JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

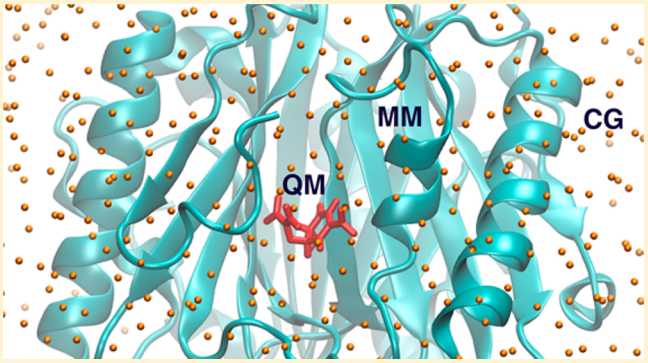
Hybrid Quantum Mechanics/Molecular Mechanics/Coarse Grained Modeling: A Triple-Resolution Approach for Biomolecular Systems

Pandian Sokkar, Eliot Boulanger, Walter Thiel, and Elsa Sanchez-Garcia*

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

Supporting Information

ABSTRACT: We present a hybrid quantum mechanics/molecular mechanics/coarse-grained (QM/MM/CG) multiresolution approach for solvated biomolecular systems. The chemically important active-site region is treated at the QM level. The biomolecular environment is described by an atomistic MM force field, and the solvent is modeled with the CG Martini force field using standard or polarizable (pol-CG) water. Interactions within the QM, MM, and CG regions, and between the QM and MM regions, are treated in the usual manner, whereas the CG−MM and CG−QM interactions are evaluated using the virtual sites approach. The accuracy and efficiency of our implementation is tested for two enzymes, chorismate mutase (CM) and *p*-hydroxybenzoate hydroxylase (PHBH). In CM, the QM/MM/CG potential energy scans along the reaction coordinate yield reaction energies that are too large, both for the standard and polarizable Martini CG water models, which can be attributed to adverse effects of using large CG water beads. The inclusion of an atomistic MM water layer (10 Å for uncharged CG water and 5 Å for polarizable CG water) around the QM region improves the energy profiles compared to the reference QM/MM calculations. In analogous QM/MM/CG calculations on PHBH, the use of the pol-CG description for the outer water does not affect the stabilization of the highly charged FADHOOH-pOHB transition state compared to the fully atomistic QM/MM calculations. Detailed performance analysis in a glycine−water model system indicates that computation times for QM energy and gradient evaluations at the density functional level are typically reduced by 40−70% for QM/MM/CG relative to fully atomistic QM/MM calculations.



INTRODUCTION

Hybrid quantum mechanics/molecular mechanics (QM/MM) methods enable the investigation of chemical processes in large molecular systems, for example, enzyme-catalyzed chemical reactions.[1−9] In this scheme, the region that is directly involved in the chemical process is described at the QM level, and the remainder is treated at the MM level. Although the hybrid QM/MM approach is well established and widely applied, it is still computationally expensive, which limits its application. In practice, QM/MM systems are often divided into active inner regions and frozen outer regions to reduce the number of degrees of freedom during optimization and dynamics simulations.[2,3,5,10−12]

QM computations normally account for most of the computational cost in QM/MM studies, especially when high-level QM methods are employed. In the case of electrostatic or polarizable embedding,[3,5,13] the evaluation of QM/MM interactions can also require a significant share of the computational effort, particularly when using large MM regions and lower-level QM methods. This is the motivation for attempts to reduce the computational cost associated with the treatment of MM atoms in QM/MM calculations. This can be achieved by boundary potential approaches, including the generalized solvent boundary potential (GSBP) and the solvated macromolecule boundary potential (SMBP).[14−16] In both cases, the entire system is divided into inner and outer regions. The inner region is treated atomistically. The outer region of the macromolecule, on the other hand, is represented by point charges, and the solvent in the outer region is modeled as a polarizable dielectric continuum (PDC), both of which contribute to the boundary potential acting on the inner region. Compared to standard QM/MM calculations, the application of SMBP in a QM/MM/SMBP framework reduces the QM calculation times by typically 30−60% without compromising accuracy (as shown in a validation study on two enzymatic reactions).[14] However, the QM/MM/SMBP approach may not be suitable if the macromolecule is highly flexible or if it undergoes significant conformational changes in the course of the reaction.[14]

Currently, there is considerable interest in the development of coarse-grained (CG) force fields to improve performance and sampling in classical molecular dynamics (MD) simulations.[17−19] In CG schemes, typically 3−4 heavy atoms (and any associated hydrogen atoms) are lumped into a single bead, thereby massively reducing the number of degrees of freedom

Received: October 27, 2014

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

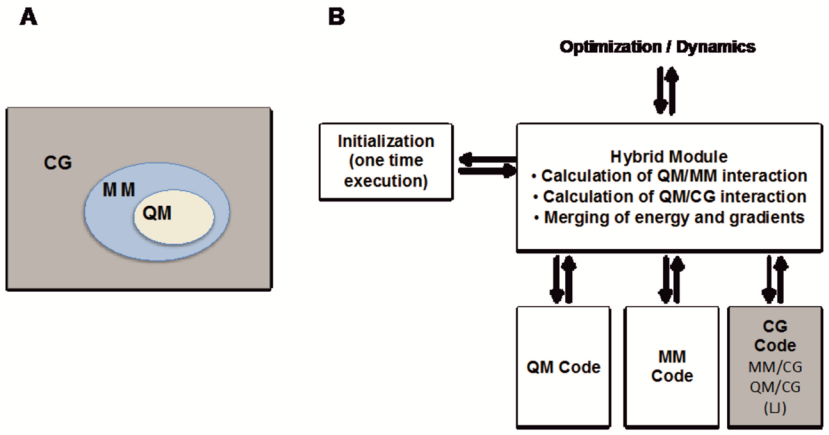


Figure 1. (A) Schematic representation of the QM/MM/CG partitioning. (B) Workflow as implemented in the ChemShell code. MM and CG denote the force field components with atomistic and coarse-grained resolution, respectively.

while still describing some properties in an acceptable manner.[18] The CG approach has been successfully applied to very large molecular systems (e.g., virus capsids), which are cumbersome for classical atomistic MD simulations.[20] However, it is not capable of reproducing fine structural details, such as fluctuations in secondary structures, arising from specific interactions that are not captured due to the coarse graining. To circumvent this problem, hybrid coarse-grained/fine-grained (CG/FG) simulation protocols have been developed.[21−25]

The CG/FG methodology is based on a similar strategy as the QM/MM approach in that the more important region is described at the atomistic level (FG) and the rest of the system is treated with lower resolution (CG level). The aim is to combine the benefits of coarse graining and atomistic resolution. Several CG/FG approaches have been reported in recent years. In serial multiscale CG/FG methods, the conformational space is sampled at CG resolution, and the final configurations are then back-mapped to atomistic resolution and sampled further.[26] In parallel approaches, these two resolutions are simultaneously present and interact with each other.[25] Adaptive resolution schemes allow the CG and FG atoms to exchange at the boundary.[27−32] In force-matching approaches, the interactions between CG and FG atoms are parametrized from fully atomistic simulations.[33]

A simple and very efficient method to couple CG/FG systems is the virtual sites approach.[21,25] In this method, virtual sites represent the FG system at the CG level. In other words, they act as virtual CG systems for the FG regions. Using this strategy, the CG/FG interaction can be represented at the CG level without any parametrization. It is thus transferable to any combination of CG and FG force fields. The forces acting on virtual sites are distributed to the constituent FG atoms weighted by their mass. This approach, originally introduced by Rzepiela et al.,[21] was adapted for the simulation of FG proteins using the nonpolarizable CG Martini force field[25] and a polarizable CG water model.[22] According to these studies, CG/FG simulations are capable of stabilizing the protein structure in a manner similar to fully atomistic FG simulations, particularly when partitioning the solvent region into an inner FG layer and an outer CG layer. The CG/FG strategy thus allows for efficient multiscale MD simulations, which are very fast (CG treatment of the outer region) and still reasonably accurate (FG description of the inner region).

