 2015; accepted 22 November 2015; published online 21 December 2015)

Accurate modeling of the molecular environment is critical in condensed phase simulations of chemical reactions. Conventional quantum mechanical/molecular mechanical (QM/MM) simulations traditionally model non-electrostatic non-bonded interactions through an empirical Lennard-Jones (LJ) potential which, in violation of intuitive chemical principles, is bereft of any explicit coupling to an atom's local electronic structure. This oversight results in a model whereby short-ranged exchange-repulsion and long-ranged dispersion interactions are invariant to changes in the local atomic charge, leading to accuracy limitations for chemical reactions where significant atomic charge transfer can occur along the reaction coordinate. The present work presents a variational, charge-dependent exchange-repulsion and dispersion model, referred to as the charge-dependent exchange and dispersion (QXD) model, for hybrid QM/MM simulations. Analytic expressions for the energy and gradients are provided, as well as a description of the integration of the model into existing QM/MM frameworks, allowing QXD to replace traditional LJ interactions in simulations of reactive condensed phase systems. After initial validation against QM data, the method is demonstrated by capturing the solvation free energies of a series of small, chlorine-containing compounds that have varying charge on the chlorine atom. The model is further tested on the $S_N2$ attack of a chloride anion on methylchloride. Results suggest that the QXD model, unlike the traditional LJ model, is able to simultaneously obtain accurate solvation free energies for a range of compounds while at the same time closely reproducing the experimental reaction free energy barrier. The QXD interaction model allows explicit coupling of atomic charge with many-body exchange and dispersion interactions that are related to atomic size and provides a more accurate and robust representation of non-electrostatic non-bonded QM/MM interactions. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4937166]

## I. INTRODUCTION

Combined quantum mechanical/molecular mechanical (QM/MM) simulations are powerful tools in the study of chemical reactions taking place while embedded within large biological scaffolds and/or in condensed phase environments. These hybrid methods are typically orders of magnitudes more efficient than fully quantum mechanical approaches, although recent advances in the development of reactive quantum mechanical force fields have greatly narrowed this gap.[1–4] QM/MM methods have been successfully applied in the study of enzymes,[5–7] catalytic RNAs,[8–10] ligand binding,[11–13] acid dissociation constants,[14–16] and small molecule reactions occurring in solution.[17–20]

Combined QM/MM approaches leverage the strengths associated with of both QM methods and MM simulations by allowing a typically small, localized reactive region requiring explicit treatment of electronic degrees of freedom to be modeled within a large, complex condensed phase environment. The overall reliability of QM/MM simulations depends critically on both the accuracy of the chosen QM and MM models as well as the treatment of the QM/MM interaction. Considerations such as the size of the reactive region and the specific chemistry occurring within it must be considered when choosing the QM method. The accuracy of MM force field model depends simultaneously on its functional form and the empirical parameters used to describe the underlying potential. The MM model is typically chosen as to best represent key features of the environment surrounding the QM region. The most advanced force field treatments include accurate descriptions of electrostatic interactions and have explicit consideration of many-body polarization. While such potentials are very promising, these force fields are typically less computationally efficient than their traditional non-polarizable counterparts and, in many cases, have parameters which either have a somewhat limited chemical scope or have not been as fully matured through critical assessment to the level of those employed by the more simple force fields. Because of these reasons, the vast majority of QM/MM simulations currently utilize conventional, point-charge force fields in the study of condensed phase reactions.

The QM and MM regions interact with each other through a QM/MM boundary. Most generally, both bonded and non-bonded interactions are treated over this division. Bonded interactions are required when the QM/MM boundary falls between two or more covalently bound atoms. A

 

variety of specialized methods; including link atoms,[21,22] pseudoatoms,[23,24] and generalized orbital methods;[25,26] have been developed to model this interaction. Non-bonded interactions can be further subdivided into electrostatic and non-electrostatic terms.

The QM/MM non-bonded electrostatic interactions can be treated using different embedding techniques that can be broadly separated into three categories: mechanical, electrostatic, or polarized embedding.[27] Electrostatic embedding enters the electrostatic potential due to fixed, static MM point charges directly into the QM Hamiltonian as an external potential to polarize the QM density that results from a self-consistent field (SCF) procedure. This scheme generally provides greater accuracy than mechanical embedding,[28] where static point charges are assigned to QM atoms and electrostatics are treated classically as Coulomb interactions between point charges. If the MM force field is itself an explicitly polarizable model then a polarized embedding[29,30] scheme can be adopted, where MM polarizable charges/multipoles are allowed to respond to the QM charge density in some way.

The QM/MM non-electrostatic non-bonded interactions include short-ranged exchange repulsion and mid/long-ranged dispersion interactions. Conventional QM/MM methods handle this term using the Lennard-Jones (LJ) model,[31–33] which is also the most commonly employed non-electrostatic non-bonded model used by MM force fields. However, unlike the electrostatic QM/MM interactions in electrostatic and polarized embedding methods, the LJ model for non-electrostatic non-bonded QM/MM interactions is not explicitly coupled to the quantum mechanical wave function, meaning these interactions cannot adjust to changes in the chemical environment. In order to account for this issue, atoms in different chemical environments are assigned different LJ interactions, and much like many MM force fields adopt the concept of an *atom type*—a prescription in which parameters for an atom are inherited based on specific local electronic and chemical bonding environment. Within this framework, each atom type is given a unique set of fixed LJ parameters.[34] Given that traditional MM force fields do not allow for the formation or destruction of chemical bonds this strategy provides a practical mechanism for modeling these interactions, provided that the scope of LJ parameters in the force field is wide enough to model a diverse range of intermolecular interactions. However, in the case of QM/MM simulations, the bonding environment can and in fact often does change. As such, adhering to the notion of assigning fixed atom types to reactive QM atoms is impractical and unsound. In careful QM/MM studies, LJ parameters for QM atoms are developed to produce an overall reasonable, balanced reaction profile that strives to eliminate artifacts due to having atoms whose assigned atom types might misrepresent changes in the chemical identity of those atoms over the course of the modeled chemical reaction. However, in doing so the accuracy of the solvation properties of individual states along the reaction pathway is sacrificed. Frequently, it is not possible to select a single LJ parameterization which can simultaneously reproduce quantitative reaction barrier data while obtaining accurate solvation free energies for the reactant and product states.

One partial solution to this issue is to use so-called adaptive QM/MM methods,[35–37] where the QM and MM representations of molecules turn on/off as they enter/leave a larger "reactive zone." In principle QM models account for changes in exchange repulsion and dispersion interactions as a function of the electronic structure. By using these methods and by pushing the QM/MM boundary further away from the reaction zone, the problems associated with static LJ potentials could be greatly reduced. While this approach is conceptually appealing and has been demonstrated to be successful if carefully applied, it is also more computationally demanding than traditional methods and will only have clear benefits if the QM model used is superior to the MM force field at representing solvation properties. For instance, the application of approximate semiempirical, Hartree-Fock, or small basis set density-functional QM models have been shown to often be inferior to a carefully parameterized MM water model for describing the physical properties of liquid water. Alternatively, one could replace the conventional, static LJ QM/MM interaction with one which is explicitly coupled to the QM electronic structure. Recent studies have pioneered this type of approach by directly linking the extent of charge delocalization to the electronic structure of the system, effectively allowing the charge of more negative species to appear more diffuse.[38,39]

This work introduces a new model that explicitly couples atomic charge with many-body exchange and dispersion interactions. The QXD (Charge-dependent eXchange and Dispersion) model allows the seamless modulation of non-electrostatic, non-bonded interactions in a chemically intuitive and meaningful way with changes in local atomic charge. As atoms become more negatively charged, they become larger and softer with regard to their interactions with surrounding MM atoms. In this way, the model is able to better represent the QM/MM interaction non-locally along the reaction coordinate.

In the present work, we examine a well-studied benchmark reaction of the chloride ion attack on methylchloride using a specific reaction parameterization semiempirical QM model[40] in a solution of non-polarizable MM water while treating electrostatic QM/MM interactions via electrostatic embedding. This combination of QM and MM models is widely used in combined QM/MM studies of biological reactions, a target application area for which the hope is that the methods introduced here will have future impact. The remainder of the manuscript will be broken into four parts: (1) Sec. II in which the QXD model is developed, (2) Sec. III which outlines details as to how the QXD model was parameterized and integrated into a QM/MM framework, and (3) Sec. IV which presents, compares, and discusses simulation results for solvation free energy calculations and reaction free energy profiles involving chloride ion attack to methylchloride, and (4) a section which summarizes the main conclusions of the work and places its importance into a broader context.

## II. BACKGROUND THEORY

The QXD model replaces the traditional LJ treatment of non-classical interactions in the QM/MM Hamiltonian for

QM atoms. It allows QM atoms to effectively "change size" in accordance to fluctuations in local charge, as determined by the single-particle density matrix. QXD does this task through a series of equations which closely follow previous work of Giese and York.[41] However, unlike the initial work where the energy was added as a post-SCF correction term, present here is a variational formulation that is integrated with the SCF procedure which provides simple expressions for the gradients needed to obtain forces when used in molecular dynamics calculations.

