

Comparison of QM-Only and QM/MM Models for the Mechanism of Tungsten-Dependent Acetylene Hydratase

Rong-Zhen Liao and Walter Thiel*

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

S Supporting Information

ABSTRACT: We report a comparison of QM-only and QM/MM approaches for the modeling of enzymatic reactions. For this purpose, we present a QM/MM case study on the formation of vinyl alcohol in the catalytic cycle of tungsten-dependent acetylene hydratase. Three different QM regions ranging from 32 to 157 atoms are designed for the reinvestigation of the previously suggested one-water attack mechanism. The QM/MM calculations with the minimal QM region **M1** (32 atoms) yield a two-step reaction profile, with an initial nucleophilic attack followed by the protonation of the formed vinyl anion intermediate, as previously proposed on the basis of QM-only calculations on cluster model **M2** (116 atoms); however, the overall QM/MM barrier with **M1** is much too high, mainly due to an overestimate of the QM/MM electrostatic repulsions. QM/MM calculations with QM region **M2** (116 atoms) fail to reproduce the published QM-only results, giving a one-step profile with a very high barrier. This is traced back to the strong electrostatic influence of the two neighboring diphosphate groups that were neglected in the QM-only work but are present at the QM/MM level. These diphosphate groups and other electrostatically important nearby residues are included in QM region **M3** (157 atoms). QM/MM calculations with **M3** recover the two-step mechanism and yield a reasonable overall barrier of 16.7 kcal/mol at the B3LYP/MM level. They thus lead to a similar overall mechanistic scenario as the previous QM-only calculations, but there are also some important variations. Most notably, the initial nucleophilic attack becomes rate limiting at the QM/MM level. A modified two-water attack mechanism is also considered but is found to be less favorable than the previously proposed one-water attack mechanism. Detailed residue interaction analyses and comparisons between QM/MM results with electronic and mechanical embedding and QM-only results without and with continuum solvation show that the protein environment plays a key role in determining the mechanistic preferences in acetylene hydratase. The combined use of QM-only and QM/MM methods provides a powerful approach for the modeling of enzyme catalysis.

1. INTRODUCTION

The use of quantum chemical methods to address enzymatic reaction mechanisms has become a booming area in enzymology.^{1–5} Currently, there are two popular approaches, the quantum mechanics-only (QM-only) method^{6–9} using a rather small model of the active site and the hybrid quantum mechanics/molecular mechanics (QM/MM) method^{10–14} using the entire protein as the model. Both methods have been successfully applied to the study of various classes of enzymes, and in many cases similar results and conclusions have been obtained.^{2,3}

In the QM-only approach, commonly also called the cluster approach,^{6–9} a model of the active site, nowadays typically in the order of 100 to 150 atoms, is designed on the basis of available crystal structures. Density functional theory methods are normally used for the calculation of the geometries and energies of all stationary points along the reaction pathways. The missing steric and electrostatic effects from the remaining part of the protein are considered by two simple procedures. The steric effects imposed by the protein matrix are taken into account by locking certain key atoms at the periphery of the cluster model. The electrostatic effects are modeled by the dielectric cavity method with a dielectric constant of 4. Recent studies of four different kinds of enzymes indicate that the solvation effects saturate at a model size of 150–200 atoms and that the particular choice of dielectric constant is then no longer of much concern.^{15–18} Zero-point vibrational effects can

be included on the basis of the calculated harmonic frequencies, while entropy effects are usually neglected because it is difficult to predict them accurately in the harmonic approximation. Since the size of cluster models is relatively small, one can have full control of the system, and it is still possible to explore all relevant stationary points even though multiple minima problems become more challenging the larger the chosen model. On the basis of extensive calculations for a large number of enzymes, the error of the cluster approach for modeling metalloenzymes has been assessed by Siegbahn to be less than 5 kcal/mol.¹⁹

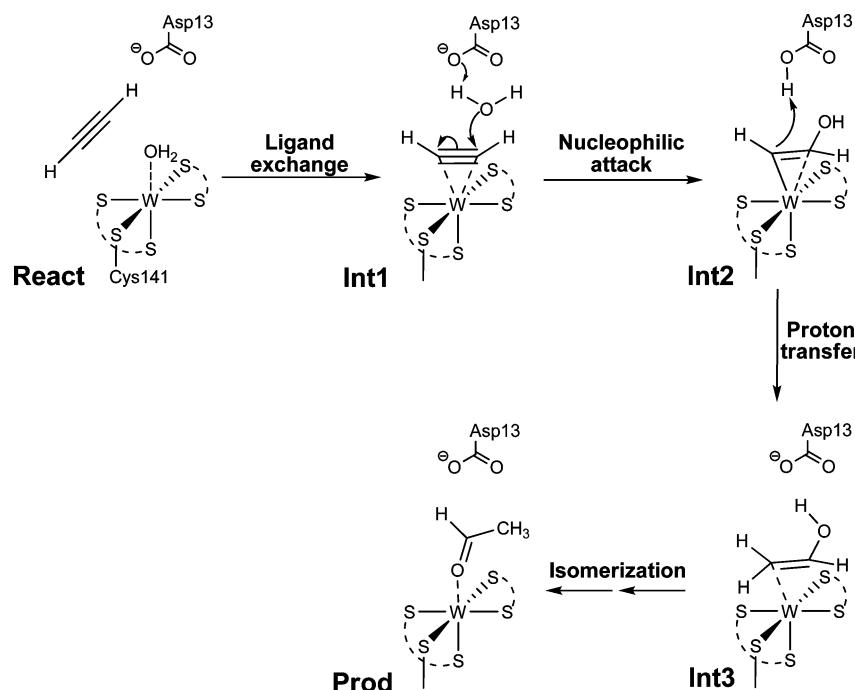
In the alternative QM/MM approach,^{10–14} the whole solvated protein is chosen as the model, thus explicitly incorporating the steric and electrostatic effects of the environment. The system is divided into two parts: the QM region where the reaction takes place, and the MM region that represents the protein environment. The most common approach to handle the QM/MM boundary is to introduce hydrogen link atoms to satisfy the valence of each bond being cut. The interactions between the QM and MM subsystems are normally treated by the electronic embedding scheme. The MM point charges are inserted into the QM Hamiltonian to capture the QM/MM electrostatic interactions, while the QM/

Special Issue: Wilfred F. van Gunsteren Festschrift

Received: January 30, 2012

Published: March 6, 2012



Scheme 1. Reaction Mechanism of Acetylene Hydratase Suggested in a Previous QM-Only Study⁴⁵

MM dispersion interactions are calculated classically, normally using standard MM parameters for the QM atoms. Free-energy profiles can be modeled by free-energy perturbation theory or related methods.¹⁴ Particular attention needs to be paid to the control of protein conformations to avoid problems arising from multiple local minima. In a recent case study on the reaction mechanism of tyrosinase,²⁰ the discrepancy between published QM-only and QM/MM results was resolved by showing that the previous QM/MM geometry optimizations suffered from convergence to an incorrect reactant state and to different local minima along the reaction pathway;²¹ once this problem was avoided by an iterative optimization procedure, both methods gave similar results.²⁰

The QM/MM approach is technically more complicated than the cluster approach. Different setups may lead to different results. Many studies have addressed the proper choice of the QM region and its influence on the QM/MM results.^{22–29} In principle, larger QM regions should give better results, but any selection of the QM/MM partition needs to be electronically balanced, with a proper screening of electrostatic interactions. The sensitivity of the QM/MM results with regard to the size of the QM region has been investigated for various properties of enzymes, including reaction energetics (relative energies of transition states, intermediates, and product),^{22–29} NMR chemical shifts,^{30–33} hyperfine coupling constants,³⁴ IR spectroscopy,^{35,36} and UV spectroscopy.^{37–39} It is generally accepted that QM regions with different sizes should be tested, if possible, to establish the internal consistency of the results.