Despite the success of CG/FG coupling schemes, reports on the use of CG models in QM/MM approaches are still scarce.[34,35] One notable exception is a study by Meier et al. that employed multiresolution MD simulations to investigate the conformational ensembles of two structurally similar enzymes (chorismate mutase from *Escherichia coli* and isochorismate-pyruvate lyase from *Pseudomonas aeruginosa*) to explain the differences in their catalytic activity.[34]

Here, we exploit the computational efficiency of CG modeling to improve the performance of QM/MM calculations. We report a general multiresolution QM/MM/CG approach in which the three resolutions are simultaneously present in the system and interact with each other (Figure 1A). Furthermore, we apply this approach to investigate the effects of CG solvent modeling on biocatalysis. Specifically, we study the factors that influence the potential energy profiles of the enzymatic reactions catalyzed by chorismate mutase (CM) and *p*-hydroxybenzoate hydroxylase (PHBH).

METHODOLOGY

Coupling Scheme. The QM/MM coupling schemes are well documented elsewhere.[3,5,13] Here, we describe the CG region by the Martini force field, which offers parameters for many different systems and yields results that are consistent with experimental findings.[19] We adopt a CG/MM coupling scheme from the literature[21,22,25] using the virtual sites approach introduced previously.[21,25] Virtual sites are defined on the MM atoms based on CG particle mapping of the Martini force field. To be specific, we describe the CG component by the nonpolarizable[19] and polarizable[36] versions of the Martini 2.1 force field and the MM component by the CHARMM22 force field.[37] The CG/QM interactions are treated analogously to the CG/MM interactions by introducing virtual sites for the QM region also. In the absence of established CG parameters for nonstandard QM residues, these QM virtual sites are defined in an approximate ad hoc manner (see Figures S1−S3 in the Supporting Information for details). The virtual sites in the MM and QM regions are used to mediate the Lennard−Jones (LJ) interactions with the CG regions. When using a polarizable CG model, electrostatic CG/MM Coulomb interactions are calculated directly, and the charges in the CG particles are included as external point charges in the QM calculation (as they are in standard QM/

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

MM methods). Thus, the Hamiltonian ($\hat{H}$) of the whole QM/MM/CG system can be expressed as

$$\hat{H} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} + \hat{H}_{CG} + \hat{H}_{MM/CG} + \hat{H}_{QM/CG} \quad (1)$$

Implementation of CG Force Fields in ChemShell. Code to handle CG force fields was implemented as an add-on function in a developmental version of the modular program package ChemShell.[10,38] The basic workflow for the hybrid QM/MM/CG method is shown in Figure 1B. The original hybrid module of ChemShell calls the QM and MM codes via interfaces and combines the results of QM and MM computations (usually energy and gradients). The CG interface has been added to the hybrid module. It includes the bonded and nonbonded potential energy functions associated with CG models and the definition of the virtual sites functions to enable the computation of CG/MM and CG/QM interactions. The new version of the hybrid module combines energy and gradients from QM, MM, and CG calculations for subsequent use in optimization or molecular dynamics engines. Detailed information about the general workflow in ChemShell can be found in the manual.[39] The chosen implementation is generic, but actual applications are currently restricted to the polarizable and nonpolarizable versions of the Martini 2.1 CG force field, particularly for the corresponding water models.

Computational Details. In this article, we employ the uncharged and polarizable Martini CG water models to describe the solvent at the CG level. For the uncharged water model, the LJ-interaction strength between CG water and all virtual sites was scaled down to 60%, as previously reported.[25] For the polarizable CG water model, the original LJ-parameters were used. We adopted relative dielectric constants of 2.5 for CG/CG screening and of 1.75 for CG/MM screening for consistency with previous work.[22] The satellite particles (type D) of the polarizable CG water model were originally parametrized to include only the Coulomb potential in their nonbonded interaction terms.[36] We assigned van der Waals radii of 2.5 Å to these particles to avoid any polarization catastrophe (see Supporting Information for further details).

For the QM/MM part, default parameters and settings were adopted unless otherwise specified.[16] The CHARMM22 force field was used for all MM calculations. Various levels of QM theory were employed, as discussed below.

RESULTS AND DISCUSSION

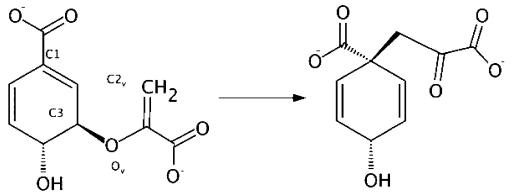
In this section, we compare the results from three-layer QM/MM/CG calculations with those from analogous standard QM/MM calculations to study the effect of CG solvent modeling on performance and accuracy. We study two biomolecular test systems for this purpose: chorismate mutase (CM) and *p*-hydroxybenzoate hydroxylase (PHBH). These two enzymes have been thoroughly investigated using various QM/MM schemes[5,6,14−16,40−45] and thus serve as suitable benchmarks to validate our methodology.

Chorismate Mutase (CM). CM catalyzes the intramolecular Claisen rearrangement of chorismate to prephenate. Previous QM/MM studies of this enzymatic reaction employed a system containing the chorismate mutase homotrimer, the chorismate substrate, and TIP3P water molecules[46,47] in a droplet of 30 Å radius (centered on chorismate).[14] Ten configurations of CM from this work[14] are used as a reference in our present study. For each of these snapshots, optimized AM1/CHARMM geometries are available for all relevant stationary points and along the reaction pathway,[14] which serve as reference geometries and starting points for the present optimizations.

To explore the potential energy surface for the rearrangement of chorismate to prephenate, we performed QM/MM/CG optimizations along a reaction coordinate defined as the difference between the C1−C2$_v$ and C3−O$_v$ distances in the substrate (Scheme 1 and eq 2).

$$\xi = d(C3{-}O_v) - d(C1{-}C2_v) \quad (2)$$

Scheme 1. Claisen Rearrangement in Chorismate Mutase



In the hybrid QM/MM/CG setup, the TIP3P water molecules were deleted from the original configurations and standard Martini CG water beads (henceforth referred to as CG water) were added to fill the 30 Å sphere surrounding the chorismate molecule (Figure 2A). CG water within 2 Å of any protein atom was deleted. This setup was subjected to energy minimization and to a short (10 ps) MD simulation using the hybrid MM/CG protocol with the protein and substrate atoms kept frozen. During the initial MD run, spherical restraints were applied to any CG beads outside the 30 Å sphere (centered on the C3 atom of the chorismate molecule). This step was performed to adjust the positions of the randomly seeded CG water beads. Thereafter, the system was divided into inner and outer regions; the inner region consisted of the chorismate molecule, the protein, and CG water atoms located within 15 Å of chorismate, whereas the remaining atoms were included in the outer region. The inner region was active during the geometry optimization, whereas the outer region was frozen.