### A. QM/MM simulations

During QM/MM simulations the total potential energy, $E$, can be written as

$$E = E_{\mathrm{QM}}(\mathbf{R},\mathbf{P}) + E_{\mathrm{MM}}(\mathbf{R},\mathbf{q}) + E_{\mathrm{QM/MM}}(\mathbf{R},\mathbf{P},\mathbf{q}), \quad (1)$$

where $\mathbf{R}$ is an array of atomic positions, $\mathbf{q}$ is a set of MM charges, and $\mathbf{P}$ is the atomic orbital (AO) representation of the QM total density matrix. This term is defined as

$$\mathbf{P} = \mathbf{P}^{\alpha} + \mathbf{P}^{\beta}, \quad (2)$$

where $\mathbf{P}^{\sigma}$ is the spin-resolved density matrix,

$$P^{\sigma}_{\mu\nu} = \sum_{\sigma,i} n^{\sigma}_{i} C^{\sigma}_{\mu i} C^{\sigma}_{\nu i}, \quad (3)$$

and $\mathbf{n}^{\sigma}$ and $\mathbf{C}^{\sigma}$ are the spin-resolved orbital occupation numbers and molecular orbital (MO) coefficients, respectively, and where $\mu$ and $\nu$ index atomic orbital basis functions.

$E_{\mathrm{QM}}(\mathbf{R},\mathbf{P})$ is the QM region's *ab initio* self-energy, $E_{\mathrm{MM}}(\mathbf{R},\mathbf{q})$ is the MM energy, and $E_{\mathrm{QM/MM}}(\mathbf{R},\mathbf{P},\mathbf{q})$ is the interaction energy of the QM region with the MM region. For a given set of atomic positions and MM charges, the total energy is minimized with respect to changes in the QM region's MO coefficients under the constraint that the MOs remain orthonormal. When the energy is extremized under these constraints, the following stationary condition (the Roothaan-Hall equation) is satisfied:

$$\mathbf{F}^{\sigma} \cdot \mathbf{C}^{\sigma} = \mathbf{S} \cdot \mathbf{C}^{\sigma} \cdot \mathbf{E}^{\sigma}, \quad (4)$$

where $S_{\mu\nu} = \int \chi_{\mu}(\mathbf{r})\chi_{\nu}(\mathbf{r}) d^3r$ is the AO overlap matrix, and $F^{\sigma}_{\mu\nu}$ is the spin-resolved Fock (or Kohn-Sham) matrix,

$$\begin{aligned} F^{\sigma}_{\mu\nu} &= \left.\frac{\partial E_{\mathrm{QM}}}{\partial P^{\sigma}_{\mu\nu}}\right|_{\mathbf{R},\mathbf{q}} + \left.\frac{\partial E_{\mathrm{QM/MM}}}{\partial P^{\sigma}_{\mu\nu}}\right|_{\mathbf{R},\mathbf{q}} \\ &= F^{\sigma,0}_{\mu\nu} + \left.\frac{\partial E_{\mathrm{QM/MM}}}{\partial P^{\sigma}_{\mu\nu}}\right|_{\mathbf{R},\mathbf{q}}, \end{aligned} \quad (5)$$

where $F^{\sigma,0}_{\mu\nu}$ is QM method's Fock matrix in the absence of the MM atoms, and $\partial E_{\mathrm{QM/MM}}/\partial P^{\sigma}_{\mu\nu}|_{\mathbf{R},\mathbf{q}}$ is an effective external electronic chemical potential that causes the QM density to respond to the MM environment. The self-consistent field procedure is as follows: 1. Given a set of trial MOs, compute the density matrix. 2. From the density matrix, compute the energy and Fock matrix. 3. If the change in the energy is small, and Eq. (4) is satisfied, then exit. 4. Construct a new set of trial MOs and go to step 1. Typically, a new guess is made by solving Eq. (4) for $\mathbf{C}^{\sigma}$.

### B. LJ QM/MM interaction potential

In conventional QM/MM implementations, $E_{\mathrm{QM/MM}}$ is expressed in terms of non-bonded electrostatic interactions and LJ potentials,

$$\begin{aligned} E_{\mathrm{QM/MM}}(\mathbf{R},\mathbf{P},\mathbf{q}) = E^{\mathrm{el}}_{\mathrm{QM/MM}}(\mathbf{R},\mathbf{P},\mathbf{q}) + E^{\mathrm{bond}}_{\mathrm{QM/MM}}(\mathbf{R},\mathbf{P}) \\ + E^{\mathrm{nb}}_{\mathrm{QM/MM}}(\mathbf{R}), \end{aligned} \quad (6)$$

where $E^{\mathrm{el}}_{\mathrm{QM/MM}}$ is the electrostatic interaction, $E^{\mathrm{bond}}_{\mathrm{QM/MM}}$ is the bonded interaction, and $E^{\mathrm{nb}}_{\mathrm{QM/MM}}$ is the non-bonded interaction term. As discussed in the introduction, $E^{\mathrm{bond}}_{\mathrm{QM/MM}}$ can be entirely avoided if the QM/MM boundary does not cross any covalent bonds. The present work will be examining systems which are entirely subsumed by the QM region and no chemical bonds intersect the boundary, so this term will be ignored to avoid unnecessary complication. Furthermore, if the LJ interaction is assumed to represent the non-classical QM/MM interactions, the total QM/MM interaction can be expressed as

$$E_{\mathrm{QM/MM}}(\mathbf{R},\mathbf{P},\mathbf{q}) = -\sum_{\substack{\mu\nu \\ c\in\mathrm{MM}}} P_{\mu\nu} q_c (\mu\nu|c) + \sum_{\substack{a\in\mathrm{QM} \\ c\in\mathrm{MM}}} E_{\mathrm{LJ},ac}(R_{ac}), \quad (7)$$

where

$$(\mu\nu|c) \equiv \int \frac{\chi_{\mu}(\mathbf{r})\chi_{\nu}(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_c|} d^3r \quad (8)$$

and $E_{\mathrm{LJ},ac}$ is the LJ potential[31,42,43]

$$E_{\mathrm{LJ},ac}(R_{ac}) = \varepsilon_{ac}\left[\left(\frac{R_{\mathrm{min},ac}}{R_{ac}}\right)^{12} - 2\left(\frac{R_{\mathrm{min},ac}}{R_{ac}}\right)^{6}\right]. \quad (9)$$

The LJ expression can be separated into two distinct, but coupled parts. The first representing exchange-repulsion and the latter models the attractive, dispersion interactions.

### C. QXD QM/MM interaction potential

The QXD model replaces the QM/MM LJ interactions with a charge-dependent functional form. The systems considered in this work do not explicitly cross the QM/MM boundary, so any specialized treatment of this type of interaction will be ignored. The total QXD energy can be written as

$$E = E_{\mathrm{QM}}(\mathbf{R},\mathbf{P}) + E_{\mathrm{MM}}(\mathbf{R},\mathbf{q}) + E_{\mathrm{QXD}}(\mathbf{R},\mathbf{P},\mathbf{q}), \quad (10)$$

where

$$\begin{aligned} E_{\mathrm{QXD}}(\mathbf{R},\mathbf{P},\mathbf{q}) = -\sum_{\substack{\mu\nu \\ c\in\mathrm{MM}}} P_{\mu\nu} q_c (\mu\nu|c) \\ + \sum_{\substack{a\in\mathrm{QM} \\ c\in\mathrm{MM}}} E_{\mathrm{XD},ac}(R_{ac}; Q_a), \end{aligned} \quad (11)$$

where $Q_a \equiv Q_a(\mathbf{P})$ is the Mulliken charge of QM atom $a$

$$Q_a(\mathbf{P}) = Z_a - \sum_{\mu\nu} P_{\mu\nu} S_{\mu\nu} W^{(a)}_{\mu\nu}, \quad (12)$$

$$W_{\mu\nu}^{(a)} = \begin{cases} 1 & \text{if } \mu\nu \in a \\ \frac{1}{2} & \text{if } \mu \in a, \nu \notin a \text{ or } \nu \in a, \mu \notin a \\ 0 & \text{otherwise} \end{cases}, \quad (13)$$

and $Z_a$ is the nuclear charge of atom $a$.

