PCCP

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 10503–10509

www.rsc.org/pccp

PAPER

Efficient solvent boundary potential for hybrid potential simulations†

Alexey Aleksandrov and Martin Field*

Received 9th December 2010, Accepted 11th February 2011
DOI: 10.1039/c0cp02828b

A common challenge in computational biophysics is to obtain statistical properties similar to those of an infinite bulk system from simulations of a system of finite size. In this work we describe a computationally efficient algorithm for performing hybrid quantum chemical/molecular mechanical (QC/MM) calculations with a solvent boundary potential. The system is partitioned into a QC region within which catalytic reactions occur, a spherical region with explicit solvent that envelops the quantum region and is treated with a MM model, and the surrounding bulk solvent that is treated implicitly by the boundary potential. The latter is constructed to reproduce the solvation free energy of a finite number of atoms embedded inside a low-dielectric sphere with variable radius, and takes into account electrostatic and van der Waals interactions between the implicit solvent and the QC and MM atoms in the central region. The method was implemented in the simulation program pDynamo and tested by examining elementary steps in the reaction mechanisms of two enzymes, citrate synthase and lactate dehydrogenase. Good agreement is found for the energies and geometries of the species along the reaction profiles calculated with the method and those obtained by previous experimental and computational studies. Directions in which the utility of the method can be further improved are discussed.

1. Introduction

Aqueous solvation is a near universal phenomenon in all known biological systems and its effects are pervasive. Thus, for example, it is essential for conformational stability and in folding and molecular recognition processes[1,2] and both solvation and desolvation play a central role in enzyme function.[3] A common approach for describing solvation in calculations is to consider explicitly a large number of solvent molecules around the solute species and to impose periodic boundary conditions. Such simulations can be very expensive because the number of solvent atoms can exceed by an order of magnitude the number of solute atoms.

Due to these limitations, a huge variety of simplified treatments have been developed, which represent some or all solvent degrees of freedom implicitly and thereby considerably reduce the computational expense. In the biomolecular domain, many successful studies employing approximations based on continuum electrostatics[4–6] have been reported that reproduce the electrostatic effects of solvation on molecular recognition. However, such methodologies are inappropriate for studies in which an explicit description of some solvent molecules is necessary due, for example, to hydrogen bonding or the participation of a solvent molecule in a chemical reaction.

To resolve these issues a different method was proposed by Beglov and Roux,[7] which represented a compromise between purely explicit and implicit solvent models. In their model a number of water molecules in the region of interest were treated explicitly, whereas the remaining infinite bulk was represented by a spherical solvent boundary potential (SSBP).

The SSBP was developed by considering a thermodynamical cycle in which: (i) a hard-sphere of a given radius is inserted in bulk solvent; (ii) neutral solute and solvent molecules are moved into the hard-sphere cavity; and (iii) the charges of the solute and solvent molecules within the cavity are switched on adiabatically. The latter term, which is the charging free energy of the solute, was determined using the continuum electrostatics formulae for a sphere derived by Kirkwood.[7,8]

The original SSBP was implemented for molecular mechanical (MM) potentials only. The goal of the present work was to extend the formalism so that it can be employed to study systems that are treated with quantum chemical (QC) methods, either exclusively or in combination with MM potentials. In this article we present the details of the method and discuss its implementation in the pDynamo simulation program.[9,10] We then illustrate the method's utility by calculations of reaction pathways in the mechanisms of two enzymes, citrate synthase

Laboratoire de Dynamique Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel (CEA, CNRS UMR5075, Université Joseph Fourier-Grenoble I), 41 rue Jules Horowitz, 38027 Grenoble, France. E-mail: martin.field@ibs.fr

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c0cp02828b

and lactate dehydrogenase. The former is an important enzyme that catalyzes the first step of the Krebs Cycle,[11] whereas the latter performs a central function in anaerobic glycolysis.[12] The reaction paths themselves were calculated using a version of the nudged elastic band (NEB) method, which was recently implemented in the pDynamo software.[13,14] The paper concludes by suggesting ways in which the method could be developed further.

2. Methods

2.1 Molecular dynamics simulations

The crystal structure of the citrate synthase (CS) homodimer was obtained from the Protein Data Bank (PDB), entry 4CSC,[15] with bound D-malate and acetyl coenzyme A (COA). The oxaloacetate (OAA) was built from the D-malate ligand by making the O2 oxygen planar to the C1 carboxylic group. The coordinates of four missing residues 82, 292, 293, and 294 were taken from another crystal structure of CS that is complexed with OAA and the inhibitor, carboxymethyldethia coenzyme (PDB entry code 1CSI).[16] The simulations included protein residues within a 24 Å sphere centered on the oxaloacetate binding site in one of the two monomers. In addition to crystal waters, a 24 Å sphere of water was overlaid, and waters that overlapped protein, crystal waters, or ligands were removed. The final model contained 332 protein residues.