Potential energy scans for the QM/MM/CG system were performed along the chosen reaction coordinate (Table 1). The optimized QM/MM/CG structures of chorismate, prephenate, and the transition state in the active site did not show significant deviations from the QM/MM reference geometries (Figure 2B).

The potential energy scan indicates that substitution of TIP3P by uncharged CG water does not greatly affect the barrier height ($\Delta\Delta E^{\ddagger}$) for the Claisen rearrangement in CM (see Table 1). However, the use of CG water makes the reaction more exothermic by 9.2 kcal/mol. This large deviation can be attributed to the different nature of TIP3P and CG water. CG water is an uncharged LJ-only water cluster that represents four atomistic water molecules. With an LJ distance of R_{min} = 5.3 Å, CG water beads are much larger in size than TIP3P water molecules. Thus, small protein cavities cannot be solvated by CG water. This leads to vacant sites in the protein that can be occupied by atomistic TIP3P water. We actually observe that there is generally only one CG water present inside the chorismate mutase (compared to nine TIP3P water molecules). This poor internal solvation will exaggerate intraprotein electrostatic interactions, and the highly polar prephenate product is thus overstabilized compared to the less

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

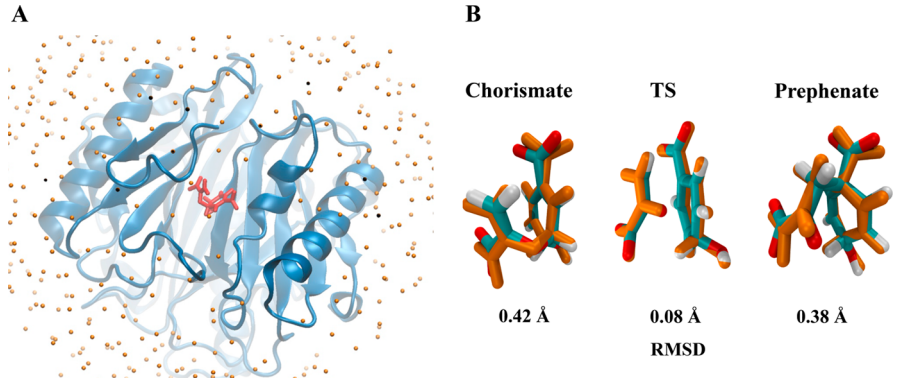


Figure 2. (A) Hybrid QM/MM/CG optimization with the chorismate molecule (red sticks) treated at the QM level (AM1), chorismate mutase enzyme (blue cartoon) treated at the MM level (CHARMM22), and water molecules (orange spheres) treated at the CG level (Martini). All CG water beads except one are outside the enzyme (see text). (B) The QM/MM/CG optimized structures of chorismate, transition state, and prephenate are shown in orange and are superposed with the corresponding QM/MM structures (colored by atom type).

Table 1. Activation Energy ($\Delta\Delta E^{\ddagger}$) and Reaction Energy ($\Delta\Delta E$) from Potential Energy Scans for Chorismate Mutase in CG Water

	standard QM/MM[a] (kcal/mol)		QM/MM/CG[b] (kcal/mol)		QM/MM/CG[c] (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	33.0	−20.0	32.9	−31.1	32.8	−19.6
2	34.8	−17.3	34.4	−31.8	33.2	−18.2
3	40.1	−17.5	37.7	−28.6	33.3	−20.7
4	38.2	−17.7	35.5	−31.2	37.7	−18.3
5	32.4	−21.9	35.6	−23.5	32.6	−22.8
6	36.2	−18.0	31.7	−31.6	34.0	−19.8
7	35.3	−19.6	33.7	−26.8	35.3	−18.6
8	35.8	−19.3	36.1	−24.0	33.4	−19.2
9	34.4	−16.3	33.1	−31.2	31.8	−19.9
10	30.6	−20.7	38.1	−20.6	31.4	−20.9
mean	35.1	−18.8	34.9	−28.0	33.5	−19.8
std. dev.	2.8	1.8	2.1	4.1	1.8	1.4

[a]QM/MM reference values (from ref 14). [b]QM/MM/CG optimization. [c]Single-point CG_{opt} energy evaluations at given QM/MM geometries with optimization of CG water positions. See text for details.

polar chorismate reactant, which is reflected in the overestimation of the reaction energy $\Delta\Delta E$ in CG water. One specific cause for the excessive stabilization of the prephenate product in CG water is the larger number of protein−prephenate hydrogen bonds: there are seven such hydrogen bonds in the QM/MM/CG structure but only three in the QM/MM structure (Figures S4 and S5 in the Supporting Information).

As a further check on the structural differences between the QM/MM/CG and QM/MM geometries, we computed root-mean-square deviations (RMSDs) for the protein component of the system (i.e., all protein residues within 15 Å of the substrate) and for five 3 Å segments (covering the range from 0 to 15 Å). The overall RMSD was 0.45 Å. The RMSDs of the individual segments did not differ significantly (data not shown), indicating that intraprotein structural deviations do not show any marked dependence on the distance from the CG water interface.

To further explore the effect of CG water on the catalytic reaction of CM, single-point QM/MM/CG energies were calculated along the previously determined QM/MM reaction pathways. This was done by replacing atomistic water by CG water in each of the QM/MM optimized geometries of the potential energy scan (when possible). Because the CG water beads were randomly seeded, straightforward single-point energy calculations led to kinks in the potential energy profile. Hence, we optimized the added CG water beads while keeping the rest of the system frozen. This CG-only optimization resulted in a smooth potential energy profile. For simplicity, the CG-only optimization will be referred to as single-point CG_{opt} energy evaluation.

The barrier heights and reaction energies obtained from these single-point CG_{opt} calculations are close to the QM/MM reference values. The differences in the computed mean values are 1.6 kcal/mol for $\Delta\Delta E^{\ddagger}$ and 1.0 kcal/mol for $\Delta\Delta E$, as opposed to 0.2 and 9.2 kcal/mol in the case of QM/MM/CG optimizations. This suggests that the water molecules do not have much influence on the barrier of the enzymatic reaction in CM (provided that they are properly placed into the predefined QM/MM geometry). The large deviations for the reaction energy in the full QM/MM/CG optimizations thus most likely originate from structural perturbations in and around the catalytic site of CM.

Chorismate Mutase in Polarizable CG Water. The polarizable Martini CG water (pol-CG) is a three-site water model with two charged particles bound to an LJ-only particle. The pol-CG model accounts for dielectric screening of bulk water and effectively reproduces the orientational polarizability of water.[36] It provides improved accuracy compared to uncharged CG water.[36] Thus, we repeated our QM/MM/CG calculations for CM in the presence of pol-CG water (Table 2) following the exact same procedures as before (see above).