The charge-dependent QXD van der Waals energy consists of exchange-repulsion and dispersion components,

$$E_{\mathrm{XD},ac}(R_{ac}; Q_a) = E_{\mathrm{X},ac}(R_{ac}; Q_a) + E_{\mathrm{D},ac}(R_{ac}; Q_a). \quad (14)$$

$E_{\mathrm{X},ac}$ treats the exchange-repulsion as being proportional to the overlap of atomic densities, which Eq. (18) models as the overlap of Slater functions $S(r;\zeta) = (\zeta^3/8\pi)e^{-\zeta r}$. The approximate relationship between exchange-repulsion and atomic overlap was recognized 40 years ago[44] and then further developed into an anisotropic repulsion model by Wheatley.[45] Recently, similar formalisms have gained popularity in the GEM force field.[46] The model used in the present manuscript is based on previous work by Giese,[41] which replaces the rigorous treatment of the electron density with an empirical model that can be tuned for high-accuracy while providing a means for incorporating charge-dependence. Specifically, the charge dependence is introduced via the Slater exponents,

$$\zeta_a \equiv \zeta_a(Q_a) = \zeta_{0,a} e^{-\zeta_{q,a} Q_a}. \quad (15)$$

$E_{\mathrm{D},ac}$ is a damped $R^{-6}$ dispersion model, whose $C_6$ dispersion coefficient is proportional to the atomic dipole polarizability, $\alpha$, and an "effective number of electrons," $N_{\mathrm{eff}}$. Charge dependence is introduced by making $\alpha$ and $N_{\mathrm{eff}}$ a function of atomic charge,

$$N_{\mathrm{eff},a} \equiv N_{\mathrm{eff},a}(Q_a) = N_{\mathrm{eff},0,a} - Q_a, \quad (16)$$

$$\alpha_a \equiv \alpha_a(Q_a) = \alpha_{0,a} e^{-\alpha_{q,a} Q_a}. \quad (17)$$

The QXD model depends on the empirical parameters: $s_a$, $\zeta_{0,a}$, $\zeta_{q,a}$, $N_{\mathrm{eff},0,a}$, $\alpha_{0,a}$, $\alpha_{q,a}$. The values of $N_{\mathrm{eff},0,a}$, which on atomic number, were developed by Pellenq and Nicholson[47] and are used directly without modification, that is, $N_{\mathrm{eff},0,a}$ is not treated as a parameter to be optimized. The expressions for $E_{\mathrm{X},ac}$ and $E_{\mathrm{D},ac}$ are summarized as follows:

$$E_{\mathrm{X},ac}(R_{ac}; Q_a) = s_a s_c \zeta_{ac} (\Delta_{ac} - \Delta_{ca}), \quad (18)$$

$$E_{\mathrm{D},ac}(R_{ac}; Q_a) = -S_{6,ac} \frac{C_{6,ac}}{R_{ac}^6}, \quad (19)$$

$$\Delta_{ac} = \frac{\zeta_c e^{-\zeta_a R_{ac}}}{R_{ac}} [4\zeta_a + R_{ac}(\zeta_a^2 - \zeta_c^2)], \quad (20)$$

$$\zeta_{ac} = \frac{1}{8\pi} \frac{\zeta_a^3 \zeta_c^3}{(\zeta_a^2 - \zeta_c^2)^3}, \quad (21)$$

$$C_{6,ac} = \frac{3}{2} \frac{\eta_a \eta_c}{\eta_a + \eta_c} \alpha_a \alpha_c, \quad (22)$$

$$\eta_a = \sqrt{\frac{N_{\mathrm{eff},a}}{\alpha_a}}, \quad (23)$$

$$S_{6,ac} = 1 - e^{-b_{ac} R_{ac}} \sum_{k=0}^{6} \frac{(b_{ac} R_{ac})^k}{k!}, \quad (24)$$

$$b_{ac} = (\Delta_{ac} - \Delta_{ca})^{-1} \frac{d}{dR_{ac}} (\Delta_{ca} - \Delta_{ac}). \quad (25)$$

In the limit $\zeta_c \to \zeta_a$, Eq. (18) reduces to

$$E_{\mathrm{X},ac}(R_{ac}; Q_a) = s_a s_c \frac{e^{-\zeta_a R_{ac}} \zeta_a^3}{192\pi} (3 + 3R_{ac}\zeta_a + R_{ac}^2 \zeta_a^2). \quad (26)$$

The Fock matrix correction for the QXD QM/MM model [Eq. (11)] is

$$\left.\frac{\partial E_{QXD}}{\partial P_{\mu\nu}^{\sigma}}\right|_{\mathbf{R},\mathbf{q}} = \sum_a \left.\frac{\partial E_{QXD}}{\partial Q_a}\right|_{\mathbf{R},\mathbf{q}} \left.\frac{\partial Q_a}{\partial P_{\mu\nu}^{\sigma}}\right|_{\mathbf{R},\mathbf{q}} - \sum_{c \in \mathrm{MM}} q_c(\mu\nu|c), \quad (27)$$

where

$$\left.\frac{\partial Q_a}{\partial P_{\mu\nu}^{\sigma}}\right|_{\mathbf{R},\mathbf{q}} = -S_{\mu\nu} W_{\mu\nu}^{(a)} \quad (28)$$

and $\partial E_{\mathrm{QXD}}/\partial Q_a|_{\mathbf{R},\mathbf{q}}$ is obtained from elementary chain-rule derivatives of Eqs. (14)–(25). Modern charge-fitting procedures could be used to more accurately represent local atomic charge, such as ESP fitting[48–52] or CMX[53–57] charges. However, early tests indicated that regardless of which charge model was used the end results were not significantly different, but rather lead to slightly different optimized parameters, as the changes in charge were highly correlated and internally consistent within the semiempirical framework. Similar results have been demonstrated previously for Mulliken and CM2 charge changes calculated with linear-scaling semiempirical methods.[58] These findings, combined with the simplicity of the Mulliken charge mapping into the density matrix and ease of incorporation into the Fock matrix for variational self-consistency, led to use of these charges as a basis for the semiempirical QXD model. It is expected that the parameters developed here are not likely transferable to other quantum methods, and the use of Mulliken charges might not be appropriate for other quantum methods which might use large, delocalized basis sets.

Upon reaching SCF convergence, the atomic gradients of atom $a$ are

$$\left.\frac{\partial E}{\partial X_a}\right|_{\mathbf{P}} = \frac{dE_{\mathrm{MM}}}{dX_a} + \left.\frac{\partial E_{\mathrm{QM}}}{\partial X_a}\right|_{\mathbf{P}} + \left.\frac{\partial E_{\mathrm{QXD}}}{\partial X_a}\right|_{\mathbf{P},\mathbf{q}} - \sum_{\mu\nu} Q_{\mu\nu} \frac{dS_{\mu\nu}}{dX_a}, \quad (29)$$

where

$$Q_{\mu\nu} = \sum_{\sigma} \sum_{k} n_k^{\sigma} E_{kk}^{\sigma} C_{\mu k}^{\sigma} C_{\nu k}^{\sigma}. \quad (30)$$

## III. COMPUTATIONAL METHODS

This section describes computational details involving the implementation, parameterization, and validation of the QXD model for interaction energies, molecular dynamics simulations, and free energy calculations. First, we outline how the QXD model can interact with traditional force fields. Second, a protocol for comparing LJ and QXD QM/MM interactions to high-level quantum data using gas-phase scans is provided. Third, MD simulation protocols to calculate solvation free energies used in the parameterization of QXD are shown. Finally, the MD protocol for free energy profile

simulations used to validate the QXD QM/MM interaction model is described.

### A. The QXD interface with modern force fields

The first step in implementing the QXD correction is to interface it with existing QM/MM frameworks. The QXD model would be inconvenient to use if special parameters would have to be redeveloped for all MM atom types. Fortunately, we have found that the LJ parameters used in current MM force fields can be suitably "mapped" into QXD parameters through a series of approximate expressions. These expressions are purely empirical, and details, including their motivation, are described in the supplementary material, Section S1.[97]

Therefore, the QXD parameters of the MM atoms are obtained directly from their LJ parameters using the prescription described below, whereas the QXD parameters of the QM atoms are parameterized for the specific QM Hamiltonian used. The description of the QM atom QXD parameters is discussed in Sec. IV C.

The prescription for mapping the MM atom LJ parameters to QXD parameters begins by removing their charge-dependence within the QXD model $\zeta_{q,c} = 0$ and $\alpha_{q,c} = 0$. The approximate relations between the remaining adjustable parameters in the QXD model to the $\varepsilon_c \equiv \sqrt{\varepsilon_{cc}}$ and $R_{\min,c} \equiv (1/2)R_{\min,cc}$ LJ parameters are

$$s_c = A\varepsilon_c^B R_{\min,c}^C, \tag{31}$$

$$\zeta_{0,c} = D\varepsilon_c^E R_{\min,c}^F, \tag{32}$$

$$\alpha_{0,c} = G_c\left(\varepsilon_c R_{\min,c}^6\right)^{2/3}. \tag{33}$$

Parameters $A$-$F$ neither depend on atom-type nor atomic number and can be used universally (the numerical values are listed in atomic units): $A = 9.4423$, $B = 0.4111$, $C = 2.8208$, $D = 3.7893$, $E = -0.0192$, and $F = -0.7249$. The $G_c$ parameter depends on atomic number, as to keep the mapping expressions consistent with the Pellenq and Nicholson values of $N_{\text{eff},0,c}$ used within the dispersion potential.[47] Values for $G_c$ have been parameterized for elements H, C, N, O, and Cl by fitting the QXD van der Waals interaction energies to reproduce the LJ energy for a large number of atom-type pairs. The numerical values of these parameters are (listed in atomic units): $G_H = 32.7324$, $G_C = 19.0422$, $G_N = 18.2066$, $G_O = 17.5986$, and $G_{Cl} = 15.1066$. A detailed comparison of the mapped QXD interactions with respect to the LJ potential is presented in Sec. IV A.

