An Extensible Interface for QM/MM Molecular Dynamics Simulations with AMBER

Andreas W. Götz,*[a] Matthew A. Clark,[a] and Ross C. Walker*[a,b]

We present an extensible interface between the AMBER molecular dynamics (MD) software package and electronic structure software packages for quantum mechanical (QM) and mixed QM and classical molecular mechanical (MM) MD simulations within both mechanical and electronic embedding schemes. With this interface, *ab initio* wave function theory and density functional theory methods, as available in the supported electronic structure software packages, become available for QM/MM MD simulations with AMBER. The interface has been written in a modular fashion that allows straight forward extensions to support additional QM software packages and can easily be ported to other MD software. Data exchange between the MD and QM software is implemented by means of files and system calls or the message passing interface standard. Based on extensive tests, default settings for the supported QM packages are provided such that energy is conserved for typical QM/MM MD simulations in the microcanonical ensemble. Results for the free energy of binding of calcium ions to aspartate in aqueous solution comparing semiempirical and density functional Hamiltonians are shown to demonstrate features of this interface. © 2013 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.23444

Introduction

Hybrid quantum mechanical and molecular mechanical (QM/MM) approaches are used extensively to study local electronic events in large molecular systems with a diverse area of applications ranging from enzymatic catalysis to properties of materials systems.[1–13] In QM/MM schemes, part of the system that includes the chemically relevant region is treated quantum mechanically while the remainder, often referred to as environment, is treated at the classical level using MM force fields. This multiscale approach reduces the computational cost significantly as compared to a QM treatment of the entire system and makes simulations possible that otherwise would not be feasible. At the same time, the numerical results obtained from QM/MM simulations should converge to full QM results if the QM region is sufficiently large such that the effect of artifacts at the QM/MM boundary is minimized and if the MM force field affords an adequate representation of the environment.

The AMBER[14,15] software package for biomolecular simulations supports QM/MM approaches that use semiempirical neglect of diatomic differential overlap type Hamiltonians[16] as well as density functional tight binding Hamiltonians.[17] These QM methods have the advantage of being computationally efficient, facilitating sampling sufficient phase space during molecular dynamics (MD) simulations. On the downside, the approximate nature of semiempirical Hamiltonians limits their accuracy and transferability, often requiring specific parameterizations for a given problem.[18,19] In addition, most semiempirical Hamiltonians are only available for selected elements of the periodic table. It thus is frequently desirable to use more accurate and generally applicable *ab initio* wave function theory or density functional theory (DFT) methods in the QM region.

Combining existing software packages for classical MD simulations with electronic structure programs is an effective approach to enable *ab initio* wave function theory and DFT based QM/MM MD simulations. It avoids duplication of programming effort and exploits the functionality and performance that are offered by the interfaced programs which frequently are the result of many years of software development. It also immediately benefits the existing user base of the simulation package who can continue to use their software infrastructure such as automated workflow schemes that rely on established input and output syntax. Consequently, several such interfaces have been developed and described in the literature.[20–32] With the exception of PUPIL[33] and the scripting environment ChemShell,[34,35] however, these are mostly limited to support only one specific electronic structure program. In addition, some interfaces are either not well maintained or have not entered the main release branch of the simulation software package and are thus not available to the end user.

[a] A. W. Götz, M. A. Clark, Ross C. Walker
 San Diego Supercomputer Center, University of California San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0505
 E-mail: agoetz@sdsc.edu
[b] Ross C. Walker
 Department of Chemistry and Biochemistry, University of California San Diego, 9500 Gilman Drive, La Jolla, California, 92093-0505
 E-mail: ross@rosswalker.co.uk

In this work, we present a versatile and easily extensible interface for QM/MM simulations within mechanical and electronic embedding schemes that supports a wide range of electronic structure software packages. This interface has been integrated into the MD engine SANDER of the AMBER[14,15] software package and has been made available with release version 12 in April 2012. The interface is written in Fortran90 using a modular fashion, which makes it easily extensible to include support for additional electronic structure software as well as portable to be included into MD software engines other than SANDER. The AMBER implementation supports the link atom approach that is available for semiempirical QM/MM simulations[16] as well as the full range of advanced sampling and free energy methods that are available in SANDER. This manuscript serves as a reference for the new interface and begins with a review of the QM/MM theory before describing features and technical details of the implementation and integration with AMBER. The numerical accuracy of the implementation is then shown by analyzing geometry optimizations of the water dimer and the energy conservation during constant energy QM/MM MD simulations of N-methylacetamide (NMA) and alanine dipeptide (ADP) in explicit solvent followed by a short discussion of typical time scales that are accessible with ab initio or DFT based QM/MM MD simulations. We finally demonstrate features of the new QM/MM interface in AMBER using the problem of calcium binding by proteins as an example. We compare results for the free energy of binding of calcium ions to aspartate in aqueous solution obtained from MD simulations using both a classical MM potential as well as QM/MM potentials using the semiempirical PM6[36] Hamiltonian and DFT before summarizing with concluding remarks.

QM/MM Theory

The total energy in a QM/MM system can be written in an additive way as

$$E = E_{QM} + E_{MM} + E_{QM/MM},\tag{1}$$

where the three terms represent the QM energy E_{QM} of the QM region in absence of perturbations due to the MM environment, the classical MM energy E_{MM} of the MM region, and the QM/MM interaction energy $E_{QM/MM}$ between the QM and the MM region. In addition to the QM and MM methods used, a QM/MM calculation thus also requires a choice for the form of the interaction energy $E_{QM/MM}$.

The simplest approach is to neglect any electronic coupling between the QM and the MM system and treat all nonbonded interactions, that is, van der Waals (vdW) and electrostatic, at the level of the classical MM force field. This is useful to impose steric constraints on the embedded QM system and commonly referred to as mechanical embedding. It can become problematic if reactive events are studied that involve significant charge transfer within the QM region because the atom types and thus both the vdW

parameters and the charges remain constant during the course of the simulation. As a consequence, the interaction between the QM and MM region of the transition and product states is typically not properly described. To improve this situation, some mechanical embedding implementations use point charges for the QM region atoms that are derived from the electronic structure calculation at each step of a simulation.

In many cases, it is also important to allow for polarization of the embedded QM region due to the electric field of the surrounding MM environment which is referred to as electronic embedding. In this case, the QM/MM interaction energy for a system consisting of N_{QM} atoms in the QM region and N_{MM} atoms in the MM region is given as

$$E_{QM/MM} = \sum_k^{N_{MM}} \int d\mathbf{r} \frac{\rho_{QM}(\mathbf{r})Q_k}{|\mathbf{r}-\mathbf{R}_k|} + \sum_A^{N_{QM}} \sum_k^{N_{MM}} \varepsilon_{Ak} \left[\left(\frac{\sigma_{Ak}}{R_{Ak}}\right)^{12} - \left(\frac{\sigma_{Ak}}{R_{Ak}}\right)^6 \right].\tag{2}$$

Here, the first term is the Coulomb interaction between the total charge density ρ_{QM} of the QM region (which consists of the electron density and in general nuclear point charges) and the fixed MM point charges Q_k. The second term is the classical vdW interaction between the atoms in the QM and MM region as given by the underlying MM force field in terms of an empirical Lennard-Jones (LJ) potential. In mechanical embedding, ρ_{QM} in the first term of eq. (2) is replaced with fixed point charges that are typically taken from the corresponding MM force field or derived from the electronic structure calculations on the fly. It is worth mentioning that electronic embedding is not always superior to mechanical embedding and an extensive study comparing different QM/MM approaches can be found for example in the work by Hu et al.[37]

If the QM/MM boundary crosses covalent bonds, the QM/MM interaction energy $E_{QM/MM}$ additionally includes bonded terms from the classical MM force field accounting for corresponding bond stretch, angle and dihedral forces between the QM and MM subsystems.

The forces acting on the atoms in a QM/MM calculation are given in terms of derivatives of the total energy expression eq. (1) with respect to the Cartesian coordinates of the atoms,

$$\mathbf{F} = -\nabla E_{QM} - \nabla E_{MM} - \nabla E_{QM/MM}.\tag{3}$$

The first two terms are the standard gradient expression for the QM method and the classical MM force field that are used in the QM and MM regions, respectively. What remains are the forces acting on the QM atoms A and the MM atoms k due to the QM/MM interaction term for which we obtain

$$\nabla_A E_{QM/MM} = \sum_k^{N_{MM}} Q_k \int d\mathbf{r} \frac{\nabla_A \rho_{QM}(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_k|} + \sum_k^{N_{MM}} \nabla_A V_{Ak}^{LJ},\tag{4}$$

where we have introduced V_{Ak}^{LJ} for the LJ potential between QM atom A and MM atom k, and

$$\nabla_k E_{QM/MM} = Q_k \int d\boldsymbol{r} \frac{\rho_{QM}(\boldsymbol{r}) \times (\boldsymbol{R}_k - \boldsymbol{r})}{|\boldsymbol{r} - \boldsymbol{R}_k|^3} + \sum_A^{N_{QM}} \nabla_k V_{Ak}^{LJ}$$

$$= -Q_k \boldsymbol{E}_{QM}(\boldsymbol{R}_k) + \sum_A^{N_{QM}} \nabla_k V_{Ak}^{LJ}, \qquad (5)$$

where $\boldsymbol{E}_{QM}$ is the electric field due to the QM charge density ρ_{QM}. The forces thus obtained can then be used for geometry optimizations or to propagate the system coordinates during MD simulations within the QM/MM framework.