The barrier heights and reaction energies for the catalytic reaction in the presence of pol-CG water, as obtained from QM/MM/CG optimizations, do not differ much from those using uncharged CG water. Single-point CG_{opt} energy calculations in the presence of pol-CG water along the predetermined reaction pathway show that the barrier height is not affected, as was the case for uncharged CG water. The average reaction energy obtained from the single-point CG_{opt} calculations is again reasonable, although it is slightly larger than the QM/MM value (by 2.9 kcal/mol) and the QM/MM/CG result with uncharged CG water (by 1.9 kcal/mol).

Although superior to uncharged CG water in terms of dielectric screening and polarizability,[36] the pol-CG water

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 2. Activation Energy ($\Delta\Delta E^{\ddagger}$) and Reaction Energy ($\Delta\Delta E$) from Potential Energy Scans for Chorismate Mutase in pol-CG Water

	standard QM/MM[a] (kcal/mol)		QM/MM/CG[b] (kcal/mol)		QM/MM/CG[c] (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	33.0	−20.0	40.2	−32.4	30.2	−19.4
2	34.8	−17.3			30.7	−22.4
3	40.1	−17.5	36.8	−27.2	34.5	−22.8
4	38.2	−17.7	32.0	−26.0	31.1	−25.6
5	32.4	−21.9	26.9	−30.5	34.3	−23.1
6	36.2	−18.0	33.6	−34.0	37.7	−19.1
7	35.3	−19.6			34.1	−23.5
8	35.8	−19.3	31.3	−31.8	37.5	−16.6
9	34.4	−16.3	27.2	−27.3	32.6	−21.2
10	30.6	−20.7	33.1	−27.9	31.5	−22.7
mean	35.1	−18.8	32.6	−29.6	33.4	−21.7
std. dev.	2.8	1.8	4.5	2.9	2.7	2.6

[a]QM/MM reference values (from ref 14). [b]QM/MM/CG optimization. [c]Single-point CG_{opt} energy evaluations at given QM/MM geometries with optimization of CG water positions. See text for details.

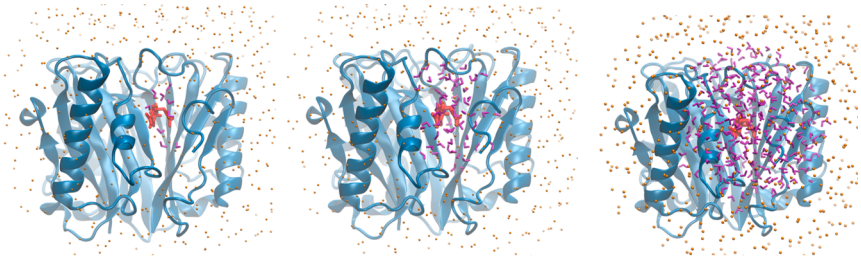
model thus does not improve the description of CM catalysis. This suggests that the larger size of CG water beads is a major concern that needs to be addressed in hybrid CG schemes. It has previously been demonstrated that uncharged CG water unnaturally promotes higher incidents of intraprotein hydrogen bonds and salt bridges during hybrid FG/CG simulations.[25] CG/MM studies using the pol-CG water models from Riniker et al.[48] and from the Martini force field[22] also reported similar trends in promoting intraprotein electrostatic interactions in hybrid simulations. These observations also point to the size of the CG water beads as an important factor that influences the accuracy of hybrid simulations.

Atomistic Water Layer in QM/MM/CG. To circumvent the problems associated with the use of CG water, we included a small layer of atomistic water molecules around the QM region. We anticipate that this will preserve the active-site conformation and overall enzyme geometry and give accurate short-range electrostatics comparable to the reference QM/MM system. We tested three different thicknesses (5, 10, and 15 Å) for the TIP3P water layer around the chorismate molecule. All relevant geometries were fully optimized.

The results are summarized in Tables 3 and 4. In the case of uncharged CG water (Table 3), the addition of a 5 Å atomistic water layer already improves the computed reaction barrier and reaction energy. With a 10 Å atomistic water layer, the potential energy profile for the chorismate−prephenate rearrangement reaction is comparable to that from the QM/MM reference calculations. The same is true for the 15 Å water layer. When using pol-CG water (Table 4), a 5 Å atomistic water layer is sufficient to produce results close to those from the standard QM/MM calculations. This confirms that the pol-CG water model describes the outer solvent layer better than the uncharged CG model, particularly with regard to long-range interactions that influence the structure in the inner region and around the active site.

It is well-known that AM1/MM calculations strongly overestimate the activation barrier for the chorismate−prephenate transformation in CM[14] (experimental value: 12.7 kcal/mol).[49] To improve the computed barrier, we performed single point QM(B3LYP/6-31G*)/MM energy calculations at the AM1/MM/pol-CG geometries (pol-CG water with 10 Å MM water, Table S1 in the Supporting Information). This gave

Table 3. QM/MM/CG Optimizations of CM with Uncharged CG Water to Test the Effect of an Atomistic TIP3P Water Layer around the QM Region in CM[a]



5 Å TIP3P 10 Å TIP3P 15 Å TIP3P

	5 Å TIP3P (kcal/mol)		10 Å TIP3P (kcal/mol)		15 Å TIP3P (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	29.4	−23.3	33.5	−23.3	33.9	−21.7
2	26.1	−27.2	33.1	−19.4	33.7	−17.6
3	33.0	−24.9	40.9	−17.8	39.3	−18.6
4	35.9	−23.1	38.8	−29.4	34.5	−24.7
5	34.3	−24.1	31.6	−24.4	23.1	−32.1
6	36.0	−28.9	35.9	−22.0	32.4	−24.1
7	34.6	−22.7	34.3	−22.2	35.4	−21.3
8	33.8	−37.5	35.4	−23.5	32.9	−20.9
9	33.9	−21.3	34.5	−19.7	33.8	−28.2
10	20.6	−35.3	31.1	−22.1	30.9	−19.7
mean	31.8	−26.8	34.9	−22.4	32.9	−22.9
std. dev.	5.2	5.8	3.2	3.1	2.3	3.3

[a]$\Delta\Delta E^{\ddagger}$: activation energy; $\Delta\Delta E$: reaction energy.

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 4. QM/MM/CG Optimizations of CM with Polarized CG Water to Test the Effect of an Atomistic TIP3P Water Layer around the QM Region in CM[a]

	5 Å TIP3P (kcal/mol)		10 Å TIP3P (kcal/mol)		15 Å TIP3P (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	31.1	−21.5	33.8	−20.2	32.3	−20.8
2	38.8	−14.5	32.5	−18.4	31.5	−27.4
3	39.7	−19.2	40.0	−22.7	35.8	−21.5
4	35.1	−21.9	35.4	−23.2	37.6	−20.6
5	30.6	−20.9	32.1	−20.0	32.5	−23.3
6	32.7	−21.6	32.5	−20.6	36.2	−21.6
7	31.4	−22.8	36.4	−16.7	34.8	−21.7
8	37.1	−16.7	36.8	−20.1	36.1	−18.7
9	32.4	−21.2	34.9	−17.6	36.1	−15.4
10	30.5	−18.7	29.0	−20.4	28.9	−22.6
mean	33.9	−19.9	34.4	−20.0	34.2	−21.4
std. dev.	3.5	2.6	3.1	2.0	2.7	3.1

[a] $\Delta\Delta E^{\ddagger}$: activation energy; $\Delta\Delta E$: reaction energy.