### B. Specific reaction parameterization Hamiltonian

Solvation free energies of chlorine containing compounds and a prototype reaction of a chloride anion attacking methylchloride will be examined to validate the QXD QM/MM interaction model. This reaction has been extensively studied[59–63] and is generally well understood, lending credence as a system for benchmarking new computational models. For all QM/MM calculations reported here, we employ the recently developed Specific Reaction Parameterization (SRP) AM1 Hamiltonian for $Cl^-$ attack on $CH_3Cl$ that has been demonstrated to accurately model this reaction in both the gas phase and in solution.[40] All non-QXD atoms in QM/MM simulations and calculations performed will use the LJ parameters provided in Table I.

### C. QXD validation against high-level quantum calculations

Gas-phase adiabatic scans of a water probe of various molecules of interest were carried out via constrained optimization to demonstrate the viability and flexibility of the QXD interaction model, while also achieving a base line departure point for solution phase parameters. In these scans the quantum-level water is constrained to the TIP4P-Ew[66] geometry. The electronic degrees of freedom and the geometry of the molecule of interest were allowed to relax. The water probe is oriented with the hydrogen atoms facing the particle of interest the and the axis of the scan bifurcating the hydrogen-oxygen-hydrogen angle. High-level calculations were performed with the Gaussian09 software suite[67] in conjunction with the GaussView tool.[68] using the M06-2X functional[69] at the mixed 6-31+g(d,p)//6-311++g(3df,2p) level using an ultrafine integration grid for both the geometry and energy evaluations, with tight SCF and optimization criteria. Gas phase QXD scans were performed in the AMBER12 software package[70] with the water probe represented by an MM TIP4P-Ew model and an infinite non-bonded cutoff was employed. QXD parameters were optimized via steepest descent optimization in conjunction with a direction set minimization routine, utilizing a Boltzmann weighted chi-squared protocol with additional weighting for reproducing the zero and minimum of the interaction energy.

### D. QXD parameterization and thermodynamic integration

Solution phase Thermodynamic Integration (TI) simulations were performed with the AMBER12 software package[70] in TIP4P-Ew[66] to parameterize the QXD model for condensed phase simulations. All simulations use periodic boundary

TABLE I. Lennard-Jones Parameters. Parameters used to determine the non-electrostatic, non-bonded interactions in conventional QM/MM and other molecular dynamics simulations. The $LJ_{Cl^-}$ chlorine parameters are taken from the work of Joung and Cheatham[64] and are specifically tailored to be used with a TIP4P-Ew solvent, while the $LJ_{CH_3Cl}$ were independently parameterized to capture the solvation free energy of methylchloride. Other parameters are from the ff99 AMBER force field.[65] $R_{\min,c}$ and $\varepsilon_c$, values are in Å and kcal/mol, respectively.

| | Element | $R_{\min,c}$ | $\varepsilon_c$ |
|---|---|---|---|
| Solvent | O | 1.775 931 | 0.162 75 |
| | H | 0.000 0 | 0.000 0 |
| Solute | C | 1.908 0 | 0.086 0 |
| | H | 1.487 0 | 0.015 7 |
| | $LJ_{Cl^-}$: Cl | 2.760 0 | 0.011 661 5 |
| | $LJ_{CH_3Cl}$: Cl | 1.448 9 | 0.457 866 2 |

conditions (PBCs) with a particle mesh Ewald treatment for long-range electrostatics, the QM/MM long-range interaction switch fix, tight density convergence, and a SCF convergence criteria set to $1.0 \times 10^{-10}$. A tighter density and SCF convergence criteria was desirable due to the importance of accurately modeling the electronic density as it now more directly impacted QXD interactions with the surrounding MM environment.

Parameters are refined through the use of dual-topology, single coordinate TI[71–74] simulations to accurately calculate solvation free energies using the thermodynamic cycle shown in Figure 1. TI mixes Hamiltonians of two different potential states ($V_0$ and $V_1$) to calculate the free energy between those states by moving the system along an artificial, alchemical pathway,

$$\Delta G = G(\lambda = 1) - G(\lambda = 0) = \int_0^1 \langle \partial V / \partial \lambda \rangle_\lambda d\lambda, \quad (34)$$

where

$$V(\lambda) = (1 - \lambda)^k V_0 + [1 - (1 - \lambda)^k] V_1. \quad (35)$$

Simulation protocol was tested for accuracy, examining box size and long-range electrostatic effects. The protocol was adjusted accordingly to eliminate any systematic errors. Following the thermodynamic cycle in Fig. 1, contributions from each leg can be analyzed separately for their contributions to the total solvation energy. Of particular note; the gas phase legs of the cycle involving adjusting non-bonded terms can be ignored as, by design, the contribution to the overall solvation free energy is zero. The gas phase MM to QM leg, designated as $\Delta G_6$, is required to correct for differences in arbitrary zeros of the QM and MM representations. Parameterization of QXD is hastened through intelligent division of the cycle and by minimizing the free energy differences experienced in each individual step. To this end, trial LJ parameters were created for each species to approximate their solvation free energy accurately, allowing for the $\Delta G_4$ step to be repeatedly run to test small adjustments of QXD parameters without needing additional simulations of other areas in the thermodynamic cycle. Each leg of the thermodynamic cycle is divided into eleven evenly spaced windows from $\lambda = 0.0$ to $\lambda = 1.0$ with a linear coupling scheme ($k = 1$) with the exclusion of the calculation of $\Delta G_1$. In this transformation a softcore potential scheme[75] is used

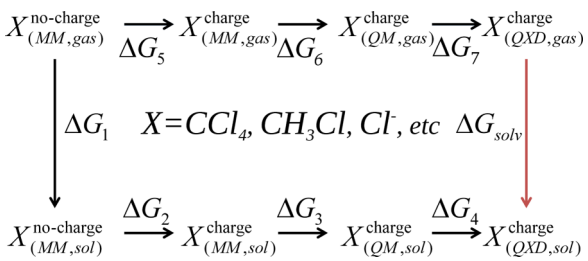


FIG. 1. Thermodynamic cycle for QXD parameterization. Shown is the cycle used for calculating the solvation free energies for the parameterization of the charge-dependent QM/MM interaction model, QXD. Cycle legs shown in black are simulated using alchemical perturbation (where needed) and the leg in red is the target free energy. For cycles not using QXD, those legs are ignored.

to avoid the "end-point catastrophe" when disappearing a LJ sphere in solution. The endpoint windows in these legs are moved to $\lambda = 0.05$ and $\lambda = 0.95$. Data were processed using a smoothing Akima spline (and in softcore cases, also extrapolated to the endpoint values of 0.0 and 1.0) to the free energy $\partial V/\partial \lambda$ plots and then those splines were integrated with respect to $\lambda$ to calculate the free energy contribution from each change. NPT equilibration of the system was run for 250 ps to allow for solvent and box size relaxation. Afterwards, each $\lambda$ window was NPT equilibrated for an additional 250 ps. Finally, NVE production was ran for 250 ps, analyzed and used in parameterization of the QXD model.

Only the $s_i$, $\zeta_i(0)$, $\zeta_{q,i}$, $\alpha_i(0)$, and $\alpha_{q,i}$ were treated as free parameters, the $N_{\text{eff},i}$ parameters being taken from literature values.[47] Data used to parameterize the QXD model were separated into training and testing sets. The training set consisted of solvation free energies for the chloride anion, methylchloride, and carbon tetrachloride molecules, as this set represented diverse charge states of chlorine. The testing set consisted of solvation free energies for the other molecules listed in Table II, as well as the free energy profile for chloride ion attack to methylchloride. Due to the length of the simulations required to obtain solvation free energies in accord with the thermodynamic cycle shown in Figure 1, a step-wise procedure was adopted whereby initial values for the QXD parameters for Cl were taken from fitting to gas-phase QM interaction energy curves at the M06-2X 6-31+G(d,p)//6-311++G(3df,3p) level. In order to reduce the degrees of freedom and conditioning of the fitting problem, this initial curve fitting was performed using LJ parameters which were then mapped into charge-static QXD parameters in accord with Equations (31)–(33), and then a single charge-dependent QXD parameter set was created to capture the three charge-static curves. These parameters were then used for optimization by exploring parameter space in discrete increments ( 5% or less of the initial values) in order refine individual training set molecule solvation free energies. Once QXD interaction curves were determined for each molecule, a single QXD parameterization was created by direction set refinement to these curves and TI simulation to obtain solvation free energy

TABLE II. Comparison between traditional and QXD models. Calculated QM/MM solvation free energies for a series different chlorine-containing compounds, using two parameterizations of the traditional Lennard-Jones model, one made to produce the chloride solvation free energy and the other to capture the $CH_3Cl$ solvation free energy, and then a single parameterization of the QXD density-dependent interaction model. Free energies provided are in kcal/mol.

| | ΔG, solvation free energy | | | |
|---|---|---|---|---|
| Compound | Experimental[79–84] | $LJ_{Cl^-}$ | $LJ_{CH_3Cl}$ | QXD |
| $Cl^-$ Anion | −89.1 | −89.21 | −125.67 | −89.64 |
| $CH_3Cl$ | −1.02 to 0.31 | 5.64 | 0.34 | 0.53 |
| $CH_2Cl_2$ | −1.90 to −0.65 | 8.11 | 0.14 | −1.37 |
| $CHCl_3$ | −1.78 to −0.38 | 11.43 | 0.62 | 0.33 |
| $CCl_4$ | −0.64 to 1.10 | 14.66 | 2.37 | 1.20 |
| $CCl_3CHCl_2$ | −2.32 to −0.24 | 17.63 | 1.51 | 0.28 |
| $CCl_3CCl_3$ | −1.57 to −0.93 | 33.84 | 4.53 | 2.94 |

results. During optimization of the QXD parameters for Cl, it was found that small adjustment of the charge-dependent QXD parameters for C and H, which were needed to balance the charge dependence of Cl, was required to get very good fits.

### E. Free energy profiles with QXD

The manuscript will then examine the attack of a chloride ion on methylchloride as an example reaction where there is a large amount of charge transfer along the reaction coordinate (RxC).[76] General simulation protocol is identical to the QM legs of the thermodynamic cycle which were outline in the previous section. Initial NPT simulations for solvent and box size equilibration were ran for 250 ps followed by another 150 ps of NPT equilibration at each different RxC window. Finally, 150 ps of NVE production was ran and used for analysis. The RxC is defined as a linear combination of orthogonal degrees of freedom, $R_1$-$R_2$, where $R_1$ and $R_2$ are defined as the distance between the attacking/leaving chloride and the carbon atoms. Umbrella windows were evenly spaced every 0.2 Å between RxC values of −3.5 Å and 3.5 Å with force constants ranging from 60 kcal/(mol/Å) and 24 kcal/(mol/Å) depending on location in the RxC. Additional windows between −0.8 Å and 0.8 Å with force constants ranging between 200 kcal/(mol/Å) and 184 kcal/(mol/Å) were used to ensure adequate sampling about the transition state of the reaction. Simulation results were analyzed using vFEP[77,78] to obtain the solution phase free energy profile.

## IV. RESULTS AND DISCUSSION

In this section the main results the manuscript are discussed. First, the performance of the interface between QXD and traditional QM/MM models will be reviewed and discussed. Next, comparisons will be drawn between the QXD model and the LJ model in their ability to reproduce high-level gas phase interaction curves. Third, the parameterization of QXD for reproducing experimental solvation free energies of various chlorine containing species will be assessed. Fourth, free energy profile results of the previous parameterization for the attack of chloride on methylchloride will be examined. Finally, an overall perspective of the QXD model and its future directions are discussed.

### A. Integration of the QXD model into conventional QM/MM frameworks using LJ mapping relations

Integration of QXD into conventional QM/MM frameworks is key to facilitate wide-spread use of the model. Toward this end, an empirical model (Eqs. (31)–(33), see Sec. III) was developed that, as seen in Figure 2, is capable of transforming traditional LJ parameters into charge-static terms which operate within a QXD framework. Differences between the mapped QXD and the LJ interaction curves were tested with thermodynamic integration simulations and found to be negligible with respect to statistical uncertainties typically on the order of 0.1 kcal/mol or less. With this empirical mapping, quantum mechanical atoms that have fully charge-dependent QXD parameters can interact with MM atoms that have static (fixed-charge) QXD parameters derived from their LJ parameters.

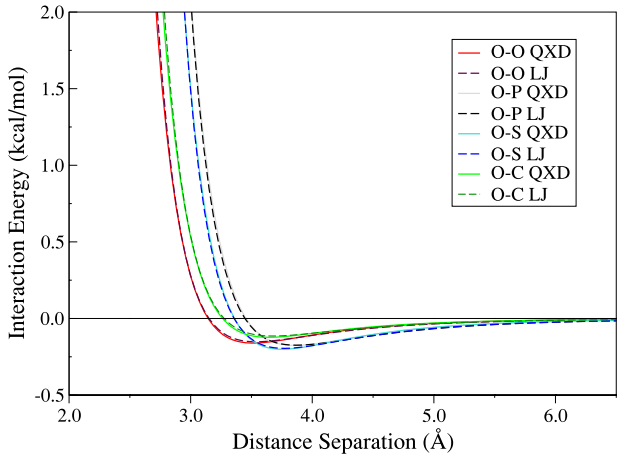


FIG. 2. Lennard-Jones interactions compared with corresponding mapped parameter QXD interactions. Shown are various atom-atom (X–O; X = O, P, S, C) interactions for both traditional Lennard-Jones and QXD models, with QXD charge-independent parameters generated from the interfacing equations which allow for on-the-fly conversion of LJ interactions into a form which can be used with the new QXD charge-dependent interaction model.

### B. Robustness of the QXD model for gas-phase interactions

High-level adiabatic QM scans were created to demonstrate the flexibility of the QXD model in reproducing interactions of systems through various charge-states. Initially, two different sets of LJ parameters were optimized to reproduce the high-level scans of the chloride anion and of the methylchloride molecule, respectively. Then, one set of QXD parameters, with full charge-dependent interactions, was optimized to best reproduce all three of the reference potentials shown. Potentials from these optimizations and the resulting scans are shown in Figure 3.

The charge-dependent QXD model stands as a marked improvement when compared to the conventional LJ model in its ability to capture the high-level gas-phase data. The two different LJ 12-6 potentials each reasonably model the interaction for which they were parameterized to reproduce; however, each fails to simultaneously capture the other interactions within a single parameterization. When trying to recapitulate the $Cl^-$ interaction, the LJ potential drastically under-stabilizes the neutral chlorinated species. Likewise, when the LJ potential is parameterized to model the methylchloride interaction the neutral particles ($CH_3Cl$ and $CCl_4$) interaction curves reasonably well; however, the parameterization significantly over-stabilizes the chloride anion. Without the ability to respond to changes in local electronic environment, the LJ potential lacks the flexibility to represent atoms over a wide range of charge states. The QXD model, on the other hand, is able to accurately capture all of the high-level quantum data with the small exception that there is a slight deviation in the $Cl^-$/water interaction curve.

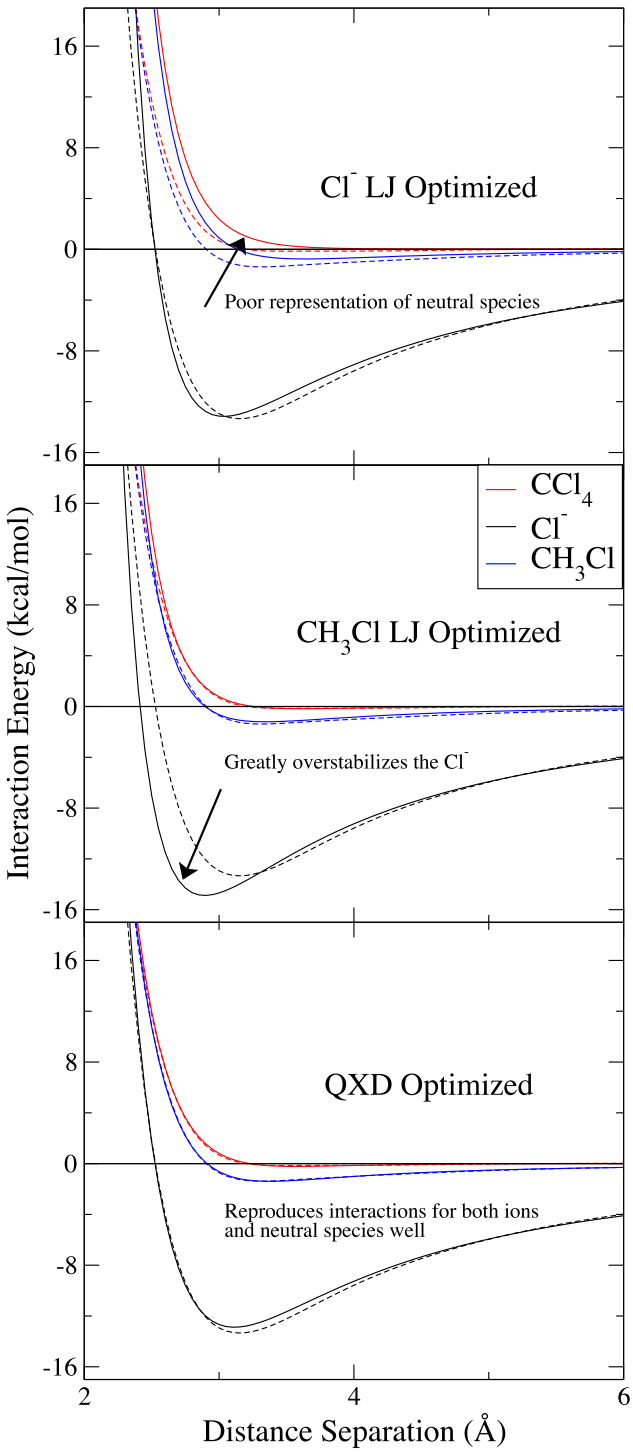


FIG. 3. Comparative X–$H_2O$ gas phase LJ and QXD interactions. Gas phase QM/MM adiabatic energy curves (solid lines) of chlorine containing compounds (X = $Cl^-$, $CH_3Cl$, and $CCl_4$) interacting with a water probe are compared to a high-level quantum benchmark (dashed lines). Water geometry was fixed to emulate a TIP4P-Ew water. Shown are the results from interactions with LJ parameters optimized to reproduce the $Cl^-$ reference energy (top), LJ parameters optimized to reproduce the $CCl_4$ reference energy (middle), and QXD with full charge dependent interactions (bottom).

