THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

The Mechanism of Citryl-Coenzyme A Formation Catalyzed by Citrate Synthase

Alexey Aleksandrov,*[,†] Elena Zvereva,[‡] and Martin Field*[,§]

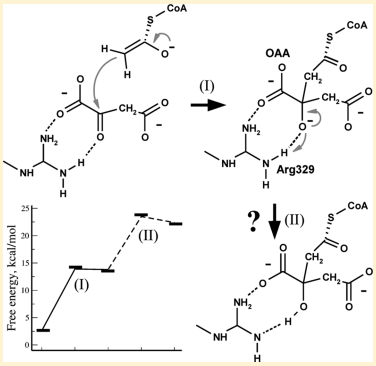
[†]Laboratoire de Biochimie (CNRS UMR7654), Department of Biology, Ecole Polytechnique, 91128 Palaiseau, France

[‡]A.E. Arbuzov Institute of Organic and Physical Chemistry, Kazan Scientific Centre, Russian Academy of Sciences, Arbuzov Street 8, 420088 Kazan, Russia

[§]DYNAMO/DYNAMOP, CEA/Grenoble/Bâtiment IBS2 (CEA, CNRS UMR5075, Université Joseph Fourier - Grenoble I), CAMPUS EPN, 6 rue Jules Horowitz, 38042 Grenoble, France

Supporting Information

ABSTRACT: The enzyme citrate synthase is used by all living cells to catalyze the first step of the citric acid cycle. In this work, we have investigated the enolization and condensation steps catalyzed by citrate synthase, using ab initio (B3LYP/def2-TZVP and MP2/aug-cc-pVDZ) quantum chemical/molecular mechanical hybrid potentials in conjunction with reaction-path-location algorithms and molecular dynamics free energy simulations. The results of the latter indicate that the catalytic His238 residue is in its neutral form, and also argue strongly for the presence of a water molecule in the enzyme's catalytic center. Such a water is observed in some, but not all, of the experimentally resolved structures of the protein. The mechanism itself starts with an enolization that proceeds via an enolate intermediate rather than the enol form, which is much more unstable. This is in agreement with the results of other workers. For the condensation step, we investigated two mechanisms in which there is a direct nucleophilic attack of the enolate intermediate on the oxaloacetate carbonyl carbon, and found the one in which there is no proton transfer from the neighboring arginine to be preferred. Although this residue, Arg329, is not implicated directly in the reaction, it helps to stabilize the negative citryl-CoA formed during the condensation step.



1. INTRODUCTION

Citrate synthase (CS) catalyzes the conversion of oxaloacetate (OAA) and acetyl-coenzyme A (acetyl-CoA) into citrate and coenzyme A (CoA), during the first step of the citric acid cycle.[1] CS has been extensively studied both experimentally[2−4] and computationally[5] and its mechanism of reaction appears to be universal across all (S−)citrate synthases.[6,7] The active site of this complex is shown in Figure 1. The reaction starts with deprotonation of acetyl-CoA to form a nucleophilic intermediate,[8] which subsequently attacks the carbonyl carbon of OAA. Asp375 (pig CS numbering) has been shown to be the catalytic base for the first deprotonation step,[2,9] as mutations at this position significantly diminish the enzyme's activity.[2,3] It has also been shown that the conserved His274 and a water molecule stabilize the enolate product during the first step of the reaction, and that the enol intermediate is unstable.[10,11]

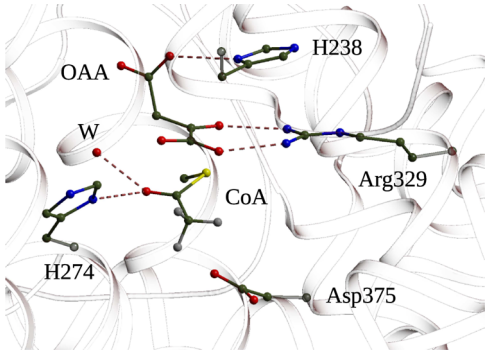


Figure 1. The active site pocket of the citrate−synthase complex (PDB entry code 1CSI).[17]

The enolate intermediate formed in the first step subsequently performs a nucleophilic attack on the carbonyl carbon of OAA. This produces citryl-CoA which is then hydrolyzed to citrate and CoA.[12] It is believed that the formation of citryl-CoA requires protonation of the carbonyl oxygen of OAA.[1,5,9] The mechanism of the condensation step was recently studied by hybrid potential simulations,[5] which suggested that Arg329 donates the proton, as this hydrogen bonds to the carbonyl oxygen of OAA.

Most existing simulation studies[8,13−15] of the reaction mechanism of CS have employed the crystal structure of chicken CS (PDB entry code 4CSC[16]) within which an important residue, His238, appears to be biprotonated. However, other crystal structures exist with a water molecule present in the catalytic center (see, for example, ref 17). This

Received: December 17, 2013
Revised: April 8, 2014
Published: April 10, 2014

interacts with His238, suggesting that His238 could, in fact, be in its neutral form. As far as we are aware, only one study has investigated the protonation state of residue His238, and it concluded that His238 should be biprotonated since the computed barrier with a singly protonated His238 was too high compared to the experimental estimate.[8] Nevertheless, despite this result, workers have assumed both singly protonated[5,15,18] and biprotonated forms for His238[8,14,19] in previous work.

Due to the uncertainty in the protonation state of His238, and the importance of the reaction, we decided to reconsider the various proposals for the mechanism in the literature and to investigate alternative possibilities using molecular simulation techniques. We employed molecular dynamics free energy (MDFE) simulations to determine the protonation states of residues that might be implicated in the reaction, and high-level hybrid quantum chemical (QC)/molecular mechanical (MM) potentials in conjunction with a nudged-elastic-band (NEB) reaction-path-finding algorithm to map out the reaction. In addition, we also analyzed the role that CS residues play in the catalysis.

The outline of this paper is as follows. Section 2 describes in detail the simulation approaches that were used, section 3 presents the results that were obtained, and section 4 concludes.

2. METHODS

2.1. Molecular Dynamics Simulations. The crystal structure of the CS homodimer was obtained from the Protein Data Bank (PDB), entry 4CSC,[16] with bound D-malate and acetyl-CoA. 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 carboxymethyldethia coenzyme inhibitor (PDB entry code 1CSI).[17] Hydrogens were constructed with ideal stereochemistry. All histidines were modeled in a neutral state, except residue 246. Residues 136, 249, and 320 were protonated on Nϵ, and residues 235, 238, 274, and 377 were protonated on Nδ. The orientations of asparagine and glutamine residues were checked by visual inspection but no changes were made with respect to those coming from the crystal structure.