Table 5. Comparison between AM1/MM/CG and AM1/MM/SMBP Results for CM[a]

	CG water (kcal/mol)		pol-CG water (kcal/mol)		SMBP[b] (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	32.4	−21.8	33.2	−20.5	33.1	−19.3
2	35.6	−18.9	35.1	−17.5	34.8	−17.2
3	39.6	−18.8	40.0	−17.5	40.4	−17.1
4	39.1	−19.5	37.6	−19.2	38.3	−17.7
5	32.8	−23.5	31.3	−22.1	32.5	−21.8
6	37.6	−19.6	36.1	−18.0	36.3	−17.8
7	37.1	−20.6	35.4	−20.3	35.4	−19.5
8	36.3	−18.8	35.4	−20.0	35.7	−19.3
9	33.6	−18.2	34.4	−17.4	34.5	−16.1
10	32.3	−21.7	33.2	−20.5	30.6	−20.7
mean	35.6	−20.1	35.2	−19.3	35.2	−18.6
std. dev.	2.7	1.7	2.4	1.6	2.8	1.8

[a] $\Delta\Delta E^{\ddagger}$: activation energy; $\Delta\Delta E$: reaction energy. [b] AM1/MM/SMBP reference values from ref 14.

an activation barrier of 12.0 kcal/mol, which is in close agreement with the experimental value, confirming that the discrepancy in the AM1/MM case is due to the use of the AM1 method.

The better performance of the pol-CG water model is probably due to the inclusion of electrostatic interactions. The pol-CG water bead representing four water molecules contains one negative and one positive site, which is less elaborate than the fully atomistic TIP3P description (with four negatively and eight positively charged sites) but clearly superior to the uncharged CG water model (only LJ interactions, no electrostatics). In summary, the use of a buffering atomistic MM layer for the inner solvent layer is generally recommended in QM/MM/CG hybrid schemes, and it is advisable to describe the outer solvent layer by a polarized CG model to capture electrostatic effects in at least an approximate manner.

As previously reported,[14] AM1/MM/SMBP calculations with an explicit MM solvent layer of 19 Å accurately reproduced the AM1/MM results for CM. To allow for direct comparisons, we performed AM1/MM/CG calculations with the same 19 Å TIP3P water layer (Table 5). As expected, the QM/MM/CG results are further improved for both uncharged CG and pol-CG water by extending the explicit TIP3P water layer (Table 5 vs Tables 3 and 4). The standard AM1/MM results (Tables 1 and 2) for the mean barrier (35.1 kcal/mol) and the mean reaction energy (−18.8 kcal/mol) are closely reproduced by the AM1/MM/CG and AM1/MM/SMBP results in Table 5 with differences of at most 1.3, 0.5, and 0.2 kcal/mol for the uncharged CG, pol-CG, and SMBP models, respectively. In practice, however, the use of such large explicit MM water layers will normally be avoided in QM/MM/CG calculations for the sake of computational efficiency.

In the case of CM, 33−52% of the computation time was saved on average in every QM energy and gradient evaluation in the AM1/MM/CG calculations (with MM water layers from 0 to 19 Å) compared to AM1/MM calculations (Table S2 in the Supporting Information). The savings amounted to 19−33% in AM1/MM/pol-CG computations. They were less pronounced (19−31% for CG water and 11−23% for pol-CG water) when the QM region was treated at the density functional theory (DFT) level (B3LYP/6-31G*).

Chorismate mutase is an excellent test case for studying the effect of CG models in QM/MM calculations because there is abundant theoretical work on the reaction mechanism of this enzyme for comparison. However, the Claisen rearrangement of chorismate to prephenate does not involve significant charge transfer, and thus this system is not ideally suited for studying the effect of long-range electrostatics on the reaction mechanism. In fact, an earlier QM/MM study suggested that totally neglecting the outer regions of this enzyme does not critically affect the reaction pathways.[14] For this reason, we also validated our QM/MM/CG method using the enzyme *p*-hydroxybenzoate hydroxylase (PHBH).

PHBH Results. PHBH catalyzes the hydroxylation of *p*-hydroxybenzoate (pOHB), which involves the transfer of a formal OH^+ unit from the flavin adenine hydroperoxide cofactor (FADHOOH) to the pOHB substrate. This reaction is associated with a significant charge transfer; hence, it may serve as a test system to assess the effect of long-range electrostatics. PHBH has been used as benchmark system for several QM/MM methods.[5,14,15,42−45] We focus on the effect of long-range electrostatics from pol-CG water on the catalytic mechanism.

As reference, we used four QM/MM configurations from previous works[14,15] in which the QM region was described by the B3LYP functional[50−52] and the 6-31G* basis set. Here, we employ the AM1 semiempirical method for the sake of simplicity and for consistency with our CM validation. As before, the QM region comprised the pOHB substrate and the FADHOOH cofactor up to the first methylene group of the ribityl side chain.[14,15] The rest of the cofactor, the enzyme, and the crystal water molecules were treated at the MM level with the CHARMM22 force field. The active site of PHBH is buried inside the protein and thus not accessible to bulk water. All TIP3P water molecules within 15 Å of the distal oxygen atom of the hydroperoxide group were kept as such in the hybrid QM/MM/CG setup. In addition, an outer 12 Å pol-CG water layer was included around PHBH.

We performed QM/MM and QM/MM/CG potential energy scans for the hydroxylation of pOHB along a standard reaction coordinate defined as the difference between the O−O_d and O_d−C3 distances (Figure 3). The activation barriers and reaction energies obtained by fully atomistic AM1/MM

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

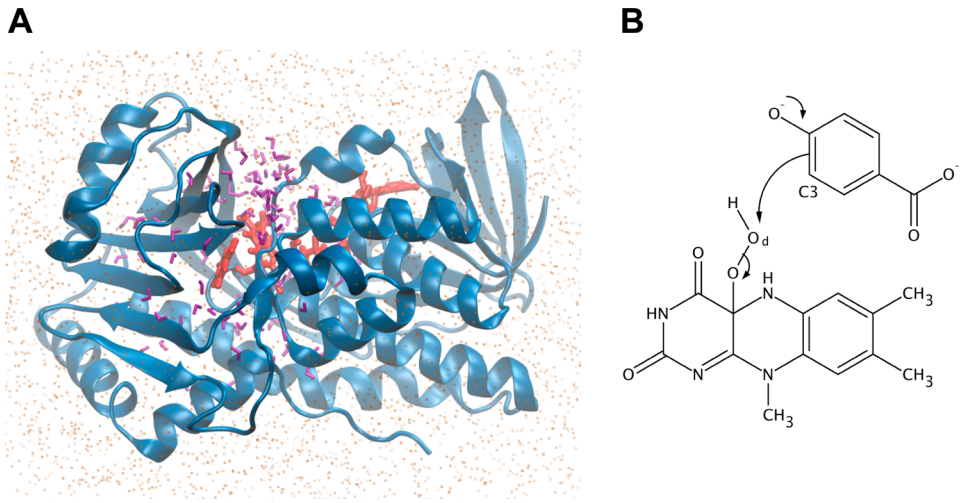


Figure 3. (A) Setup for the QM/MM/CG optimization with PHBH (cartoon), FADHOOH and pOHB (sticks), and pol-CG water (dots). The TIP3P water molecules surrounding the QM region are shown in purple. (B) QM region.