For lactate dehydrogenase (LDH) we started from the crystal structure of *Thermus thermophilus* LDH with bound NADH and oxamate, resolved at 2.1 Å resolution, PDB entry code 2V7P.[17] The nitrogen of oxamate was converted into a sp^3-hybridized carbon to create the pyruvate ligand. We considered protein residues within a 24 Å sphere centered on the pyruvate ligand of one monomer of the LDH tetramer. In addition to crystal waters, a 24 Å sphere of water was overlaid, and waters that overlapped protein, crystal waters, or ligands were removed. The final model contained 279 protein residues.

For both systems hydrogens were constructed with ideal stereochemistry. Protonation states of histidines were assigned by visual inspection and then pK_a's were computed with the PROPKA program[18] to verify the assignment of protonation states. Orientations of His, Asn, and Gln side chains in the active site were taken from the crystal structure and verified by visual inspection.

All simulations were done with the SSBP solvent model,[7,19] which treats the region outside the 24 Å sphere as a uniform dielectric continuum, with a dielectric constant of 80. This is reasonable, since most of the outer region is water. Newtonian dynamics was used for the innermost region, within 20 Å of the sphere's center and Langevin dynamics for the outer part of the sphere, with a thermal bath at 292 K. Protein atoms between 20 Å and 24 Å from the sphere's center were harmonically restrained to their experimental positions. The CHARMM27 force field was used for the protein[20,21] and the TIP3P model for water.[22] Acetyl coenzyme A and oxaloacetate were specifically parametrized (details of the parametrization and force field parameters are given in the ESI†). Electrostatic interactions were computed without any cutoff, using a multipole approximation for distant groups.[23]

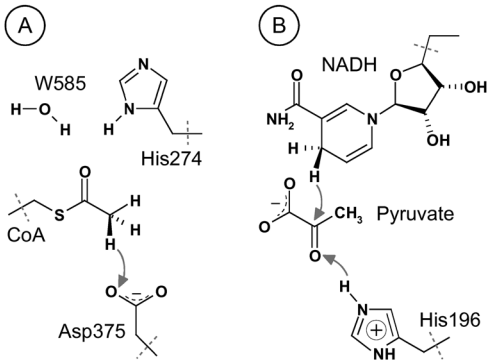

Fig. 1 The QC regions used in calculations on (A) the CS system and (B) the LDH system.

Calculations were done with the CHARMM program.[24] The complexes were equilibrated for 2 ns following 200 ps of thermalization.

2.2 Hybrid potential simulations

For the QC/MM calculations, the citrate synthase system was partitioned between QC and MM regions as depicted in Fig. 1. The QC region had 30 atoms, and the MM region contained 6970 atoms. For lactate dehydrogenase, the system was partitioned between QC and MM regions as depicted in Fig. 1. The dihydronicotinamide group, the ribose ring of NADH, pyruvate, and the sidechain of H196 were included in the QC region, which had 50 atoms, and the MM region contained 7110 atoms.

For both systems the CHARMM27 force field was employed for the MM region whereas the AM1 method was used for the QC region. All geometry optimizations and reaction path calculations were performed with this semiempirical QC/MM potential. Calculations were done with SSBP switched on and off. Single point calculations on the optimized structures were performed with a density functional theory (DFT) method using the BLYP functional and a 6-31G basis with polarization functions on all atoms except hydrogen.

The starting coordinates of the complexes were taken after 2 ns of MD simulation described above. The QC/MM calculations were performed with the pDynamo software.[10] No truncation or cut-off was employed to calculate non-bonded interactions. The reaction path was optimized with the nudged elastic band (NEB) method[13,14,25] implemented in pDynamo. The utility of NEB is that it does not require a predefined set of reaction coordinate variables and makes no assumptions as to how the reaction proceeds. The only bias in the calculation comes from the structures employed in the starting guess for the NEB pathway.

2.3 Solvent boundary potential

The implementation of the SSBP method in pDynamo follows closely that given in ref. 7 and 26. The method is derived by formally separating the multidimensional solute–solvent configurational integral into terms corresponding to a group of inner solvent molecules, nearest to the solute, and the remaining outer bulk solvent molecules. Roux and co-workers then showed that equilibrium statistical properties representative of

the infinite bulk system can be obtained from the finite system by considering a thermodynamical cycle in which: (i) a hard-sphere of a given radius is inserted in bulk solvent; (ii) neutral solute and solvent molecules are moved into the hard-sphere cavity; and (iii) the charges of the solute and solvent molecules within the cavity are switched on adiabatically. The energies corresponding to these operations are denoted $\mathscr{W}_{\mathrm{hsr}}$, $\mathscr{W}_{\mathrm{cav}}$ and $\mathscr{W}_{\mathrm{Kirkwood}}$, respectively. In addition it was found necessary to add two additional terms, $\mathscr{W}_{\mathrm{ang}}$ and $\mathscr{W}_{\mathrm{emp}}$, that improve the simulated properties of water molecules at the boundaries of the reduced SSBP system, and when the cavity is very small. The full SSBP energy, $\mathscr{W}$, is thus:

$$\mathscr{W} = \mathscr{W}_{\mathrm{ang}} + \mathscr{W}_{\mathrm{cav}} + \mathscr{W}_{\mathrm{emp}} + \mathscr{W}_{\mathrm{hsr}} + \mathscr{W}_{\mathrm{Kirkwood}} \quad (1)$$

Of the five terms in eqn (1), the first four are identical in the MM and QC/MM cases as they are independent of how the charge distribution or electronic state of the atoms are treated. This leaves the electrostatic-charging term, $\mathscr{W}_{\mathrm{Kirkwood}}$, which Beglov and Roux[7] approximated using the expression derived by Kirkwood for the energy of a set of point charges within a sphere immersed in a medium of a given dielectric constant.[8] The expression they used was:

$$\mathscr{W}_{\mathrm{Kirkwood}} = -\frac{1}{2}\sum_{lm}\frac{4\pi|Q_{lm}|^2}{2l+1}\frac{1}{R^{2l+1}}\left\{\frac{\varepsilon_{\mathrm{bulk}}-1}{\varepsilon_{\mathrm{bulk}}+l/(l+1)}\right\} \quad (2)$$

where the moments of the charge distribution, Q_{lm}, are defined as:

$$Q_{lm} = \sum_i q_i Y^{\dagger}_{lm}(\theta_i,\phi_i) r_i^l \quad (3)$$

In these equations, R is the radius of the hard sphere, $\varepsilon_{\mathrm{bulk}}$ is the dielectric constant of the bulk, Y_{lm} is a complex spherical harmonic, q_i is the partial charge of atom i and r_i, θ_i and ϕ_i are an atom's spherical polar coordinates relative to the centre of the sphere.

Eqn (2) is quadratic in the charges of the atoms within the sphere. To adapt it to QC/MM potentials, we separate out the terms that depend upon the charges of the QC atoms. The resulting expression is:

$$\mathscr{W}_{\mathrm{Kirkwood}} = \mathscr{W}^{\mathrm{MM/MM}}_{\mathrm{Kirkwood}} + \mathscr{W}^{\mathrm{QC/MM}}_{\mathrm{Kirkwood}} + \mathscr{W}^{\mathrm{QC/QC}}_{\mathrm{Kirkwood}} \quad (4)$$

The first term is the same as eqn (2) except that the sums in eqn (3) go over MM atoms only. The second and third terms are linear and quadratic in the QC atom charges, respectively, and take the forms:

$$\mathscr{W}^{\mathrm{QC/MM}}_{\mathrm{Kirkwood}} = \sum_u q_u \Phi_u \quad (5)$$

$$\mathscr{W}^{\mathrm{QC/QC}}_{\mathrm{Kirkwood}} = \sum_{uv} q_u q_v \Omega_{uv} \quad (6)$$

where the subscripts u and v refer to QC atoms only. The functions Φ_u and Ω_{uv} follow straightforwardly from eqn (1) and (3). Their forms will not be given here, but Φ_u depends upon the coordinates and charges of all the MM atoms as well as the coordinates of atom u, whereas Ω_{uv} depends upon the coordinates of atoms u and v only.

The QC charges q_u are variable and change during the self-consistent field procedure (SCF) that is used to determine the QC wavefunction at each energy calculation. To ensure that this wavefunction is consistent with the SSBP approximation, terms derived from eqn (4) must be included in the SCF equations. Supposing that the element of the Fock or Kohn–Sham matrix for the QC basis functions, μ and ν, is denoted $F_{\mu\nu}$ and that $P_{\mu\nu}$ is the corresponding element of the one-particle density matrix, one has to evaluate additional terms at each SCF iteration of the form:

$$\begin{aligned} F^{\mathrm{SSBP}}_{\mu\nu} &= \sum_u \frac{\partial q_u}{\partial P_{\mu\nu}}\frac{\partial \mathscr{W}_{\mathrm{Kirkwood}}}{\partial q_u} \\ &= \sum_u \frac{\partial q_u}{\partial P_{\mu\nu}}\left\{\Phi_u + 2\sum_v \Omega_{uv} q_v\right\} \end{aligned} \quad (7)$$

The cost of this extra step is negligible during an SCF calculation as the functions Φ_u and Ω_{uv} are precalculated at the beginning of the energy calculation along with the terms required for the energy $\mathscr{W}^{\mathrm{MM/MM}}_{\mathrm{Kirkwood}}$. In addition, our experience has shown that they have no significant adverse effect on the convergence rate of solution of the SCF equations.

The derivatives of $\mathscr{W}_{\mathrm{Kirkwood}}$ with respect to atomic positions are obtained by differentiation of eqn (4) once the converged QC energy and wavefunction have been obtained. When derivatives are required and QC atoms are present, it is more efficient to delay calculation of the derivatives of $\mathscr{W}^{\mathrm{MM/MM}}_{\mathrm{Kirkwood}}$ as well and do both QC and MM derivatives together. When there are no QC atoms the MM/MM energy and gradient can be done simultaneously.

It is to be noted that the formalism described above is also appropriate for cases in which the SSBP system consists solely of QC atoms and no MM atoms occur. One simply omits the MM terms from eqn (4). Likewise, the approach is completely general and independent of the QC method that is employed as long as there is a suitable way of writing the QC atomic partial charges, q_u, in terms of the density matrix elements, $P_{\mu\nu}$. In pDynamo, we use Mulliken charges for the MNDO-type semi-empirical QC methods that are implemented, whereas for DFT QC methods a number of alternative approaches are available, including Mulliken, Löwdin and density-fitted charges.[9,10]

As a final point, we remark that we have kept the point charge approximation for the QC atoms that is implicit in the treatment of the electrostatic energy by the Kirkwood formula. It would be possible to go beyond this by, for example, using a higher-order multipole representation of the charge distribution,[27] but whether this would significantly enhance accuracy is open to question as point-charge approximations are commonly employed in other QC and QC/MM implicit solvation models both at the *ab initio* and semi-empirical levels (see, for example, ref. 28 and 29).

3. Results

3.1 Reaction mechanism in citrate synthase

Citrate synthase (CS) catalyzes the formation of citrate from acetyl-coenzyme A (CoA) and oxaloacetate during the first

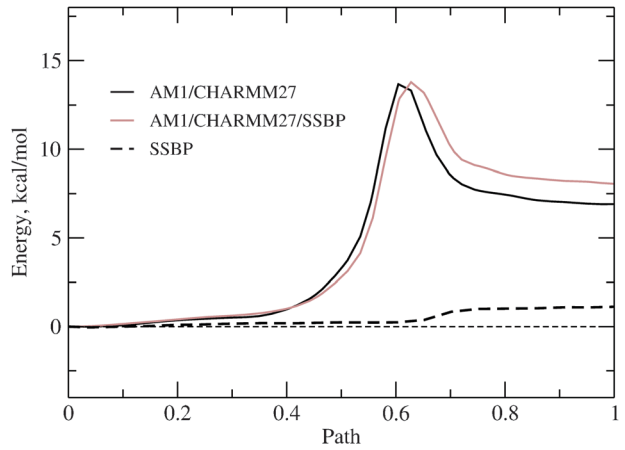

Fig. 2 Optimized nudged elastic band pathways for the enolization step of the reaction catalyzed by citrate synthase. The solid gray and black lines show the energy profile computed with and without the solvent boundary potential (SSBP), respectively. The SSBP contribution is shown as a dashed line.

step of the Krebs Cycle. In eukaryotes and gram-positive bacteria CS occurs as a homodimer. The enzyme has been extensively studied, both experimentally and theoretically, and it is well established that the reaction of acetyl-CoA with oxaloacetate is a Claisen-type condensation.[11] The condensation reaction starts with deprotonation of acetyl-CoA by a base of the enzyme to form the enolate form of the substrate, thereby creating the nucleophile that attacks oxaloacetate. This reaction step is of particular interest as it is apparently rate-limiting for the complete CS reaction.[16,30] Enolate formation has been well studied in a number of previous works[30–33] and these indicate that Asp-375 (the numbering corresponds to the chicken CS) acts as an acid for the reaction. It has also been shown that the charged enolate intermediate rather than the neutral enol is formed during the reaction.

Table 1 Energies for the reactant, product, and TS state along the NEB profile for the CS reaction computed with and without SSBP

	Potential energy	
State	Without SSBP	With SSBP
Reactant	0.00	0.00
TS	13.68	13.80
Product	6.91	8.06

Energies in kcal mol^{-1}. TS is the transition state for the enolization step of the reaction catalyzed by citrate synthase. The reactant state is chosen as a reference.

In the present work, we examined the reaction pathway for formation of enolate in CS by NEB calculations. The aim is to demonstrate the utility of the solvent boundary potential and to elucidate the influence of the infinite bulk solvent beyond the explicit 24 Å sphere on the reaction.

Details of calculations are given in the Methods section. It was found in a previous study[31] that hydrogen bonds from Wat585 and His-274 provide specific stabilization of the enolate, and thus, play an important role in the reaction. For this reason both the His-274 side-chain, up to and including Cβ, and Wat-585 were included in the QC region.

Fig. 2 shows the variation of the potential energy of the system along the reaction pathway optimized using the NEB method. The energy profiles for the explicit 24 Å spherical system were computed with and without the spherical boundary potential. The barrier to the reaction was computed to be 13.68 kcal mol^{-1} without SSBP, and 13.80 kcal mol^{-1} with SSBP (Table 1). The activation energies are very similar in both cases, which indicates that in the case of the CS enzyme the bulk solvent outside 24 Å does not significantly contribute to the observed activation barrier. Both values are comparable to the previous theoretical studies[30,34,35] and to the apparent, experimental activation energy of 14.7 kcal mol^{-1} for the overall reaction.[36]

The transition state (TS) computed by NEB is shown in Fig. 3. It generally agrees well with the TS structures computed in other studies.[30,34,35] The NEB pathways optimized with and without the solvent boundary potential are very similar, with root mean square deviation (RMSD) between structures at the same distance along pathways of less than 0.1 Å. Thus, in the case of CS, SSBP neither changes the reaction pathway nor the mechanism of the reaction. Contribution of the solvent boundary potential to the total potential energy is shown in Fig. 2 by the dashed line. Overall SSBP increases the energy of the product structure by 1.12 kcal mol^{-1} relative to the reactant state, whereas its influence on the energy of TS is just 0.12 kcal mol^{-1}. Fig. 2 shows that the increase in SSBP energy occurs after formation of the TS, which explains the similar energies for the TS state computed with and without SSBP. Although the barrier for the forward reaction is very similar with and without SSBP, the barrier for the backward reaction is decreased by 1.03 kcal mol^{-1} with SSBP.

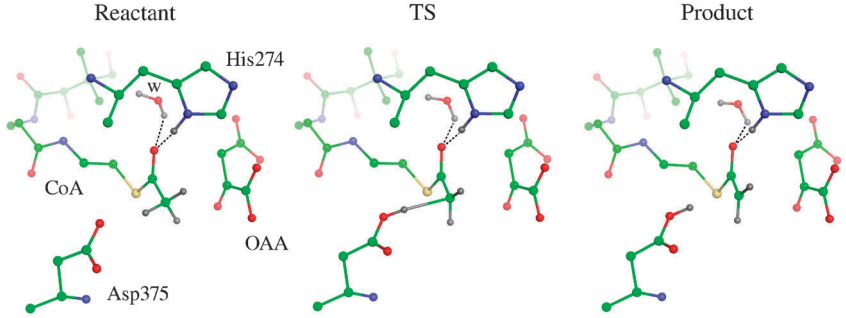

Fig. 3 Enolate formation step in citrate synthase. Three structures are shown: the reactant, product, and transition state (TS) computed by the NEB method.

The NEB method does not assume any predefined set of reaction coordinates. Thus His-274 was allowed to directly participate in the reaction as its side-chain was included in the QC region. The distance between atoms Hε-His274 and O-CoA gradually changes along the NEB pathway from 2.0 Å in the reactant state to 1.9 Å in the product state. Thus the enol never forms during the reaction computed by NEB. This supports the result of previous work,[31] which demonstrated that His-274 and CoA create a normal hydrogen bond rather than a low-barrier hydrogen bond (LBHB).

3.2 Reaction mechanism in lactate dehydrogenase

L-Lactate dehydrogenase (LDH) is present in a wide variety of organisms. It catalyses the reversible interconversion of pyruvate to lactate, and the simultaneous transformation of the nicotinamide adenine dinucleotide cofactor from the NADH state to its NAD^+ form. Although the rate-limiting step in the overall reaction corresponds to conformational changes upon substrate binding,[37,38] the chemical step of LDH has been extensively studied by both experimental[39,40] and theoretical approaches.[41–45] The chemical step involves a proton transfer from a diprotonated histidine in the enzyme catalytic center to the carbonyl oxygen of pyruvate, and a hydride transfer from the NADH cofactor to the carbonyl group of pyruvate. While a stepwise mechanism is more likely than a concerted one, results from different computational studies differ in the order in which the hydride transfer and proton transfer steps occur.[42,43,46] Indeed, one study has highlighted how sensitive the mechanism that is found is to the details of the simulation model that is employed.[45]

In the present work, we examined the reaction pathway of LDH by NEB calculations and compared the results to those of previous semiempirical studies.[43,46] The reaction in LDH has also been studied recently by the transition path sampling method,[42,47] which demonstrated that the synchronous motions of enzyme residues in the catalytic center are an integral component of the complex reaction coordinate.

The end-point structures of the reaction as well as the transition state (TS) computed by NEB are shown in Fig. 6. The latter generally agrees well with the TS structures computed in other studies.[43,46] The NEB pathways optimized with and without the solvent boundary potential are very similar, with root mean square deviation (RMSD) between structures at the same distance along the pathway of less than 0.1 Å.

The structural parameters along the NEB pathway indicate that the hydride-ion transfer precedes proton transfer, with formation of deprotonated lactate as a reaction intermediate. This result agrees well with the other semi-empirical studies[43,46] as well as with results obtained with higher levels of theory.[41]

The computed energy profiles using the AM1 semi-empirical method are shown in Fig. 4. Overall the hydride transfer step is the rate-limiting step in the chemical reaction, in agreement with other studies.[43,44] The AM1 activation barrier for the forward reaction is 22.75 kcal mol^{-1} computed without SSBP and 22.98 kcal mol^{-1} including SSBP. The energy barrier for the proton transfer from His-196 is very small, less than 0.5 kcal mol^{-1} with or without the solvent boundary potential.

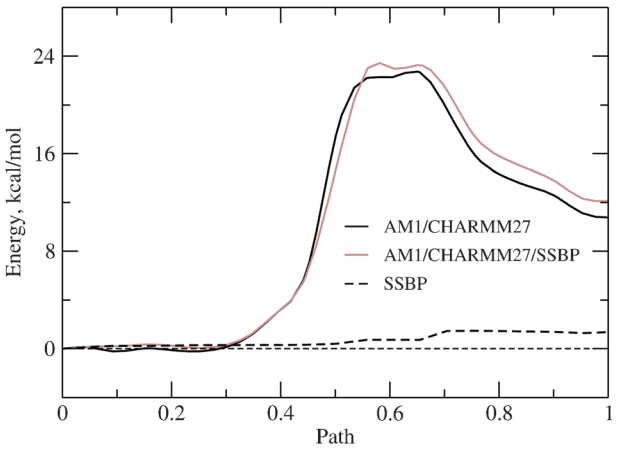

Fig. 4 Optimized nudged elastic band pathways for the mechanism of the reaction catalyzed by lactate dehydrogenase. The solid gray and black lines show the energy profile computed with and without the solvent boundary potential (SSBP), respectively. The SSBP contribution is shown as a dashed line.

Table 2 Energies for the reactant, product, and TS state along the NEB profile for the LDH reaction computed with and without SSBP

	Potential energy	
State	Without SSBP	With SSBP
Reactant	0.00	0.00
TS/HT	22.30	23.45
IS	22.28	22.98
TS/PT	22.75	22.98
Product	10.77	12.11

Energies in kcal mol^{-1}. TS/HT, TS/PT are the transition states for the hydride-ion transfer step, and for the subsequent proton transfer, respectively. IS corresponds to the intermediate state with deprotonated lactate. The reactant state is chosen as a reference.

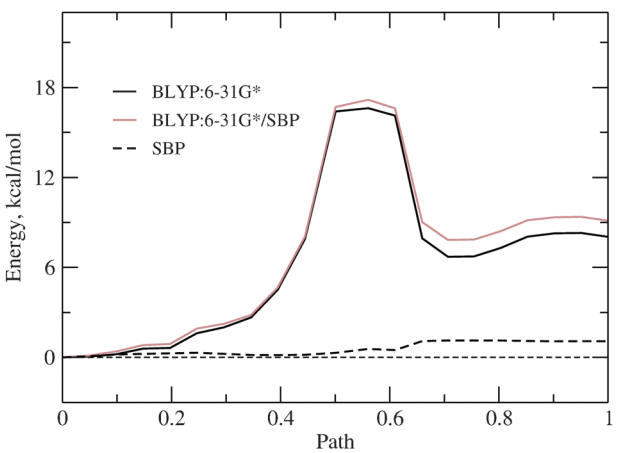

Fig. 5 Single-point calculations with a BLYP/6-31G* DFT/MM hybrid potential along the AM1/MM NEB-optimized pathway for the lactate dehydrogenase reaction. The solid gray and black lines show the energy profile computed with and without the solvent boundary potential (SSBP), respectively. The SSBP contribution is shown as a dashed line.

The reactant and product states are separated by 10.77 kcal mol^{-1} as computed without SSBP, and 12.11 kcal mol^{-1} as computed with SSBP (Table 2). Again this result is in

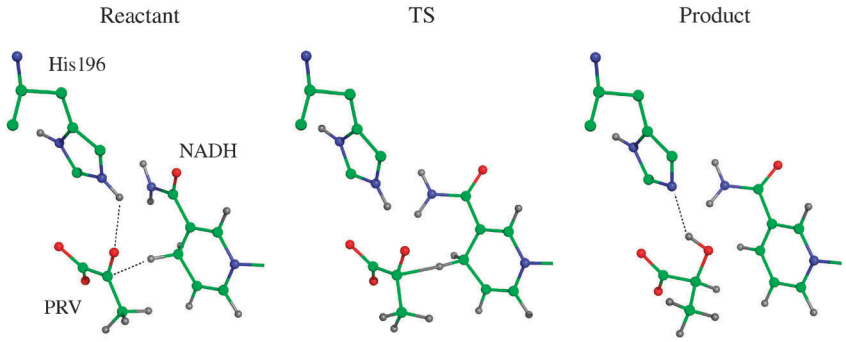

Fig. 6 Reaction catalyzed by lactate dehydrogenase. Three structures are shown: the reactant, product, and transition state (TS) for the hydride transfer step computed by NEB.

accord with the previous study, in which this energy difference was estimated to be 12.5 kcal mol^{-1},[43] and also with the observed reversibility of the reaction.

The SSBP contribution to the potential energy with the AM1 potential is shown by a dashed line in Fig. 4. The solvent boundary potential increases the energy of the product complex by 1.34 kcal mol^{-1} with respect to the reactant state. Also the TS for the hydride transfer is 1.15 kcal mol^{-1} higher with SSBP than without, while the TS for the proton transfer is 0.23 kcal mol^{-1} higher for the pathway computed with SSBP. The intermediate state, with deprotonated lactate is shifted by 0.7 kcal mol^{-1}. Overall, SSBP makes the forward reaction less likely since its contribution never becomes negative relative to the reactant state.

We performed single point calculations at the AM1 optimized structures using a DFT QC/MM method employing the BLYP functional and the 6-31G* basis set to test the SSBP dependence on the QC level of description. The energy profiles are shown in Fig. 5. The activation energies with and without SSBP were computed to be 17.18 kcal mol^{-1} and 16.62 kcal mol^{-1}, respectively. These results differ substantially from the AM1 results, and appear to be closer to the experimental estimate of 15 kcal mol^{-1} for the upper bound to the energy barrier.[48] This difference between semi-empirical and DFT results has also been observed in a previous study.[46] A step-wise mechanism was predicted with a semi-empirical potential and a concerted reaction with a higher level of theory. Although the absolute energies differ, the SSBP contribution is not very different in both cases. At the BLYP:6-31G* level of theory, SSBP increases the product state energy by 1.07 kcal mol^{-1} which is close to the 1.34 kcal mol^{-1} seen with the AM1 potential. Thus, in this test the SSBP contribution to the reaction energy profile appears to be relatively independent of the QC potential.

4. Conclusions

In this paper, we have described a simulation protocol for hybrid quantum chemical/molecular mechanical potential calculations with a solvent boundary potential. The overall implementation of the SSBP method in pDynamo follows closely that of Roux and co-workers,[7,26] although the protocol has been extended so that systems treated with a quantum chemical or QC/MM potential can be handled. Our experience has shown that SSBP has no significant adverse effect on the convergence rate of solution of the QC SCF equations and also does not increase significantly the time required for the QC energy calculation. The derivatives of the SSBP with respect to atomic positions are straightforwardly obtained, and can be used directly in minimization or molecular dynamics simulations.

Simulation tests of the protocol for the reaction mechanisms in two enzymes show reasonable agreement with experimental values, and are consistent with the results of previous computational studies. In both cases SSBP does not change the mechanism of the reaction, although the energetics are slightly different. For the two enzymes studied in this work, incorporation of SSBP leaves the energy of the TS almost unchanged with respect to the reactants but increases the energy of the products. Overall the effect of SSBP on the reaction is smaller or comparable to other uncertainties in the simulation models, such as those arising from the choice of QC region and the level of QC theory.[32] Nevertheless, it is important to have a physically rigorous scheme available to treat long-range effects, thereby eliminating one source of uncertainty in the model. We note that both our tests were performed on relatively large systems, and that the charge rearrangements involved in the studied reactions occur quite locally in the center of the explicit spheres. In other cases, the effects could be more profound, such as, for example, in smaller systems or in reactions involving larger spatial charge rearrangements.

Another method, related to SSBP, has also been proposed by Roux and co-workers[49] that allows the effects of solute atoms outside of the central explicit region to be taken into account. It employed a modified finite-difference Poisson–Boltzmann solver for the calculation of the electrostatic reaction field energies. This method, referred to as the generalized solvent boundary potential (GSBP), was recently extended for use with QC/MM potentials by Cui and co-workers[50] and Benighaus and Thiel.[51] The protocol presented in the current study complements these approaches.

The SSBP approach is quite general and permits further development. These include: (i) adaptations to other geometries, such as planes that can be used for simulations of membrane systems; (ii) extensions that treat explicitly atom charges outside the low-dielectric sphere; and (iii) inclusion of non-zero ionic strength terms[8] that could be advantageous for simulating highly charged complexes such as the ribosome.[52,53]

In summary, we expect that the current algorithm and its implementation in the pDynamo software will be of general use for simulations with hybrid quantum chemical/molecular mechanics potential and explicit solvent in an efficient and accurate way.

Acknowledgements

We thank T. Simonson and B. Roux for useful discussions and the French National Research Agency (ANR-08-COSI-015-04) for partial support.

References

1 H. Böhm and G. Schneider, *Protein–Ligand Interactions: From Molecular Recognition to Drug Design*, Wiley-VCH, New York, 2003.
2 C. Tanford, *The Hydrophobic Effect*, John Wiley and Sons, New York, 1980.
3 A. Fersht, *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, Freeman, New York, 1999.
4 T. Simonson, *Curr. Opin. Struct. Biol.*, 2001, **11**, 243–252.
5 A. Aleksandrov, D. Thompson and T. Simonson, *J. Mol. Recognit.*, 2009, **23**, 117–127.
6 P. Kukic and J. E. Nielsen, *Future Med. Chem.*, 2010, **2**, 647–666.
7 D. Beglov and B. Roux, *J. Chem. Phys.*, 1994, **100**, 9050–9063.
8 J. G. Kirkwood, *Chem. Rev.*, 1936, **19**, 275–307.
9 M. J. Field, *A Practical Introduction to the Simulation of Molecular Systems*, Cambridge University Press, 2007.
10 M. J. Field, *J. Chem. Theory Comput.*, 2008, **4**, 1151–1161.
11 G. Wiegand and S. Remington, *Annu. Rev. Biophys. Biophys. Chem.*, 1986, **15**, 97–117.
12 D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, W. H. Freeman, 4th edn, 2004.
13 R. Crehuet and M. J. Field, *J. Chem. Phys.*, 2003, **118**, 9563–9571.
14 I. F. Galvan and M. J. Field, *J. Comput. Chem.*, 2008, **29**, 139–143.
15 M. Karpusas, D. Holland and S. Remington, *Biochemistry*, 1991, **30**, 6024–6031.
16 K. Usher, S. Remington, D. Martin and D. Drueckhammer, *Biochemistry*, 1994, **33**, 7753–7759.
17 N. Coquelle, E. Fioravanti, M. Weik, F. Vellieux and D. Madern, *J. Mol. Biol.*, 2007, **374**, 547–562.
18 H. Li, A. D. Robertson and J. H. Jensen, *Proteins: Struct., Funct., Bioinf.*, 2005, **61**, 704–721.
19 T. Simonson, *J. Phys. Chem. B*, 2000, **104**, 6509–6513.
20 A. Mackerell, D. Bashford, M. Bellott, R. Dunbrack, J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, B. Roux, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
21 A. D. Mackerell, M. Feig and C. L. Brooks, *J. Comput. Chem.*, 2004, **25**, 1400–1415.
22 W. Jorgensen, J. Chandrasekar, J. Madura, R. Impey and M. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
23 R. Stote, D. States and M. Karplus, *J. Chim. Phys. Phys.-Chim. Biol.*, 1991, **88**, 2419–2433.
24 B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545–1614.
25 R. Crehuet, A. Thomas and M. J. Field, *J. Mol. Graphics Modell.*, 2005, **24**, 102–110.
26 Y. Deng and B. Roux, *J. Phys. Chem. B*, 2004, **108**, 16567–16576.
27 M. J. Schnieders and J. W. Ponder, *J. Chem. Theory Comput.*, 2007, **3**, 2083–2097.
28 C. J. Cramer and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 760–768.
29 E. Pellegrini and M. J. Field, *J. Phys. Chem. A*, 2002, **106**, 1316–1326.
30 A. J. Mulholland and W. G. Richards, *Proteins: Struct., Funct., Bioinf.*, 1997, **27**, 9–25.
31 A. Mulholland, P. Lyne and M. Karplus, *J. Am. Chem. Soc.*, 2000, **122**, 534–535.
32 M. W. van der Kamp, J. Ziurek, F. R. Manby, J. N. Harvey and A. J. Mulholland, *J. Phys. Chem. B*, 2010, **114**, 11303–11314.
33 O. Donini, T. Darden and P. A. Kollman, *J. Am. Chem. Soc.*, 2000, **122**, 12270–12280.
34 M. van der Kamp, F. Perruccio and A. Mulholland, *Proteins: Struct., Funct., Bioinf.*, 2007, **69**, 521–535.
35 S. Bjelic, B. Brandsdal and J. Aqvist, *Biochemistry*, 2008, **47**, 10049–10057.
36 G. Alter, J. Casazza, W. Zhi, P. Nemeth, P. Srere and C. Evans, *Biochemistry*, 1990, **29**, 7557–7563.
37 A. D. Waldman, K. W. Hart, A. R. Clarke, D. B. Wigley, D. A. Barstow, T. Atkinson, W. N. Chia and J. J. Holbrook, *Biochem. Biophys. Res. Commun.*, 1988, **150**, 752–759.
38 C. R. Dunn, H. M. Wilks, D. J. Halsall, T. Atkinson, A. R. Clarke, H. Muirhead and J. J. Holbrook, *Philos. Trans. R. Soc. London, Ser. B*, 1991, **332**, 177–184.
39 T. Atkinson, D. Barstow, W. Chia, A. Clarke, K. Hart, A. Waldman, D. Wigley, H. Wilks and J. Holbrook, *Biochem. Soc. Trans.*, 1987, **15**, 991–993.
40 M. Kotik and H. Zuber, *Biochemistry*, 1992, **31**, 7787–7795.
41 S. Ferrer, I. Tunon, S. Marti, V. Moliner, M. Garcia-Viloca, A. Gonzalez-Lafont and J. Lluch, *J. Am. Chem. Soc.*, 2006, **128**, 16851–16863.
42 S. L. Quaytman and S. D. Schwartz, *J. Phys. Chem. A*, 2009, **113**, 1892–1897.
43 S. Ranganathan and J. Gready, *J. Phys. Chem. B*, 1997, **101**, 5614–5618.
44 A. Yadav, R. Jackson, J. Holbrook and A. Warshel, *J. Am. Chem. Soc.*, 1991, **113**, 4800–4805.
45 V. Moliner and I. H. Williams, *Chem. Commun.*, 2000, 1843–1844.
46 S. Ferrer, E. Silla, I. Tunon, M. Oliva, V. Moliner and I. H. Williams, *Chem. Commun.*, 2005, 5873–5875.
47 S. L. Quaytman and S. D. Schwartz, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 12253–12258.
48 A. Clarke, D. Wigley, W. Chia, D. Barstow, T. Atkinson and J. Holbrook, *Nature*, 1986, **324**, 699–702.
49 W. Im, S. Bernèche and B. Roux, *J. Chem. Phys.*, 2001, **114**, 2924–2937.
50 P. Schaefer, D. Riccardi and Q. Cui, *J. Chem. Phys.*, 2005, **123**, 014905.
51 T. Benighaus and W. Thiel, *J. Chem. Theory Comput.*, 2008, **4**, 1600–1609.
52 A. Aleksandrov and T. Simonson, *J. Am. Chem. Soc.*, 2008, **130**, 1114–1115.
53 J. Sund, M. Ander and J. Aqvist, *Nature*, 2010, **465**, 947–950.