The simulations included protein residues within a 24 Å sphere, centered on the OAA binding site. In addition to crystal waters, a 26 Å sphere of water was overlaid and waters that overlapped ligands or crystal waters were removed. Throughout the molecular dynamics (MD) simulations, protein atoms between 20 and 26 Å from the sphere's center were harmonically restrained to their experimental positions. Simulations were done with the SSBP solvent model,[20,21] which treats the region outside the 26 Å sphere as a uniform dielectric continuum with a dielectric constant of 80. 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 temperature bath at 298 K. The CHARMM27 force field was used for the protein[22,23] and the TIP3P model for water.[24] Electrostatic interactions were computed without any cutoff, using a multipole approximation for distant groups.[25] Calculations were done with the CHARMM program.[26]

2.2. Hybrid Potential Simulations. We tested and employed a number of QC/MM partitionings when simulating the enolization and condensation steps. Our reaction path optimizations of the enolization step used the same QC region of 33 atoms as Mulholland et al.[15] This comprised the methylthioester part of acetyl-CoA, the side chains of Asp375 and His274, and the water molecule interacting with the acetyl moiety of acetyl-CoA. Single point calculations on the optimized structures were also done with an enlarged QC region comprising the methylthioester part of acetyl-CoA up to the first amide group; the side chains of Asp375 and His274 and the water molecule interacting with the acetyl moiety of acetyl-CoA; a water molecule interacting with Asp375; the side chain of His235; the oxaloacetate ligand with the side chains of interacting residues Arg329, Arg401, Arg421; and two water molecules interacting with oxaloacetate. This large QC region contained 123 atoms and was electrically neutral. For both partitionings, a hydrogen-link atom model was employed to saturate bonds that crossed the QC/MM boundary, giving 3 and 9 link atoms for the small and large QC regions, respectively.

For the condensation step, we optimized reaction paths with two QC/MM partitionings. The smaller QC region was the same as that proposed by van der Kamp et al.[5] and comprised the OAA ligand, the methylthioester part of acetyl-CoA, and the side chains of Asp327, Asp375, and Arg329. The larger region contained 115 atoms, had a total charge of −1, and comprised the methylthioester part of acetyl-CoA up to the first amide group, the side chains of residues Arg329, Arg401, Arg421, Asp327, Asp375, and His320; the oxaloacetate ligand; and three water molecules.

The atoms in the MM region were represented with the CHARMM27 force field,[22,23] whereas a density functional theory (DFT) method was used for the QC region with the BLYP functional and Ahlrichs's split valence basis with polarization functions (SVP).[27,28] Single point calculations on the optimized structures were done with two higher-level QC methods: (i) the B3LYP functional with an empirical van der Waals correction to account for dispersion forces[29] and a def2-TZVP basis set,[30] which has polarization functions on all atoms; and (ii) an MP2 method with the large aug-cc-pVDZ basis set.[31]

All geometry optimizations and reaction path calculations were performed with the QC(DFT)/MM potential implemented in the pDynamo software,[32] and its interface to the ORCA program.[33] To include the system beyond the explicit 26 Å sphere, we used our recent QC/MM implementation of the stochastic solvent boundary potential,[34] which treats the region outside the explicit sphere as a dielectric continuum, and includes its effect in the self-consistent field procedure of the QC calculation. No truncation or cutoff was employed to calculate nonbonding interactions.

Reaction paths were optimized with the NEB method[35] implemented in pDynamo. The utility of the NEB method 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. After a regular NEB calculation was converged, the climbing image method (CI-NEB) was applied to the image with the highest energy in the path. Previously we showed that application of the CI-NEB is required for fully refined saddle point structures.[35]

The starting coordinates of the complexes for reaction were taken after 200 ps of MD simulation on the systems described in section 2.1. These structures were then geometry optimized

dx.doi.org/10.1021/jp412346g | J. Phys. Chem. B 2014, 118, 4505−4513

and used to obtain models of the reaction end-points using a constrained optimization approach, followed by unrestrained optimization. Initial paths for the mechanisms were generated from these structures by linear interpolation using 19 path structures for mechanisms I–III and 39 structures for mechanism IV. This is in agreement with the approach that we employed in our previous study.[35] After path determination, the identity of the critical points along the reaction paths were verified using normal-mode analysis to ensure that they indeed corresponded to either minima or saddle points.

2.3. Alchemical Molecular Dynamics Free Energy Simulations. To determine the protonation state of the catalytic residue His238, we performed MDFE simulations,[36,37] using the thermodynamic cycle shown in Figure 2. The MDFE

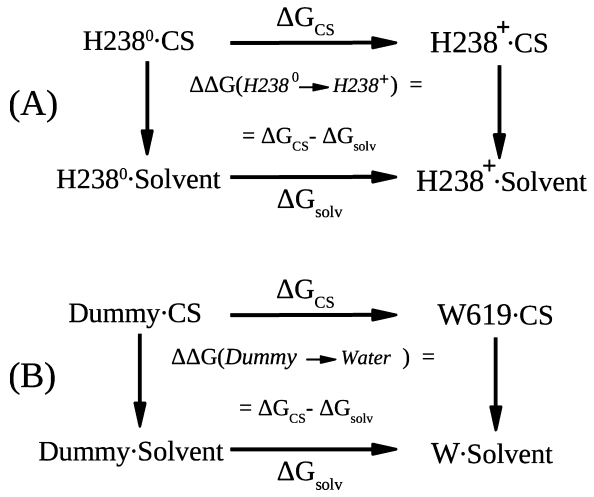


Figure 2. (A) The thermodynamic cycle used to calculate the protonation free energy difference for His238 in the citrate–synthase complex and in solution. Horizontal legs represent the protonation of His238 on its Nδ, either in the protein (above) or in solution (below). (B) The thermodynamic cycle used to calculate the binding free energy of the water molecule W619. Horizontal legs represent the switching on and off of interactions between the water molecule in the citrate–synthase complex and in solution. The MDFE simulations follow the horizontal legs in all cases.

method follows the horizontal legs of the cycle. Protonated histidine is reversibly transformed into its deprotonated form during a series of MD simulations and the corresponding work is derived from a thermodynamic integration formula.[38,39] For the lower leg of the thermodynamic cycle, we simulated methyl-imidazole in solution, which served as the reference molecule, whereas for the upper leg we simulated a portion of the CS complex. Both the CS and the solution transformations were modeled as rearrangements of atomic charges on the imidazole group. See refs 37 and 40 for details. The systems were solvated in a 70 Å box of water and periodic boundary conditions were imposed so that the entire box was replicated periodically in all directions. All long-range electrostatic interactions were computed efficiently by the particle mesh Ewald method,[41] and the appropriate number of sodium counterions were included to render the system electrically neutral. MD simulations were performed at constant room temperature and pressure, after 200 ps of thermalization. The CHARMM27 force field was used for the protein[22,23] and the TIP3P model for water.[24] Calculations were done with the NAMD program.[42]

During the MDFE simulations, the energy function can be expressed as a linear combination of terms associated with the neutral and protonated forms of His238:

$$U(\lambda) = U_0 + (1 - \lambda)U(\mathrm{His238^+}) + \lambda U(\mathrm{His238^\circ}) \quad (1)$$

in which λ is a coupling parameter and U_0 represents interactions between parts of the system other than the histidine. The free energy derivative with respect to λ has the form

$$\frac{\partial G}{\partial \lambda}(\lambda) = \langle U(\mathrm{His238^\circ}) - U(\mathrm{His238^+})\rangle_\lambda \quad (2)$$

where the brackets indicate an average over the MD trajectory with the energy function $U(\lambda)$.[38] We gradually mutated His238$^+$ into His238° by changing λ from zero to one. The successive values of λ were 0.001, 0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.99, and 0.999. The free energy derivatives were computed at each λ value from a 100 ps MD simulation, or window. The last 80 ps of each window were used for averaging. A complete mutation run corresponded to 12 windows and 1.2 ns of simulation. Five runs were performed in each direction (His238$^+$ into His238° and the reverse), and the uncertainty was calculated as the standard deviation of the individual runs obtained from the five forward/backward pairs,[39] totaling 12.0 ns. This simulation length was necessary to correctly sample the important degrees of freedom and to obtain a low statistical uncertainty.

To compare relative stability of a water molecule in CS, we used the thermodynamic cycle in Figure 2, panel B. The MDFE method follows the horizontal legs of the cycle within which interactions between the water molecule and the protein or solvent are gradually switched on. Otherwise, we used the same simulation setup as described above for the free energy simulations of the protonation of His238.

2.4. Poisson–Boltzmann Solvation Free Energy Calculations. To estimate the electrostatic contribution to the solvation free energy we employed the Poisson–Boltzmann (PB) implicit solvent model. The contribution was obtained by subtracting the electrostatic free energies of the complex computed with a high dielectric constant for solvent and a low dielectric constant for the protein. The electrostatic potential was calculated by numerically solving the PB equation, using a cubic grid and a finite-difference algorithm.[26] The calculations were done at physiological ionic strength with a solvent dielectric constant of 80. For the solute dielectric constant, values of 1 and 2 were used. In principle, structural reorganization is accounted for in the NEB algorithm, which would suggest that a dielectric constant of 1 was most appropriate. However, calculations with a constant of 2 provide an indication of how sensitive the results of the calculations are to changes in the model parameters. For the PB calculations, all water molecules were removed, except for five water molecules that interact with Asp375 (1), His238 (1), and the OAA and CoA ligands (3). CHELPG[43] charges computed with the B3LYP/6-311+G* DFT method were used for the atoms in the QC region and CHARMM27 charges for those in the MM region. The total free energy was obtained by adding the PB solvation energy to the internal QC/MM energy for the system obtained in the absence of the explicit solvent environment.

2.5. QC/MM Reaction Path Free-Energy Perturbation Calculations. We estimated free energies of the structures along the NEB pathways using the free energy perturbation (FEP) method of Kästner et al.,[44] which itself is similar to that

dx.doi.org/10.1021/jp412346g | J. Phys. Chem. B 2014, 118, 4505–4513

of Zhang and co-workers.[45] The system setup was similar to that described in section 2.1, except that the atoms in the QC region and the link atoms were fixed at their corresponding optimized geometries from the NEB profile. CHELPG[43] charges, computed with the B3LYP/6-311+G* DFT method, were used for the atoms in the QC region and CHARMM27 charges for those in the MM region. This approximation was shown to give a good result in comparison to the results obtained with the full self-consistent field iterations in each MD step.[44]

For each image of the NEB profiles, a MD simulation was performed at constant room temperature and pressure for 500 ps. This gave ∼10 ns of simulation for mechanisms I−III and ∼20 ns for mechanism IV. The free-energy change between image i and $i + 1$ due to the QC/MM interactions was calculated according to the formula:

$$\Delta A^{i\rightarrow i+1} = -k_B T \ln\langle \exp(-E_{pert}^{i\rightarrow i+1}/k_B T)\rangle_{MM,i} \quad (3)$$

where $E_{pert}^{i\rightarrow i+1}$ is the energy of perturbation computed with the MM atoms coordinates from the MD simulation of image i, and the QC positions of image $i + 1$. Following the previous work,[44,45] estimates of the free energies of the atoms in the QC region were obtained using a rigid-rotor harmonic-oscillator approximation after normal mode analysis. However, we found that these contributed negligibly ($\ll$1 kcal mol^{-1}) to the calculated free energy differences between the different structures.

3. RESULTS

3.1. His238 Protonation Free Energy Simulations and a Missing Water Molecule. A necessary preliminary to studies of mechanism is the determination of the protonation states of residues that might be implicated in the reaction, as these can have a profound effect on the mechanisms and energetics that are found.[46] Those of most residues were straightforward to deduce, but there was some uncertainty about the ϵ-protonation of His238 in the CS complex. To resolve this, we employed MDFE simulations[36,37,40] to calculate the free energy of protonation from the thermodynamic cycle shown in Figure 2. To effect the protonations of the histidine in solution and in the CS complex, we performed a series of MD simulations in which the neutral and charged forms of His238 were reversibly transformed in both directions. This technique has no adjustable parameters and gives good results for acid/base reactions.[37]

The computed free energies in solvent and in the protein are −1.8 (±0.3) and 9.1 (±0.7) kcal mol^{-1}, respectively, which gives a protonation free energy for His238 in the CS of 10.9 (±0.8) kcal mol^{-1}. This free energy value clearly favors the neutral form of His238. The large destabilization of a positively charged His238 may be understood by the interactions that it has with the adjacent positively charged residue Arg239.