Table 6. Activation Barriers ($\Delta\Delta E^{\ddagger}$) and Reaction Energies ($\Delta\Delta E$) in PHBH-Mediated pOHB Hydroxylation[a]

	AM1/MM[b] (kcal/mol)		AM1/MM/CG[b] (kcal/mol)		DFT/MM[c] (kcal/mol)		DFT/MM/CG[c] (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	30.2	−55.0	31.5	−52.0	7.6	−26.9	8.8	−26.1
2	32.9	−52.6	33.5	−50.7	10.1	−26.6	10.3	−25.6
3	29.6	−51.0	29.8	−50.1	8.1	−25.1	8.2	−25.1
4	32.9	−47.1	33.3	−46.6	10.8	−21.6	11.2	−22.4
mean	31.4	−51.4	32.0	−49.8	9.1	−25.0	9.6	−24.8
std. dev.	1.8	3.3	1.8	2.3	1.5	2.4	1.4	1.7

[a]QM/MM results with TIP3P solvent and QM/MM/CG results with pol-CG solvent (see text for details). [b]Optimizations used AM1 for the QM region. [c]Single-point energy calculations using B3LYP/6-31G* for the QM region.

optimizations are listed in Table 6. In comparison with the previously reported B3LYP/MM calculations on the same snapshots,[14] AM1 strongly overestimates the reaction energy. However, it has been previously shown that the optimized geometries obtained from these two methods are rather similar.[42] To ascertain this, we performed single-point energy calculations at the QM(B3LYP/6-31G*)/MM level on the QM(AM1)/MM optimized structures along the reaction coordinate. The energy profile thus obtained is indeed similar to the results from the literature.[14] For comparison, we performed AM1/MM/CG optimizations in which all TIP3P water molecules farther than 15 Å away from the O_d atom (Figure 3, right) were removed and replaced by pol-CG water molecules. The positions of the pol-CG water molecules were relaxed by energy minimization followed by a 10 ps AM1/MM/CG MD simulation while keeping the fine-grained atoms frozen.

Comparison of the AM1/MM and AM1/MM/CG results indicates minor differences in the mean values of the activation barriers and reaction energies of TIP3P vs pol-CG water (0.6 and 1.6 kcal/mol, respectively). Single-point calculations at the DFT/MM/CG level of theory reveal even smaller deviations (0.5 and 0.2 kcal/mol, respectively). Going from a TIP3P to a pol-CG description of the outer solvent layer leads to an increase in the activation barrier of ∼0.5 kcal/mol.

For the sake of clarity, we emphasize that the DFT/MM energies in Table 6 were obtained at AM1/MM optimized geometries. They differ slightly from the DFT/MM energies reported previously[14] for the same configurations, which had been obtained at DFT/MM optimized geometries (e.g., 9.7 and −25.2 kcal/mol for the mean barrier and reaction energy, respectively). Furthermore, a DFT/MM/SMBP study of bulk solvent effects resulted in a decrease of the mean activation barrier by 1.3 kcal/mol (compared with the DFT/MM reference values), indicating that the PDC treatment of bulk water in the SMBP approach may provide a more complete account of long-range interactions than is captured at the QM/MM level for a finite water solvent ball.[14]

In the case of PHBH, QM computation times using AM1/MM/pol-CG were found to be 48% smaller on average than AM1/MM, and the savings for DFT/MM/pol-CG were 43% compared to DFT/MM (Table S3 in the Supporting Information). These large savings can be attributed to the large number of outer explicit MM water molecules being replaced by pol-CG water.

Multiscale MM/CG MD Simulation of PHBH. In the QM/MM/CG computations reported in the preceding section, the outer-layer CG water molecules were randomly seeded, and their positions were then adjusted by short (10 ps) QM/MM/CG MD simulations. In the case of pol-CG water, this procedure may not be sufficient to closely reproduce the electrostatic potential of the original outer-layer MM water molecules. We thus decided to pre-equilibrate the system using longer initial MM/CG simulations and to investigate whether this could affect the computed QM/MM/CG reaction pathway.

For this purpose, we performed 5 ns MD simulations of the MM region consisting of the PHBH system with a 15 Å TIP3P water layer centered on the O_d atom, which was embedded in a

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

cubic box of pol-CG water serving as the CG region. GROMACS 4.5.6 code[53] was used for these hybrid CG/MM simulations under conditions analogous to those reported in earlier work.[22] Briefly, virtual sites were defined on the PHBH protein and the cofactor, which mediate the LJ interaction between pol-CG water and the MM system. Electrostatic interactions between the MM and CG components were calculated directly. The PHBH protein was first equilibrated with harmonic restraints on backbone atoms, water oxygen atoms, FADHOOH, and pOHB. In the final production run, harmonic positional restraints were applied only on the cofactor and on the MM water oxygen atoms to avoid their diffusion into the bulk CG water. Eight conformations from the last 50 ps were selected as snapshots for subsequent QM/MM/CG optimizations in which all pol-CG water molecules beyond a cutoff distance of 12 Å from the protein were removed.

The activation barriers and reaction energies obtained from the optimization of the MM/CG snapshots are listed in Table 7. The mean barrier height (AM1/MM/CG) is now very similar to the AM1/MM reference result, whereas the mean reaction energy differs by a similar amount as before (Table 6). The single point DFT/MM/CG results for the mean activation barrier and mean reaction energy are within the statistical error limit of the reference DFT/MM results (Table 6).

Table 7. Activation Barriers ($\Delta\Delta E^{\ddagger}$) and Reaction Energies ($\Delta\Delta E$) in PHBH-Mediated pOHB Hydroxylation[a]

	AM1/MM/CG (kcal/mol)		DFT/MM/CG (kcal/mol)	
snapshot	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$	$\Delta\Delta E^{\ddagger}$	$\Delta\Delta E$
1	31.0	−54.2	7.6	−32.8
2	30.0	−45.3	8.8	−20.8
3	31.2	−46.8	8.7	−25.2
4	32.6	−48.4	10.0	−25.5
5	31.6	−53.4	8.6	−32.3
6	32.0	−48.7	10.2	−25.2
7	32.7	−48.3	11.0	−26.7
8	29.0	−51.1	8.2	−26.9
mean	31.3	−49.5	9.1	−26.9
std. dev.	1.3	3.1	1.1	3.9

[a]Optimized AM1/MM/CG results with pol-CG solvent with starting geometries taken from snapshots of the final stage of a 5 ns MM/CG simulation (see text); the DFT(B3LYP/6-31G*)/MM/CG results are obtained at AM1/MM/CG geometries.