There is an ongoing discussion which of the two approaches is preferable for studying enzymatic reactions. If the QM region can be chosen large enough to cover all relevant short- and long-range interactions, both approaches should give essentially the same results. Therefore, one central question is at which size of the QM system the energies and other properties (see above) converge in both approaches. Ochsenfeld and co-workers studied the convergence of isomerization energies with respect to the size of the QM region for three proteins.^{40,41}

They showed that the QM/MM approach converges considerably faster than the QM-only approach. For example, in their study of a twin arginine pair in adenovirus Ad11, QM/MM calculations with 437 (84) QM atoms gave a difference of only 3 (15) kJ/mol compared to the QM/MM results with 1035 QM atoms, while the corresponding deviation for the QM-only approach was about 30 (110) kJ/mol.⁴¹ Ryde et al. investigated the reaction energies of a proton transfer in [Ni,Fe] hydrogenase with different sizes of the QM regions in both QM-only and QM/MM calculations.^{42,43} They also found a better convergence behavior for the QM/MM method when the junctions between the QM and MM parts were properly selected.

In this paper, we present a comparison of the QM-only and QM/MM models for the mechanism of tungsten-dependent acetylene hydratase.⁴⁴ The cluster approach has previously been applied to elucidate the reaction mechanism of this enzyme.⁴⁵ The mechanism involves direct coordination of the acetylene to the tungsten ion, followed by nucleophilic attack by a water molecule concomitant with a proton transfer to an ionized second-shell aspartate, which then delivers the proton back to the substrate to form vinyl alcohol (Scheme 1). Thus a fast isomerization from vinyl alcohol to acetaldehyde takes place with the help of the tungsten ion and the ionized aspartate. The suggested mechanism has been used to rationalize the chemoselectivity of this enzyme, which is not capable of hydrating propyne, ethylene, and acetonitrile.⁴⁶ Interestingly, a similar mechanism has also been advocated for acetylene hydration promoted by a biomimetic tungsten complex.⁴⁷ Here we are mainly interested in the methodological aspects related to enzyme modeling, and we will thus focus only on the vinyl alcohol formation process (from React to Int3) in the proposed mechanism. Different QM regions ranging from 32 to 157 atoms are used in the QM/MM calculations. A modified mechanism involving an additional water molecule for the proton transfer is also considered.

2. METHODS

2.1. System Preparation. The initial coordinates were taken from the X-ray structure (PDB entry: 2E7Z, resolution: 1.26 Å) obtained for *Pelobacter acetylinicus*.⁴⁴ The missing residues (Met1-Ala2-Ser3) were added with the Swiss-PdbViewer.⁴⁸ The protonation states of the titratable residues (His, Asp, Glu) were determined on the basis of their pK_a values calculated by the empirical PROPKA⁴⁹ program and verified by visual inspection of the local hydrogen-bond networks.

Missing hydrogen atoms in the crystal structure were added using the CHARMM program package.⁵⁰ Parameters for the tungsten–pterin complexes were adopted from previous calculations on molybdenum-dependent aldehyde oxidoreductase and xanthine oxidase^{51–53} and modified if necessary (charges were recalculated). The parameters for acetylene were taken from the literature.⁵⁴ All nonstandard parameters are documented in the Supporting Information.

The system was hydrated using the droplet model with a 40 Å sphere of equilibrated TIP3 water molecules centered at tungsten, and all water molecules with their oxygen atoms within 2.8 Å of any protein heavy atoms were deleted. Energy minimization and a 50 ps molecular dynamics (MD) simulation at 320 K (optimum temperature for activity at 50 °C⁵⁵) were performed using the CHARMM force field⁵⁶ as implemented in the CHARMM program.⁵⁰ A stochastic boundary potential⁵⁷ was applied to maintain the overall structure of the water sphere, and all nonwater molecules were kept frozen during both the energy minimization and equilibration. This hydration procedure was performed iteratively until the number of added water molecules became roughly constant. To neutralize the system, 10 magnesium counterions were added at the protein surface by employing the CHARMM tool. All Mg²⁺ ions were located more than 25 Å away from the tungsten ion in the active site. During the 500 ps MD simulation, residues more than 30 Å away from the tungsten atom and the nonhydrogen atoms of the cofactor tungsten complex were fixed. The equilibrated system was composed of 30 942 atoms, including 6461 TIP3 water molecules.

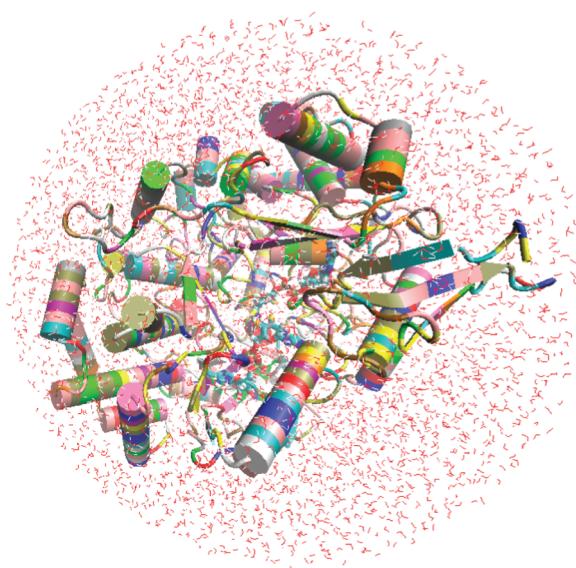


Figure 1. Acetylene hydratase with solvent shell (water ball with radius of 40 Å).

Five snapshots (AS100, AS200, AS300, AS400, and AS500) were selected for the QM/MM calculations of the ligand exchange step. The optimized **Int1** structure of Snap500 (AS500) was further equilibrated for 500 ps to check for protein structural changes during the formation of the enzyme–substrate complex. Three snapshots from this MD run (BS100, BS300, and BS500) were chosen for the study of the vinyl alcohol formation.

2.2. QM/MM Calculations. The selected snapshots from the MD trajectories were taken as the initial structures for the QM/MM calculations using the ChemShell^{58,59} package, where the TURBOMOLE⁶⁰ program was used for the QM treatment and DL_POLY⁶¹ for handling the MM part using the CHARMM topology and parameter data. The electronic embedding scheme was employed for the QM region. No electrostatic cutoff was introduced for the MM/MM and QM/MM interactions. Hydrogen link atoms in combination with the charge-shift scheme were applied to treat the QM/MM boundary. Geometry optimizations were carried out using the HDLC optimizer,⁶² and an active region was selected including the QM atoms and all residues and water molecules in the MM region within 15 Å around the tungsten center (Fe₄S₄ excluded, see Supporting Information for details).

Three different QM regions ranging from 32 to 157 atoms were devised for the QM/MM calculations. The QM region **M1** with 32 atoms (total charge of -2) represents the smallest model, including the tungsten-cofactor model [W(S₂C₂H₂)-(SCH₃)(OH₂)]⁻, the second-shell Asp13 residue (CH₃COO⁻), and the substrate acetylene (C₂H₂). The QM region **M2** is exactly the same as that in the previous QM-only calculations,⁴⁵ containing 116 atoms with a total charge of -1. The largest QM region **M3** is composed of 157 atoms (total charge of -3), selected on the basis of the magnitude of the electrostatic contribution of each active-site residue on the reaction energetics in the QM/MM calculations with QM region **M1** (see Supporting Information for details). This model is labeled as **M3a** or **M3b** depending on whether one or two water molecules participate in the reaction, respectively.

In order to efficiently handle also the largest QM region (157 atoms), the QM part for all models was treated by the BP86^{63–65} density functional employing the resolution of the identity^{66,67} approximation and the def2-SVP⁶⁸ basis set. For QM regions **M1** and **M2**, final energies were evaluated by performing single-point QM/MM calculations using the functionals BP86, B3LYP^{69,70} (with VWN functional V as in the Gaussian09⁷¹ program), and TPSS⁷² in combination with the larger basis set used in the previous QM-only work⁴⁵ (LANL2TZ(f)⁷³ pseudopotential for W and 6-311+G(2d,2p) for S, O, N, C, and H elements, labeled as BS1). For QM regions **M3a** and **M3b**, the LANL2TZ(f) basis set was used for W and the 6-311+G(d,p) basis set for all other atoms (labeled as BS2), for the sake of computational efficiency. To better understand the influence of the environment on models **M3a** and **M3b**, solvation effects were also calculated at the B3LYP/BS2 level using the SMD⁷⁴ solvation model ($\epsilon = 4$) implemented in the Gaussian09⁷¹ program (as in the previous QM-only study).