The discrepancy between QXD and the high-level reference data at the minimum of the interaction curve is likely due to the limited electronic degrees of freedom in the semiempirical and MM models. Despite the high-level water having its geometry fixed to that of the TIP4P-Ew water model, differences between the reference and the QXD scans arise because the electronic density about the atoms of the high-level reference are still allowed to fully polarize. The QM/MM simulations do not allow the electronic density about the chloride anion (and, clearly, around the MM water) to relax as freely as the high-level reference and thus the quantum scans gain additional stabilization beyond what the QXD model can compensate for. Nonetheless, the QXD model demonstrates greater flexibility and robustness in capturing interactions at varied local charge state than the traditional LJ model and represents a significant improvement over the LJ potential in recapitulating high-level quantum data.

### C. QXD and LJ QM/MM solvation free energies

QM/MM solvation free energies were calculated through the use of TI and were used in the parameterization of chlorine QXD atoms for solution phase simulations. Parameters for carbon and hydrogen atoms were provided from the mapping equations with additional charge dependence assigned to these atoms to prevent artificial charge transfer between atoms. Chlorine parameters were adjusted to best reproduce the experimental solvation free energy data; the resulting parameters can be found in Table III.

The QXD model for non-bonded QM/MM interactions shows promise in accurately capturing experimentally known solvation free energies over a series of chlorine-containing compounds and represents a marked improvement over the results obtained when employing a traditional LJ model. While individual sets of parameters in either model could be generated to reproduce the solvation free energy for a single given species, the QXD model is able to accurately replicate experimental solvation free energies of several different species in a variety of local electronic states, while the LJ potential is incapable of this task. For comparison, data from QXD and two sets of LJ parameters can be found in Table II. The LJ parameter sets used where chosen to represent two different chlorine "atom types": one set that accurately reproduces the solvation free energy of a chloride anion, $LJ_{Cl^-}$, and another set that captures the methylchloride solvation free energy, $LJ_{CH_3Cl}$.

The $LJ_{Cl^-}$ set of parameters accurately reproduces the chloride anion solvation free energy, at −89.21 kcal/mol; however, for all other species it consistently undersolvates the charge neutral species. This is especially notable on molecules containing several chlorine atoms, such as hexachloroethane which is undersolvated by nearly 35 kcal/mol. This behavior

TABLE III. QXD parameters for chlorine containing compounds. Shown are the QXD parameters used in the present work to accurately predict the solvation free energy for the series of chlorine containing compounds, and the free energy profile for chloride attack to methyl chloride. The optimization procedure is described in detail in Sec. III.

| Type | $s_i$ | $\zeta_i(0)$ | $\zeta_{q,i}$ | $\alpha_i(0)$ | $\alpha_{q,i}$ | $N_{\text{eff},i}$ |
|---|---|---|---|---|---|---|
| $Cl/Cl^-$ | 11.300 | 1.970 | −0.225 | 26.31 | 0.300 45 | 5.551 |
| C | 21.308 | 2.487 | −0.215 | 8.553 | 0.000 00 | 2.657 |
| H | 5.2423 | 3.078 | −0.205 | 1.746 | 0.000 00 | 0.824 |

is expected for this parameterization as a strong repulsive potential is required to overcome the extremely favorable charge/partial-charge interactions that the chloride ion would experience in solution. As the traditional LJ model cannot respond to changes in the local electronic density, the "size" of the chlorine atoms represented in the non-ionic species are the same as those for the chloride anion. From a physical standpoint, given that the local atomic charge around a chlorine atom in a neutral species is considerably less than that of a chloride anion with full −1 charge, the interactions should reflect the smaller, harder character and have a steeper repulsive wall and closer contact distance. Thus, the $LJ_{Cl^-}$ parameter set causes the chlorine atoms in the molecular compounds to interact as if they were too large and soft, reducing the ability of water to effectively solvate.

The $LJ_{CH_3Cl}$ set of parameters captures the solvation free energy of methylchloride, reproducing the upper experimental range at 0.34 kcal/mol. Also, because the methylchloride is electronically more similar to several other molecules in the test set than the chloride anion, the $LJ_{CH_3Cl}$ parameterization generally out performs the $LJ_{Cl^-}$ set as a whole. However, it still fails dramatically is several cases. Most predominately, the chloride anion is oversolvated by approximately 35 kcal/mol. Furthermore, while the methylchloride has the appropriate amount of solvation, the hexachloroethane molecule remains undersolvated by approximately 5.5 kcal/mol. Similar arguments could be made as in the previous paragraph as to why these results are not unexpected. The static nature of the LJ potential means that if the atom being modeled is in a different local electronic environment, then new parameters must be generated to accurately model the non-bonded interactions.

The QXD model, however, is able to capture the solvation free energies for nearly all species with errors of approximately 1.0 kcal/mol or less. The QXD model provides a much more accurate representation of the non-bonded interactions, reproducing solvation free energies for both the chloride anion, at −89.64 kcal/mol, and the methylchloride molecule, at 0.53 kcal/mol. Error can be seen in the hexachloroethane molecule, undersolvating the species by about 4 kcal/mol; however, this number is still in better agreement with experiment than either of the individual LJ models. QXD overcomes the limitations of the LJ model by allowing the non-bonded interaction to adjust to changes in the local electronic density and to behave in a more physically realistic manner. As seen in Table II, more accurate solvation results can be achieved when using the QXD model within a single parameterization than with the tradition LJ potential. The accuracy gained through QXD model does have some computational overhead associated with the evaluation of the modified exchange repulsion and dispersion interactions that must be recomputed at each step of the SCF procedure, and in our current implementation, this introduces a factor of around two in the simulation time relative to the conventional LJ QM/MM model. The ability of the QXD model to adjust to different electronic environments may ultimately allow the need for atom types to be eliminated in QM/MM calculations, which will have particular impact in studies of chemical reactions where local charge changes along the reaction coordinate.

### D. QXD and LJ free energy profiles

Free energy profile simulations were performed in order to test the effects of the QXD model on the attack barrier of a chloride anion reacting with methylchloride. In previous studies,[40] the traditional LJ model was shown to accurately capture this reaction barrier while using the $LJ_{Cl^-}$ parameters set. The QXD parameter set used for these simulations is the set used to recapitulate experimental solvation free energies outlined in Subsection IV C. Results of these simulations, and comparisons to the standard LJ model with both the $LJ_{Cl^-}$ and $LJ_{CH_3Cl}$ parameter sets, can be seen in Figure 4.

The QXD model predicts the attack barrier for the reaction of a chloride anion on methylchloride to be 25.4 kcal/mol, roughly 1 kcal/mol less than the experimentally estimated value of 26.5 kcal/mol. The $LJ_{Cl^-}$ and $LJ_{CH_3Cl}$ parameter sets bracket the QXD reaction barrier: the $LJ_{Cl^-}$ set, which as discussed above produces the worst solvation free energies for neutral chlorine-containing compounds, matches very closely the experimental $Cl^-$ attack barrier, whereas the $LJ_{CH_3Cl}$ set underestimates the barrier by about 3 kcal/mol. The reason for the somewhat fortuitous agreement of the $LJ_{Cl^-}$ parameter set with the experimentally estimated barrier is discussed in more detail below.

### E. Discussion of errors

While the results of the current study are encouraging, it is nonetheless important to point out limitations that remains within the scope of the current work. The QXD model provides an improved representation of QM/MM non-electrostatic non-bonded interactions by allowing the character of atoms, in terms of their contribution to exchange repulsion and dispersion interactions, to adjust based on the underlying charge distribution. Nonetheless, the applied QM/MM protocol has additional shortcomings not directly

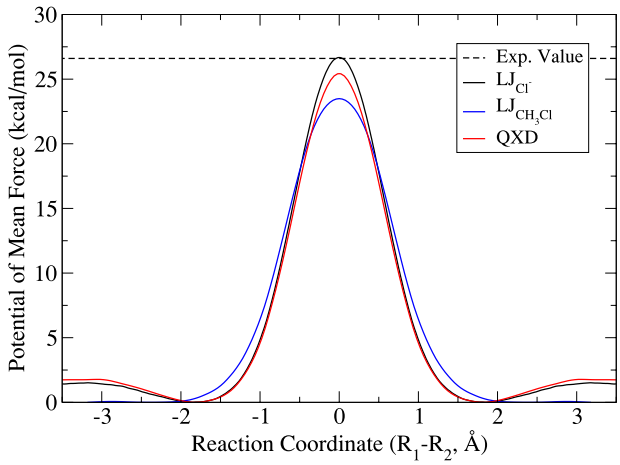


FIG. 4. Comparative LJ and QXD free energy profiles for the attack of chloride on methylchloride. Free energy profiles were run to examine the effects of the different parameter sets for the traditional LJ potential and the QXD charge-dependent potential had on the attack barrier of the symmetric reaction of $Cl^-$ with methylchloride. The simulated free energy barriers for the $LJ_{Cl^-}$, $LJ_{CH_3Cl}$, and QXD (26.5, 23.5, and 25.4 kcal/mol) parameterizations are compared to the experimental value (26.5 kcal/mol estimated from the rate constant).

associated with the QM/MM interaction that arise from (1) the lack of explicit many-body response of the MM solvent, and (2) the inflexibility inherent in the minimal valence basis set used in the semiempirical quantum model. Both of these limitations factor directly into the parameterization of the models presented here and thus warrant discussion in order to properly interpret results.

Without the appropriate many-body response of the MM solvent, differently charged QM species will experience electrostatic interactions with the same set of unpolarized MM charges on the water molecules. Further, the minimal valence basis set used in the semiempirical quantum model systematically underestimates the electronic polarization response, and this error becomes worse as anionic charge increases.[41] In the extreme cases, such as that of a $Cl^-$ anion in isolation, the orbitals are fully occupied, leaving no virtual orbitals available for polarization. Each of these effects requires compensating adjustment of the non-electrostatic non-bonded parameters (i.e., LJ or QXD) to overcome their shortcomings in an effort to reproduce experimental solvation free energies.

There is a large body of the literature[85–90] that specifically examines the importance of polarization and quantum many-body effects for modeling a solvated chloride anion. While the degree to which polarization plays a roll in correctly predicting the solvation free energy and long-range solvent structure of the chloride anion is still debated, it is clear that some amount of polarization (in both the solute and the solvent) is required if both of these observables are to be accurately be represented. The mismatch of the polarization effect on $Cl^-$ and $CH_3Cl$ at different stages along the reaction coordinate is more difficult to correct within the scope of the current work. For instance, in Sec. IV D, the QXD model predicts the attack barrier for the reaction of a chloride anion on methylchloride to be 25.4 kcal/mol, roughly a kcal/mol under the experimentally observed value of 26.5 kcal/mol. The $LJ_{Cl^-}$ parameter set, on the other hand, agrees with experiment to within the statistical error of the calculation which, in retrospect, may not be all that surprising because the LJ terms are consistent with the solvation free energy of a non-polarizable MM $Cl^-$ anion, and due to the minimal valence basis set, the QM $Cl^-$ anion also lacks polarization. The lower barrier of the $LJ_{CH_3Cl}$ parameter set, which models the Cl atoms as being smaller, reflects the preferential stabilization of the transition state where both Cl atoms carry a significant charge (approximately 0.89 $e$ in the simulation, see supplementary material[97]).

As mentioned in the introduction, despite its limitations, the present QXD model is introduced within a QM/MM framework that uses a specific reaction parameter semiempirical QM model in combination with a traditional non-polarizable MM force field. This type of QM/MM model is frequently used to study of biological reactions that often require treatment of a fairly large quantum region, on the order of 50-200 atoms, necessitating a fast QM model. The extensive MM region surrounding the active site is most often modeled by an established static charge force field. It is the hope that the QXD model will ultimately lead to improved QM/MM models for this important application area. In moving toward that goal, it should be recognized that the errors inherent in this choice of QM/MM model become coupled with the parameterization of the QXD model for the QM/MM interaction. However, we argue that introducing a model whereby the exchange repulsion and dispersion interactions of QM atoms are allowed to adjust based on the underlying charge distribution is, at least, and incremental improvement to the alternative of the ubiquitous assumption that these interactions do not depend on electronic structure. Further, the present approach is a systematic first step in the sense that we first consider a QM/MM model using a non-polarizable MM force field such that only atoms in the QM region change their charge distribution and require charge-dependent QXD parameters. It is likely that the use of polarizable force fields, where the charge distribution is also changing in the MM region, may also require the non-electrostatic non-bonded interactions to adjust accordingly.

Future work has promise to overcome both of the limitations described herein. Recently, methods have been recently introduced[41,91] that use "chemical potential equalization"[92] to overcome the problem of poor modeling of electronic response properties with approximate QM models without increasing the size of the minimal valence basis set (and thus remaining highly efficient). Further, the ongoing development of polarizable force fields[93–95] and quantum mechanical force fields (QMFFs)[3,4,96] poses a possible solution to seamlessly allow mutual polarization and other many-body quantum effects to be modeled for very large systems with practical efficiency.

## V. CONCLUSION

Herein we develop a charge-dependent QXD model for exchange and dispersion interactions in QM/MM simulations that is demonstrated to accurately capture the reaction barrier of the attack of a chloride anion on methylchloride (within approximately 1 kcal/mol) while simultaneously predicting experimental solvation free energies for a series of chlorine-containing compounds. This represents a feat which conventional LJ models fail to accomplish. Additionally, QXD has demonstrated superior flexibility over LJ models by being able to accurately recapitulate high-level gas-phase intermolecular interactions.

As it currently stands, the QXD QM/MM interaction model offers an attractive alternative to traditional LJ interactions for chemical reactions or other processes where changes in local atomic charge occur. An advantage of the QXD model is that it surmounts the problem of pre-assigning non-bonded parameters based on "atom types" corresponding to a particular chemical environment. The QXD model may therefore be useful in simulations of $pK_a$ values and $pK_a$ shifts, allowing for the correct response in the non-classical terms of key residues in large biopolymers at different catalytic steps in a reaction pathway. Furthermore, the QXD methodology could be applied to next-generation quantum mechanical force fields, which still do not explicitly couple the non-bonded, non-electrostatic interactions directly to the underlying electronic structure.

## ACKNOWLEDGMENTS

The authors are grateful for financial support provided by the National Institutes of Health (No. GM107485 to D.Y.). Computational resources were provided by the Minnesota Supercomputing Institute for Advanced Computational Research (MSI) and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant No. OCI-1053575, with Project No. TG-MCB110101 (D.Y.).

[1]A. C. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard III, J. Phys. Chem. A **105**, 9396 (2001).
[2]W. Xie and J. Gao, J. Chem. Theory Comput. **3**, 1890 (2007).
[3]T. J. Giese, M. Huang, H. Chen, and D. M. York, Acc. Chem. Res. **47**, 2812 (2014).
[4]T. J. Giese, M. T. Panteva, H. Chen, and D. M. York, J. Chem. Theory Comput. **11**, 451 (2015).
[5]H. M. Senn and W. Thiel, Curr. Opin. Chem. Biol. **11**, 182 (2007).
[6]I. A. Topol, R. E. Cachau, A. V. Nemukhin, B. L. Grigorenko, and S. K. Burt, Biochim. Biophys. Acta **1700**, 125 (2004).
[7]L. S. Devi-Kesavan, M. Garcia-Viloca, and J. Gao, Theor. Chem. Acc. **109**, 133 (2003).
[8]T.-S. Lee, C. Silva Lopez, G. M. Giambaşu, M. Martick, W. G. Scott, and D. M. York, J. Am. Chem. Soc. **130**, 3053 (2008).
[9]K.-Y. Wong, T.-S. Lee, and D. M. York, J. Chem. Theory Comput. **7**, 1 (2011).
[10]P. Banáš, P. Jurečka, N. G. Walter, J. Šponer, and M. Otyepka, Methods **49**, 202 (2009).
[11]A. E. Cho, V. Guallar, B. Berne, and R. Friesner, J. Comput. Chem. **26**, 915 (2005).
[12]W. L. Jorgensen and J. Tirado-Rives, J. Comput. Chem. **26**, 1689 (2005).
[13]J. Antony, S. Grimme, D. G. Liakos, and F. Neese, J. Phys. Chem. A **115**, 11210 (2011).
[14]S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, J. Phys. Chem. B **113**, 1253 (2009).
[15]J. Gao, N. Li, and M. Freindorf, J. Am. Chem. Soc. **118**, 4912 (1996).
[16]J. H. Jensen, H. Li, A. D. Robertson, and P. A. Molina, J. Phys. Chem. A **109**, 6634 (2005).
[17]B. A. Gregersen, X. Lopez, and D. M. York, J. Am. Chem. Soc. **126**, 7504 (2004).
[18]X. Lopez, D. M. York, A. Dejaegere, and M. Karplus, Int. J. Quantum Chem. **86**, 10 (2002).
[19]M. W. van der Kamp and A. J. Mulholland, Nat. Prod. Rep. **25**, 1001–1014 (2008).
[20]B. K. Radak, M. E. Harris, and D. M. York, J. Phys. Chem. B **117**, 94 (2013).
[21]I. Antes and W. Thiel, J. Phys. Chem. A **103**, 9290 (1999).
[22]N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, J. Phys. Chem. A **104**, 1720 (2000).
[23]Y. Zhang, T.-S. Lee, and W. Yang, J. Chem. Phys. **110**, 46 (1999).
[24]Z. Zhang, L. Xu, G. Wang, C. Shao, and L. Jiang, Comput. Theor. Chem. **1043**, 38 (2014).
[25]J. Pu, P. Jiali Gao, P. Donald, and G. Truhlar, ChemPhysChem **6**, 1853 (2005).
[26]J. Pu, J. Gao, and D. G. Truhlar, J. Phys. Chem. A **108**, 632 (2004).
[27]H. M. Senn and W. Thiel, Top. Curr. Chem. **268**, 173 (2007).
[28]H. Lin and D. G. Truhlar, Theor. Chem. Acc. **117**, 185 (2007).
[29]J. Gao, J. Comput. Chem. **18**, 1061 (1997).
[30]Y. Zhang, H. Lin, and D. G. Truhlar, J. Chem. Theory Comput. **3**, 1378 (2007).
[31]J. E. Lennard-Jones, Proc. Phys. Soc. **43**, 461 (1931).
[32]M. Freindorf and J. Gao, J. Comput. Chem. **17**, 386 (1996).
[33]M. Freindorf, Y. Shao, T. R. Furlani, and J. Kong, J. Comput. Chem. **26**, 1270 (2005).
[34]H. M. Senn and W. Thiel, Angew. Chem., Int. Ed. **48**, 1198 (2009).
[35]R. E. Bulo, B. Ensing, J. Sikkema, and L. Visscher, J. Chem. Theory Comput. **5**, 2212 (2009).
[36]S. O. Nielsen, R. E. Bulo, P. B. Moore, and B. Ensing, Phys. Chem. Chem. Phys. **12**, 12401 (2010).
[37]K. Park, A. W. Götz, R. C. Walker, and F. Paesani, J. Chem. Theory Comput. **8**, 2868 (2012).
[38]G. Hou, X. Zhu, M. Elstner, and Q. Cui, J. Chem. Theory Comput. **8**, 4293 (2012).
[39]D. Riccardi, P. Schaefer, Y. Yang, H. Yu, N. Ghosh, X. Prat-Resina, P. König, G. Li, D. Xu, H. Guo *et al.*, J. Phys. Chem. B **110**, 6458 (2006).
[40]E. R. Kuechler and D. M. York, J. Chem. Phys. **140**, 054109 (2014).
[41]T. J. Giese and D. M. York, J. Chem. Phys. **127**, 194101 (2007).
[42]J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).
[43]L. Verlet, Phys. Rev. **159**, 98 (1967).
[44]S. Kita, K. Noda, and H. Inouye, J. Chem. Phys. **64**, 3446 (1976).
[45]R. J. Wheatley and S. L. Price, Mol. Phys. **69**, 507 (1990).
[46]J. Piquemal, G. Cisneros, P. Reinhardt, N. Gresh, and T. A. Darden, J. Chem. Phys. **124**, 104101 (2006).
[47]R. Pellenq and D. Nicholson, Mol. Phys. **95**, 549 (1998).
[48]C. M. Breneman and K. B. Wiberg, J. Comput. Chem. **11**, 361 (1990).
[49]B. H. Besler, K. M. Merz, Jr., and Peter A. Kollman, J. Comput. Chem. **11**, 431 (1990).
[50]L. E. Chirlian and M. M. Francl, J. Comput. Chem. **8**, 894 (1987).
[51]U. C. Singh and P. A. Kollman, J. Comput. Chem. **5**, 129 (1984).
[52]F. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, and P. Cieplak, Phys. Chem. Chem. Phys. **12**, 7821 (2010).
[53]J. D. Thompson, C. J. Cramer, and D. G. Truhlar, J. Comput. Chem. **24**, 1291 (2003).
[54]J. M. Brom, B. J. Schmitz, J. D. Thompson, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. A **107**, 6483 (2003).
[55]J. Li, J. Xing, C. J. Cramer, and D. G. Truhlar, J. Chem. Phys. **111**, 885 (1999).
[56]J. Li, B. Williams, C. J. Cramer, and D. G. Truhlar, J. Chem. Phys. **110**, 724 (1999).
[57]J. Li, T. Zhu, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. A **102**, 1820 (1998).
[58]J. Khandogin and D. M. York, Proteins **56**, 724 (2004).
[59]D. F. DeTar, J. Org. Chem. **45**, 5174 (1980).
[60]X. G. Zhao, S. C. Tucker, and D. G. Truhlar, J. Am. Chem. Soc. **113**, 826 (1991).
[61]W.-P. Hu and D. G. Truhlar, J. Am. Chem. Soc. **117**, 10726 (1995).
[62]B. J. Gertner, R. M. Whitnell, K. R. Wilson, and J. T. Hynes, J. Am. Chem. Soc. **113**, 74 (1991).
[63]M. A. van Bochove and F. M. Bickelhaupt, Eur. J. Org. Chem. **4**, 649 (2008).
[64]I. S. Joung and T. E. Cheatham III, J. Phys. Chem. B **112**, 9020 (2008).
[65]J. Wang, P. Cieplak, and P. A. Kollman, J. Comput. Chem. **21**, 1049 (2000).
[66]H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, J. Chem. Phys. **120**, 9665 (2004).
[67]M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson *et al.*, GAUSSIAN 09, Revision A.02, Gaussian, Inc., Wallingford, CT, 2009.
[68]R. Dennington II, T. Keith, J. Millam, K. Eppinnett, W. L. Hovell, and R. Gilliland, Gaussview, Version 3.09, Semichem, Inc., Shawnee Mission, KS, 2003.
[69]Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).
[70]D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz *et al.*, *AMBER 12* (University of California, San Francisco, CA, 2012).
[71]T. P. Straatsma and J. A. McCammon, Annu. Rev. Phys. Chem. **43**, 407 (1992).
[72]D. A. Pearlman, J. Chem. Phys. **98**, 8946 (1993).
[73]J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[74]M. J. Mitchell and J. A. McCammon, J. Comput. Chem. **12**, 271 (1991).
[75]T. Steinbrecher, D. L. Mobley, and D. A. Case, J. Chem. Phys. **127**, 214108 (2007).
[76]J. P. Bergsma, B. J. Gertner, K. R. Wilson, and J. T. Hynes, J. Chem. Phys. **86**, 1356 (1987).
[77]T.-S. Lee, B. K. Radak, A. Pabis, and D. M. York, J. Chem. Theory Comput. **9**, 153 (2013).
[78]T.-S. Lee, B. K. Radak, M. Huang, K.-Y. Wong, and D. M. York, J. Chem. Theory Comput. **10**, 24 (2014).
[79]R. Schmid, A. M. Miah, and V. N. Sapunov, Phys. Chem. Chem. Phys. **2**, 97 (2000).
[80]A. C. Chamberlin, C. J. Cramer, and D. G. Truhlar, J. Phys. Chem. B **112**, 3024 (2008).
[81]D. Mackay and W. W. Shiu, J. Phys. Chem. Ref. Data **10**, 1175 (1981).
[82]R. Sander, Atmos. Chem. Phys. Discuss. **14**, 29615 (2014).
[83]M. H. Hiatt, J. Chem. Eng. Data **58**, 902 (2013).
[84]H. Kang, H. Choi, and H. Park, J. Chem. Inf. Model. **47**, 509 (2007).
[85]S. J. Stuart and B. J. Berne, J. Phys. Chem. **100**, 11934 (1996).
[86]W. L. Jorgensen and D. L. Severance, J. Chem. Phys. **99**, 4233 (1993).

[87]L. Perera and M. L. Berkowitz, J. Chem. Phys. **95**, 1954 (1991).
[88]P. Jungwirth and D. J. Tobias, J. Phys. Chem. A **106**, 379 (2002).
[89]D. Tobias, P. Jungwirth, and M. Parrinello, J. Chem. Phys. **114**, 7036 (2001).
[90]A. Serr and R. R. Netz, Int. J. Quantum Chem. **106**, 2960 (2006).
[91]S. Kaminski, T. J. Giese, M. Gaus, D. M. York, and M. Elstner, J. Phys. Chem. A **116**, 9131 (2012).
[92]D. M. York and W. Yang, J. Chem. Phys. **104**, 159 (1996).
[93]P. E. M. Lopes, B. Roux, and A. D. MacKerell, Jr., Theor. Chem. Acc. **124**, 11 (2009).
[94]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. DiStasio, Jr. *et al.*, J. Phys. Chem. B **114**, 2549 (2010).
[95]R. Chaudret, N. Gresh, C. Narth, L. Lagardère, T. A. Darden, G. A. Cisneros, and J.-P. Piquemal, J. Phys. Chem. A **118**, 7598 (2014).
[96]T. J. Giese, H. Chen, M. Huang, and D. M. York, J. Chem. Theory Comput. **10**, 1086 (2014).
[97]See supplementary material at http://dx.doi.org/10.1063/1.4937166 for all materials stated to be in the supplementary materials.