Performance Analysis for QM/MM/CG Using the Glycine−Water System. One important goal of the triple-layer QM/MM/CG model is to reduce the computational cost of QM/MM calculations without compromising accuracy. We analyzed the performance of QM/MM/CG calculations for a glycine−water test system consisting of one glycine residue (QM region) solvated in a 15 Å TIP3P water sphere, which served as the inner region for all calculations. The outer solvent region comprised a 15 Å layer of TIP3P water (QM/MM) and uncharged CG or pol-CG water (QM/MM/CG). The average time for single QM energy and gradient calculations was measured for different DFT functionals and basis sets. All calculations were performed in serial mode on an Intel Xeon 2.9 GHz CPU with 12 GB RAM. Compared to the fully atomistic QM/MM model, the QM/MM/CG average computation times for a single energy and gradient evaluation are reduced by 44−73% for uncharged CG water and by 44−61% for pol-CG water (Table 8 and Figure 4).

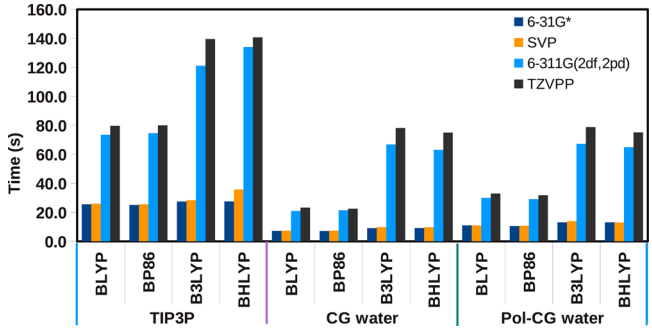


Figure 4. Average time for one single QM energy and gradient evaluation (see text and Table 8).

The corresponding savings are even somewhat larger when considering only the gradient evaluation (Table S4 in the Supporting Information). The improved performance of QM/MM/CG compared to QM/MM calculations is caused by the strong reduction in the number of point charges in the QM Hamiltonian. The computational savings are even larger for semiempirical QM methods (Table S5 in the Supporting Information) in which the evaluation of the electrostatic QM/

Table 8. Average Time for One Single QM Energy and Gradient Evaluation as well as the Computational Savings for QM/MM/CG Relative to Full Atomistic Calculations

		time (s) and computational savings (%, in parenthesis)			
		6-31G*	SVP	6-311G(2df,2pd)	TZVPP
atomistic	BLYP	25.5	25.9	73.4	79.6
	BP86	25.1	25.5	74.6	80.0
	B3LYP	27.5	28.3	121.0	139.5
	BHLYP	27.5	35.7	134.0	140.6
15 Å CG water	BLYP	7.2 (72)	7.3 (72)	20.9 (72)	23.3 (71)
	BP86	7.1 (72)	7.4 (71)	21.4 (71)	22.5 (72)
	B3LYP	9.1 (67)	9.8 (65)	66.8 (45)	78.1 (44)
	BHLYP	9.2 (67)	9.7 (73)	63.1 (53)	75.0 (47)
15 Å pol-CG water	BLYP	11.0 (57)	11.0 (58)	30.0 (59)	32.9 (59)
	BP86	10.6 (58)	10.8 (58)	29.1 (61)	31.8 (60)
	B3LYP	13.1 (52)	13.9 (51)	67.2 (44)	78.7 (44)
	BHLYP	13.1 (52)	12.9 (64)	64.9 (52)	75.1 (47)

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

MM interactions represents a dominant component of the computational effort.

Conclusions. We have implemented the three-layer multiresolution model QM/MM/CG. For validation, we applied this hybrid approach to investigate two well-known enzymes, chorismate mutase and *p*-hydroxybenzoate hydroxylase. We partitioned these systems into: (i) substrates or cofactors that are crucial to the chemical reaction and must be treated at the QM level; (ii) protein atoms or cofactors that are not directly involved in the reaction and can thus be treated at the MM level; and (iii) bulk solvent molecules that can be described at the CG level. This approach reduces the computational costs arising from the large number of outer-layer solvent molecules, which are expected to play only a very minor role in the reaction.

QM/MM/CG-optimized potential energy profiles along the reaction coordinate of the chorismate−prephenate transformation in CG water show a substantial overestimation of the reaction energy, both for uncharged and polarizable CG water models. However, single-point energy evaluations with CG water along the predetermined QM/MM reaction pathway give energy profiles closer to the QM/MM results. This indicates that the CG water models affect the geometries along the reaction pathway by perturbing the active-site conformation, which is probably caused by the poor solvation ability of the large CG water beads. We thus recommend the use of a buffering atomistic inner layer of solvent in QM/MM/CG hybrid schemes.

We also tested the effects of long-range electrostatics of bulk pol-CG water in the PHBH system. The AM1/MM/CG results indicate that bulk pol-CG water slightly underestimates the long-range interactions in the highly charged FADHOOH-pOHB transition state, and thus slightly overestimates the activation barrier, compared to the AM1/MM reference results. The AM1/MM/pol-CG barriers move slightly closer to the AM1/MM reference results when starting the optimization from snapshots taken from a preparatory 5 ns MM/pol-CG MD simulation. Single-point energy evaluations at the DFT/MM/pol-CG level yield close agreement with the reference DFT/MM results.

Detailed performance analysis on a glycine−water system shows that the three-layer DFT/MM/CG approach reduces the average computational time for DFT energy and gradient evaluations by typically 40−70% compared to the standard DFT/MM approach. For the CM and PHBH systems, the savings are typically around 40%, ranging from 11 to 52%, depending on the size of the MM water layer and the QM level of theory.

■ ASSOCIATED CONTENT

Supporting Information

Virtual sites mapping, active site geometries of selected snapshots, activation barriers and reaction energies from single point energy calculations, computing times, and force field parameters for type D particles in pol-CG water. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: esanchez@mpi-muelheim.mpg.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

E.S.-G. acknowledges a Liebig stipend from the Fonds der Chemischen Industrie, Germany. This work was supported by the Cluster of Excellence RESOLV (EXC 1069) and the Collaborative Research Center (SFB 1093), both funded by the German Research Council (DFG).

■ REFERENCES

(1) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions − Dielectric, Electrostatic and Steric Stabilization of Carbonium-Ion in Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227−249.

(2) Friesner, R. A.; Guallar, V. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389−427.

(3) Senn, H. M.; Thiel, W. QM/MM Methods for Biological Systems. *Top. Curr. Chem.* **2007**, *268*, 173−290.

(4) Hai, L.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185−199.

(5) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(6) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. A practical guide to modeling enzyme-catalysed reactions. *Chem. Soc. Rev.* **2012**, *41*, 3025−3038.

(7) van der Kamp, M. W.; Mulholland, A. J. Combined quantum mechanics /molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* **2013**, *52*, 2708−2728.

(8) Nanhao, C.; Zhou, J.; Li, J.; Xu, J.; Wu, R. Concerted cyclization of lanosterol C-ring and D-ring under human oxidosqualene cyclase catalysis: an ab initio QM/MM MD study. *J. Chem. Theory Comput.* **2014**, *10*, 1109−1120.