Implementation

We have developed a self-contained, easily extensible interface for QM/MM calculations in which the geometry optimization or MD simulation is driven by a classical MM program. As such, the purpose of the interface is to provide the driving MM program with the QM contribution to the energy and the forces, eqs. (1–5), as obtained from an external electronic structure program. The interface is written in Fortran 90 with a simple application programming interface (API) that makes it easy to be linked with the MM program at the source code level while communication between the interface and the electronic structure programs is implemented via either file based data exchange or alternatively, as initially implemented within TeraChem,[38–40] via a client/server model for data exchange based on version 2 of the message passing interface[41] (MPI-2) standard. The interface has been integrated into the MD engine SANDER of the AMBER[14,15] software package for biomolecular simulations and was released with version 12 of AMBER in April 2012. The usefulness of this interface has already been demonstrated for DFT based QM/MM MD simulations of aqueous systems[42] and the simulation of electronic absorption spectra of the photoactive yellow protein.[40] An overview of the capabilities of the interface and details of its implementation are given in the remainder of this section.

Features

The interface supports both mechanical and electronic embedding. For the latter, the electronic structure program has to support QM calculations in an external electric field of point charges including the ability to calculate either the electric field due to the QM charge density at the position of the MM point charges or directly the forces exerted on the MM atoms that arise due to electrostatic interaction with the QM charge density, see eq. (5). At the time of writing the following electronic structure, programs are supported for mechanical embedding:

- ADF[43–45]
- GAMESS[46,47]
- NWChem[48]

and the following programs are supported for mechanical and electronic embedding:

- Gaussian[49]
- Orca[50]
- TeraChem[38,39]

This represents a set of widely used programs, both commercial and freely available, each with its own strengths for different electronic structure methods and computing platforms ranging from desktop workstations to supercomputers. In the case of TeraChem, this also includes accelerator hardware in the form of graphics processing units. Most researchers, both academic and industrial, will have access to one or the other of these software packages. The development version of AMBER also includes support for Q-Chem[51] and it is our intention to add support for additional electronic structure software with future releases, including plane wave DFT codes for materials science related QM/MM simulations.

The implementation within AMBER's MD engine SANDER builds upon the existing QM/MM functionality for semiempirical QM methods[16] and thus inherits all of its features with the exception of approaches that are either not available for ab initio wave function theory and DFT methods or that would require changes to the electronic structure software. For example, the automatic link atom setup[16] for simulations in which the QM/MM boundary crosses covalent bonds is available in exactly the same fashion as for the built-in semiempirical methods. However, both the generalized Born (GB) solvent models[52,53] and, in the case of simulations with periodic boundary conditions (PBCs) and electronic embedding, the treatment of long-range electrostatic interactions between QM and MM regions and electrostatic interactions of the QM region with its own periodic images via the particle mesh Ewald (PME) approach[16,54,55] are not available. Instead, electrostatic interactions, the first term in eq. (2), are truncated beyond a cutoff that is defined as the minimum distance between an MM point charge and any atom of the QM region. For nonperiodic simulations, this is generally not problematic as they can be run without truncation by increasing the cutoff beyond the system size, thus enabling energy-conserving MD. However, under PBCs this is not possible and it is current practice to use a cutoff that is as large as possible while using a thermostat to dissipate the heat that is introduced in MD simulations due to discontinuities in the potential resulting from the truncation of the electrostatic interactions.

The AMBER implementation supports QM/MM geometry optimizations, standard MD simulations, as well as the advanced sampling and free energy methods such as umbrella sampling[56,57] that are available in the SANDER MD program. Parallelization of replica exchange[58,59] MD (REMD) simulations and the various quantum dynamics approaches such as path integral[60,61] MD (PIMD) simulations are supported with each replica or bead running concurrently via the MPI implementation of SANDER. Parallelization of the individual QM calculations is available as provided by the corresponding electronic structure software.

A comprehensive set of regression tests covering all supported external QM software packages and different simulation options, including geometry optimizations, standard MD

simulations and advanced sampling methods like REMD and PIMD has been included with the release of AMBER 12. These tests not only help to maintain code integrity and functionality with future releases of AMBER and the interfaced electronic structure programs but also serve as examples for a user on how to use the interface.

User interface

One of the driving forces when integrating the interface with AMBER was that it should be easy to use for anybody who is familiar with the AMBER MD software package. No expert knowledge with the external QM software package is required as long as the most commonly used electronic structure methods such as DFT and second order Møller–Plesset perturbation theory (MP2) are used. For a user it is thus irrelevant which of the supported QM software packages is installed as long as it supports the electronic structure method the user wishes to use. Only minor modifications to an input file that would be used for semiempirical QM/MM simulations with AMBER are required.

Simulation setup and QM region selection. The setup of QM/MM simulations using the new interface follows the same scheme as for the built-in semiempirical models.[16] For completeness, we summarize the main steps involved, details can be found in the AMBER user manual. A QM/MM simulation with AMBER requires initially setting up input files for an MM simulation, including parameters, topology, and coordinates. This can be useful in itself for example to equilibrate a system at the MM level, however, requires providing MM parameters for nonstandard residues if these are not available in the AMBER force field library. In particular, it is important to realize that vdW parameters and in the case of mechanical embedding also charges of atoms in the QM region will be used during the QM/MM simulations, see eq. (2). This also holds for covalent force field terms that cross the QM/MM boundary if the QM and MM regions are covalently linked.

Once an MM simulation has been set up, the required modifications to the input file are minimal. The user only needs to specify the atoms that make up the QM region, the QM region charge and spin multiplicity, as well as the QM method to be used. Setting of details of the QM/MM embedding scheme that deviate from the default electronic embedding and cutoff for the real-space electrostatic interactions between QM and MM region are optional.

From this point on the QM/MM implementation takes care of everything else automatically, making sure that the QM and MM codes calculate the required contributions to the energy and forces, eqs. (1)–(5). This is straightforward if the QM region is not covalently linked to the MM region. In this case, the MM force field is modified by deleting the covalent force field terms (bond, angle, dihedral terms) and, in the case of electronic embedding, by deleting the atom point charges for all atoms in the QM region. A link atom scheme[16] is used if the QM/MM boundary crosses a covalent bond to saturate the dangling bond of the QM region, the important point of which is that it does not introduce any additional degrees of

```
&cntrl
  ...
/
&qmmm
  ...
  qm_theory = 'EXTERN',
/
&orc
  method = 'b3lyp',
  basis  = '6-31g*',
/
```

Figure 1. Example of minimal modifications (bold face) required to an AMBER input file to perform QM/MM MD simulations with B3LYP/6–31G* using the Orca program via the new interface instead of semiempirical methods implemented in AMBER.

freedom into the simulation. The link atom is added automatically for the QM calculations without user intervention but details can be controlled by the user. The point charge on the MM region atom whose bond crosses the QM/MM boundary is set to zero in the case of electronic embedding to avoid overpolarization of the QM region and any residual charge due to this procedure is evenly distributed among all remaining atoms in the MM region to maintain charge neutrality. This behavior can also be modified by the user. Finally, bonded force field terms crossing the QM/MM boundary are retained if at least one atom is part of the MM region.

QM program and method selection. Figure 1 shows the relevant parts of an AMBER control input file `mdin` that uses the QM program Orca[50] for a QM/MM simulation with the B3LYP/6-31G* method in the QM region. To specify that an external software package shall be used for the QM/MM calculation, instead of the built-in semiempirical QM methods of AMBER, it is sufficient to set the `qmmm` namelist variable `qm_theory` to 'EXTERN'. The settings for the QM approach to be used by Orca, in this case the B3LYP density functional method with the 6-31G* basis set, need to be provided in the `orc` namelist. If instead, for example, the Gaussian program is to be used, then `orc` needs to be replaced with `gau` and similar for other electronic structure software packages.

Default parameters for Hartree-Fock and DFT calculations are provided for all supported QM programs such that the forces are computed with sufficient accuracy for good energy conservation during MD simulations in the microcanonical ensemble. Specifically these are the self-consistent field (SCF) convergence and associated integral neglect thresholds as well as grid size parameters for the numerical quadrature of the exchange-correlation (XC) potential and energy in DFT calculations. If possible, the interface instructs the QM program to use the converged wave function from the previous geometry optimization or MD step as the initial guess for the present step. This extrapolation results in substantial computational savings but leads to an energy drift for MD simulations in the microcanonical ensemble.[62,63] However, the resulting energy drift is typically small enough to be acceptable for many applications, in particular if tight SCF convergence thresholds are
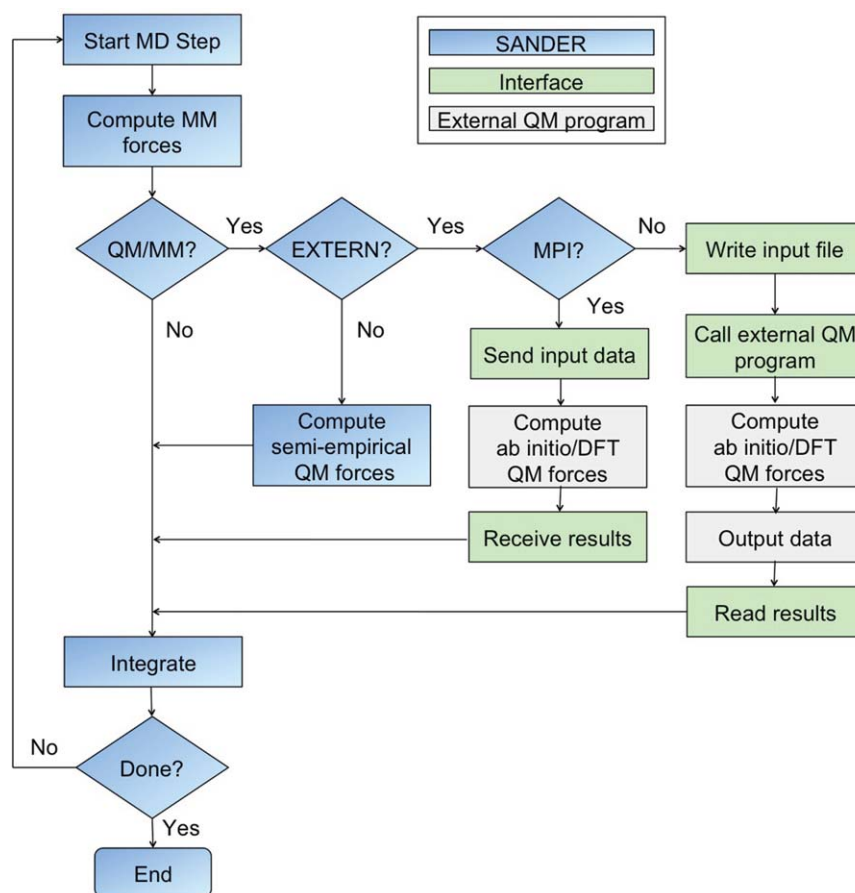
Figure 2. Flow chart for a QM/MM simulation with the AMBER MD program SANDER using the interface to external QM programs. Data exchange is either based on files and system calls or, in the client/server model, proceeds by sending and receiving the required data between SANDER and the QM program using the MPI-2 standard. The client/server model requires corresponding changes to the QM software package and is currently supported by TeraChem. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

used and a simulation is coupled to a thermostat. The user can modify these settings for a given program via its corresponding namelist in the AMBER control input file `mdin`.

It is not possible for a simple interface to account for all input options that are available in the supported QM software packages. Furthermore, new options may be added to the external software package which cannot be anticipated. Therefore, the control options have been restricted to the most likely usage scenarios, that is DFT and MP2 since most other electronic structure methods are computationally too expensive to run routine QM/MM MD simulations. An expert user may wish to use advanced input options for the supported external QM software packages that go beyond what is supported with the present version of the interface. To this end the interface also supports input for the external software packages via user-provided template files. Such a template file needs to contain all information that is required to fully specify the QM method that shall be used for a simulation (such as density functional and basis set), including accuracy settings (such as SCF convergence thresholds), that deviate from default settings of the electronic structure program. The AMBER interface will then use the information provided in the template file and supplement it with the missing data: coordinates for atoms in the QM region; coordinates of point

charges within the specified cutoff (in the case of electronic embedding); instructions to perform a single point energy or gradient calculation as required for postprocessing snapshots of MD trajectories or performing geometry optimizations and MD simulations.

Technical details

One of the key design goals of the interface has been modularity and extensibility. To achieve this, the interface was written entirely in Fortran 90 with data types, subroutines and functions for each of the supported electronic structure packages collected into separate modules. Only the driver subroutine for exchanging relevant information with the MD program (QM region atom types and coordinates, charge, spin multiplicity, MM region point charges and coordinates, QM contribution to the energy and forces) is exposed, all other functionality that handles communication with the QM program is private to the modules. A separate module collects utility routines that are common to all QM program-specific modules, such as debug and printing functions.

Two different communication methods are implemented for data exchange between the interface and the QM programs (see Fig. 2). Communication via files and system calls is

available for all supported software packages. In this case, upon each call of its driver routine, the interface proceeds by

1. writing input files for the QM program that contain the current QM region atomic numbers and coordinates and MM point charges and coordinates,
2. executing the QM program via a system call, and
3. parsing the output files of the QM program to retrieve the energy and forces

If requested by the user the interface will also store the dipole moment and atomic partial charges of the QM region along an MD trajectory. The interface stores the input and output files for the QM calculation of the present and the last geometry optimization or MD step. This simplifies debugging in case of program crashes.

Data exchange is also implemented via a client/server model based on MPI-2. This is currently supported only by Tera-Chem[40] as it requires source code changes to the QM program. However, the interface is very generic and support could be easily added to other electronic structure software. For use with the MPI-2 interface, the QM program is started in server mode at the beginning of a simulation. The interface then connects as a client to the QM program and all subsequent data exchange proceeds via standard MPI send and receive calls. At the end of the simulation, the interface sends a signal for the QM program to quit and disconnects.

The client/server model of data exchange has several advantages over communication that proceeds through files and system calls. It avoids the initialization of the QM program that is otherwise required during each geometry optimization or MD step and thus reduces computational overhead that can be significant for small QM regions. More importantly this avoids any loss of precision during the data transfer that typically occurs when formatted input and output files are used. In addition, data exchange between the MM and QM program is possible during the optimization of the wave function or electron density in the SCF procedure. With corresponding modifications in the QM program and the interface this could be useful for more advanced QM/MM coupling schemes including GB solvation models or the treatment of long-range QM/MM eletrostatics under PBCs via PME approaches similar to the ones that are available for semiempirical QM methods.[16] In comparison to the file based interface, the MPI-2 based client/server interface is easier to maintain as it is robust to changes in the format of either input or output files. Alternative interprocess communication methods relying on different protocols, for example using sockets, could also be envisioned.

For the implementation into the MD engine SANDER, the existing QM/MM code in AMBER has been refactored and if an external electronic structure program is used for a QM/MM calculation an AMBER-specific driver routine for the new interface is called instead of the built-in semiempirical code. Additional details about the implementation including the API of the interface are available in the Supporting Information.

Computational Details

The software base used for all simulations in this work was a development version of AMBER 14. The executables were built under the Rocks Cluster Distribution 5.4.3 (based on CentOS 5.6) with the Intel compiler and MKL library version 12.1.1.256 and the MVAPICH2 MPI-2 implementation version 1.8a1p1. QM/MM calculations were either performed with the PM6[36] semiempirical model as implemented by us in AMBER or using the Gaussian 09[49] electronic structure program. Standard MNDO[64]-type expressions are used in AMBER with semiempirical Hamiltonians for QM/MM interactions between point-charges and electrons and between point-charges and QM cores (nuclei plus core electrons). The exponential damping function of the point-charge core interaction uses a value of 5.0 for the exponent corresponding to the point charge and the value that has been optimized for the corresponding Hamiltonian for the exponent corresponding to the QM atom. The PM6 Hamiltonian uses the PM3[65] exponents where available. No PM3 parameters are available for Ca^{2+} and thus we used exponents of 1.3 and 2.0 in this work, denoted as PM6/a and PM6/b, respectively. The SCF was considered converged when the energy difference between two consecutive SCF cycles dropped below 10^{-10} kcal/mol for the PM6 implementation in AMBER or when the root-mean-square of the difference density matrix elements between two SCF steps dropped below 10^{-8} in the Gaussian calculations. Other than this, default Gaussian 09 settings were used. DFT calculations were performed using the BP86,[66,67] BLYP[66,68] and B3LYP[69] XC functionals and the 6-31G*[70,71] or 6-311G**,[72] TZVP[73] and aug-cc-pVQZ[74,75] basis sets and MP2 calculations were performed using the cc-pVDZ[74] basis set. All simulations have been set up with the tleap program of AmberTools.

Geometry optimizations of the water dimer were performed with AMBER using either the TIP3P[76] or the SPC[77] rigid three-site point-charge water models and a combination of different QM/MM methods with electronic embedding and a truncated Newton conjugate gradient algorithm with a termination threshold of 10^{-2} kcal/mol/Å for the root-mean-square of the gradient and a maximum of 100 optimization steps.

NMA and ADP were solvated in a droplet of SPC/Fw[78] flexible three-site point-charge water molecules of 15 Å radius (408 and 403 water molecules, respectively). A soft half-harmonic restraining potential was used beyond this radius. The ff99SB force field[79] was used for NMA and ADP in MM calculations. QM/MM calculations used electronic embedding with NMA in the QM region. A QM/MM boundary crossing covalent bonds was tested with ADP, selecting the QM region such that peptide bonds were not cut: the acetyl capping group and its adjacent nitrogen atom as well as the methyl group on the N-methyl capping group were kept in the MM region, leaving a total of 12 atoms in the QM region including two hydrogen link atoms that are automatically placed along the broken bonds between carbon and nitrogen. A time step of 0.5 fs was used for all NMA and ADP simulations. Non-bonded interactions were not truncated. The system was equilibrated for 20 ps at the MM level with Langevin dynamics[80]

at 300 K using a collision frequency of 5 ps^{-1} before switching to constant energy QM/MM simulations. Energy drifts are obtained from a linear regression of total energies along the trajectory.

Geometry optimizations of the Ace–Asp–NMe peptide (acetyl and N-methyl capped aspartate) and Ca^{2+} ion were performed with AMBER using the ff99SB force field and a limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm with a termination threshold of 10^{-2} kcal/mol/Å for the root-mean-square of the gradient. In order to keep the Ca^{2+} ion at a fixed distance from the carboxylate group, two strong restraints were used: an angle restraint was added to keep the Ca^{2+} ion aligned with the bond between the carboxyl atom and the β-carbon atom of aspartate, that is, to keep the angle C$_\beta$–C$_{carboxyl}$–Ca^{2+} at 180°; the distance $R_{C,Ca^{2+}}$ between the Ca^{2+} ion and the carboxyl carbon atom was restrained to values ranging from 2 to 6 Å with a spacing of 0.1 Å. QM single point calculations were performed at the geometries obtained from the MM geometry optimizations.

For the MD simulations, the Ace-Asp-NMe peptide and Ca^{2+} ion were solvated with TIP3P[76] rigid three-site point-charge water molecules. A rectangular box of approximately 63.3×63.9×58.9 Å^3 (6384 water molecules) was used for MM simulations and a water droplet with soft half-harmonic potential beyond 20 Å radius (1006 water molecules) was used for both MM and QM/MM simulations. The ff99SB[79] force field was used for the MM simulations and the QM/MM calculations used electronic embedding with the peptide and the Ca^{2+} ion in the QM region. A time step of 2.0 fs was used for all simulations with bond distances to hydrogen atoms constrained using the SHAKE[81,82] algorithm. Nonbonded interactions were not truncated for the water droplet simulations. Simulations using PBCs (MM only) were performed with a cutoff of 8 Å for the real-space nonbonded interactions and the PME algorithm[54] to account for long-range electrostatics beyond the cutoff. The water droplet was equilibrated for 100 ps at the MM level using Langevin dynamics at 300 K with a collision frequency of 5 ps^{-1}. The position of the carbon atom of the carboxyl group and the position of the calcium ion were restrained with a harmonic potential with a force constant of 100 kcal mol^{-1} Å^{-2} during the equilibration to avoid diffusion toward the droplet boundaries. The water box was equilibrated using the same protocol using constant volume Langevin dynamics followed by another 100-ps equilibration using constant pressure Langevin dynamics with the Berendsen barostat[83] with a target pressure of 1 bar and a pressure relaxation time of 1 ps.

All subsequent simulations to determine the potential of mean force (PMF) for calcium binding to the aspartate carboxyl group were performed using Langevin dynamics at 300 K and constant volume in the case of periodic boundaries. The position restraint on the carboxyl carbon atom was retained and a harmonic angle restraint with force constant 300 kcal mol^{-1} Å^{-2} rad^{-2} was added to keep the Ca^{2+} ion aligned with the bond between the carboxyl carbon atom and the β-carbon atom of aspartate, as described above for the geometry optimizations. The reaction coordinate chosen for the

biased MD simulations[57] was the distance $R_{C,Ca^{2+}}$ between the Ca^{2+} ion and the carboxyl carbon atom, ranging from 2 to 6 Å with a window spacing of 0.1 Å. A harmonic biasing potential with a force constant of 300 kcal mol^{-1} Å^{-2} was used and initial configurations along the reaction coordinate were generated by equilibrating for 50 ps using MM. QM/MM simulations were equilibrated for another 4 ps starting from the MM equilibrated configurations. The data from the biased MD simulations were collected for 20 ps and unbiased using the weighted histogram analysis method[56,84,85] with a bin size of 0.05 Å and a stringent tolerance of 10^{-4} kcal/mol on every point in the PMF.

Figures were generated with VMD[86] version 1.8.7 and gnuplot[87] version 4.4.

Numerical Accuracy and Performance

Water dimer geometry optimization

We have chosen the water dimer to benchmark our QM/MM implementation as this is a standard test system and reference data is available for comparison. Table 1 shows results for MM, QM, and QM/MM geometry optimizations using the TIP3P and SPC classical water models and DFT (BP86/TZVP, BLYP/aug-cc-pVQZ, B3LYP/TZVP) as QM method as well as experimental data. In the QM/MM calculations, either the hydrogen bond donor molecule (D) or acceptor molecule (A) is in the QM region (see Fig. 3). All geometry optimizations result in C$_S$ symmetry which we have tested also starting from distorted geometries.

We first note that in general our results are in very good agreement with previously published data with numerical differences likely due to details of the used DFT implementations and geometry optimization algorithms. Loferer et al.[27] and Lev et al.[30] used density fitting (also called RI- J approximation) for their BP86/TZVP and BLYP/aug-cc-pVQZ calculations which slightly affects energetics and geometries while Meier et al.[32] used a simple steepest descent algorithm with the energy as convergence criterium for their geometry optimizations. The latter can be problematic as can be seen from the purely classical SPC results for which we obtain a binding energy and D–A distance in agreement with Jorgensen et al.[76] and the similar TIP3P model while Meier et al.[32] report a much longer D–A distance and smaller bond angle a(OH · ·O). We confirmed that a steepest-descent geometry optimization starting from the geometry reported by Meier et al.[32] remains in the vicinity of the starting point. The values reported by Meier et al.[32] thus have to be interpreted with care.

The DFT calculations result in binding energies and geometries that are close to the experimental values, underestimating the hydrogen bond distance by 0.1 Å. The TIP3P and SPC water models, being parameterized to reproduce bulk water properties, overestimate the binding energy by 1 kcal mol^{-1} and underestimate the hydrogen bond distance over 0.2 Å. Compared to the reference QM calculations, the QM/MM calculations all result in increased binding energies and reduced

Table 1. Binding energy (kcal/mol), distances (Å) and angles (°) of a water dimer in vacuum using different QM, MM and QM/MM (hydrogen bond donor D or acceptor A in QM region) Hamiltonians with electronic embedding.

D-A	Reference	E	$d(H\cdot\cdot O)$	$d(O\cdot\cdot O)$	$a(O\text{-}H\cdot\cdot O)$
Experiment	Refs. [88–90]	5.44	N/A	2.98	
					174 ± 20
Full QM					
BP86	this work	5.72	1.91	2.89	174.5
	Meier et al.[a],[32]	N/A	1.91	2.89	174.0
	Loferer et al.[b],[27]	5.68	1.89	2.86	165.5
B3LYP	this work	5.99	1.93	2.90	175.6
	Meier et al.[b],[32]	N/A	1.93	2.90	176.0
Full MM: TIP3P water model					
TIP3P	this work	6.55	1.79	2.75	174.0
	Jorgensen et al.[76]	6.50	N/A	2.74	N/A
	Loferer et al. [27]	7.08	1.75	2.73	176.3
	Lev et al.[b],[30]	6.14	1.83	2.81	178.7
Full MM: SPC water model					
SPC	this work	6.61	1.75	2.75	176.1
	Jorgensen et al.[76]	6.59	N/A	2.75	N/A
	Meier et al.[a],[32]	N/A	1.93	2.99	164.9
QM/MM: TIP3P water model					
BP86-TIP3P	this work	8.09	1.65	2.65	179.9
	Loferer et al.[b],[27]	7.97	1.73	2.72	178.0
BLYP-TIP3P	this work	8.18	1.64	2.63	178.9
	Lev et al.[b],[30]	8.34	1.69	2.68	179.0
TIP3P-BP86	this work	6.62	1.82	2.78	177.7
	Loferer et al.[b],[27]	6.87	1.78	2.76	177.9
TIP3P-BLYP	this work	5.56	1.86	2.82	176.7
	Lev et al.[b],[30]	6.06	1.84	2.81	178.3
QM/MM: SPC water model					
BP86-SPC	this work	7.79	1.68	2.67	179.7
	Meier et al.[b],[32]	N/A	1.68	2.65	166.8
B3LYP-SPC	this work	7.72	1.70	2.68	179.6
	Meier et al.[a],[32]	N/A	1.70	2.66	166.7
SPC-BP86	this work	6.66	1.79	2.79	179.6
	Meier et al.[a],[32]	N/A	1.98	2.97	172.2
SPC-B3LYP	this work	6.72	1.79	2.79	179.9
	Meier et al.[a],[32]	N/A	1.97	2.97	171.8

BP86 and B3LYP calculations were performed with the TZVP basis set, BLYP calculations with the aug-cc-pVQZ basis set. [a] Steepest descent geometry optimization with energy as convergence criterium. [b] DFT calculations employed density fitting (RI-J approximation).

D–A distances while the bond angle $a(OH\cdot\cdot O)$ remains close to linear. The hydrogen bond is shorter with a larger binding energy if the hydrogen bond donor water molecule is in the QM region. The QM/MM calculations are thus closer to the reference QM results and experimental data if the QM water is the hydrogen bond acceptor.

MD energy conservation for NMA and ADP in explicit water

The forces obtained from electronic structure software packages using default settings are in some instances not accurate enough for reasonable energy conservation during constant energy MD simulations, although the numerical accuracy may be sufficient for standard quantum chemical applications such as explorative geometry optimizations with loose convergence criteria. We thus have established default settings for Hartree–Fock and DFT calculations for all supported electronic structure programs that are used by the QM/MM interface to reduce the numerical noise such that the energy is conserved

to a high degree during MD simulations in the microcanonical ensemble. This typically involves tightening the default SCF convergence criteria and associated integral neglect thresholds as well as increasing the accuracy of the numerical quadrature grid for the XC potential and energy in the case of DFT calculations. As stated earlier, by default the converged wave function or electron density is used as the initial guess in the subsequent MD step to speed up the SCF convergence. This can lead to an energy drift,[62,63] however, using a tight SCF
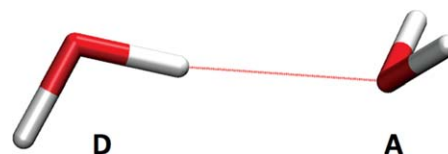

Figure 3. Water dimer as optimized with B3LYP/TZVP. The hydrogen bond donor (D) is on the left, the hydrogen bond acceptor (A) is on the right. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
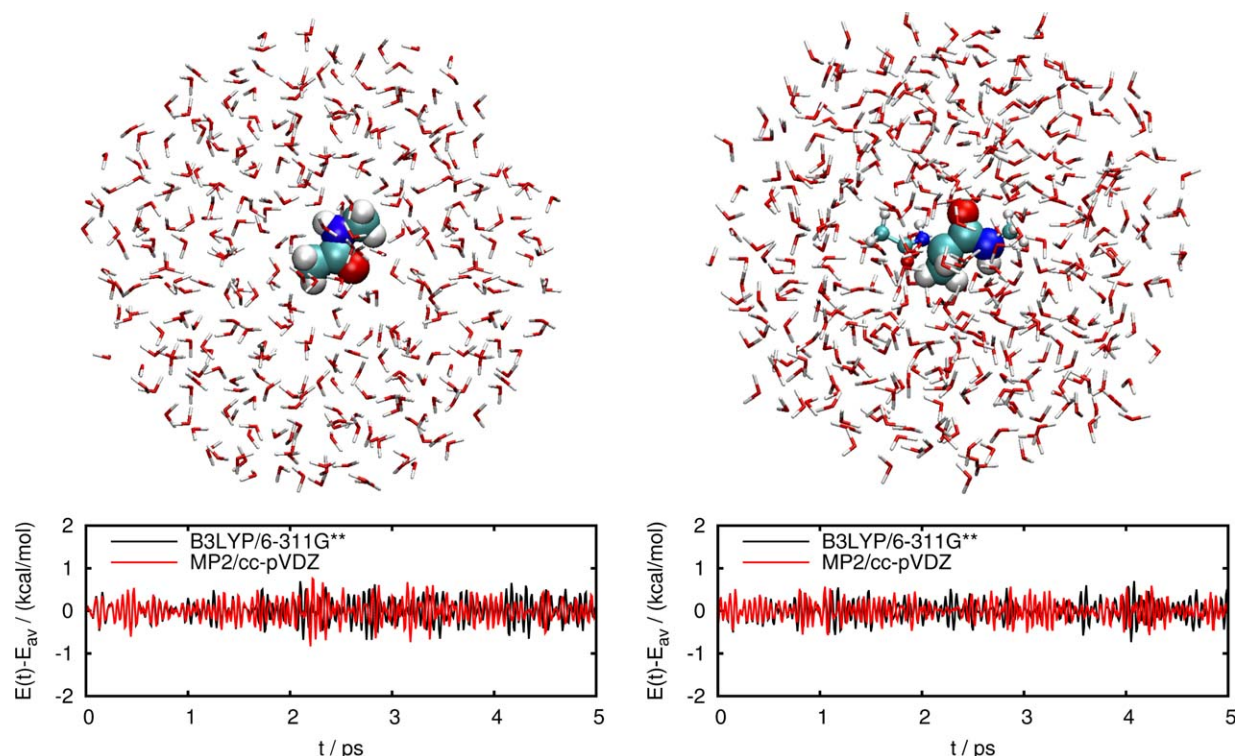
Figure 4. Energy conservation during constant energy QM/MM MD simulations of NMA (left) and ADP (right) in a droplet of 408 and 403 SPC/Fw water molecules, respectively, with electronic embedding at 300 K using a time step of 0.5 fs. The QM region is highlighted and consists of the entire NMA molecule but only part of ADP with the QM/MM boundary crossing covalent bonds such that peptide bonds are left intact. Hydrogen link atoms are used to saturate dangling bonds in the QM region. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

convergence threshold that minimizes numerical noise in the gradients, reduces this substantially.

Figure 4 shows the total energy during constant energy QM/MM MD simulations of NMA and ADP in a droplet of SPC/Fw water molecules. The trajectories were started after an initial equilibration with MM MD at 300 K, used electronic embedding, a time step of 0.5 fs, and the default settings of the QM/MM interface. Only part of ADP is in the QM region which is chosen such that peptide bonds do not cross the QM/MM boundary. Hydrogen link atoms are used to saturate dangling bonds in the QM region. The energy conservation is excellent both without (NMA) and with link atoms (ADP), with an energy drift of 2.5×10^{-3} kcal/mol/ps (which is equivalent to 1.1×10^{-6} kT/dof/ps, where dof is degrees of freedom) for the B3LYP/6-311G** simulation of NMA, an energy drift of 3.5×10^{-3} kcal/mol/ps (equivalent to 1.6×10^{-6} kT/dof/ps) for the MP2/cc-pVDZ simulation of NMA, an energy drift of 3.8×10^{-3} kcal/mol/ps (equivalent to 1.7×10^{-6} kT/dof/ps) for the B3LYP/6-311G** simulation of ADP, and an energy drift of -3.6×10^{-3} kcal/mol/ps (equivalent to -1.6×10^{-6} kT/dof/ps) for the MP2/cc-pVDZ simulation of ADP.

Typical computational throughput

The computational throughput that can be achieved with QM/MM MD simulations depends on many parameters, in particular the QM method and basis set used, and to some extent also the electronic structure software and the available hard-

ware. It is not our intention to benchmark the performance of the different electronic structure codes that are supported by the QM/MM interface and each of the different software packages has its own strengths and advantages, both in terms of available QM methods and in terms of computational performance on different hardware. However, it is useful to have an idea of the order of magnitude of the time scales that are currently accessible with QM/MM MD. Of particular interest is DFT since it has an excellent cost/accuracy ratio. For QM/MM MD simulations with a QM region size of 50 to 100 atoms using a time step of 0.5 fs and running on 16 state-of-the art CPU cores we have observed a computational throughput of around 0.1–0.5 ps/day using hybrid DFT methods with split valence basis sets and polarization functions on all atoms. If constraining the bond distances to hydrogen atoms using the SHAKE[81,82] algorithm does not affect the simulation results, a time step of 2.0 fs can be used also for QM/MM MD simulations, with a corresponding increase of the computational throughput by a factor of four as compared to a time step of 0.5 fs. For some applications a larger degree of numerical noise in the forces than provided with the default settings of the QM/MM interface may be acceptable, in particular if one is not interested in dynamical quantities but average statistics and a thermostat or stochastic dynamics are used. In this case, the SCF convergence threshold and XC quadrature grid parameters may be loosened which would lead to a corresponding speedup, however, the order of magnitude of accessible time scales would remain.
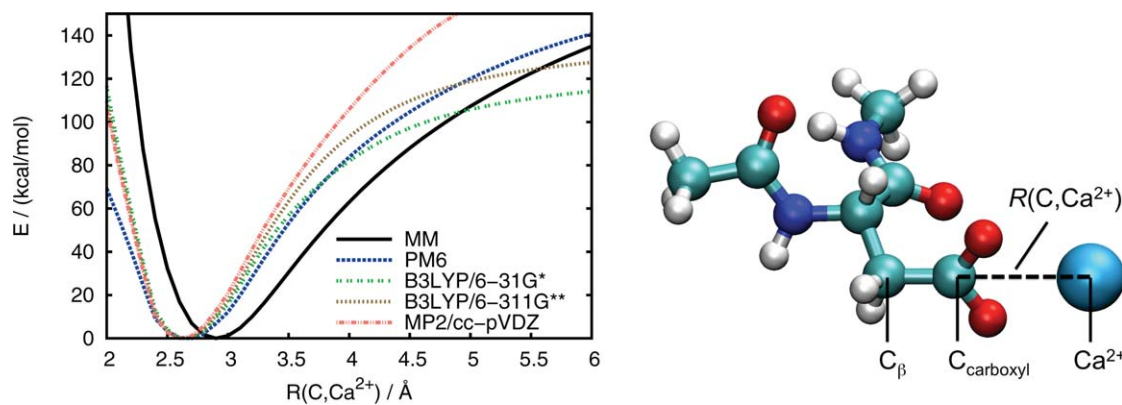
Figure 5. Potential energy profile of calcium ion coordination to the carboxyl group in acetyl and *N*-methyl capped aspartate in vacuum. QM energies are evaluated at MM geometries optimized with the AMBER ff99SB force field. The angle C_β–$C_{carboxyl}$–Ca^{2+} was kept at $180°$.

Calcium Coordination to Aspartate

Binding of calcium ions to carboxylate groups is of importance for biological function, for example in cell signaling based on ion exchange proteins,[91] but also for technological applications such as ion exchangers based on polyelectrolytes.[92] Here we present studies of calcium coordination to the carboxyl group in acetyl and N-methyl capped aspartate as a model system for ion/protein interactions in aqueous solution. We compare results from MM simulations and QM/MM simulations that serve to demonstrate the functionality of the new interface in combination with the advanced sampling techniques that are available in AMBER.

Binding energy curve in vacuum

Figure 5 shows the potential energy profile for binding of a Ca^{2+} ion by the carboxylate group of Ace–Asp–NMe in vacuum as obtained with different MM and QM potentials at geometries optimized with the AMBER ff99SB force field. The reaction coordinate chosen is the distance between the carbon atom of the carboxylate group and the calcium ion which was restrained during the geometry optimizations. The binding curves obtained with the various methods are distinctively different. However, all QM models show an encouraging agree-

ment around the minimum of the binding curve which is found at a reaction coordinate value of approximately 2.6 Å. The MM binding curve has its minimum at a larger distance of approximately 2.9 Å, indicating that the classical vdW potential in the AMBER ff99SB force field is too repulsive. At large separation, the binding curves will be dominated by the classical 1/R behavior of the electrostatic interaction between the two ions. At intermediate distances, however, dispersion interactions are of importance. The latter are not properly accounted for in the DFT models used here which explains the discrepancy between MP2 and DFT results with increasing ion separation. Based on the potential energy profiles presented here, one would expect results for the free energy profile of this coordination process in aqueous solution that differ between MM and QM/MM models. It is reasonable to expect that similar results should be obtained with the semiempirical PM6 model and DFT, which is not the case as shown below.

Binding free energy curve in explicit water

Figure 6 shows the PMF for the reaction coordinate defined above as obtained from MM and QM/MM MD simulations in aqueous solution. This PMF is an upper bound to the PMF that would be obtained if the Ca^{2+} ion were allowed to move
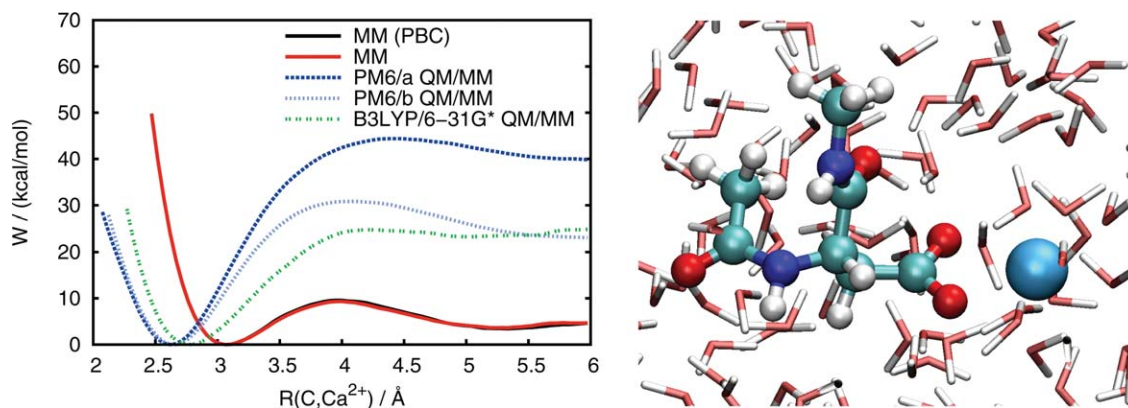

Figure 6. Free energy profile of calcium ion coordination to the carboxyl group in acetyl and *N*-methyl capped aspartate solvated with TIP3P water. Results are presented for PBCs and a water droplet. The peptide and Ca^{2+} ion are treated with the AMBER ff99SB force field in the MM simulations and quantum mechanically in the QM/MM simulations. PM6/a and PM6/b use different semiempirical QM/MM interaction potentials.

freely instead of restraining the angle C_β–$C_{carboxyl}$–Ca^{2+} to 180°. However, this additional restraint greatly aids in converging the simulations and the effect is expected to be rather small (see also below).

Results from MM simulations. MM simulations have been performed both with the peptide and ion solvated in a water box using PBCs as well as a water droplet. MM based MD simulations often use PBCs which is computationally very efficient because a cutoff is used for the real-space nonbonded interactions while long-range electrostatics beyond the cutoff are accounted for with the PME algorithm.[54] DFT based QM/MM simulations under PBCs, however, have to apply a cutoff for the electrostatic interactions between the QM and MM regions, first term on the right hand side of eq. (2). There is also no advantage in terms of computational efficiency since the QM calculation dominates the computational effort. Figure 6 clearly shows that the results obtained from the PBC and water droplet simulations are indistinguishable, thus justifying the use of a water droplet instead of PBCs for the QM/MM simulations.

In a study of calcium binding to polyacrylates[92] using classical MM force fields, the free energy gain for binding to a single carboxylate group was found to be approximately 6 kcal/mol with a barrier for detachment of approximately 11 kcal/mol. This data is in good agreement with the MM results for calcium binding to Ace-Asp-NMe presented here which are 4 kcal/mol for the free energy of binding, and 9 kcal/mol for the detachment barrier using the AMBER ff99SB force field and TIP3P water. This good agreement also justifies the use of the distance based reaction coordinate in conjunction with the angle restraint as discussed above. The minimum of the binding curve is found at approximately 3.1 Å, a slightly larger value than obtained from the geometry optimizations with restrained reaction coordinate (Fig. 5), in agreement with the intuitive picture that the presence of a polar solvent facilitates ion dissociation.

Results from QM/MM simulations. In all QM/MM simulations presented here, both the Ace–Asp–NMe peptide and the Ca^{2+} ion are treated quantum mechanically while the TIP3P water model is retained for the surrounding water droplet. This is a rather drastic approximation as it neglects all charge transfer between the ions and the solvent, however, is a useful model to compare semiempirical and DFT methods. The semiempirical simulation setup requires some additional explanation. In AMBER, the electrostatic interaction between QM cores and MM point charges is modeled with the standard MNDO[64]-type core repulsion function that uses atom-specific parameters for the exponents in its damping function. The PM6 Hamiltonian,[36] however, uses a core repulsion function with an explicit atom pair-wise parameterization and parameters for an interaction between QM cores and MM point charges are not available. Such parameters could certainly be optimized, however, for the present work we have chosen to retain the MNDO-type expression with parameters borrowed from the PM3[65] Hamiltonian, where available. As PM3 parameters are not available for calcium, we have tested two different values

for the corresponding exponent to be used in conjunction with the PM6 Hamiltonian: (a) a value of 1.3 Å^{-1} (denoted as PM6/a) which is close to the value for magnesium and (b) a value of 2.0 Å^{-1} (denoted as PM6/b) which reduces the magnitude of the interaction between the QM core and the point charges. The choice of this parameter will thus clearly have an effect on the solvation behavior of the Ca^{2+} ion.

From Figure 6, we can see that the parameterization used for the electrostatic QM/MM interaction of the semiempirical QM/MM Hamiltonian has a pronounced effect on the free energy of binding and the corresponding barrier for ion dissociation. Conversely, there is virtually no effect close to the equilibrium binding distance. We also note that the minimum for the PM6 binding curve remains approximately at the same value as obtained from the static calculations (Fig. 5). The PM6/b results lead to a much lower binding energy and barrier for ion dissociation than the PM6/a results which can be understood in terms of the discussion above: the larger exponent used in the core repulsion function for QM/MM core/point charge interactions for PM6/b leads to an improved hydration of the Ca^{2+} ion which counterbalances the energy loss upon ion dissociation. This lowers the barrier to approximately 31 kcal/mol which is much closer to the DFT results than the PM6/a results.

Unlike the semiempirical models, DFT based QM/MM MD simulations do not depend on any additional parameters for QM/MM interactions and are uniquely defined through eqs. (1) and (2) and the choice of the MM force field and the QM Hamiltonian. The PMF obtained with B3LYP/6-31G* in the QM region shows large differences from the purely classical MM result (see Fig. 6). In particular, the minimum of the binding curve is at a shorter distance of approximately 2.8 Å. This could be expected based on the results from the static calculations (Fig. 5) which also have the DFT minimum at a shorter distance than the MM minimum. Similar to the MM results, the minimum in the DFT based QM/MM free energy profile in aqueous solution is at larger distance than in gas phase which again can be rationalized in terms of the polar solvent facilitating ion dissociation. Note that this is not the case for the QM/MM calculations with the semiempirical PM6 Hamiltonian, which, compared to the DFT results, have the minimum at a distance that is too short. The free energy barrier for ion dissociation with B3LYP/6-31G* is approximately 25 kcal/mol, which is lower than the 31 kcal/mol obtained with PM6/b, but much larger than the 9 kcal/mol obtained from the MM simulations.

Remaining error sources. There are several potential sources of error for the QM/MM MD simulations of ion dissociation in solution as presented in this work. In the case of the PM6 simulations, the semiempirical Hamiltonian itself puts a strict limitation on the attainable numerical accuracy and, as shown above, distinctively different results are obtained with a more sophisticated DFT model such as B3LYP/6-31G*.

A major source of error for all QM/MM simulations is the QM/MM boundary. In the simulations presented here, the QM/MM boundary must have a pronounced effect on the PMF for ion dissociation in solution because it is situated right next to

the atoms involved in the dissociation process and crosses coordination bonds between the calcium ion and the solvent. For one, the QM/MM vdW interactions[93,94] can be expected to significantly affect the reaction free energies and barrier heights for ion association or dissociation processes. At least as important, however, is the fact that charge transfer between the ions and the solvent is neglected and that the TIP3P water model lacks polarizability, both of which can be expected to be important for the stabilization of highly ionic systems such as the ones studied here.

The effect of the QM/MM boundary can often be minimized by increasing the size of the QM region such that the QM/MM boundary is sufficiently far from the region of interest. This is possible for example for studies of reactive events in active sites of enzymes that are deeply buried within the protein and thus have an environment that does not change during the course of a simulation. For reactive events in solution such as the ion dissociation studied here, however, this is not possible with conventional QM/MM schemes that require a selection of atoms belonging to the QM and MM region at the beginning of a simulation. Alternative approaches, such as adaptive QM/MM[42,95] (A. W. Götz, K. Park, R. E. Bulo, F. Paesani, R. C. Walker, in preparation) that allow for a diffusion of solvent molecules into and out of the QM region, are thus required to include solvent surrounding the ions into the QM region to improve upon the results presented in this work.

Conclusions and Outlook

We have presented a versatile and easily extensible QM/MM interface that supports a wide range of electronic structure software packages. This interface has been integrated with the AMBER MD software package, enabling *ab initio* wave function theory and DFT based QM/MM geometry optimizations and MD simulations within both mechanical and electronic embedding schemes. The implementation supports all of AMBER's advanced sampling techniques and has been designed to be easy to use for anybody who is familiar with classical MD simulations, requiring not much more than a straight forward selection of the QM region and the QM method in addition to the classical MD simulation setup. An automated link atom setup is used for simulations in which the QM/MM boundary crosses covalent bonds.

The interface controls the required data exchange between the MD software and the electronic structure software which is implemented in two fashions: (a) traditional, file based data exchange with system calls is available for all supported QM software packages (at the time of writing ADF, GAMESS-US and NWChem for mechanical embedding and Gaussian, Orca and TeraChem for mechanical and electronic embedding) (b) a client/server model based on the MPI-2 standard that increases performance and portability but requires corresponding modifications to the electronic structure code. This interface is currently supported by TeraChem. Additional electronic structure software can use the interface defined by the API of the MPI-2 client/server model without modification of the present implementation of the interface or its integration with AMBER.

The QM/MM interface defines default parameters for the electronic structure programs that lead to good energy conservation during MD simulations in the microcanonical ensemble, which we have shown with constant energy QM/MM MD simulations of NMA and ADP in a water droplet using both DFT and MP2 in the QM region. Results for QM and QM/MM geometry optimizations of the water dimer are in good agreement with published data.

We have furthermore demonstrated geometry optimizations with restraints and QM/MM free energy calculations of a calcium ion binding to the carboxylate group of acetyl and *N*-methyl capped aspartate in aqueous solution as a model for ion/protein interactions, comparing the semiempirical PM6 Hamiltonian to DFT with the B3LYP XC functional and the 6-31G* basis set. We have shown that the PM6 results depend strongly on the semiempirical parameters chosen for the QM/MM core/point charge interactions and that the DFT based QM/MM simulation predicts an equilibrium binding distance that lies in between the PM6 and MM results. The free energy of binding and corresponding dissociation barrier obtained from the QM/MM simulations is too large as compared to MM results. The QM/MM results will need to be improved with an appropriate description of the solvent in the vicinity of the ion and the carboxylate group, for example through inclusion into the QM region via adaptive QM/MM methods.

The new interface will be useful for such investigations as well as applications toward electronic events in large biomolecular systems, for example the photophysics of chromophores embedded in proteins or enzymatic reactions.

Acknowledgments

[1] A. Warshel, M. Levitt, *J. Mol. Bio.* **1976**, *103*, 227.
[2] U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1986**, *7*, 718.
[3] M. J. Field, P. A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700.
[4] D. Bakowies, W. Thiel, *J. Phys. Chem.* **1996**, *100*, 10580.
[5] P. Sherwood, in Modern Methods and Algorithms of Quantum Chemistry Proceedings, Vol. 3 of NIC Series; J. Grotendorst, Eds.; John von Neumann Institute for Computing, Jülich, **2000**; pp. 285–305.
[6] J. Gao, D. G. Truhlar, *Annu. Rev. Phys. Chem.* **2002**, *53*, 467.
[7] R. A. Friesner, V. Guallar, *Annu. Rev. Phys. Chem.* **2005**, *56*, 389.
[8] H. Lin, D. G. Truhlar, *Theor. Chem. Acc.* **2007**, *117*, 185.
[9] W. Thiel, in Multiscale Simulation Methods in Molecular Sciences, Vol. 42 of NIC series; J. Grotendorst, N. Attig, S. Blügel, D. Marx, Eds.; Institute for Advanced Simulation, Forschungszentrum Jülich: Jülich, **2009**; pp. 203–214.

[10] H. M. Senn, W. Thiel, *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.

[11] N. Bernstein, J. R. Kermode, G. Csányi, *Rep. Prog. Phys.* **2009**, *72*, 1.

[12] R. Zhang, B. Lev, J. E. Cuervo, S. Y. Noskov, D. R. Salahub, in Advances in Quantum Chemistry, Vol. *59*, Chapter. 10; J. R. Sabin, E. Brändas, Eds.; Elsevier Academic Press Inc.: San Diego, **2010**; pp. 353–400.

[13] S. F. Sousa, P. A. Fernandes, M. J. A. Ramos, *Phys. Chem. Chem. Phys.* **2012**, *14*, 12431.

[14] 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, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, Amber 12; University of California: San Francisco, **2012**.

[15] R. Salomon-Ferrer, D. A. Case, R. C. Walker, *WIREs Comput. Mol. Sci.* **2013**, *3*, 198.

[16] R. C. Walker, M. F. Crowley, D. A. Case, *J. Comput. Chem.* **2008**, *29*, 1019.

[17] G. de M. Seabra, R. C. Walker, M. Elstner, D. A. Case, A. E. Roitberg, *J. Phys. Chem. A* **2007**, *111*, 5655.

[18] I. Rossi, D. G. Truhlar, *Chem. Phys. Lett.* **1995**, *233*, 231.

[19] K. Nam, Q. Cui, J. Gao, D. M. York, *J. Chem. Theory Comput.* **2007**, *3*, 486.

[20] D. Wei, D. Salahub, *Chem. Phys. Lett.* **1994**, *224*, 291.

[21] R. V. Stanton, L. R. Little, K. M. Merz, *J. Phys. Chem.* **1995**, *99*, 17344.

[22] R. V. Stanton, D. S. Hartsough, K. M. Merz, Jr., *J. Comput. Chem.* **1995**, *16*, 113.

[23] U. Ryde, *J. Comput. Aided Mol. Des.* **1996**, *10*, 153.

[24] P. D. Lyne, M. Hodošcek, M. Karplus, *J. Phys. Chem. A* **1999**, *103*, 3462.

[25] M. Eichinger, P. Tavan, J. Hutter, M. Parinello, *J. Chem. Phys.* **1999**, *110*, 10452.

[26] A. Laio, J. VandeVondele, U. Rothlisberger, *J. Chem. Phys.* **2002**, *116*, 69416947.

[27] M. J. Loferer, H. H. Loeffler, K. R. Liedl, *J. Comput. Chem.* **2003**, *24*, 1240.

[28] P. K. Biswas, V. Gogonea, *J. Chem. Phys.* **2005**, *123*, 164114.

[29] H. L. Woodcock, III, M. Hodošcek, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer, III, B. R. Brooks, *J. Comput. Chem.* **2007**, *28*, 1485.

[30] B. Lev, R. Zhang, A. de la Lande, D. Salahub, S. Y. Noskov, *J. Comput. Chem.* **2010**, *31*, 1015.

[31] T. Okamoto, K. Yamada, Y. Koyano, T. Asada, N. Koga, M. Nagaoka, *J. Comput. Chem.* **2011**, *32*, 932.

[32] K. Meier, N. Schmid, W. F. van Gunsteren, *J. Comput. Chem.* **2012**, *33*, 2108.

[33] J. Torras, E. Deumens, S. B. Trickey, *J. Comput.-Aid. Mater. Des.* **2006**, *13*, 201.

[34] ChemShell, a Computational Chemistry Shell, Available at: www.chem-shell.org. Last accessed date 9.17.13.

[35] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, et al. *J. Mol. Struct. (Theochem)* **2003**, *632*, 1.

[36] J. J. P. Stewart, *J. Mol. Model.* **2007**, *13*, 1173.

[37] L. Hu, P. Söderhjelm, U. Ryde, *J. Chem. Theory Comput.* **2011**, *7*, 761.

[38] I. S. Ufimtsev, T. J. Martinez, *J. Chem. Theory Comput.* **2009**, *5*, 1004.

[39] I. S. Ufimtsev, T. J. Martinez, *J. Chem. Theory Comput.* **2009**, *5*, 2619.

[40] C. Isborn, A. W. Götz, M. A. Clark, R. C. Walker, T. M. Martínez, *J. Chem. Theory Comput.* **2012**, *8*, 5092.

[41] Message Passing Interface Forum, MPI: A Message-passing Interface Standard, Version 2.2; High-Performance Computing Center Stuttgart, University of Stuttgart, Nobeliusstr. 19, 70550 Stuttgart, Germany, **2009**.

[42] K. Park, A. W. Götz, R. C. Walker, F. Paesani, *J. Chem. Theory Comput.* **2012**, *8*, 2868.

[43] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. F. Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931.

[44] C. F. Guerra, J. Snijders, G. te Velde, E. J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391.

[45] ADF2012, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, Available at: http://www.scm.com. Last accessed date 9.17.13.

[46] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, Jr., *J. Comput. Chem.* **1993**, *14*, 1347.

[47] M. S. Gordon, M. W. Schmidt, in Theory and Applications of Computational Chemistry, the first forty years, Chapter 41; C. E. Dykstra, G. Frenking, K. S. Kim, G. E. Scuseria, Eds.; Elsevier: Asterdam, **2005**; pp. 1167–1189.

[48] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, W. A. de Jong, *Comput Phys. Commun.* **2010**, *181*, 1477.

[49] 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, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, Hasegawa J., M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, Vreven T., J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, Normand J., K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, Tomasi J., M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G.A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision C.01; Gaussian, Inc.: Wallingford CT, **2010**.

[50] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73.

[51] Y. Shao, L. Fusti-Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, Jr., R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Daschel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C.-P. Hsu, G. Kedziora, R. Z. Khaliullin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, III, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, III, J. Kong, A. I. Krylov, P. M. W. Gill, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.

[52] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127.

[53] E. Pellegrini, M. J. Field, *J. Phys. Chem. A* **2002**, *106*, 1316.

[54] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10090.

[55] K. Nam, J. Gao, D. M. York, *J. Chem. Theory Comput.* **2005**, *1*, 2.

[56] B. Roux, *Comput. Phys. Commun.* **1995**, *91*, 275.

[57] J. Kästner, *WIREs Comput. Mol. Sci.* **2011**, *1*, 932.

[58] A. Mitsutake, Y. Sugita, Y. Okamoto, *Biopolymers* **2001**, *60*, 96.

[59] H. Nymeyer, S. Gnanakaran, A. E. Gárcia, *Methods Enzymol.* **2004**, *383*, 119.

[60] D. Ceperley, *Rev. Mod. Phys.* **1995**, *67*, 279.

[61] B. J. Berne, D. Thirumalai, *Annu. Rev. Phys. Chem* **1986**, *37*, 401.

[62] P. Pulay, G. Fogarasi, *Chem. Phys. Lett.* **2004**, *386*, 272.

[63] J. M. Herbert, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3269.

[64] M. Dewar, W. Thiel, *J. Am. Chem. Soc.* **1977**, *99*, 4899.

[65] J. J. P. Stewart, *J. Comput. Chem.* **1989**, *10*, 209.

[66] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.

[67] J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822.

[68] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[69] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.

[70] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257.

[71] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213.

[72] R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 650.

[73] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829.

[74] T. H. Dunning, Jr., *J. Chem. Phys.* **1989**, *90*, 1007.

[75] R. A. Kendall, T. H. Dunning, Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.

[76] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.

[77] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, Interaction Models for Water in Relation to Protein Hydration; Reidel, Dordrecht, The Netherlands, **1981**; pp. 331–342.

[78] Y. Wu, H. L. Tepper, G. A. Voth, *J. Chem. Phys.* **2006**, *124*, 024503.

[79] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins* **2006**, *65*, 712–725.

[80] G. S. Grest, K. Kremer, *Phys. Rev. A* **1986**, *33*, 3628.

[81] J. -P. Ryckaert, G. Ciccotti, H. J. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.

[82] S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*, 952.

[83] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.

[84] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, J. M. Rosenberg, *J. Comput. Chem.* **1992**, *13*, 1011.

[85] A. Grossfield, "WHAM: the Weighted Histogram Analysis Method", version 2.0.6, Available at: http://membrane.urmc.rochester.edu/content/wham. Last accessed date 9.17.13.

[86] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33.

[87] T. Williams, C. Kelley. Gnuplot: An Interactive Plotting Program; Available at: http://www.gnuplot.info. Last accessed date 9.17.13.

[88] T. Dyke, J. Muenter, *J. Chem. Phys.* **1974**, *60*, 2929.

[89] L. Curtiss, D. Frurip, M. Blander, *J. Chem. Phys.* **1979**, *60439*, 2703.

[90] K. S. Kim, B. J. Mhin, U. -S. Choi, K. Lee, *J. Chem. Phys.* **1992**, *97*, 6649.

[91] J. Liao, H. Li, W. Zeng, D. B. Sauer, R. Belmares, Y. Jiang, *Science* **2012**, *335*, 686.

[92] R. E. Bulo, D. Donadio, A. Laio, F. Molnar, J. Rieger, M. Parrinello, *Macromolecules* **2007**, *40*, 3437.

[93] Y. Tu, A. Laaksonen, *J. Chem. Phys.* **1999**, *111*:, 7519–7525.

[94] D. Riccardi, G. Li, Q. Cui, *J. Phys. Chem. B* **2004**, *108*, 6467.

[95] R. E. Bulo, B. Ensing, J. Sikkema, L. Visscher, *J. Chem. Theory Comput.* **2009**, *5*, 2212.