Further support for the neutral state of His238 comes from the observation that the position of the histidine during the free energy MD simulations is in much better agreement with the crystal structure when it is neutral. In the latter, the distance between the Oδ of Asp327 and the Nϵ of His238 is 4.6 Å, whereas in the simulations, the average distances are 2.7 and 4.7 Å with biprotonated and neutral His238, respectively. The shorter distance with the charged form is because Asp327 prefers to hydrogen bond with the charged His238. As a result, His238 departs from its crystal position, so that its protonated Nϵ can interact with the Oδ of Asp327.

A crystal structure of CS complexed with the carboxymethyldethia coenzyme inhibitor (PDB entry code 1CSI)[17] has a water molecule (residue W619) in the catalytic center that is missing from the structure 4CSC. This water molecule interacts with residues His238 and Asp327 as shown in Figure 3. To test

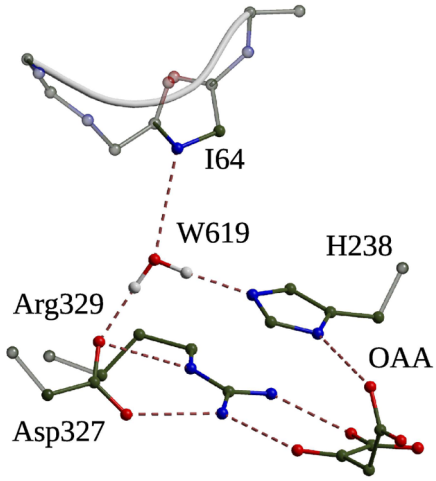


Figure 3. The structure around the water W619 that interacts with His238 (PDB entry code 1CSI).[17]

for the presence of this water molecule in the CS structure we employed MDFE simulations using the thermodynamic cycle shown in Figure 2, panel B. We performed a series of MD simulations in which the interactions between the water molecule and the rest of the system were reversibly switched on and off in solvent and in the protein. Note that in these simulations His238 was in its preferred neutral form.

The computed free energies in solvent and in the protein are −6.1 (±0.2) and −16.5 (±0.2) kcal mol^{-1}, respectively, which gives a free energy difference for W619 in CS of −10.4 (±0.3) kcal mol^{-1}, thereby clearly favoring the presence of W619 in the CS structure. The large stabilization of W619 may be understood by the interactions that it has with the Nϵ of His238, the Oδ of Asp327, and with the backbone amide group of Ile64.

3.2. Reaction Path Calculations. As previous simulation studies of CS catalysis have employed a biprotonated His238 and have lacked the water molecule, W619, in the active site, we reinvestigated the enolization and condensation steps of the reaction with a neutral His238 and in the presence of W619. Schematics of the mechanisms that were considered are shown in Figure 4. Potential energies for the transition state (TS) and product structures resulting from the NEB calculations are summarized in Table 1 with ancillary data obtained with different QC/MM partitionings and QC levels of theory in Table S1. Table 1 also contains estimates of the free energies of the reaction path structures obtained with a PB implicit solvent model and via QC/MM FEP calculations with explicit solvent. A pictorial representation of the reaction path energies is given in Figure 5, whereas details of the normal mode frequencies and structures of the critical points along the reaction paths are given in Table S2 and Figures S1 and S2, respectively.

3.2.1. Enolization Step. We studied two mechanisms, I and II, of enolization (Figure 4). Mechanism I is deprotonation of the acetyl-CoA substrate by residue Asp375, which results in

dx.doi.org/10.1021/jp412346g | J. Phys. Chem. B 2014, 118, 4505−4513

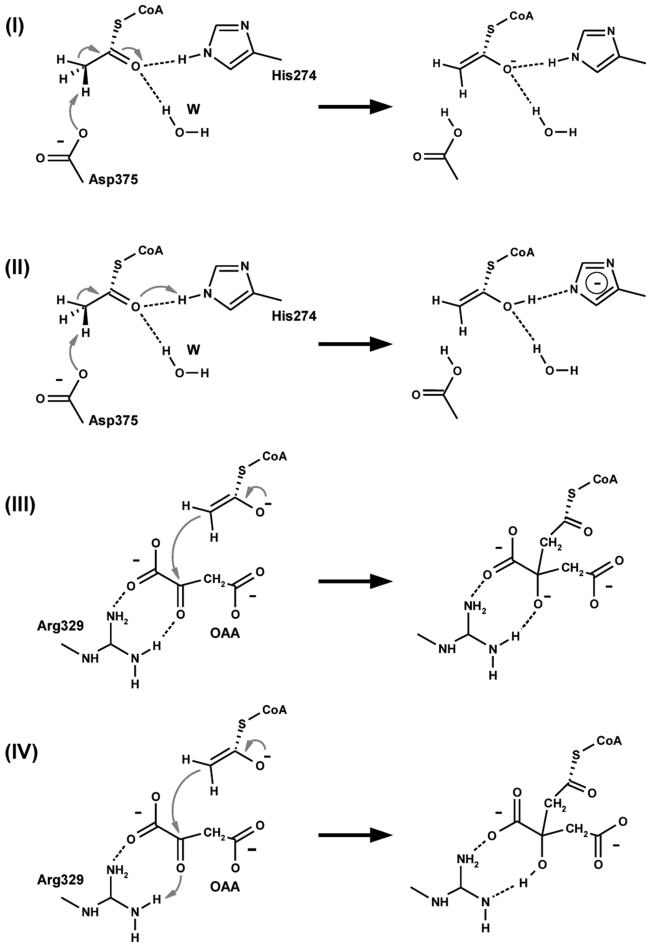


Figure 4. Schematics of the mechanisms investigated in the hybrid potential QC/MM simulations. Mechanisms I and II correspond to the enolization step and mechanisms III and IV to the condensation step.

Table 1. Energies of the TS and Product Structures from the Proposed Mechanisms of Enolization and Condensation That Are Shown in Figure 4[a]

mechanism	structure	QC small	QC large	FEP	Poisson−Boltzmann ϵ_p = 1, 2
I	TS	11.2	8.2	6.5	6.2/10.7
	Product	10.0	5.5	2.6	3.1/7.5
II	Product	25.2	18.9	16.7	15.6/17.8
III	TS	-	15.0	13.9	9.5/7.9
	Product	18.6	14.1	13.8	8.0/6.0
IV	TS	19.4	21.1	23.5	11.8/11.5
	Product	18.0	20.7	22.4	10.1/8.7

[a]The QC small and QC large energies are potential energies and were obtained with the explicit solvent B3LYP/def2-TZVP QC/MM hybrid potentials that are defined in the text. The FEP and Poission−Boltzmann energies are free energies determined with the same large QC region model. The two Poisson−Boltzmann values correspond to those using a protein dielectric constant, ϵ_p, of 1 and 2, respectively. All energies are in kcal mol^{-1} and are relative to the reactant species.

the formation of an enolate intermediate that is stabilized by His274 and the water molecule. The barrier for this step is 8.2 kcal mol^{-1} with the product being 2.7 kcal mol^{-1} more stable than the TS with the large QC region. With the smaller QC region, the barrier is 11.2 kcal mol^{-1} and the product is only 1.2

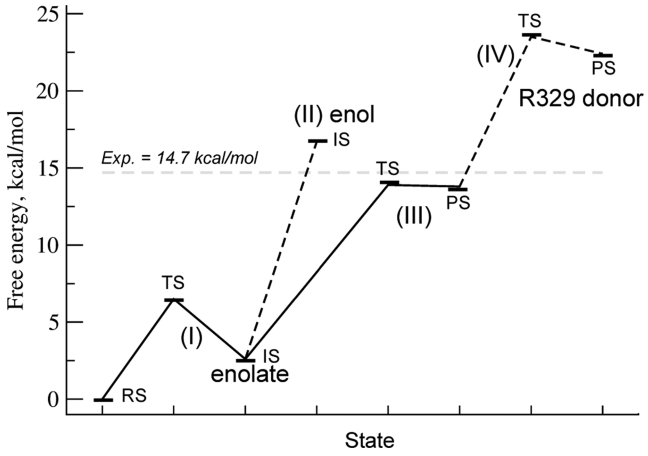


Figure 5. Free energies of the TS and product structures for the proposed mechanisms shown in Figure 4. The free energies were obtained from FEP reaction path simulations with the large QC region B3LYP/def2-TZVP hybrid potential.

kcal mol^{-1} more stable than the TS. The latter values are very similar if a small QC region MP2/aug-cc-pVDZ hybrid potential is employed (Table S1) and are also in reasonable agreement with the previous study of van der Kamp et al.[15] who reported energies of 13.1 and 9.2 kcal mol^{-1} for the TS and product, respectively, using the same, smaller QC region and a SCS-MP2/aug-cc-pVDZ QC/MM hybrid potential on top of B3LYP/6-31+G(d) QC/MM optimized geometries. The FEP and PB (ϵ_p = 1) free energy methods lower the TS and product energies by ∼2−3 kcal mol^{-1}, whereas the PB (ϵ_p = 2) method increases them slightly, although to values that are still lower than those obtained with the small QC region.

Mechanism II is similar to Mechanism I but also involves a proton transfer from the Nδ of His274 to acetyl-CoA. The computed NEB profile shows that the reaction is not concerted and passes via the enolate intermediate, thereby making the first step of this reaction identical to mechanism I. The second step of the reaction, formation of the enol, is uphill and barrierless with the enol product being 13.4 kcal mol^{-1} less stable than the enolate, and 18.9 kcal mol^{-1} higher in energy than the reactants. The free energy methods reduce this value somewhat with the most pronounced reduction being to 15.6 kcal mol^{-1} for the PB (ϵ_p = 1) method. Nevertheless, formation of the enol is strongly disfavored with respect to the enolate, in agreement with previous simulation studies of this reaction.

3.2.2. Formation of Citryl-CoA during the Condensation Step. For the condensation step catalyzed by CS, we also considered two mechanisms. As in the previous section, residue His238 was in its neutral form and the water molecule, W619, was present. For the optimizations, we employed both the small and large QC regions that are described in the Methods section. We took the starting structures in both mechanisms to be the product structure of mechanism I (Figure S1) but used different end structures (Figure S2). No assumptions were made about the order of the events in either reaction path, which were optimized independently of each other.

Mechanism III is a nucleophilic attack on the carbonyl carbon of OAA by the enolate intermediate formed in the enolization step. The computed energies of the TS and product relative to the reactant for this mechanism with the larger QC region are 15.0 and 14.1 kcal mol^{-1}, respectively. The free energy values of these structures are lower by ∼1 kcal mol^{-1}

dx.doi.org/10.1021/jp412346g | J. Phys. Chem. B 2014, 118, 4505−4513

with the FEP approach and between 4 and 5 kcal mol^{-1} with the PB methods. With the smaller QC region used previously by van der Kamp et al.,[5] the reaction is uphill with the product for this step being higher in energy than the products by 18.6 kcal mol^{-1}. This is in agreement with the previous study,[5] but disagrees with the result obtained using the larger QC region. This shows the importance of including the charged residues that interact with the ligand in the QC region so as account properly for the strong polarization effects that are present. We note, in fact, that the "product" structure for mechanism III obtained with the small QC region hybrid potential model is a TS (see Table S2) and that we were unable to find a stable product minimum at this level of theory.

Mechanism IV, in contrast to mechanism III, requires a proton transfer from residue Arg329 to citryl-CoA, in addition to nucleophilic attack on the carbonyl carbon of OAA by the enolate intermediate. It is essentially the same mechanism as the one investigated by van der Kamp et al.[5] With the larger QC region, the reaction is stepwise. The first step is equivalent to mechanism III and involves nucleophilic attack on the carbonyl carbon of OAA by the enolate intermediate, whereas the second step is proton transfer from residue Arg329 to citryl-CoA. The computed barrier for the proton transfer is 7.0 kcal mol^{-1} with respect to the intermediate from the first step and the product is just 0.4 kcal mol^{-1} more stable than the TS. The overall barrier is thus 21.1 kcal mol^{-1} and the product is a high-energy conformation, 20.7 kcal mol^{-1} less stable than reactants. The FEP free energy method increases both these values by ∼2 kcal mol^{-1}, whereas the PB approaches lower them quite substantially, by approximately 10 kcal mol^{-1}, although without changing the relative ordering of the mechanisms.

In contrast to the large QC region results, the small QC region NEB profile indicates a concerted reaction, due to the fact that the product of mechanism III is not stable with this potential. The barrier in this case is 19.4 kcal mol^{-1} with the product being 18.0 kcal mol^{-1} higher in energy than the reactants. The MP2/aug-cc-pVDZ small region QC energy (Table S1) for the TS is similar to that with the B3LYP/def2-TZVP model, whereas the product energy is 4 kcal mol^{-1} less, thereby mirroring the behavior of the product energies with the same potentials in mechanism III.

Consideration of the range of models in Tables 1 and S1 indicates that the barrier for mechanism IV is higher than that for mechanism III and also higher than the experimental estimate to the overall barrier of 14.7 kcal mol^{-1}.[3,12] In our calculations, this discrepancy is due to the proton transfer. In solution, this transfer is favored as the pK_a of citryl-CoA in solvent is about 16,[47] whereas the pK_a of arginine is 3.5 units lower at 12.5.[48] Nonetheless, in the protein, this reaction appears to us unlikely, as the positive charge of Arg329 is stabilized by interactions with the negative residue Asp327, with the distance between the Oδ of Asp327 and NH2 of Arg329 being 3.1 Å in the product of mechanism III. In addition the oxygen of citryl-CoA, which makes a hydrogen bond with the hydrogen of Arg329, is not in the proximity of any negatively charged residues in the protein, which could destabilize the oxyanion and favor proton transfer.

This proton transfer was hypothesized to be important for the change from ligase to hydrolase activity of the CS enzyme,[5] as the weaker interaction between Arg329 and citrate would allow for the reopening of the protein that has been postulated to be required for hydrolysis.[49] We note that no work has established that conformational changes take place when ligands are bound and which convert the function of CS from a ligase to a hydrolase. In addition, a recent crystal structure (PDB reference code 2R9E) of CS complexed with a citryl-CoA analogue did not show any conformational changes.

3.3. Effect of the Protonation of His238 and the Presence of the Water Molecule on the Mechanism. To see the effect of the protonation of His238, we reoptimized the NEB pathways for mechanisms I, III, and IV but without the water molecule, W619, and with a biprotonated His238. To evaluate the effect of the presence of the water molecule, W619, we took the optimized NEB pathways from sections 3.2.1 and 3.2.2 that were optimized in the presence of W619 and a neutral His238, and performed single-point calculations after W619 was removed. Table 2 summarizes the results and shows that the choice of the His238 protonation state affects the energies significantly.

Table 2. A Comparison of the Energies of the TS and Product Structures from the Mechanisms of Enolization (I) and Condensation (III, IV) Calculated with Neutral His238 and the Water W619 and with Cationic His238 and No W619.[a]

mechanism	structure	H238$^+$	H238δ	H238δ/W619
I	TS	6.3	8.1	8.2
	Product	2.0	5.6	5.5
III	TS	5.1	15.8	15.0
	Product	2.3	15.0	14.1
IV	TS	5.7	21.0	21.1
	Product	0.8	19.8	20.7

[a]All energies are in kcal mol^{-1} and are relative to the reactant species. All calculations were done with the large QC region B3LYP/def2-TZVP hybrid potential.

In mechanism I, the computed barrier for the first step is 1.9 kcal mol^{-1} lower with a biprotonated His238 than when it is neutral, and the enolate product is also 3.5 kcal mol^{-1} more stable due to the favorable interactions with the extra positive charge. As a comparison, we also computed the barrier for the enolization step with zeroed His238 charges in the presence of the water molecule, W619. The computed barrier is just 0.5 kcal mol^{-1} higher than that computed with the neutral His238 charges (data not shown).

The protonation of His238 has a more profound effect on the condensation step. The biprotonated His238 stabilizes the TS for this step by 9.9 kcal mol^{-1} and the product by 11.8 kcal mol^{-1}. This is because the positive charge of the histidine favors transfer of the negative charge of the enolate intermediate to citryl-CoA, which hydrogen bonds to the Nδ of His238.

The barrier for the enolization step computed in the absence of W619 and a neutral His238 is 0.1 kcal mol^{-1} lower than the barrier computed in its presence. Likewise, the enolate intermediate is less stable by 0.1 kcal mol^{-1} without the water. Therefore, the presence of the water molecule W619 does not affect the energies significantly for the enolization step.

For mechanism III of the condensation step of the reaction, the TS and product are 0.8 and 0.9 kcal mol^{-1} more stable in the presence of the water, respectively, whereas for mechanism IV, the TS and product are actually slightly less stable in its presence, with values of 0.1 and 0.9 kcal mol^{-1}, respectively. Overall, however, the water appears to have relatively little effect on the energetics of the critical points.

dx.doi.org/10.1021/jp412346g | J. Phys. Chem. B 2014, 118, 4505−4513

3.4. Catalytic Role of the Residue R329. Residue R329 has been postulated to have various roles in the reaction.[4,5] To analyze its electrostatic influence, we performed calculations on the structures previously obtained for the preferred mechanisms' critical points, but with the atomic charges of R329 set to zero and its side chain included in the MM region. The resulting values are given in Table 3.

Table 3. Energies of Selected TS and Product Structures from Mechanisms I (Enolization) and III (Condensation) of the CS Reaction in the Presence of a Neutral His238 and the Water W619[a]

		energies	
mechanism	structure	full charges on R329	zero charges on R329
I	TS	8.4	12.1
	Product	5.8	10.7
III	TS	17.5	35.9
	Product	17.8	41.0

[a]All energies are in kcal mol^{-1} and are relative to the reactant species. Energies were determined in explicit solvent using the large QC region B3LYP/def2-TZVP hybrid potential with and without charges on the atoms of the residue R329.

For both the enolization and condensation mechanisms, I and III, it can be seen that a charged R329 actually significantly decreases the barriers to reaction and also strongly stabilizes the products that are formed by 4.9 kcal mol^{-1} for the enolate product of mechanism I, and by 23.2 kcal mol^{-1} for the product of the condensation reaction of mechanism III, due to the strongly favorable long-range electrostatic interactions between the negative citryl-CoA and the positive Arg329. Overall, therefore, it appears that R329 has an important stabilizing role on the species arising during the catalysis although it does not participate directly in the reaction.

4. DISCUSSION

In this paper, we have studied the mechanisms of enolization and condensation catalyzed by the enzyme, citrate synthase, using a mixture of simulation techniques, including MDFE simulations and high-level QC/MM hybrid potentials in conjunction with a NEB reaction-path-finding algorithm and free energy perturbation calculations.

Our MDFE simulations strongly suggest that the important active site residue, His238, is in its neutral form and that a water molecule, which is present in some CS structures, should also be present. Nevertheless, in our reaction path calculations, we examined mechanisms in which His238 was in its cationic as well as its neutral states and with the water absent as well as present. For the enolization step, the favored mechanism with a neutral His238 was mechanism I, which leads to the charged enolate intermediate. Mechanism II, in which the enol is formed by a proton transfer from His274, is highly unfavorable due to the instability of the product.

The favored mechanism for the condensation step is mechanism III, in which there is direct nucleophilic attack of the enolate on OAA. The residue Arg329 is not directly implicated in this reaction as there is no proton transfer to the citryl-coenzyme A product which remains in its deprotonated form. However, the arginine plays a significant role by stabilizing both the TS and the product of the condensation reaction. Likewise, analysis shows that the barrier to condensation is actually lower with a biprotonated His238, although this is not supported by the available experimental data for which the overall barrier to reaction is 14.7 kcal mol^{-1}.[3,12] This latter value is consistent with the values of 6.5 and 13.9 kcal mol^{-1} that we find for the free energies of the enolization and condensation steps of the reaction, respectively, with the condensation step being rate-limiting.

Since the mechanism of the reaction appears to be universal across all (S−)citrate synthases,[6,7] it is likely that the pathways of enolization and condensation that we have elucidated here are of general relevance. However, there is clearly scope for future work with the hydrolysis step of the reaction deserving particular attention. Thus, it would be helpful to determine whether this step occurs with a deprotonated citryl-coenzyme A as starting structure, as predicted here, or with a protonated form, and what role, if any, conformational changes of the protein might play in initiating hydrolyase activity.

■ ASSOCIATED CONTENT

Supporting Information

Details of the normal mode frequencies and structures of the critical points along the reaction paths; ancillary data obtained with different QC/MM partitionings and QC levels of theory. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: alexey.aleksandrov@polytechnique.fr.
*E-mail: martin.field@ibs.fr.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

A.A. would like to thank Adrian Mulholland for clarifications concerning aspects of his and his colleagues' simulations of the CS reaction mechanism.

■ REFERENCES

(1) Remington, S. Structure and Mechanism of Citrate Synthase. *Curr. Top. Cell. Regul.* **1992**, *33*, 209−229.
(2) Alter, G.; Casazza, J.; Zhi, W.; Nemeth, P.; Srere, P.; Evans, C. Mutation of Essential Catalytic Residues in Pig Citrate Synthase. *Biochemistry* **1990**, *29*, 7557−7563.
(3) Evans, C. T.; Kurz, L. C.; Remington, S. J.; Srere, P. A. Active Site Mutants of Pig Citrate Synthase: Effects of Mutations on the Enzyme Catalytic and Structural Properties? *Biochemistry* **1996**, *35*, 10661−10672.
(4) Man, W. J.; Li, Y.; O'Connor, C. D.; Wilton, D. C. Conversion of Citrate Synthase into Citryl-CoA Lyase as a Result of Mutation of the Active-Site Aspartic Acid Residue to Glutamic Acid. *Biochem. J.* **1991**, *280*, 521−526.
(5) van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. High-level QM/MM Modelling Predicts an Arginine as the Acid in the Condensation Reaction Catalysed by Citrate Synthase. *Chem. Commun.* **2008**, 1874−1876.
(6) Sutherland, K. J.; Henneke, C. M.; Towner, P.; Hough, D. W.; Danson, M. J. Citrate Synthase from the Thermophilic Archaebacterium Thermoplasma Acidophilum. *Eur. J. Biochem.* **1990**, *194*, 839−844.
(7) Russell, R. J. M.; Ferguson, J. M. C.; Hough, D. W.; Danson, M. J.; Taylor, G. L. The Crystal Structure of Citrate Synthase from the Hyperthermophilic Archaeon Pyrococcus Furiosus at 1.9 Å Resolution. *Biochemistry* **1997**, *36*, 9983−9994.

(8) Mulholland, A. J.; Richards, W. G. Acetyl-CoA Enolization in Citrate Synthase: A Quantum Mechanical/Molecular Mechanical (QM/MM) Study. *Proteins* **1997**, *27*, 9−25.
(9) Karpusas, M.; Branchaud, B.; Remington, S. J. Proposed Mechanism for the Condensation Reaction of Citrate Synthase: 1.9-Å Structure of the Ternary Complex with Oxaloacetate and Carboxymethyl Coenzyme A. *Biochemistry* **1990**, *29*, 2213−2219.
(10) Mulholland, A.; Lyne, P.; Karplus, M. Ab Initio QM/MM Study of the Citrate Synthase Mechanism. A Low-Barrier Hydrogen Bond Is Not Involved. *J. Am. Chem. Soc.* **2000**, *122*, 534−535.
(11) Donini, O.; Darden, T.; Kollman, P. A. QM-FE Calculations of Aliphatic Hydrogen Abstraction in Citrate Synthase and in Solution: Reproduction of the Effect of Enzyme Catalysis and Demonstration that an Enolate Rather than an Enol Is Formed. *J. Am. Chem. Soc.* **2000**, *122*, 12270−12280.
(12) Kurz, L. C.; Constantine, C. Z.; Jiang, H.; Kappock, T. J. The Partial Substrate Dethiaacetyl-Coenzyme A Mimics All Critical Carbon Acid Reactions in the Condensation Half-Reaction Catalyzed by Thermoplasma Acidophilum Citrate Synthase. *Biochemistry* **2009**, *48*, 7878−7891.
(13) van der Kamp, M.; Perruccio, F.; Mulholland, A. Substrate Polarization in Enzyme Catalysis: QM/MM Analysis of the Effect of Oxaloacetate Polarization on Acetyl-CoA Enolization in Citrate Synthase. *Proteins* **2007**, *69*, 521−535.
(14) Bjelic, S.; Brandsdal, B.; Aqvist, J. Cold Adaptation of Enzyme Reaction Rates. *Biochemistry* **2008**, *47*, 10049−10057.
(15) van der Kamp, M. W.; Ziurek, J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Testing High-Level QM/MM Methods for Modeling Enzyme Reactions: Acetyl-CoA Deprotonation in Citrate Synthase. *J. Phys. Chem. B* **2010**, *114*, 11303−11314.
(16) Karpusas, M.; Holland, D.; Remington, S. 1.9-A Structures of Ternary Complexes of Citrate Synthase with D- and L-Malate: Mechanistic Implications. *Biochemistry* **1991**, *30*, 6024−6031.
(17) Usher, K.; Remington, S.; Martin, D.; Drueckhammer, D. A Very Short Hydrogen Bond Provides only Moderate Stabilization of an Enzyme-Inhibitor Complex of Citrate Synthase. *Biochemistry* **1994**, *33*, 7753−7759.
(18) Yang, W.; Drueckhammer, D. G. Computational Study of the Citrate Synthase Catalyzed Deprotonation of Acetyl-Coenzyme A and Fluoroacetyl-Coenzyme A: Demonstration of a Layered Quantum Mechanical Approach. *J. Phys. Chem. B* **2003**, *107*, 5986−5994.
(19) van der Kamp, M. W.; Perruccio, F.; Mulholland, A. J. Ab Initio QM/MM Modelling of Acetyl-CoA Deprotonation in the Enzyme Citrate Synthase. *J. Mol. Graphics Modell.* **2007**, *26*, 676−690.
(20) Beglov, D.; Roux, B. Finite Representation of an Infinite Bulk System: Solvent Boundary Potential for Computer Simulations. *J. Chem. Phys.* **1994**, *100*, 9050−9063.
(21) Simonson, T. Electrostatic Free Energy Calculations for Macromolecules: a Hybrid Molecular Dynamics/Continuum Electrostatics Approach. *J. Phys. Chem. B* **2000**, *104*, 6509−6513.
(22) Mackerell, A.; Bashford, D.; Bellott, M.; Dunbrack, R.; Evanseck, J.; Field, M.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; et al. An All-Atom Empirical Potential for Molecular Modelling and Dynamics Study of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(23) Foloppe, N.; MacKerell, A. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(24) Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(25) Stote, R.; States, D.; Karplus, M. On the Treatment of Electrostatic Interactions in Biomolecular Simulation. *J. Chem. Phys.* **1991**, *88*, 2419−2433.
(26) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(27) Gill, P. M. W.; Johnson, B. G.; Pople, J. A.; Frisch, M. J. The Performance of the Becke-Lee-Yang-Parr (B-LYP) Density Functional Theory with Various Basis Sets. *Chem. Phys. Lett.* **1992**, *197*, 499−505.
(28) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571−2577.
(29) Grimme, S. Semiempirical GGA-type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(30) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(31) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(32) Field, M. J. The pDynamo Program for Molecular Simulations Using Hybrid Quantum Chemical and Molecular Mechanical Potentials. *J. Chem. Theory Comput.* **2008**, *4*, 1151−1161.
(33) Neese, F. ORCA—an Ab Initio, Density Functional and Semiempirical Program Package, Version 2.6; University of Bonn, 2008.
(34) Aleksandrov, A.; Field, M. Efficient Solvent Boundary Potential for Hybrid Potential Simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10503−10509.
(35) Aleksandrov, A.; Field, M. A Hybrid Elastic Band String Algorithm for Studies of Enzymatic Reactions. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12544−12553.
(36) Sham, Y.; Chu, Z.; Warshel, A. Consistent Calculations of pK_a's of Ionizable Residues in Proteins: Semi-Microscopic and Microscopic Approaches. *J. Phys. Chem. B* **1997**, *101*, 4458−4472.
(37) Simonson, T.; Carlsson, J.; Case, D. A. Proton Binding to Proteins: pK_a Calculations with Explicit and Implicit Solvent Models. *J. Am. Chem. Soc.* **2004**, *126*, 4167−4180.
(38) Simonson, T. Macromolecular Electrostatics: Continuum Models and Their Growing Pains. *Curr. Opin. Struct. Biol.* **2001**, *11*, 243−252.
(39) Aleksandrov, A.; Thompson, D.; Simonson, T. Alchemical Free Energy Simulations for Biological Complexes: Powerful but Temperamental. *J. Mol. Recognit.* **2010**, *23*, 117−127.
(40) Aleksandrov, A.; Proft, J.; Hinrichs, W.; Simonson, T. Protonation Patterns in Tetracycline:Tet Repressor Recognition: Simulations and Experiments. *ChemBioChem* **2007**, *8*, 675−685.
(41) Darden, T. In *Computational Biochemistry & Biophysics*; Becker, O., Mackerell, A., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; Chapter 4.
(42) Phillips, J.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(43) Breneman, C. M.; Wiberg, K. B. Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *J. Comput. Chem.* **1990**, *11*, 361−373.
(44) Kästner, J.; Senn, H. M.; Thiel, S.; Otte, N.; Thiel, W. QM/MM Free-Energy Perturbation Compared to Thermodynamic Integration and Umbrella Sampling: Application to an Enzymatic Reaction. *J. Chem. Theory Comput.* **2006**, *2*, 452−461.
(45) Zhang, Y.; Liu, H.; Yang, W. Free Energy Calculation on Enzyme Reactions with an Efficient Iterative Procedure to Determine Minimum Energy Paths on a Combined Ab Initio QM/MM Potential Energy Surface. *J. Chem. Phys.* **2000**, *112*, 3483−3492.
(46) Warshel, A.; Naray-Szabo, G. *Computational Approaches to Biochemical Reactivity*; Kluwer Academic Publishers: New York, 1997.
(47) Takahashi, S.; Cohen, L. A.; Miller, H. K.; Peake, E. G. Calculation of the pK_a Values of Alcohols from σ^* Constants and from the Carbonyl Frequencies of Their Esters. *J. Org. Chem.* **1971**, *36*, 1205−1209.

(48) McElroy, W.; Glass, B. *Phosphorus Metabolism*; Johns Hopkins University Press: Baltimore, MD, 1951; Vol. *1*.
(49) Man, W.; Li, Y.; O'Connor, C.; Wilton, D. The effect of Replacing the Conserved Active-Site Residues His-264, Asp-312 and Arg-314 on the Binding and Catalytic Properties of Escherichia coli Citrate Synthase. *Biophys. J.* **1994**, *300*, 765−770.