(9) Mlýnský, V.; Banáš, P.; Šponer, J.; van der Kamp, M. W.; Mulholland, A. J.; Otyepka, M. Comparison of ab initio, DFT, and semiempirical QM/MM approaches for description of catalytic mechanism of hairpin ribozyme. *J. Chem. Theory Comput.* **2014**, *10*, 1608−1622.

(10) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1−28.

(11) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J. Chem. Theory Comput.* **2006**, *2*, 815−826.

(12) Hu, H.; Yang, W. Free energies of chemical reactions in solution and in enzymes with ab initio QM/MM methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573−601.

(13) Bakowies, D.; Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **1996**, *100*, 10580−10594.

(14) Benighaus, T.; Thiel, W. Long-range electrostatic effects in QM/MM studies of enzymatic reactions: application of the solvated macromolecule boundary potential. *J. Chem. Theory Comput.* **2011**, *7*, 238−249.

(15) Benighaus, T.; Thiel, W. A general boundary potential for hybrid QM/MM simulations of solvated biomolecular systems. *J. Chem. Theory Comput.* **2009**, *5*, 3114−3128.

(16) Boulanger, E.; Thiel, W. Solvent boundary potentials for hybrid QM/MM computations using classical Drude oscillators: a fully polarizable model. *J. Chem. Theory Comput.* **2012**, *8*, 4527−4538.

(17) Tozzini, V. Coarse-grained models for proteins. *Curr. Opin. Struc. Biol.* **2005**, *15*, 144−150.

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(18) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(19) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI coarse-grained force field: extension to proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.
(20) Saunders, M. G.; Voth, G. A. Coarse-graining of multiprotein assemblies. *Curr. Opin. Struct. Biol.* **2012**, *22*, 144−150.
(21) Rzepiela, A. J.; Louhivuori, M.; Peter, C.; Marrink, S. J. Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437−10448.
(22) Wassenaar, T. A.; Ingólfsson, H. I.; Prieß, M.; Marrink, S. J.; Schäfer, L. V. Mixing MARTINI: electrostatic coupling in hybrid atomistic−coarse-grained biomolecular simulations. *J. Phys. Chem. B* **2013**, *117*, 3516−3530.
(23) Stansfeld, P. J.; Sansom, M. S. From coarse grained to atomistic: a serial multiscale approach to membrane protein simulations. *J. Chem. Theory Comput.* **2011**, *7*, 1157−1166.
(24) Riniker, S.; Eichenberger, A. P.; van Gunsteren, W. F. Structural effects of an atomic-level layer of water molecules around proteins solvated in supra-molecular coarse-grained water. *J. Phys. Chem. B* **2012**, *116*, 8873−8879.
(25) Sokkar, P.; Choi, S. M.; Rhee, Y. M. Simple method for simulating the mixture of atomistic and coarse-grained molecular systems. *J. Chem. Theory Comput.* **2013**, *9*, 3728−3739.
(26) Kamerlin, S. C. L.; Vicatos, S.; Dryga, A.; Warshel, A. Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. *Annu. Rev. Phys. Chem.* **2011**, *62*, 41−64.
(27) Zavadlav, J.; Melo, M. N.; Marrink, S. J.; Praprotnik, M. Adaptive resolution simulation of an atomistic protein in MARTINI water. *J. Chem. Phys.* **2014**, *140*, 054114/1−4.
(28) Abrams, C. F. Concurrent dual-resolution Monte Carlo simulation of liquid methane. *J. Chem. Phys.* **2005**, *123*, 234101/1−13.
(29) Praprotnik, M.; Site, L. D.; Kremer, K. Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly. *J. Chem. Phys.* **2005**, *123*, 224106/1−14.
(30) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. Energy conservation in adaptive hybrid atomistic/coarse-grain molecular dynamics. *J. Chem. Theory Comput.* **2007**, *3*, 1100−1105.
(31) Heyden, A.; Truhlar, D. G. Conservative algorithm for an adaptive change of resolution in mixed atomistic/coarse-grained multiscale simulations. *J. Chem. Theory Comput.* **2008**, *4*, 217−221.
(32) Nielsen, S. O.; Bulo, R. E.; Moore, P. B.; Ensing, B. Recent progress in adaptive multiscale molecular dynamics simulations of soft matter. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12401−12414.
(33) Lu, L.; Dama, J. F.; Voth, G. A. Fitting coarse-grained distribution functions through an iterative force-matching method. *J. Chem. Phys.* **2013**, *139*, 121906/1−10.
(34) Meier, K.; Choutko, A.; Dolenc, J.; Eichenberger, A. P.; Riniker, S.; van Gunsteren, W. F. Multi-resolution simulation of biomolecular systems: a review of methodological issues. *Angew. Chem., Int. Ed.* **2013**, *52*, 2820−2834.
(35) Pezeshki, S.; Lin, H. Recent developments in QM/MM methods towards open-boundary multi-scale simulations. *Mol. Simul.* **2015**, *41*, 168−189.
(36) Yesylevskyy, S. O.; Schäfer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comput. Biol.* **2010**, *6*, e1000810/1−17.
(37) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(38) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell—a modular software package for QM/MM simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 101−110.
(39) ChemShell User Manual. http://www.cse.scitech.ac.uk/ccg/software/chemshell/manual/ (accessed on October 24, 2014).
(40) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of chorismate mutase catalysis by QM/MM modelling of enzyme-catalysed and uncatalysed reactions. *Org. Biomol. Chem.* **2011**, *9*, 1578−1590.
(41) Polyak, I.; Boulanger, E.; Sen, K.; Thiel, W. A microiterative intrinsic reaction coordinate method for large QM/MM systems. *Phys. Chem. Chem. Phys.* **2013**, *15*, 14188−14195.
(42) Ridder, L.; Harvey, J. N.; Rietjens, I. M. C. M.; Vervoort, J.; Mulholland, A. J. Ab initio QM/MM modeling of the hydroxylation step in *p*-hydroxybenzoate hydroxylase. *J. Phys. Chem. B* **2003**, *107*, 2118−2126.
(43) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H. J. High-accuracy computation of reaction barriers in enzymes. *Angew. Chem., Int. Ed.* **2006**, *118*, 7010−7013.
(44) Mata, R. A.; Werner, H. J.; Thiel, S.; Thiel, W. Toward accurate barriers for enzymatic reactions: QM/MM case study on *p*-hydroxybenzoate hydroxylase. *J. Chem. Phys.* **2008**, *128*, 025104/1−8.
(45) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. QM/MM free-energy perturbation compared to thermodynamic integration and umbrella sampling: application to an enzymatic reaction. *J. Chem. Theory Comput.* **2006**, *2*, 452−461.
(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(47) Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954−9960.
(48) Riniker, S.; van Gunsteren, W. F. A simple, efficient polarizable coarse-grained water model for molecular dynamics simulations. *J. Chem. Phys.* **2011**, *134*, 084110/1−12.
(49) Kast, P.; Asif-Ullah, M.; Hilvert, D. Is chorismate mutase a prototypic entropy trap? Activation parameters for the *Bacillus subtilis* enzyme. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.
(50) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(51) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle−Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785−789.
(52) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(53) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845−854.

DOI: 10.1021/ct500956u
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX