JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Improving Efficiency in SMD Simulations Through a Hybrid Differential Relaxation Algorithm

Claudia L. Ramírez,[†,§] Ari Zeida,[†,§] Gabriel E. Jara,[†,§] Adrián E. Roitberg,[∥] and Marcelo A. Martí*[,‡,§]

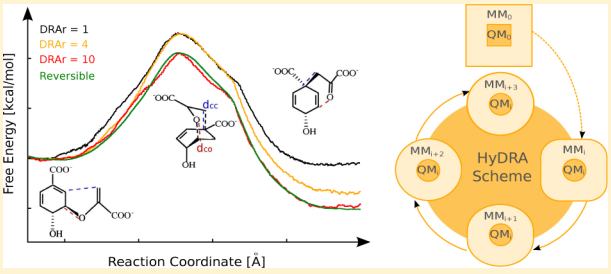
[†]Departamento de Química Inorgánica, Analítica y Química Física and [‡]Departamento de Química Biológica, FCEN, UBA, Buenos Aires C1428EGA, Argentina

[§]Instituto de Química Física de los Materiales, Medio Ambiente Y Energía (INQUIMAE), UBA-CONICET, Buenos Aires C1428EGA, Argentina

[∥]Quantum Theory Project and Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, United States

Supporting Information

ABSTRACT: The fundamental object for studying a (bio)-chemical reaction obtained from simulations is the free energy profile, which can be directly related to experimentally determined properties. Although quite accurate hybrid quantum (DFT based)-classical methods are available, achieving statistically accurate and well converged results at a moderate computational cost is still an open challenge. Here, we present and thoroughly test a hybrid differential relaxation algorithm (HyDRA), which allows faster equilibration of the classical environment during the nonequilibrium steering of a (bio)chemical reaction. We show and discuss why (in the context of Jarzynski's Relationship) this method allows obtaining accurate free energy profiles with smaller number of independent trajectories and/or faster pulling speeds, thus reducing the overall computational cost. Moreover, due to the availability and straightforward implementation of the method, we expect that it will foster theoretical studies of key enzymatic processes.



1. INTRODUCTION

Computer simulation of molecular systems is currently an invaluable and ubiquitous tool for the study of chemical and biochemical reactions and their underlying mechanisms. For those reactions occurring in solution (or in an enzyme) at near room temperature, the central object to be determined is the free energy profile (FEP) along the reaction (or the potential of mean force), since it can be directly related to the experimentally determined properties such as reaction rates and equilibrium constants, and thus their kinetic and thermodynamic behavior. Even though quite accurate computation of enthalpies for a given system configuration is straightforward, determination of free energies requires extensive sampling to account for the large number of meaningful conformations, in particular, in complex environments such as solvated biomolecules.

To properly describe chemical reactivity, the reactive system must be described using quantum mechanics (QM) methods, such as Density Functional Theory (DFT). Due to the nonlinear increase of their computational cost with system size, the treatment of systems larger than a few hundreds of atoms at the QM level is computationally expensive and an area of ongoing development. A possible way of overcoming this drawback is to use cheaper QM methods or the so-called divide-and-conquer type of strategies, resulting in linear scaling and thus allowing to treat an entire protein at the QM level.[1−4] Another, and possibly the most widely used strategy to overcome the size limit but nonetheless retain the ability to describe large molecular systems (such as an enzymatic reaction or a chemical reaction in aqueous solution), is to employ hybrid quantum mechanical/molecular mechanical (QM-MM) schemes. These allow the detailed investigation of chemical events that take place in a certain region of a large system, which is modeled using ab initio techniques (QM region), while the remainder of the system (MM region) is treated at the less expensive classical, force-field based level. Although this type of methods has been extensively and successfully used in the last decades,[5−7] it is important to note that due to the large degree of configurational sampling that is necessary to obtain accurate free energies and the high computational expense of the QM-MM implementations (mainly required to compute the QM energy and forces), the development and validation of free energy determination strategies that properly balance accuracy and computational efficiency is a very active research area.

To determine the FEP along a given selected reaction coordinate (here called λ), either using a QM, QM-MM, or pure classical method, enhanced or biased sampling strategies are usually used. The oldest most common and well-known of these is possibly the umbrella sampling strategy,[8] but more recently, other methods such as Metadynamics,[9] Adaptive

Received: July 25, 2014

   dx.doi.org/10.1021/ct500672d | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Biasing Force,[10] Free Energy Perturbation,[11] and Orthogonal Space Random Walk[12] have also been applied to these type of studies. A very potent quite novel strategy for this purpose is based on performing multiple steered molecular dynamics (MSMD) simulations, which drive the system along the desired reaction coordinate under nonequilibrium conditions, using an external force. For each MSMD, the work performed by the force is determined ($W_i(\lambda)$), and they are combined using Jarzynski's Relationship (JR)[13] (eq 1) to determine the corresponding FEP ($G(\lambda)$),

$$G(\lambda) = -\beta^{-1}\ln\langle e^{-\beta W_i(\lambda)}\rangle \quad (1)$$

where $G(\lambda)$ represents the FEP as a function of the reaction coordinate, $\beta = 1/k_bT$, where k_b is Boltzmann's constant and T is the system temperature, and the brackets represent the average of the function within them. The JR thus relates a system's nonequilibrium dynamics to its equilibrium properties and has been validated both using computer simulations and single molecule pulling experiments using optical tweezers.[14]

The MSMD strategy has been described in more detail elsewhere and will be presented here only briefly, focusing on the key aspects related to the present work. The first important note is that the brackets in eq 1 represent an average taken over the ensemble of molecular dynamics trajectories each starting at different snapshots extracted from an equilibrated run at the initial value of the reaction coordinate. The equality holds when the average is properly converged, which is formally true only when an infinite number of trajectories or works have been performed. In practice, a finite sampling is performed (defined by the number of independent performed trajectories, NT), and thus, the free energy is a statistical estimator. The other key aspect of JR is that, in theory, there is no requirement as to how fast the system is driven from reactants to products. In the simplest implementation, the external force, $\vec{F}$ is applied as a harmonic potential whose minimum moves at a constant velocity (ν), along the reaction coordinate λ, according to eq 2,

$$\vec{F} = -k(\lambda - \lambda_0 - \nu\cdot dt) \quad (2)$$

where λ_0 is the starting value of the reaction coordinate, which together with ν defines the center of the well of the total harmonic potential in each MD step. The velocity, also called pulling speed, thus determines the amount of steps (and computational cost) needed to perform each trajectory. Using this type of external force, the work is simply computed by numerically integrating the force.

To understand how pulling speed and number of trajectories are related to the convergence of the exponential average and the free energy profile estimate, it is useful to consider some extreme cases. First, we can expect that the obtained distribution of work values is roughly Gaussian (this requirement is not needed but makes explanations clearer). Under near-equilibrium conditions (i.e., extremely slow pulling), the distribution of work values is very nearly a delta function and, thus, the external work equals the free energy, as required by the second Law of thermodynamics. However, when the pulling speed is increased, the average work gets larger and the width of the work distribution increases. Since to obtain the free energy using JR requires a converged exponential average, the works that need to be sampled are those in the lower tail of the Gaussian distribution. Thus, faster conversion rates (larger ν) result in the need of a larger number of trajectories to be performed, in order to properly converge the nonexponential average. Another important aspect of the work distribution when performing MSMD is that, at any given pulling speed, the distribution is wider the further from the starting conditions, which were at equilibrium. Thus, when using a constant number of trajectories, more accurate free energy estimates are obtained closer to the starting conformations. Usually, both forward and backward transformations are performed starting from equilibrated reactants and product simulations, and the complete free energy profile is obtained by combining both estimates.[15,16]

Wider work distributions for higher transformation speeds, and further from the start values, are due to the system having been driven further away from equilibrium. Thus, any strategy that would retain the system closer to its equilibrium at each point along the transformation is expected to result in a narrow distribution of work values and hence a better free energy estimate with a smaller number of trajectories. In the present work, we develop such a strategy by allowing differential relaxation of the MM environment in a QM-MM MSMD free energy determination.

The use of differential sampling of the phase space in QM-MM simulations was implemented several years ago within the PAW ab initio molecular dynamics package using the Carr Parrinello method[17] by Woo et al.,[18] based on the multiple-time-step algorithm of Tuckerman et al.,[19,20] and is based on the use of different masses and time steps for the QM and MM subsystems. The MM subsystem was sampled with a time step of dt/n, with dt being the dt used for the propagation of the QM system, and n was set to 20. Masses of the classical atoms were scaled using different schemes in the 1/10 to 1/100 range. The authors showed that the multiple time step method method generated the same trajectory as a standard MD simulation and allowed faster (about 2 times) convergence of the average force applied on a geometrical restraint, which could be used in a thermodynamic integration scheme. No free energy profile was computed.

In the present work, we have developed, validated, and applied a Hybrid QM-MM Differential Relaxation Algorithm—HyDRA—which allows efficient equilibration of the classical environment during a steering process that drives the QM system along a desired reaction at faster than the equilibrium (or reversible) required velocity. Our results show the method results in accurate free energy profiles estimates, which require a smaller number of independent nonequilibrium simulations and/or can be steered faster than those using conventional QM-MM schemes. Thus, the HyDRA (hybrid differential relaxation algorithm) method allows for the determination of QM-MM reaction free energy profiles at a smaller computational cost, especially in those cases where the relative QM to MM calculation cost is large.

In this article, we present the method, describe the condition under which it works, and make an argument as to why it works, while also testing and applying it to a number of systems.

2. METHODOLOGY

2.1. Theoretical Basis. In the present method, we present an approach inspired on the works performed by Woo et al.[18] and Ozer et al.[21] We use a differential relaxation scheme that allows the MM system to relax for a given number of MM steps, while the QM region remains fixed (or frozen), and then, both systems move together for another joint (or QM-MM) step. The ratio between the MM and the joint steps is called the

Differential Relaxation Algorithm ratio (DRAr), and thus, a DRAr ratio of 1 is equivalent to conventional QM-MM MD simulation, while a DRAr of 4 implies that for each QM step, 4 MM steps are performed. The key to this implementation is the correct calculation of the forces and what actually represents a given structure/conformation after a joint QM-MM or only MM relaxation step. The proposed strategy is shown graphically in scheme 1 for a DRAr = 4. When the MD starts both the QM and MM systems are synchronized. For this conformation (subscript 0), all contribution to forces are computed as usual in QM-MM simulations. The forces acting on the QM atoms have contributions from the atoms in the QM subsystem in the initial conformation (upper figure) and also from the atoms in the MM region in this conformation. In our force notation, these are described as $F_{QM}(QM_0, MM_0)$. The equivalent is true for the forces acting on the MM subsystem, $F_{MM}(QM_0, MM_0)$. With this set of forces, the program performs a first dynamics step, leading to conformation i (right figure). Then, a pure MM step is performed: pure MM forces and QM-MM forces acting on MM atoms are computed as usual, considering the frozen QM coordinates in conformation i, thus the MM system moves in a fixed QM field arriving to $i + 1$ conformation (Figure 1). Forces

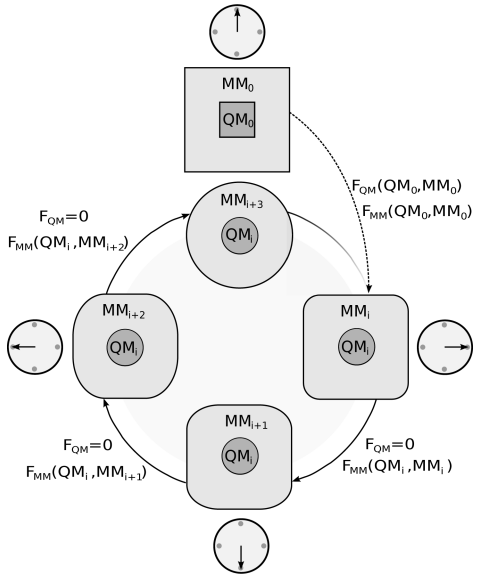


Figure 1. Hybrid Differential Relaxation Algorithm scheme for a DRAr of 4.

acting on the QM subsystem are zero-ed out, and thus, the QM subsystem remains fixed in the i conformation. After the MM subsystem has evolved, another two MM only steps are executed, arriving finally at MM_{i+3} conformation (circular figure). Now, the system performs again a joint step. The forces acting on the system for this step are calculated considering the $i + 3$ structure for the MM subsystem and the i structure for the QM subsystem. At this point, the cycle has started once again. In this algorithm, each structure/conformation after a joint step is similar to any given QM-MM MD snapshot, while any structure after all MM steps have been performed represents a structure where the MM regions was allowed to relax or accommodate slightly to the frozen QM conformation.

An important point to discuss is what type of approximation we have made with this differential relaxation algorithm, and when will it hold and when will it fail. From Jarzynski's original ideas, it is clear that only work done onto the system by an external force is to be counted toward the free energy estimator, while heat dissipated or added should not. Only instances when the coordinate of choice is being moved should count toward the work associated with the free energy estimator. Instances where the bath coordinates move but the reaction coordinate does not count as heat and hence do not matter for the free energy calculation. This means that if the HyDRA procedure was applied to a single distance with everything moving, followed by a small number of steps where said distance is kept fixed while the rest moves; then, the procedure would be exact and no approximation would have been made. This has already been shown analytically by Ozer.[21] In the present manuscript, we are in a slightly different regime. While relaxing the bath, we keep fixed not just the reaction coordinate but the whole QM region. This is an approximation, which we need to study. The results in the next sections validate our approach. This approximation works well when most of the dissipative work is associated with the MM region (e.g., solvent) and not due to local strains in the QM region.

2.2. Implementation. The hybrid differential relaxation algorithm was implemented in the SANDER module of the AMBER 12[22] program package. The corresponding code was modified to allow computing the QM-MM contribution to the forces acting on the MM atoms, at the fixed QM structure $F_{MM}(QM_i)$ every DRAr steps, and subsequently allow propagating only the MM subsystem for the corresponding steps. The modified subroutine files and user's instructions are available upon request to the author.

2.3. Classical Force Field and Parameters. All classical parameters for both, pure classical equilibration simulations and the MM part in the QM-MM dynamics were taken from the Amber force field, ff99SB for the 20 amino acid residues[23] and TIP3P for the water molecules.[24] Classical parameters for the reactants (chorismate, prephenate, methylthiol, and hydrogen-peroxide) were taken from previous works from our group.[25−27] Simulations were performed at constant temperature using the Berendsen thermostat as implemented in SANDER, using periodic boundary conditions and either constant volume or constant pressure. Ewald Sums were used to treat long-range electrostatics in the chorismate−prephenate systems, while a direct cutoff method was used for the methyltiol reaction performed at the DFT-PBE level (see below), where Ewald sums are not available.

2.4. QM Parameters and QM-MM Scheme. The QM-MM simulations were carried out using two different theory levels: Density Functional Tight Binding (DFTB)[28,29] as implemented in the sander module of AMBER,[22] and a pure DFT method with the Perdew, Burke, and Erznheof (PBE) exchange and correlation functional, as developed in our group[30−33] and implemented in the SANDER module.[34]

2.5. Set up of the System, Equilibration, and Simulation Strategy. *2.5.1. Chorismate in Solution.* The system consists of chorismate solvated with an octahedral box of 1483 water molecules. First, we performed a complete geometry optimization using the AMBER default algorithm followed by a pure classical Molecular Dynamics (MD) simulations to equilibrate the system density at the desired temperature (300K). For this sake, the system was first gently heated to 300 K during 1 ns at constant volume (i.e., NVT ensemble), followed by another 1 ns long constant pressure (i.e., NPT ensemble) simulation, both using a time step of 1 fs and without the addition of any restraints. Subsequently, a

dx.doi.org/10.1021/ct500672d | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

hybrid QM-MM MD—considering the solute as quantum—was performed to equilibrate the system with the new Hamiltonian, which consisted of first a 200 ps NVT simulation, followed by a 2 ns long production simulation, which was performed restraining the chosen reaction coordinate at their desired equilibrium values (see below). All QM-MM MD simulations were performed using a 0.5 fs time step. From this run, 20 different starting structures were selected for the MSMD production runs. The chosen reaction coordinate was $d_{CC} - d_{CO}$ (Figure 2), which has been shown to adequately

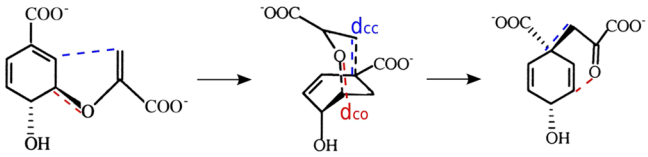


Figure 2. Employed reaction coordinate for the chorismate to prephenate reaction in solution and in the chorismate mutase enzyme. The distance between the two carbons, d_{CC}, is shown in blue, while d_{CO} is shown in red.

represent the process in previous work[25,26,35−37] and was varied from −3.26 to +1.99 Å. In order to obtain the desired free energy profiles and thus analyze the HyDRA efficiency, different set of production runs were performed, varying the pulling speed (v), the MM steps to QM steps ratio (called the DRAr), the number of trajectories, and the initial conditions.

2.5.2. Chorismate/Prephenate in Bacillus subtilis Chorismate Mutase. To study the chorismate to prephenate reaction in *Bacillus subtilis* chorismate mutase (BsCM), we began with the crystal structure of the corresponding wild type enzyme trimer bound to prephenate (PDB ID 1COM),[38] also employed in our previous works on QM/MM studies of *Bacillus subtilis* chorismate mutase reaction mechanism using SIESTA HYBRID method.[25,26]

Hydrogen atoms were added with the tleap module of the Amber Program package. Standard protonation states were assigned to titrable residues (D and E are negatively charged, K and R positively charged). Histidine protonation was assigned favoring formation of hydrogen bonds in the crystal structure. The resulting protein complex was immersed in a truncated octahedral box of TIP3P water, and consisted of ca. 15 000 atoms. The system was first optimized using a conjugate gradient algorithm for 2000 steps, followed by 200 ps long constant volume pure classical MD where the temperature of the system was slowly raised to 300 K. The heating was followed by additional 200 ps long constant temperature and constant pressure MD simulation to equilibrate the systems density. During this equilibration, the α carbons of the protein were restrained with an harmonic potential to their initial position using a 1 kcal/mol·Å^2 force constant. Finally, the system was subjected to an NPT run for 5 ns, in a pure classical MD without restraints. Snapshots of this run were selected as starting points for QM/MM equilibrium simulations. One of the prephenate molecules was selected as the QM subsystem. First, a QM/MM optimization of the QM subsystem was performed to allow the system to adapt to the new Hamiltonian smoothly. This was followed by 50 ps long QM/MM equilibration simulation at constant temperature and constant volume. Finally, 500 ps long QM/MM production simulations were performed to select snapshots for the Pre to Cho MSMD simulations. To obtain equilibrated snaphots of the BsCM bound to chorismate, the last structure of a QM/MM MSMD simulation, where the Pre to Chor reaction occurred, was subjected to 50 ps long QM/MM equilibration simulation at constant temperature and constant volume and then to another 500 ps long QM/MM production simulations to select snapshots for the MSMD simulations of the Chor to Pre MSMD simulations.

2.5.3. Thiol Oxidation by Hydrogen Peroxide. The thiol oxidation by hydrogen peroxide in aqueous solution was recently studied using QM/MM methods and the umbrella sampling strategy.[27] The quantum solute consisting of the negatively charged methyl thiolate and hydrogen peroxide (CH_3S^- + H_2O_2) was embedded in a box containing 4290 classical TIP3P water molecules. We first performed a pure classical equilibration of the system temperature and density followed by a 50 ns QMMM MD where the solute was treated at the DFTB semiempirical level. Selected snapshots for the MSMD simulations were relaxed for 1 ps QM/MM MD using a DFT level of theory (PBE) as described above.

The chosen reaction coordinate consists in the difference between the O1−O2 and the S−O1 distances, as depicted in Figure 3. The reaction was sampled from −1.7 to +1.2Å. Details about the strategy used for the final determination of the FEP using MSMD are described in the Results section.

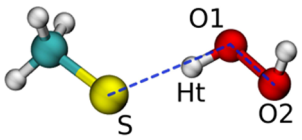


Figure 3. Reacion coordiante for the thiol oxidation by hydrogen peroxide.

3. RESULTS

The results are organized as follows. First, we study the chorismate to prephenate reaction in order to test and validate our methodology. We calculated several FEP for the Chor to Pre reaction in solution and in the enzyme. Results for the forward (Chor to Pre) reaction are shown in the next sections. Information about the backward (Pre to Chor) reaction can be found in the Supporting Information. We also model the thiol oxidation by hydrogen peroxide at the hybrid MM/DFT-PBE level, computing several FEP employing JR. Finally, to test the acceleration of the method, we performed hybrid MM/DFT-PBE MD for several systems with varying number of QM and MM atoms, previously studied in our group,[31,39,40] and compared their computational cost.

3.1. Chorismate to Prephenate Reaction. We use the chorismate to prephenate reaction, at the DFTB/MM level of theory, as test case since it allow us to generate a quite accurate reference free energy profile to set as reference. We can then study many different protocols (varying the key HyDRA parameters) at a moderate computational cost and analyze its accuracy and relative computational cost.

3.1.1. Reference Free Energy Profile in Solution. We begin our analysis by determining a reference FEP for the Chor to Pre conversion in solution with the QM(DFTB)/MM methodology. During the equilibrium simulation, the CO distance is almost constant (since it corresponds to a chemical bond), while the CC distance is free to explore a wider range. Previous studies on Chor have shown that the reactive

dx.doi.org/10.1021/ct500672d | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

structure, called the diaxial conformation, displays a CC distance of ca. 4.7Å; thus, we restrained the reaction coordinate at −3.0Å position in the equilibrium simulations of Chor. We performed three very slow conversions (quasistatic) under these conditions. Figure 4 shows the corresponding work

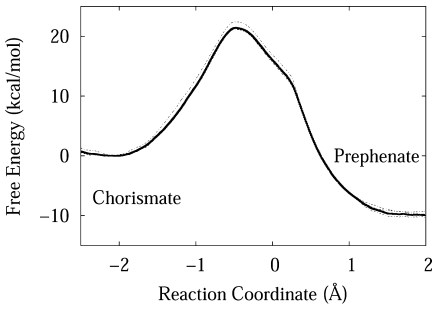


Figure 4. Work and FEP for the Chor to Pre conversion at the QM(DFTB)/MM in solution. The thin lines corresponds to work profiles obtained at a pulling speed of 0.02 Å/ps repeated three times from different initial conditions. The thick line corresponds to Jarzynski's estimator.

profiles (thin lines, repeated three times from different initial conditions) and the FEP (thick line, computed through the JR) vs reaction coordinate, determined at a pulling speed of 0.02 Å/ps. The results shows that the three work profiles barely differ (less than 1 kcal/mol) and thus were performed close to equilibrium rate, yielding a quite accurate estimation of the FEP. The resulting FEP (shown in bold in Figure 4) will be considered for the remaining of the present work as the reference FEP for this reaction, with a ΔG^{o} of −10 kcal/mol and an activation free energy barrier, $\Delta G^{\ddagger}$ of 21 kcal/mol. The values are in good agreement with experimental data.

We will now analyze the accuracy of the FEP determination for the chorismate to prephenate conversion reaction using nonequilibrium MSMD simulations and JR, and if (and how) they improve with the HyDRA scheme. Usually, in this approach, each transformation is performed 10 to 500 times faster than the time required to perform the reaction in quasi equilibrium conditions and 5 to 40 independent trajectories are performed. Given the aim of the present work, we will systematically analyze how the different simulation parameters (number of trajectories, pulling speed, and DRA ratio) affect the accuracy of the obtained FEP.

3.1.2. Effect of the DRA Ratio. In order to analyze how the DRA ratio affects the obtained FEP we performed 10 nonequilibrium trajectories, with pulling speed of 1 Å/ps (50 times faster than our reference), with DRA ratios of 1, 4, 10, and 20, respectively. Figure 5a presents the results, together with the reference FEP, while Figure 5b shows the barrier ($\Delta G^{\ddagger}$) and reaction free energy (ΔG^{o}) values as a function of DRAr.

The results show, as expected, that nonequilibrium FEP significantly over estimates both ΔG^{o} and $\Delta G^{\ddagger}$ when compared to the reference. Interestingly, these results clearly show that when using the HyDRA method with a DRAr greater than one and the same number of QM steps, more accurate FEPs are obtained ($\Delta G^{\ddagger}$ and ΔG^{o} closer to reference values). Noteworthy is the fact that increasing the DRA ratio from 4 to 10 improves the accuracy, although further increase (to 20) does not produce a significant effect. This observation is also not unexpected since there must be a limit to the classical

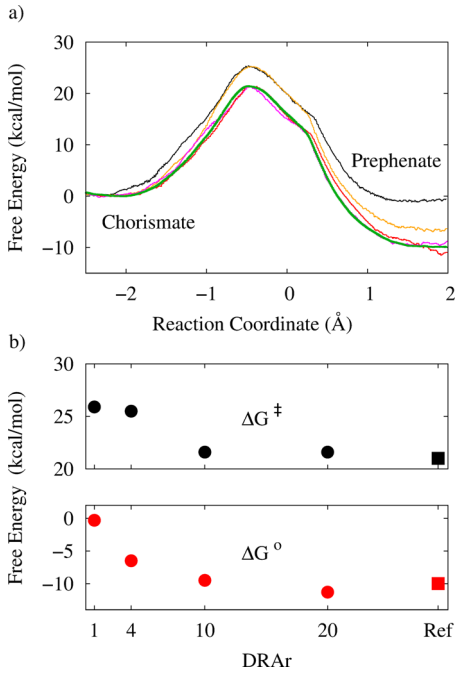


Figure 5. FEP for the Chor to Pre reaction, obtained through 10 nonequilibrium dynamics with a pulling speed of 1 Å/ps, and reference FEP. (a) Free energy profiles for the reaction in solution with a DRA ratio of 1, 4, 10, and 20, shown in black, orange, magenta, and red curves, respectively. The green curve is reference FEP. (b) $\Delta G^{\ddagger}$ and ΔG^{o} for the FEP shown in part a. Circles represent values for FEP with different DRAr, and the squares are the values for the reference profile.

relaxation, for a given amount of perturbation performed to the QM subsystem.

The use of HyDRA also affects the standard deviation between work traces for different trajectories. As DRAr increases, the trajectories become similar in terms of work values, which decreases the standard deviation and improves convergence (data shown in the Supporting Information).

3.1.3. Effect of Pulling Speed. To further analyze the potential of the DRA strategy, we repeated the FEP determinations, using the same four DRA ratios, but using faster and slower pulling speeds. The first thing to be noted is that for any given speed larger DRAr results in smaller (more accurate) barriers and lower reaction free energies. The other key observation is that DRA effect is larger for faster speeds. For example, for a pulling speed of 0.5 Å/ps (25 times faster than the reference run) DRAr of 1 slightly overestimates (3 kcal/mol) the barrier and reaction free energy and higher DRAr only moderately improve these results, while for a pulling speed 100 times faster that the reference reaction (2 Å/ps) the DRAr effect (with values of 10 −20) is notorious, being able to obtain barrier and free energy values in reasonable agreement with the reference, while for regular dynamics—DRAr = 1—the depicted behavior in even thermodynamically wrong (ΔG^{o} > 0) and the barrier is overestimated by more than 10 kcal/mol. Thus, the faster the pulling speed, the bigger the impact of using the HyDRA strategy. It is also interesting to compare the FEPs obtained through different pulling speeds but equal DRAr. As expected for DRAr = 1, the improvement in the accuracy of the FEP when decreasing the pulling speed is notorious going from a completely wrong to a reasonable prediction. For a DRAr = 4, too fast pulling speeds still results

dx.doi.org/10.1021/ct500672d | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

in too high barriers and wrong prediction of the reaction free energy, but moderate pulling speeds (1A/ps) already yield meaningful results. Higher DRA ratios (10 or 20) strikingly show almost no difference between the FEPs obtained with different pulling speeds yielding pretty good results even at fast speeds. In summary, DRA allows to significantly increase the pulling speed used to drive the reaction, being nonetheless able to yield accurate FEP (see Table 1).

Table 1. Relevant Values of the FEP for the Chor to Pre Reaction in Solution, Obtained with Different Methodologies[a]

pulling speed	trajectories	DRAr	$\Delta G^{\ddagger}$	ΔG^{o}
0.02	3	1	21	−10
0.5	10	1	24.6	−7.0
		4	18.4	−10.3
		10	22.6	−8.1
		20	20.6	−12.8
1	5	1	25.1	−0.6
		4	25.1	−7.8
		10	21.1	−9.9
		20	22.6	−9.9
	10	1	25.9	−0.9
		4	25.5	−6.5
		10	21.6	−9.5
		20	21.6	−11.3
	15	1	25.6	−0.9
		4	25.1	−7.5
		10	21.1	−9.9
		20	21.4	−11.5
	20	1	25.2	−1.6
		4	25.1	−7.0
		10	21.1	−9.9
		20	21.1	−11.9
2	10	1	33.8	8.4
		4	29.0	−0.2
		10	23.6	−6.3
		20	23.4	−9.1

[a]Pulling speed in Å/ps and ΔG in kcal/mol.

3.2. FEP for the Chorismate to Prephenate Reaction in *Bacillus subtilis* Chorismate Mutase Enzyme. To further explore the potential of the presented strategy, we now turn our attention to the Cho to Pre interconversion reaction performed by the *Bacillus subtilis* chorismate mutase (BsCM). The resulting key energetic values obtained from the corresponding FEPs are presented in Table 2. As for the case in solution, we computed a reference FEP using a very slow pulling speed of 0.02 A/ps. Nonequilibrium simulations with and without HyDRA were performed 50 to 200 times faster. The results show the same tendency as in solution, although differences seem slightly smaller. The value of ΔG^{o}, for example, decreases systematically when using HyDRA, even at the slowest pulling speeds (1 and 2 Å/ps), while for the fastest the difference is more notorious.

For $\Delta G^{\ddagger}$, the increment of the value of DRAr leads, as expected, to lower free energy barriers. Indeed, it is possible to obtain a profile showing very similar barrier compared to the reference for all speeds using DRAr of 10, while for non-DRA simulations it is overestimated by several kcal/mol.

We can compare the obtained results together with those of Table 1 with available experimental and theoretical data from the literature. Experimental studies of the chorismate to prephenate conversion in solution showed that reaction has an activation free energy of 24.5 kcal/mol (Andrews et al.[41]). In BsCM the activation free energy is reduced to 15.4 kcal/mol (Kast et al.[36]), thus resulting in a barrier lowering of 9.1 kcal/mol, which makes BsCM a very efficient enzyme. Kast et al.[35] also determined the reaction free energy of the reaction yielding a value −13.4 kcal/mol.

Table 2. Relevant Values of the FEP for the Reaction of Chorismate to Prephenate in BsCM Enzyme, Obtained with 10 Trajectories and Different Pulling Speed and DRAr[a]

pulling speed	trajectories	DRAr	$\Delta G^{\ddagger}$	ΔG^{o}
0.02	3	1	9.6	−16.5
1	10	1	12.2	−11.4
		4	10.1	−15.4
		10	8.3	−15.6
2	10	1	15.2	−9.5
		4	12.5	−12.6
		10	9.9	−12.4
4	10	1	15.4	−5.2
		4	11.8	−7.5
		10	9.6	−12.0

[a]Pulling speed in Å/ps and ΔG in kcal/mol.

From the computed reaction free energy presented in Table 2, we see that slow speeds produce values close to the experimental values, even without DRA. However, for faster speeds, the effect of DRA is crucial. Concerning the barrier, it is well-known that DFTB tends to underestimate their magnitude;[25,26] thus, the best way to asses the accuracy of the FEPs is comparing the change in the free energy barriers between protein and solution ($\Delta\Delta G^{\ddagger}$). Taking our best solution estimate for the barrier of 21 kcal/mol, and for BsCM of 9.6 kcal/mol, we obtain an enzyme efficiency of 11.4 kcal/mol, thus slightly overestimating the experimental value.

The data in Table 2 show that, for the slowest speed (without DRA), the simulations already yields values in that range ($\Delta\Delta G^{\ddagger}$ of 12.4 kcal/mol), while using DRA simulations the value gets closer to the that obtained using the reference FEP for slow and moderate speeds and is only slightly larger (14 kcal/mol) for the fastest speeds and DRAr of 10.

In summary, the results for the study of CM reaction show that the use of the DRA strategy result in lower and more accurate barriers and reaction free energies and more so at fast pulling speed. These results are important, since they show the potential of the DRA strategy for the study of enzymatic reactions.

3.3. Thiol Oxidation by Hydrogen Peroxide in Aqueous Solution. To have another and possibly more stringent test of the potential of the HyDRA strategy, we used it to compute the free energy profile of the methanethiolate oxidation reaction by hydrogen peroxide in aqueous solution

$$CH_3S^- + H_2O_2 \rightarrow CH_3SO^- + H_2O$$

which was recently studied using the umbrella sampling strategy.[27]

The reaction involves significant changes in the solvation properties of the participating groups as well as a proton transfer, and thus, we expect to significantly benefit from the differential relaxation of the classical environment. The reaction was studied using a high level of theory, taking advantage of our own QM-MM implementation, based on a pure GGA-DFT,

dx.doi.org/10.1021/ct500672d | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

that uses the Exchange Correlation functional proposed by Perdew, Burke, and Erznheof (PBE), which has been extensively used to study chemical reactions[30−33] but requires significantly more computational cost compared to the DFTB semiempirical method used for the previous reaction. Table 3

Table 3. Relevant Values of the FEP for the Methanethiolate Oxidation Reaction by Hydrogen Peroxide in Aqueous Solution, Obtained with 10 Trajectories and Different Pulling Speed and DRAr[a]

pulling speed	DRAr	$\Delta G^{\ddagger}$	$\sigma^{\ddagger}$	ΔG^{o}	σ^{o}
0.5	1	12.7	6.1	−35.9	11.4
	4	12.0	2.5	−34.8	8.6
	10	10.1	1.1	−35.9	5.6
1	1	14.3	6.6	−27.0	9.5
	4	11.7	5.6	−31.0	9.4
	10	12.2	4.5	−31.9	9.4

[a]Pulling speed in Å/ps and ΔG in kcal/mol.

shows the relevant values of the free energy profiles obtained with JR, combining 10 trajectories, for the corresponding reaction using DRA ratios of 1, 4, and 10 with pulling speeds of 1 Å/ps (left), and with DRA ratio of 1, 4, and 10 with a pulling speed of 0.5 Å/ps (right), and their corresponding standard deviations. The results again show that increasing the DRA ratio, results in lower barriers and reaction free energies. Indeed, DRA ratios of 4 and 10 show a barrier for the fastest pulling, which is slightly lower than that obtained using half the velocity and ratio of 1. The standard deviation presented shows that the SD diminishes with increasing DRAr. What is more interesting is that for the slower pulling (0.5 Å/ps), even when the results in the FEP do not differ significantly for different DRAr, the SD value analysis reveals that larger DRAr possibly result in better convergence and thus a better estimate.

3.4. Computational Cost of HyDRA. Another important point for the evaluation of the benefits of using the HyDRA algorithm is to evaluate how much computational cost (i.e., overhead time) is added by performing the classical relaxation steps, since it must be overcompensated by the gain in accuracy or reduction in the number of trajectories and/or pulling speed. Lets perform first a theoretical analysis of the whole computational cost. Our first assumption is that $t_{MM} = \alpha \cdot t_{QM}$, where t_{MM} its the consumed time when performing only one classical step, t_{QM} is the time consumed in performing a joint (or QM/MM) step, and α is a scaling parameter ($\alpha < 1$) that relates both costs and is system dependent. In conventional non-HyDRA simulations, each step thus requires t_{QM} time. With this notation the total time consumed in one trajectory of n_s (quantum) steps is

$$T_1 = n_s[t_{QM} + (\mathrm{DRAr} - 1)t_{MM}] \tag{3}$$

where t_{QM} is the non-DRA time, and (DRAr − 1) t_{MM} is the overhead time.

The number of steps can be related with the pulling speed through $n_s = \delta v^{-1}$, where δ is the ratio between whole reaction coordinate length and the dt employed. Replacing this expression in equation number 3, and multiplying for the number of total trajectories, N_{Tr}, used to obtain the final FEP, we obtain the total simulation time, T, employed in the determination of a given FEP using JR and HyDRA, according to

$$T = \delta\left(\frac{N_{Tr}}{\nu}\right)(t_{QM}(1 + \alpha(\mathrm{DRAr} - 1))\ ) = \delta \cdot \mathrm{SS} \cdot \mathrm{Sy} \tag{4}$$

where we have divided the different contributions to the total time employed in three terms: δ, which is the same in all cases compared; SS, which is the value that depends on the simulation strategy (number of trajectories and speed); and Sy, which is the term that contains contributions that depend of each system (the t_{QM} consumed, the DRAr used, and α). To compare the computational cost of HyDRA with conventional simulations, we define the relative total time as

$$\tau_R = \frac{T}{T^o} = \frac{\mathrm{SS} \cdot \mathrm{Sy}}{\mathrm{S}^o \cdot \mathrm{Sy}^o} \tag{5}$$

where the superscript o indicates the values for a reference simulation. In other words, τ_R represents the relative time needed to obtain a given FEP between two different simulations schemes considering all the relevant parameters (pulling speed, number of trajectories, system, and DRAr).

In order to analyze τ_R, we will first look at those cases where SS is the same (i.e., the only variation is thus the DRAr) and determine the corresponding value. To this end, we performed plain MD for 1 ps with and without HyDRA and with different DRArs in six different systems: nitrate in water, methanethiolate in hydrogen peroxide, chorismate in water, chorismate in BsCM enzyme, myoglobin and extracellular signal-regulated kinase- displaying different number of both QM and MM atoms, and treated at the DFT-PBE level. The dynamics were performed with an Intel Core i3 processor in a single work station. The obtained average values for τ_R and S_y are summarized in Table 4.

The results show that for most systems with DRAr of 4, τ_R values are 1.1, which means that the HyDRA simulation requires ca. 10% additional computational cost compared to a

Table 4. Computational Cost of the Different QM/MM Calculations Schemes Using HyDRA at a DFT Level of Theory, This Calculation Was Performed with an Intel Core i3 Processor in a Single Work Station[a]

system	NO_3^- water	$CH_3S^- + H_2O_2$	Chor water	Chor BsCM	Mb	ERK
no. QM atoms	4	9	24	24	42	55
no. MM atoms	2070	8808	9834	40763	18758	53478
Sy^o	7.2	9.8	185.0	209.1	1059.4	776.8
τ_R(DRAr = 4)	1.1	1.1	1.2	1.1	1.1	1.1
τ_R(DRAr = 10)	1.4	1.5	1.3	1.4	1.2	1.3
τ_R(DRAr = 20)	1.9	2.2	1.6	1.7	1.5	1.5

[a]Sy^o is the average time (in seconds) required to perform a joint step in a standard (i.e., DRAr = 1) QM/MM-MD simulation. τ_R(DRAr) is the ratio between the average time (in seconds) required to perform a complete QM/MM HyDRA cycle with respect to Sy^o.

dx.doi.org/10.1021/ct500672d | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

conventional QM/MM run. For DRAr of 10 to 20, values of τ_R are still bellow 2, except for those systems having a small QM region. Generally speaking the relative increase in computational cost, τ_R, is smaller the larger the QM region, for example for a QM/MM calculation of Myoglobin, which includes in the QM system the iron contained in the heme group, only an additional 20% of computational cost is paid for DRAr of 10.

It is clear now that any improvement in the total computational cost required to calculate a FEP will depend on the concerted variation of SS and Sy. We analyzed these variations employing the results for the reaction of Chor to Pre, summarized in Tables 1 and 4. We will approach to this subject calculating τ_R for different examples.

There is no FEP that differs by less than 1 kcal/mol from the reference when using DRAr = 1. The best estimate from DRAr = 1 (i.e., 10 conventional trajectories with a pulling speed of 0.5 Å/ps) differs in approximately 3 kcal/mol from the reference and has a τ_R of 0.1. Instead, for DRAr = 10, we obtain a τ = 0.04, that is, five trajectories and a pulling speed of 1 Å/ps, which shows a clear gain in computational time, while achieving the desired accuracy.

Concerning the estimates of BsCM catalytic power ($\Delta\Delta G^{\ddagger}$ is 9.1 kcal/mol), leaving aside the solution reference FEP already discussed and considering only those profiles which yield correct trends in the reaction free energy, our best non-DRA estimate of $\Delta\Delta G^{\ddagger}$ can be performed combining values of v = 0.5 Å/ps in solution and v = 1 Å/ps in BsCM, with 10 trajectories in both cases, which yield a value for $\Delta\Delta G^{\ddagger}$ of 12.4 kcal/mol. Using DRA best estimates for DRAr of 10, v = 1 Å/ps, 5 trajectories in solution and v = 4 Å/ps and 10 trajectories in BsCM we obtain a $\Delta\Delta G^{\ddagger}$ value of 10.5 kcal/mol, which is closer to the experimental value and with a τ_R of 0.34.

This results indicate that when using HyDRA fewer trajectories or fastest pulling speed can be used to achieve the same FEP, when comparing with conventional SMD at a DFT level of theory.

We have also analyzed the time gain when employing the DFTB level of theory. Because DFTB is optimized in order to reduce the time of the QM calculations, the time consumed for a QM step is comparable with an MM step. Our results with DFTB show that the gain obtained from the SS term (previously discussed for DFT) does not compensate the overhead that is due Sy (that is DRAr times bigger that for the case with no HyDRA), resulting in no gain in total computational cost (this data can be found in Supporting Information).

4. DISCUSSION

The present work presents our Hybrid Differential Relaxation Algorithm (HyDRA), which allows for selective relaxation of the classical environment in QM/MM simulations of chemical and/or enzymatic reactions, thus resulting in faster (in terms of computational time) convergence of free energy profiles in the context of Jarzynski's relationship. We study the chorismate to prephenate conversion, where HyDRA allowed us to obtain accurate reaction free energies at less than 10% the computational cost. In essence, these observations reflect the expected behavior of using HyDRA, which results in trajectories that are closer to the equilibrium even for relatively fast (over 100 faster than quasi-static) pulling speeds.

From a statistical-mechanics point of view, our results also highlight some interesting properties of the behavior of nonequilibrium trajectories and their work profiles in relation to common simulation strategy parameters, such as pulling speed and number of trajectories. The present results and our previous experience in the subject[25,26,42,43] systematically show that it is very difficult to get accurate FEP when the pulling speed is too fast, no matter how many trajectories are performed (unless very big numbers are considered). Thus, it seems clear that if the system evolves far from equilibrium and the standard deviation (σ) of the work values is much larger than k_bT, it is very difficult to get meaningful results. This effect has been theoretically analyzed by Pohorille et al.[44] who estimated that if $\sigma = k_bT$, $^1/_6$ of the trajectories samples values close to the real $G(r)$ and thus allow good convergence and accurate results. The value decreases to $^1/_{40}$ trajectories for $\sigma = 2k_bT$ and to 1/3.000.00 if $\sigma = 5k_bT$, strongly biasing the work values and creating very poor convergence. The HyDRA results in closer to equilibrium trajectories even at faster pulling speeds, significantly reducing the standard deviation of the work values and thus significantly increasing average accuracy and convergence.

It is also interesting to discuss other sampling schemes based on JR, which allow for faster or better convergence of the FEP seeking to reduce the spread, or standard deviation, of the work trajectories, that could be applied to QM-MM simulations. In one approach, introduced by Amzel's group,[45] the reaction is divided in segments. After performing the pulling in the first segment and computing the corresponding work profile, additional equilibration MD is performed to allow the system to relax, and then, a new segment trajectory is performed and the corresponding work profile computed. Each segment work profile is then joined combinatorially, generating many more complete reaction work profiles, which are used to compute the final FEP using JR. Another such approach is the Adaptative Steered Molecular Dynamics (ASMD) as developed by Ozer et al.[21] In the ASMD method, the range of the reaction coordinate is divided into stages, but instead of performing equilibrium dynamics, which require computational cost, the following set of trajectories corresponding to the adjacent segment are begun from the trajectory displaying the lowest work (i.e., that which is expected to be closest to equilibrium). The final FEP is obtained by joining the FEP obtained for each segment using JR. So far, and to our knowledge, both these strategies have been used in pure Classical MD simulations to study biological processes, such as small peptide folding or ligand binding, and none of them was tested in a QM-MM scheme to study chemical or enzymatic reactions. In this context, it could be interesting to address which strategy is the most cost-effective for computing reactions in solutions or in enzymes using QM-MM. Moreover, the HyDRA and ASMD methods could be combined, since they exploit different pitfalls of JR, possibly resulting in a synergic effect.

Concerning the additional computational cost that is paid when increasing the pure MM steps, although the present detailed analysis of computational cost and accuracy can only be done for the extensive tested Chor to Pre system of the present work, the results from Tables 1 and 4, show that using a DRAr between 4 and 10 have the potential of yielding more accurate free energy estimates using smaller number of trajectories and faster pulling speeds, and could thus be used already from the beginning in QM/MM studies of chemical or enzymatic reactions.

5. CONCLUSION

We have presented and thoroughly tested a hybrid differential relaxation algorithm (HyDRA) to be used in QM/MM simulations, implemented in the AMBER computer simulation package, which allows obtaining accurate free energy profiles of chemical reactions in solution and enzymes at a reduced computational cost in the context of Multiple Steered Molecular Dynamics Simulations and Jarzynki's Relationship. The method power relies in its capacity to allow faster relaxation of the classical protein/solvent environment to the reacting QM subsystem. We expect that our method will foster studies of key enzymatic processes specially those requiring significant protein reorganization along the chemical step.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

All the FEPs and standard deviation for the chorismate to prephenate reaction, in solution and in the enzyme BsCM. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*Email: marcelo@qi.fcen.uba.ar

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Computer power was provided by High Performance Computer (HPC) Center at the University of Florida. Research was funded by Grants PICT-2010-416, PIP 2012-2014 #112 201101 00850 to M.A.M.

■ REFERENCES

(1) He, X.; Merz, K. M. J. *J. Chem. Theory Comput.* **2010**, *6* (2), 405411.
(2) Stewart, J. *J. Mol. Modeling* **2009**, *15* (7), 765−805.
(3) Dixon, S.; Merz, K., Jr. *J. Chem. Phys.* **1996**, *104* (17), 6643−6649.
(4) Stewart, J. *MOPAC2009*; Stewart Computational Chemistry: Colorado Springs, CO, 2009. http://openmopac.net (accessed Sept. 12, 2014).
(5) Bash, P.; Field, M.; Karplus, M. *J. Am. Chem. Soc.* **1987**, *109*, 8092−8094.
(6) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(7) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.
(8) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(9) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(10) Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904−2914.
(11) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(12) Zheng, L.; Chen, M.; Yang, W. *J. Chem. Phys.* **2009**, *130*, 234105.
(13) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(14) Liphardt, J.; Dumont, S.; Smith, S. B.; Tinoco, I. J.; Bustamante, C. *Science* **2002**, *296*, 1832−1835.
(15) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245−268.
(16) Crooks, G. E. *Phys. Rev. E* **2000**, *61*, 2361−2366.
(17) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(18) Woo, T. K.; Margl, P.; Blöchl, P. E.; Ziegler, T. *J. Phys. Chem. A* **2002**, *106*, 1173−1182.
(19) Tuckerman, M. E.; Parrinello, M. *J. Chem. Phys.* **1994**, *101*, 1302−1315.
(20) Tuckerman, M. E.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990−2001.
(21) Ozer, G.; Valeev, E. F.; Quirk, S.; Hernandez, R. *J.Chem. Theory Comput.* **2010**, *6*, 3026−3038.
(22) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; Simmerling, C. L.; Wang, J.;Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California: San Francisco, 2012.
(23) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712−725.
(24) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(25) Crespo, A.; Martí, M. A.; Estrin, D. A.; Roitberg, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 6940−6941.
(26) Crespo, A.; Scherlis, D. A.; Martí, M. A.; Ordejón, P.; Roitberg, A. E.; Estrin, D. A. *J. Phys. Chem. B* **2003**, *107*, 13728−13736.
(27) Zeida, A.; Babbush, R.; González Lebrero, M. C.; Trujillo, M.; Radi, R.; Estrin, D. A. *Chem. Res. Toxicol.* **2012**, *25*, 741−746.
(28) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(29) Seabra, G. d. M.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. *J. Phys. Chem. A* **2007**, *111*, 5655−5664.
(30) Nitsche, M. A.; Ferreria, M.; Mocskos, E. E.; González Lebrero, M. C. *J. Chem. Theory Comput.* **2014**, *10*, 959−967.
(31) González Lebrero, M. C.; Bikiel, D. E.; Elola, M. D.; Estrin, D. A.; Roitberg, A. E. *J. Chem. Phys.* **2002**, *117*, 2718−2725.
(32) González Lebrero, M. C.; Estrin, D. A. *J. Chem. Theory Comput.* **2007**, *3*, 1405−1411.
(33) Morzan, U. N.; Ramírez, F. F.; Oviedo, M. B.; Sánchez, C. G.; Scherlis, D. A.; Lebrero, M. C. G. *J. Chem. Phys.* **2014**, *140*, 164105−164114.
(34) Götz, A. W.; Clark, M. A.; Walker, R. C. *J. Comput. Chem.* **2014**, *35*, 95108.
(35) Kast, P.; Tewari, Y. B.; Wiest, O.; Hilvert, D.; Houk, K. N.; Goldberg, R. N. *J. Phys. Chem. B* **1997**, *101*, 10976−10982.
(36) Kast, P.; Asif-Ullah, M.; Hilvert, D. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.
(37) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. *J. Org. Biomol. Chem.* **2011**, *9*, 1578−1590.
(38) Chook, Y.; Gray, J.; Ke, W. H.; Lipscomb. *J. Mol. Biol.* **1994**, *240*, 476−500.
(39) Arroyo-Mañez, P.; Bikiel, D. E.; Boechi, L.; Capece, L.; Di Lella, S.; Estrin, D. A.; Martí, M. A.; Moreno, D. M.; Nadra, A. D.; Petruk, A. A. *Biochim. Biophys. Acta* **2011**, *1814*, 1054−1064.
(40) Turjanski, A. G.; Hummer, G.; Gutkind, J. S. *J. Am. Chem. Soc.* **2009**, *131*, 6141−6148.
(41) Andrews, P. R.; Smith, G. D.; Young, I. G. *Biochemistry* **1973**, *18*, 3492−3498.
(42) Xiong, H.; Crespo, A.; Marti, M.; Estrin, D.; Roitberg, A. *Theor. Chem. Acc.* **2006**, *116*, 338−346.
(43) Torras, J.; De Seabra, G.; Deumens, E.; Trickey, S.; Roitberg, A. *J. Comput. Chem.* **2008**, *29*, 1564−1573.
(44) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(45) Echeverria, I.; Amzel, L. M. *Proteins* **2010**, *78*, 1302−1310.