To validate the use of the RI-BP86/def2-SVP structures for the relevant stationary points, further QM/MM geometry optimizations with QM region **M1** were performed using the B3LYP functional together with the same basis set as in the previous QM-only calculations⁴⁵ (LANL2TZ(f) pseudopotential for W, 6-311+G(d) for S, and 6-31G(d,p) for C, N, O, and

H, labeled as BS3). These calculations showed that RI-BP86/def2-SVP and B3LYP/BS3 give similar geometric parameters and relative energies including barriers (final energies estimated at the B3LYP/BS1 level).

3. RESULTS

In this section, we present our QM/MM results for the initial ligand exchange step and the subsequent formation of the vinyl alcohol intermediate. We compare them with the results from a recent QM-only (B3LYP/BS3) study⁴⁵ of the cluster model M2. Unless noted otherwise, the quoted QM-only energies were obtained from single-point calculations at optimized gas-phase geometries using the conductor-like polarizable continuum model to account for the environment.⁴⁵

At the beginning of the reaction, the substrate acetylene enters into the active site of the enzyme and displaces the tungsten-bound water molecule to form the Michaelis complex. This ligand exchange step is quite favorable according to the previous QM-only calculation, where the reaction energy for this step was estimated to be exothermic by 5.4 kcal/mol.⁴⁵

In the present MD simulations, the substrate acetylene is initially solvated by water molecules and is located at a distance between 10 to 15 Å away from the tungsten ion. To compute the energetics of ligand exchange at the QM/MM level, we use the same QM region M2 as in the previous QM-only study (116 QM atoms with a total charge of -1 in the QM part). QM/MM reaction energies for the first ligand exchange are listed in Table 1 for the five selected snapshots. The average

The optimized structures of intermediates and transition states for snapshot BS100 are shown in Figure 2. The QM/MM relative energies for all three snapshots are listed in Table 2.

Similar to the previous QM-only calculations, the reaction proceeds through two steps.⁴⁵ The first step is a concerted nucleophilic attack and proton transfer, leading to the vinyl anion intermediate Int2 (Figure 2). At the RI-BP86/def2-SVP level, the barrier for this step is 19.0 kcal/mol, and Int2 lies at +14.4 kcal/mol relative to Int1 (see linear transit scan, Figure S2 in Supporting Information). When a larger basis set is used in single-point calculations (see Table 2), the energy of Int2 increases further relative to that of TS1, which is also analogous to the QM-only results.⁴⁵ The critical geometry parameters of TS1 (Figure 2) obtained from QM/MM calculations are similar to those from QM-only calculations. For instance, the distances C1–O1, O1–H1, and O2–H1 from the QM-only model are 1.88, 1.05, and 1.43 Å, respectively,⁴⁵ while they are 1.88, 1.07, and 1.42 Å, respectively, using the QM/MM model.

For the following step, a carboxylic acid C–O bond rotation from Int2 to Int2b has to take place to facilitate the proton transfer to the vinyl anion intermediate. In the QM-only model (M2), Asp13 is truncated at its C_α position, and a free rotation of its C_α–N bond is viable during the second proton-transfer step.⁴⁵ When the whole enzyme is considered, the protein matrix prohibits such rotation, and thus some extra energy (6.4 kcal/mol, BS100, B3LYP/BS1) is needed to enable the conformational change from Int2 to Int2b. This effect was missed in the QM-only model (M2) used before⁴⁵ but would presumably be recovered by including more groups into the active-site model. Proton delivery from protonated Asp13 in Int2b has to overcome only a rather small barrier, since TS2 lies only 6.7 kcal/mol higher than Int2b (BS100, B3LYP/BS1). However, the overall barrier (TS2 relative to Int1) is larger than 40 kcal/mol at the B3LYP/MM level (see Table 2) compared with 23.0 kcal/mol in the QM-only model. Generally, the QM/MM barriers are highest for QM = B3LYP, followed by TPSS and BP86, but the differences are not very large (see Table 2). Furthermore, the reaction is calculated to be endothermic at the QM/MM level by more than 10 kcal/mol (for all three functionals), while it is exothermic by 11.5 kcal/mol using the QM-only method.⁴⁵

These differences between the QM/MM and QM-only results may arise from the use of different QM regions (M1 vs M2, 32 vs 116 atoms). This can be checked by QM/MM calculations with larger QM regions. To guide the choice of which residues or groups should be included, we estimated their electrostatic contributions to the relative energies of Int1 and Int2 through single-point QM/MM calculations, in which the corresponding MM point charges were set to zero. Int1 and Int2 were selected for this purpose because they show the largest charge redistribution, with Asp13 going from ionized to neutral and the charge of the tungsten core (QM part of the tungsten complex) going from -1 to -2. In the original model, with no charges switched off, the energy of Int2 is 17.4 kcal/mol relative to Int1. The energy changes upon removal of the MM charges of selected residues and groups are displayed in Figure 3 (for numerical results see Figure S5, Supporting Information). Obviously, the two diphosphate groups of the tungsten complex have the biggest effect (-20.3 kcal/mol), reflecting the strong electrostatic repulsion between the tungsten core and the diphosphates (total charge of -4). The second-shell residue Arg606, which forms a hydrogen bond to one of the dithiolene sulfur atoms, also has a significant

Table 1. QM/MM Ligand Exchange Energy (in kcal/mol) of Model M2 (116 Atoms) for Various Snapshots and Functionals

	AS100	AS200	AS300	AS400	AS500
BP86/BS1	-17.4	-20.0	-22.4	-18.6	-22.4
TPSS/BS1	-19.2	-21.6	-23.6	-20.2	-24.2
B3LYP/BS1	-10.7	-14.2	-15.7	-12.2	-16.4

ligand exchange energy is calculated to be -20.2, -21.8, and -13.8 kcal/mol, respectively, when using BP86, TPSS, and B3LYP as the QM component. The QM/MM model thus yields a larger exothermicity than the QM-only model.⁴⁵

As can be seen from Table 1, the conformational diversity of the environment has a moderate effect on the ligand exchange energy, with a maximum difference between different snapshots of about 5–6 kcal/mol for a given density functional. This should be viewed in light of the fact that the local environments of the acetylene and the exchanged water molecule vary considerably during the exchange, since acetylene has to move by more than 10 Å (see above) to reach the coordination site.

In the remainder of this section, we focus on the formation of the vinyl alcohol intermediate (from Int1 to Int3, Scheme 1) in the enzymatic hydration of acetylene. We present the QM/MM results for QM regions ranging from 32 to 157 atoms to investigate how the reaction energetics changes with the size of the QM region. For the largest model M3, we also consider an alternative mechanistic scenario, in which an additional water molecule acts as a bridge for the proton transfer for both steps. Unless noted otherwise, the numerical results given in the following text refer to snapshot BS100.

3.1. QM Region M1. We first consider a minimum model (M1, 32 atoms, total charge of -2) including the first-shell ligands, the base Asp13, one water molecule, and acetylene.

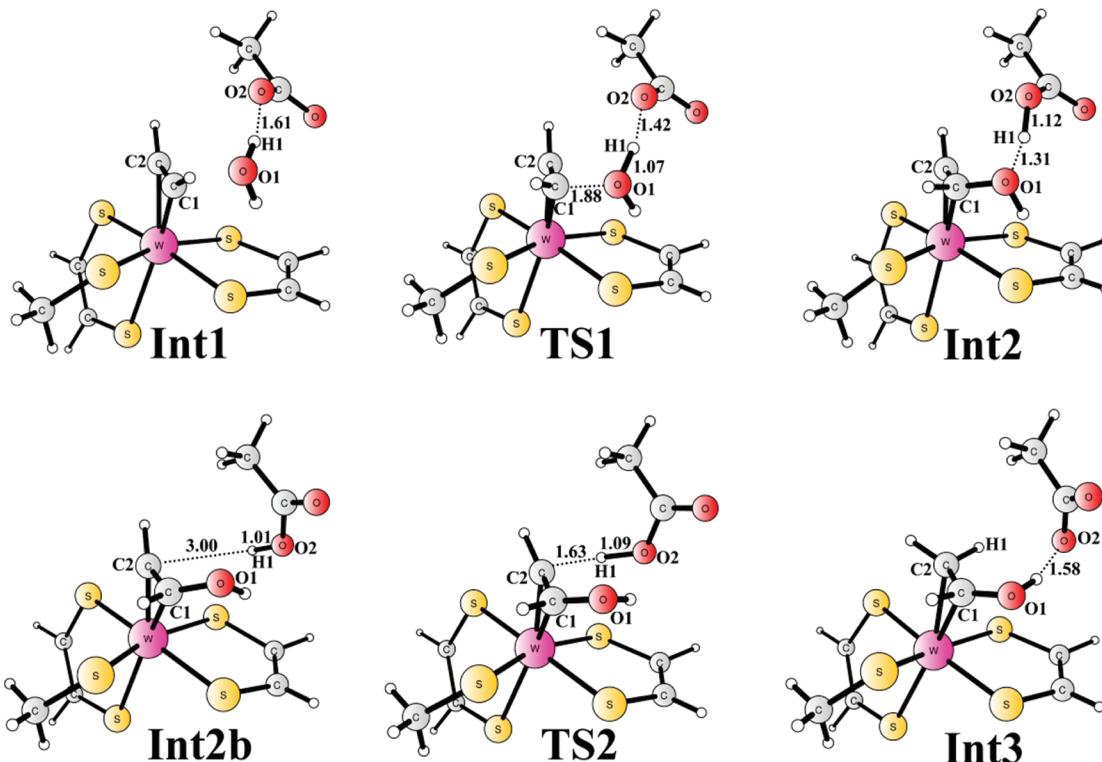


Figure 2. Optimized structures of intermediates and transition states for QM region M1, snapshot BS100, RI-BP86/MM. All bond lengths are given in Å.

Table 2. QM/MM Relative Energies (in kcal/mol)
Calculated for Three Snapshots with QM Region M1 using
Three Different Functionals and the Corresponding Average
Values

	Int1	TS1	Int2	Int2b	TS2	Int3
BS100, BP86/BS1	0	23.0	23.9	31.9	34.4	9.5
BS100, TPSS/BS1	0	23.8	25.7	33.2	36.8	10.4
BS100, B3LYP/BS1	0	28.7	29.6	36.0	42.7	13.5
BS300, BP86/BS1	0	21.8	24.2	35.7	36.8	15.8
BS300, TPSS/BS1	0	22.3	25.6	36.7	38.8	16.8
BS300, B3LYP/BS1	0	28.2	30.6	42.3	45.5	20.4
BS500, BP86/BS1	0	20.7	22.6	28.1	36.9	11.9
BS500, TPSS/BS1	0	21.1	24.0	29.2	38.7	13.2
BS500, B3LYP/BS1	0	26.7	28.7	32.1	44.7	15.1
average, BP86/BS1	0	21.8	23.6	31.9	36.0	12.4
average, TPSS/BS1	0	22.4	25.1	33.0	38.1	13.5
average, B3LYP/BS1	0	27.9	29.6	36.8	44.3	16.3

effect, stabilizing **Int2** by as much as 13.5 kcal/mol. A quantum-classical treatment might not be accurate enough to reliably describe such large electrostatic effects of the surrounding groups on the reaction energetics, and it is thus generally advisable in such a case to enlarge the QM region by including these groups. This guideline has been followed in the definition of our largest QM region **M3** (see Section 3.3). The intermediate QM region **M2** (taken from the published QM-only work)⁴⁵ includes Arg606 and some other residues from Figure 3 (e.g., dihydropyran, Cys12, and Trp179) but not the two diphosphate groups.

3.2. QM Region M2. To allow for a direct and consistent comparison with the published QM-only study, we performed QM/MM reaction pathway calculations for snapshot BS100 with the very same QM region **M2** that had been used in the

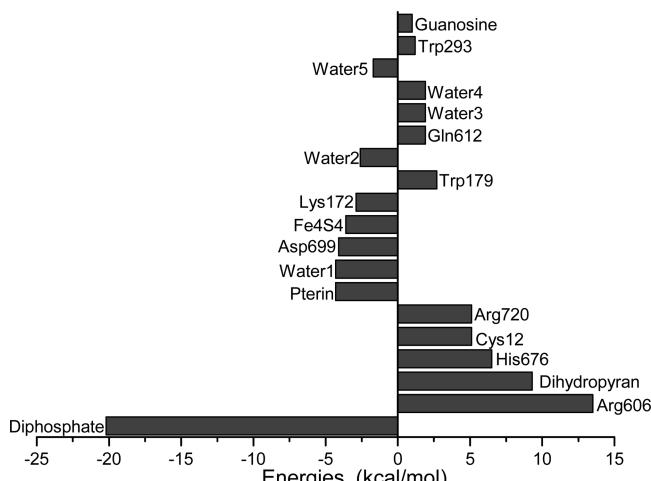


Figure 3. Changes of QM energies of **Int2** by switching off MM point charges at selected residues and groups (BS100, RI-BP86/def2-SVP). A negative (positive) value means that the energy of **Int2** goes down (up).

QM-only work.⁴⁵ Model **M2** contains several active-site residues and groups that we found to be important in our charge deletion analysis (see Section 3.1), in particular the side chain of Arg606 and the dihydropyran moieties connected to the dithiolene ligands, but it lacks the negatively charged diphosphate groups that gave the largest electrostatic interaction energies in our analysis. In the QM-only study it has apparently been assumed implicitly that these diphosphate groups are shielded sufficiently by nearby positively charged residues, like His676 and Arg720. The QM/MM results for the optimized stationary points and the associated relative energies are shown in Figure 4 and Table 3, respectively.

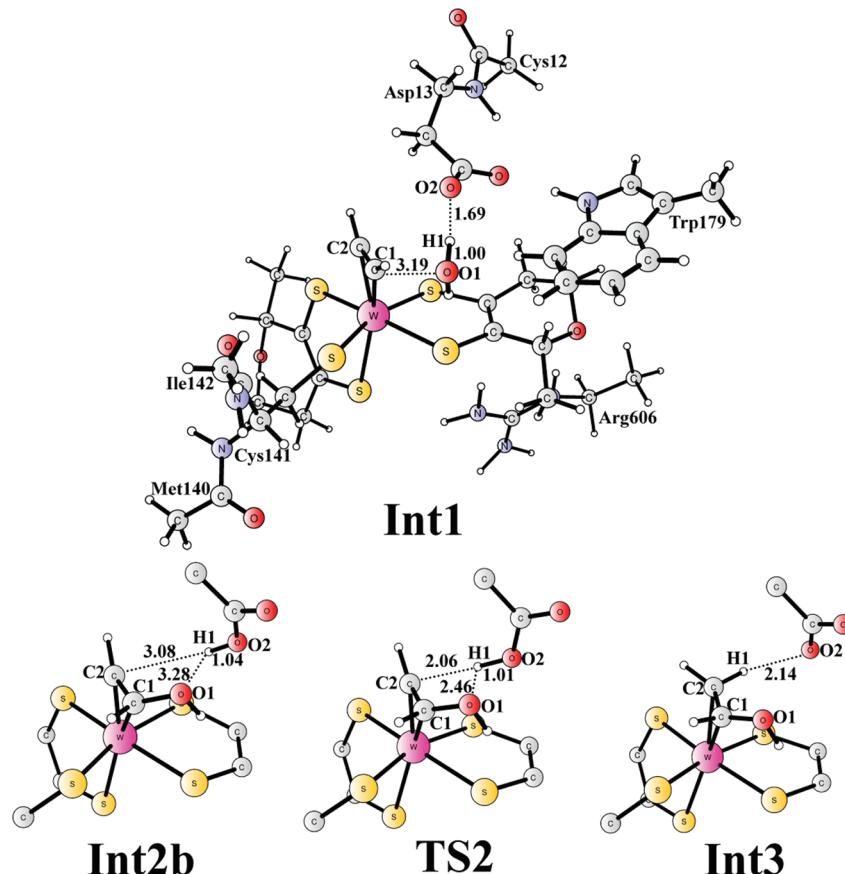


Figure 4. Optimized structures of intermediates and transition states for QM region M2 (BS100, RI-BP86/MM). All bond lengths are given in Å. For clarity, the full QM region is shown only for Int1. For the other stationary points, only the core region is shown.

According to the QM/MM calculations with QM region M2, the reaction takes place in one single concerted step. Figure 5

Table 3. QM/MM Relative Energies (in kcal/mol) Calculated for Snapshot BS100 with QM Region M2 using Three Different Functionals

	BP86/BS1	TPSS/BS1	B3LYP/BS1
Int1	0	0	0
Int2b	36.7	37.8	41.0
TS2	39.3	39.9	44.6
Int3	14.6	14.8	19.1

shows the energy profile of a linear transit scan (RI-BP86/def2-SVP) starting from **Int1** and proceeding along a reaction coordinate defined as the difference of the distances C1–O1 and O1–H1 (see Figures S2–S7, Supporting Information for linear transit scans with **M1**). The approach of the water molecule to acetylene leads to a steady increase of the energy in the scan, apparently without the formation of a stable vinyl anion intermediate (as found with QM region **M1**). The energy relative to **Int1** reaches a value of about 30 kcal/mol (RI-BP86/def2-SVP) when the C1–O1 distance is decreased to 1.45 Å. A high-energy intermediate **Int2b** was located by rotation of the carboxylic acid part of Asp13, thus breaking its hydrogen bond to the attacking water oxygen. **Int2b** lies 41.0 kcal/mol above **Int1** and easily rearranges to the vinyl alcohol intermediate **Int3**, since the required proton transfer has to overcome a barrier of only 3.6 kcal/mol. The overall reaction is endothermic by 19.1 kcal/mol (B3LYP/BS1, see Table 3).

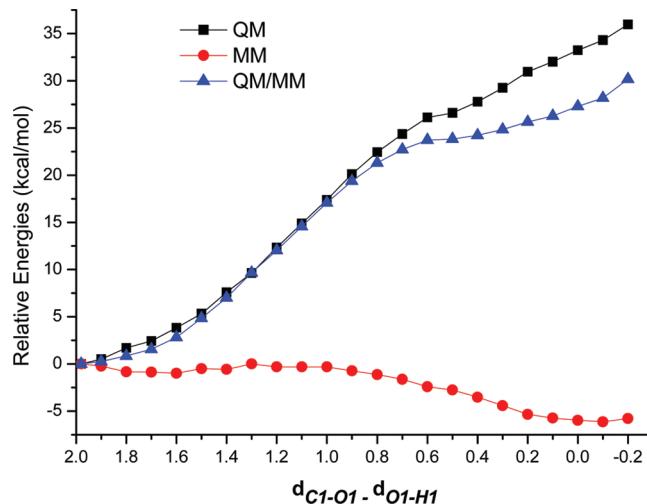


Figure 5. Potential energy profile for the nucleophilic attack using QM region M2 (BS100). Energies (RI-BP86/def2-SVP) are given in kcal/mol relative to **Int1**.

Quite surprisingly at first sight, the QM-only and QM/MM results with the same QM region **M2** thus differ significantly, with a much higher barrier and a large endothermicity in the QM/MM case. One possible reason for this discrepancy might be the influence of the diphosphate groups, which were neglected in the QM-only study and are treated at the MM level in the QM/MM calculations with QM region **M2**. Using the same charge deletion analysis as before, the diphosphate

groups are found to destabilize **Int2b** relative to **Int1** by 19.0 kcal/mol in QM/MM calculations with QM region **M2** (see Table S9, Supporting Information), in analogy to QM region **M1** (20.3 kcal/mol, see above). Such large electrostatic effects call for the use of a larger QM region that includes the diphosphate groups (**M3**, see Section 3.3).

3.3. QM Region M3a: One-Water Attack. Our largest QM region **M3** was designed on the basis of the residue interaction analysis (see Section 3.1). Compared to model **M2**, we added in **M3a** the diphosphate groups together with the neighboring positively charged residue Arg720, the pterin groups, the second-shell His676 residue, and two water molecules. In the peptide backbone, the Met140-Cys141-Ile142 residues were moved from the QM to the MM region because of their small influence on the reaction energetics and for the sake of computational efficiency. The total charge of the QM region was -3 (157 atoms, Figure 6).

In accordance with model **M1**, the reaction proceeds in two steps. However, the initial nucleophilic attack becomes rate limiting, with a barrier of 16.7 kcal/mol at the QM(B3LYP/

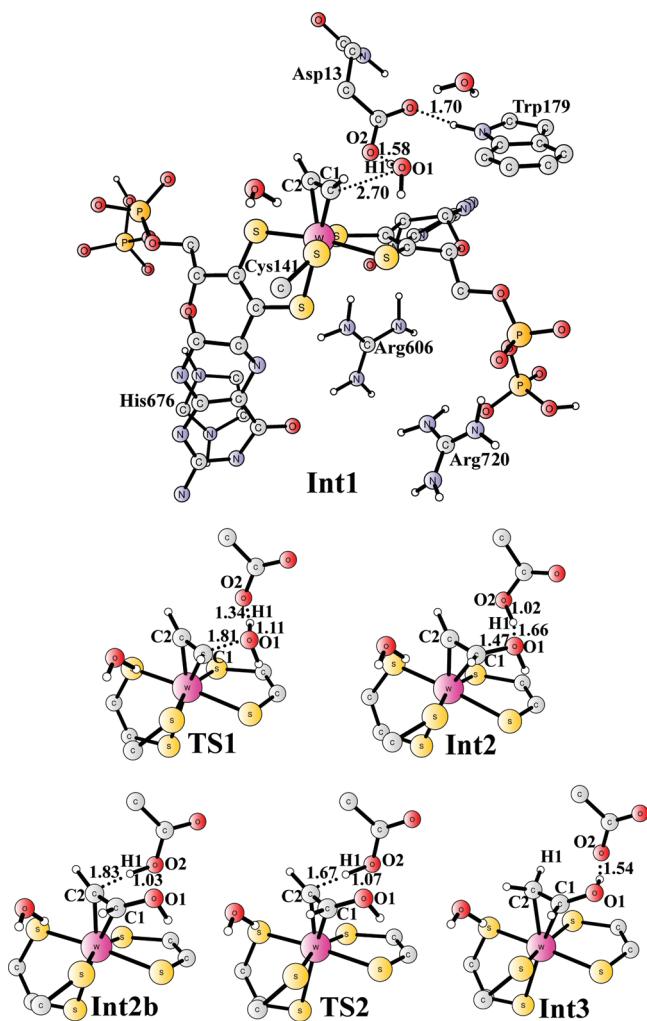


Figure 6. Optimized structures of intermediates and transition states for the one-water attack mechanism using QM region **M3a**, BS100-RI-BP86/MM. All bond lengths are given in Å. For clarity, the full QM region is shown only for **Int1** (with some unimportant hydrogen atoms omitted). For the other stationary points, only the core region is shown.

BS2)/MM level (Table 4). The vinyl anion intermediate **Int2** and the transition state **TS2** for the subsequent proton-transfer step are computed to lie 8.9 and 13.9 kcal/mol higher than **Int1**, respectively, and the whole reaction is exothermic by as

Table 4. QM/MM Relative Energies (in kcal/mol)
Calculated for Snapshot BS100 with QM Regions **M3a** and **M3b** using Three Different Functionals

	BP86/BS2 ^a	TPSS/BS2 ^a	B3LYP/BS2 ^a	B3LYP/BS2 ^b
M3a	Int1	0	0	0
	TS1	12.0	12.2	16.7
	Int2	6.8	6.4	8.9
	Int2b	7.2	7.0	10.7
	TS2	9.5	9.5	13.9
	Int3	-17.1	-17.7	-15.2
M3b	Int1	-2.9	-3.7	-3.3
	TS1	10.9	10.7	16.7
	Int2	12.6	13.3	18.6
	TS2	20.9	22.2	30.1
	Int3	-2.4	-3.0	-0.7
				-10.4

^aElectronic embedding (charge shift). ^bMechanical embedding.

much as 15.2 kcal/mol [QM(B3LYP/BS2)/MM]. Analogous two-step energy profiles are obtained when using different density functionals as the QM component (Table 4): BP86 and TPSS yield a slightly lower barrier for the initial rate-limiting step (12.0 and 12.2 kcal/mol) and a slightly higher exothermicity (17.1 and 17.7 kcal/mol, respectively).

Compared with the previous QM-only study,⁴⁵ the QM/MM calculations with QM region **M3a** provide a qualitatively similar reaction mechanism, but there are also significant variations. A conformational change is again needed to facilitate the proton transfer (as in the case of **M1**, see Section 3.1), and more importantly, the rate-limiting step is different (first step here vs second step in the QM-only work). In addition, the overall QM/MM barrier with QM region **M3a** is about 6 kcal/mol lower compared with the one from the QM-only model **M2**.⁴⁵

To further analyze the difference between the QM-only and QM/MM results, we performed QM/MM single-point calculations using mechanical embedding (Table 4) as well as QM-only single-point calculations for QM region **M3a** in the gas phase and in solution using the SMD continuum solvation model with a dielectric constant of 4, in analogy to the previous cluster model study⁴⁵ (Table 5). Obviously, electronic and mechanical embedding gives similar QM/MM energetics for **M3a**, with a mean absolute deviation of 2.6 kcal/mol between the two sets of data (Table 4); this close agreement implies that polarization effects by the MM environment must be quite small for the large QM region **M3a**. By contrast, upon complete neglect of the MM environment and its electrostatic influence, gas-phase single-point QM-only calculations at the B3LYP/BS2 level provide a different profile (Table 5). They indicate that **TS2** is rate limiting, with an overall barrier of 25.3 kcal/mol, which is close to the gas-phase barrier from the QM-only model **M2** (23.2 kcal/mol).⁴⁵ The reaction energy of -14.0 kcal/mol (B3LYP/BS2) is also compatible with the previous QM-only data for **M2** (-9.2 kcal/mol).⁴⁵ Continuum solvation does not affect the computed gas-phase relative energies for QM region **M3a** very much, with differences of typically less than 2 kcal/mol (Table 5). The results with QM = B3LYP/BS2 in Tables 4 and 5 thus imply that the MM environment of the large QM region **M3a** lowers the barrier for the second step (**TS2**) by

Table 5. QM-Only Relative Energies (in kcal/mol) Calculated for Snapshot BS100 with QM Regions M3a and M3b at Various Levels^a

	BP86/BS2		B3LYP/BS2		
	SMD ($\epsilon = 4$)		gas phase	SMD ($\epsilon = 4$)	QM/MM
	gas phase				
M3a	Int1	0	0	0	0
	TS1	14.9	14.4	19.7	19.3
	Int2	17.7	15.5	20.1	17.8
	Int2b	19.5	17.8	23.5	21.6
	TS2	19.0	18.8	25.3	23.6
	Int3	-16.2	-16.3	-14.0	-14.5
M3b	Int1	-6.3	-3.8	-4.8	-3.9
	TS1	4.4	10.9	12.1	17.1
	Int2	7.4	13.5	14.5	19.9
	TS2	13.9	20.0	24.6	29.5
	Int3	-6.7	-4.5	-4.0	-2.6

^aQM(B3LYP/BS2)/MM relative energies with electronic embedding from Table 4 are included for convenience.

around 10 kcal/mol, beyond what is obtained with continuum solvation, while the barrier for the first step (TS1) is lowered less (by around 3 kcal/mol). As a result, the rate-limiting step is switched in the QM/MM calculations with QM region **M3a**, compared with the previous QM-only model **M2** and the current single-point QM-only calculations for **M3a**. Apparently, even larger QM-only cluster models (beyond **M3a**) would be needed to get results consistent with the current best QM/MM data.

In our experience, it is unusual that the MM environment affects QM/MM energy profiles significantly when using a QM region as large as **M3a** (157 atoms). This is probably at least partly due to the fact that the investigated reactions involve proton transfers and substantial charge redistribution (Scheme 1), which will also give rise to some conformational reorganization around the active site upon geometry optimization. The relative energies of the relevant intermediates and transition states may thus be influenced by long-range electrostatic interactions between the QM region (total charge of -3) and the MM environment as well as by differences in van der Waals interactions because of geometric reorganization. Focusing on the relative energies of TS1 and TS2 in the QM/MM calculations with QM region **M3a** (see above), a detailed energy partitioning shows that the QM/MM electrostatic interactions lower TS2 more than TS1 (by 5.7 vs 0.6 kcal/mol), and the same applies to the other nonbonded QM/MM interactions (5.7 vs 2.4 kcal/mol). It is difficult, however, to pinpoint the origin of these shifts. For example, the electrostatic contributions from the most important remaining MM residues (Cys12, Asp699, Fe4S4, and Lys172) are relatively small and compensate each other to a large extent (see Table S12, Supporting Information). It thus seems reasonable to expect that QM/MM calculations with QM region **M3a** are able to describe the reaction mechanism of acetylene hydratase in a realistic manner.

3.4. QM Region M3b: Two-Water Attack. In this section, we consider an alternative mechanistic scenario, in which another active-site water molecule (located in **M3a** to the left of the acetylene moiety in Figure 6) acts as a bridge for the proton transfer in both steps (in an arrangement labeled as **M3b**, Figure 7). The intermediates and transition states from corresponding QM(RI-BP86)/MM calculations are shown in

Figure 7, and their QM/MM energies are listed in Table 4 relative to **Int1-M3a**. The starting intermediate **Int1-M3b** is

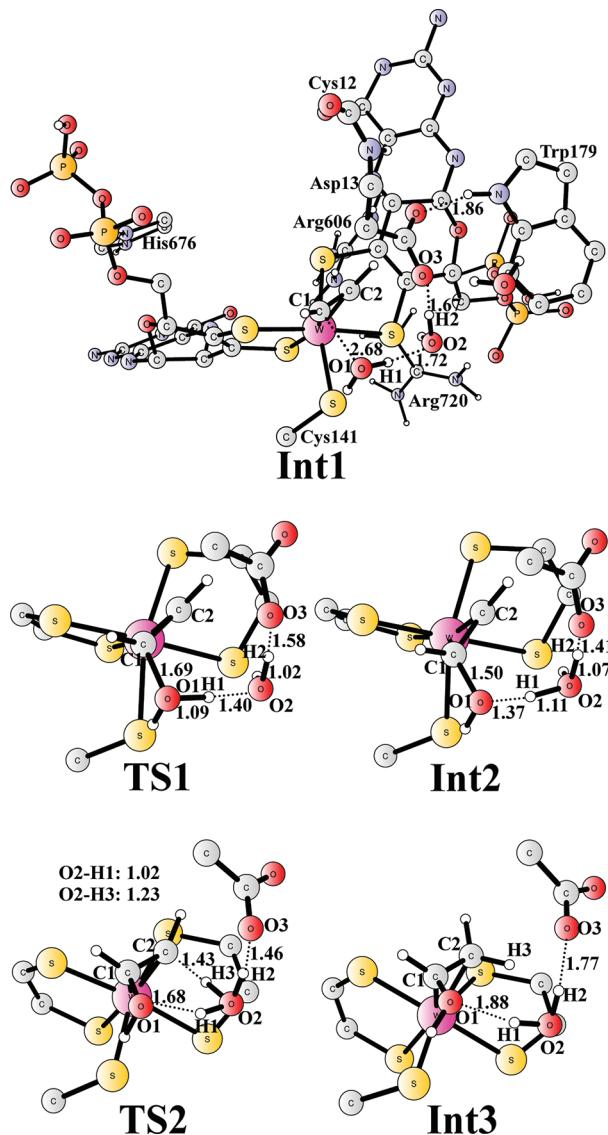


Figure 7. Optimized structures of intermediates and transition states for the two-water attack mechanism using QM region **M3b** (BS100, RI-BP86/MM). All bond lengths are given in Å. For clarity, the full QM region is shown only for **Int1** (with some unimportant hydrogen atoms omitted). For the other stationary points, only the core region is shown.

computed to be slightly more stable than **Int1-M3a**, by -3.3 kcal/mol at the single-point QM(B3LYP/BS2)/MM level with electronic embedding, suggesting that both species are thermodynamically accessible during the formation of the tungsten-acetylene adduct. The reaction occurs in two steps, with the second proton-transfer step being rate determining. Strikingly, the overall barrier is calculated to be 30.1 kcal/mol relative to **Int1-M3a** at this QM/MM level (Table 4) and thus considerably higher than that of the one-water attack mechanism. A comparison of the QM/MM results from electronic and mechanical embedding shows (Table 4) that the polarization of the QM region by the MM point charges has a significant effect on the energy of **Int1-M3b** relative to **Int1-M3a**, causing a shift of about 8 kcal/mol. A similar shift is found

for the other stationary points in the **M3b** system, so that the QM/MM energies relative to **Int1–M3b** are actually fairly similar when using electronic and mechanical embedding. The gas-phase QM-only relative energies for QM region **M3b** (Table 5) yield a qualitatively similar profile, with the second step (**TS2**) remaining rate limiting by a wide margin. Inclusion of continuum solvation raises the relative energies of **TS1**, **Int2**, and **TS2** by about 5–6 kcal/mol and those of **Int1** and **Int3** by 1–2 kcal/mol (Table 5). This results in a B3LYP/BS2-SMD energy profile that is very close to the one obtained from QM(B3LYP/BS2)/MM with electronic embedding (mean absolute deviation of 1.0 kcal/mol, see Table 5).

Comparing the rate-limiting barriers for the one- and two-water attack mechanisms (**M3a** vs **M3b**), the former is favored by 13.4 (7.8) kcal/mol in the QM/MM calculations with electronic (mechanical) embedding (Table 4). In gas-phase QM-only calculations, these barriers are almost equal (25.3 vs 24.6 kcal/mol for **M3a** vs **M3b**, Table 5). However, by lowering the rate-limiting transition state (**TS2**) for the former (**M3a**) by 2.3 kcal/mol and raising it for the latter (**M3b**) by 4.9 kcal/mol, continuum solvation again favors the one-water over the two-water attack (by 5.9 kcal/mol, Table 5).

We checked whether these two mechanisms can already be distinguished in QM/MM calculations with a minimum QM region. For this purpose, we extended QM region **M1** by including one active-site water molecule at a position analogous to **M3a** or **M3b**, thus generating QM regions **M1a** and **M1b**, respectively (35 atoms, total charge of −2, see Figures S9 and S10, Supporting Information). With this setup, the one-water attack (**M1a**) is favored by about 10 kcal/mol (**TS2**) over the two-water attack (**M1b**) for each of the three functionals applied (see Table S17, Supporting Information). The QM/MM calculations with a minimum QM region (35 atoms) thus predict the same mechanistic preference as those with the large QM region (157 atoms), even though they overestimate the rate-limiting barrier substantially.

3.5. Single-Point Calculations with QM Regions of Different Size. In a previous convergence study of QM-only and QM/MM energies, geometries were optimized using a small QM region, and single-point calculations were performed using larger QM regions.⁴² The basic assumption is that the QM/MM geometries obtained with a small QM region are reliable and close to those obtained with a large QM region. A similar assumption is commonly made in QM studies on small systems, where a medium basis set is used for geometry optimization, and a larger basis set is used for the final energy evaluation. In the QM/MM case, there may be significant changes in the optimized geometries when enlarging the QM region, which could cause adverse fluctuations of relative energies from single-point calculations at QM/MM geometries optimized with a given QM region, and it may then be difficult to converge the corresponding QM-only or QM/MM energies with regard to the size of the QM region. This calls for benchmarks on representative systems to validate this approach.

To check whether the QM/MM geometries obtained with a small QM region are presently good enough for QM/MM single-point calculations with larger QM regions, we used the optimized QM/MM geometries for **M1** and **M2** to run QM/MM single-point calculations with larger QM regions. The overall barriers and reaction energies are shown in Table 6. There are obviously quite large differences between the various entries. For example, when QM/MM geometries from the

minimal model **M1** are used for single-point calculations with the largest model **M3a**, the barrier is 7.1 kcal/mol higher than

Table 6. QM/MM Single-Point Relative Energies (in kcal/mol) Calculated at the B3LYP Level for Snapshot BS100 with Different QM Regions at Different QM/MM Optimized Geometries (One-Water Attack Mechanism)

QM region used in single-point calculations	geometries from QM region			reference energy ^c
	M1 (32 atoms)	M2 (116 atoms)	—	
M2 (116 atoms) ^a	overall barrier	36.5	—	44.6
	reaction energy	2.7	—	19.1
M3a (157 atoms) ^b	overall barrier	23.8	29.7	16.7
	reaction energy	−2.3	9.9	−15.2

^aB3LYP/BS1. ^bB3LYP/BS2. ^cUsing models **M2** and **M3a**, respectively.

that obtained from the fully optimized **M3a** geometries (23.8 vs 16.7 kcal/mol), and the reaction exothermicity is underestimated by 12.9 kcal/mol (−2.3 vs −15.2 kcal/mol).

We conclude that it is dangerous to use QM/MM geometries obtained with small QM regions in QM/MM single-point calculations with larger QM regions. Geometry optimizations seem to be required to determine consistent relative energies.

4. DISCUSSION AND CONCLUSIONS

With acetylene hydratase as a representative example, QM/MM calculations have been carried out for the purpose of comparing QM-only and QM/MM models in the study of enzymatic reactions. We investigated three different QM regions (**M1–M3**), ranging from 32 to 157 atoms and considered two mechanisms, involving either one water molecule (as in the previous QM-only study⁴⁵) or two water molecules in the nucleophilic attack and proton transfer steps.

The QM/MM calculations with the largest QM region **M3a** give a qualitatively similar mechanistic scenario as the previous QM-only model,⁴⁵ with a realistic overall barrier. However, some variations can be observed. First, in all QM/MM models for the one-water attack mechanism (**M1**, **M2**, **M3a**), a rotation of the carboxylic group at the vinyl anion intermediate has to take place prior to the proton transfer. This was not captured in the previous QM-only model (**M2**), which apparently overestimated the flexibility of this group at the periphery (due to missing steric constraints from the protein environment). This problem can presumably be solved by adopting a more extended cluster model (larger than **M2**). Second, in the QM/MM calculations with the largest QM region **M3a**, the rate-limiting step is the initial nucleophilic attack, rather than the subsequent protonation of the vinyl anion intermediate (as in the QM-only model **M2**). The calculated overall barrier is also several kcal/mol lower at the QM/MM level, but no experimental kinetic data are available for comparison. The change in the rate-limiting step is caused by the influence of the MM environment, which lowers the barrier of the second step more than that of the first one in the QM/MM calculations with region **M3a**.

The choice of the QM region turns out to be crucial for the proper description of the reaction pathway. A direct

comparison of the QM and QM/MM energetics for all QM regions is provided in the Supporting Information (Table S16 and Figure S8). QM/MM calculations with a minimum QM region (**M1**) are helpful for a preliminary understanding of the mechanism, even though the calculated barriers are overestimated. An analysis (at this level) of the electrostatic contributions from each group or residue close to the active site can provide guidance for the design of larger QM regions. At present, we selected two intermediates with significant charge relocation (**Int1** and **Int2**) as reference points for the analysis and included all groups with large electrostatic effects in our preferred large QM region (**M3**). The latter contains the two diphosphate groups, which are found to influence the relative energies of the two intermediates most strongly (by about 20 kcal/mol) but are missing in the intermediate QM region (**M2**) used in the previous QM-only study.⁴⁵ A comparison of the QM/MM results for **M2** and **M3** shows that inclusion of the two diphosphate groups in the QM region (**M3**) is essential for getting realistic relative energies and barriers, whereas the MM description of these groups (**M2**) leads to an unreasonable energy profile with a very high overall barrier. In the light of these findings, it is surprising that the previous QM-only study⁴⁵ with model **M2** arrived at a plausible mechanism with a sensible energy profile, in spite of the fact that the negatively charged diphosphate groups and their strong electrostatic effects are neglected in model **M2**. One has to assume that the previous QM-only work has benefitted from some error cancellation and has not yet been converged with regard to the size of the cluster model. This could of course be checked by QM-only calculations with larger cluster models, which are however beyond the scope of our present work. A suitable such model might be QM region **M3a** plus the connecting Met140-Cys141-Ile142 residues that are treated at the MM level in the present QM/MM calculations (174 atoms).

We have also considered an alternative two-water attack mechanism in which another active-site water molecule acts as a bridge for the proton transfer in both steps. QM/MM calculations with the largest QM region **M3** show that this mechanism is less favored than the one-water attack mechanism. QM-only calculations on the large cluster model with continuum solvation included (**M3a** vs **M3b**) reproduce this mechanistic preference qualitatively as well as QM/MM calculations with a minimum QM region (**M1a** vs **M1b**), whereas QM-only gas-phase calculations on **M3a** and **M3b** give essentially the same overall barrier for both mechanisms. The protein environment thus appears to play a decisive role, by lowering (raising) the energy of **TS2** in the one-water (two-water) attack mechanism.

On the technical side, we find that it is not advisable to adopt QM/MM geometries optimized with a small QM region in QM/MM single-point calculations with larger QM regions, because the resulting relative energies may be inconsistent and not as reliable as those obtained from fully optimized QM/MM geometries, with the same QM region being used in QM/MM geometry optimizations and energy evaluations.

Finally, the present study emphasizes that both QM-only and QM/MM methods require properly designed models (QM regions) to provide reasonable mechanistic scenarios and realistic energy profiles for enzymatic reactions. Demonstrating convergence with regard to the size of QM region is important in both cases. The two methods are to some extent complementary to each other, and their combined application

can thus be a powerful approach for the study of enzymatic reactions.

ASSOCIATED CONTENT

S Supporting Information

Setup of the systems; CHARMM parameters; active regions for QM/MM optimizations; QM, MM, and QM/MM energies for all stationary points; energy profiles along reaction coordinates; electrostatic interaction energies from charge deletion analysis; and detailed QM/MM results for **M1a** and **M1b**. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: thiel@kofo.mpg.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Max Planck Society and the Volkswagenstiftung. R.Z.L. thanks Iakov Polyak for his help in the setup of the QM/MM calculations and Prof. Adrian Mulholland for insightful suggestions. This article is dedicated to Professor Wilfred van Gunsteren on the occasion of his 65th birthday.

REFERENCES

- (1) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421–437.
- (2) Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41*, 689–698.
- (3) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949–1017.
- (4) Acevedo, O.; Jorgensen, W. L. *Acc. Chem. Res.* **2010**, *43*, 142–151.
- (5) Lonsdale, R.; Ranaghan, K. E.; Mulholland, A. J. *J. Chem. Commun.* **2010**, *46*, 2354–2372.
- (6) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729–738.
- (7) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232–240.
- (8) Himo, F.; Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2009**, *14*, 643–651.
- (9) Siegbahn, P. E. M.; Himo, F. *WIREs Comput. Mol. Sci.* **2011**, *1*, 323–336.
- (10) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.
- (11) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (12) Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, *268*, 173–290.
- (13) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (14) Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.
- (15) Sevastik, R.; Himo, F. *Bioorg. Chem.* **2007**, *35*, 444–457.
- (16) Hopmann, K. H.; Himo, F. *J. Chem. Theory Comput.* **2008**, *4*, 1129–1137.
- (17) Georgieva, P.; Himo, F. *J. Comput. Chem.* **2010**, *31*, 1707–1714.
- (18) Liao, R.-Z.; Yu, J.-G.; Himo, F. *J. Chem. Theory Comput.* **2011**, *7*, 1494–1501.
- (19) Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2006**, *11*, 695–701.
- (20) Siegbahn, P. E. M.; Borowski, T. *Faraday Discuss.* **2011**, *148*, 109–117.
- (21) Inoue, T.; Shiota, Y.; Yoshizawa, K. *J. Am. Chem. Soc.* **2008**, *130*, 16890–16897.
- (22) Schöneboom, J. C.; Lin, H.; Reuter, N.; Thiel, W.; Cohen, S.; Ogliaro, F.; Shaik, S. *J. Am. Chem. Soc.* **2002**, *124*, 8142–8151.
- (23) Cummins, P. L.; Gready, J. E. *J. Comput. Chem.* **2005**, *26*, 561–568.

- (24) Altun, A.; Shaik, S.; Thiel, W. *J. Comput. Chem.* **2006**, *27*, 1324–1337.
- (25) Stanton, C. L.; Kuo, I.-F. W.; Mundy, C. J.; Laino, T.; Houk, K. N. *J. Phys. Chem. B* **2007**, *111*, 12573–12581.
- (26) López-Canut, V.; Martí, S.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Phys. Chem. B* **2009**, *113*, 7816–7824.
- (27) van der Kamp, M. W.; Żurek, J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. *J. Phys. Chem. B* **2010**, *114*, 11303–11314.
- (28) Kumar, D.; Thiel, W.; de Visser, S. P. *J. Am. Chem. Soc.* **2011**, *133*, 3869–3882.
- (29) Tian, B.; Strid, Å.; Eriksson, L. A. *J. Phys. Chem. B* **2011**, *115*, 1918–1926.
- (30) Waller, M. P.; Bühl, M.; Geethalakshmi, K. R.; Wang, D.; Thiel, W. *Chem.—Eur. J.* **2007**, *13*, 4723–4732.
- (31) Waller, M. P.; Geethalakshmi, K. R.; Bühl, M. *J. Phys. Chem. B* **2008**, *112*, 5813–5823.
- (32) Johnson, E. R.; DiLabio, G. A. *J. Mol. Struct.: THEOCHEM* **2009**, *898*, 56–61.
- (33) Geethalakshmi, K. R.; Waller, M. P.; Thiel, W.; Bühl, M. *J. Phys. Chem. B* **2009**, *113*, 4456–4465.
- (34) Rinkevicius, Z.; Murugan, N. A.; Kongsted, J.; Frecuș, B.; Steindal, A. H.; Ågren, H. *J. Chem. Theory Comput.* **2011**, *7*, 3261–3271.
- (35) Baer, M.; Mathias, G.; Kuo, I.-F. W.; Tobias, D. J.; Mundy, C. J.; Marx, D. *ChemPhysChem* **2008**, *9*, 2703–2707.
- (36) Phatak, P.; Ghosh, N.; Yu, H.; Cui, Q.; Elstner, M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 19672–19677.
- (37) Altun, A.; Yokoyama, S.; Morokuma, K. *J. Phys. Chem. B* **2008**, *112*, 16883–16890.
- (38) Sanchez-Garcia, E.; Doerr, M.; Thiel, W. *J. Comput. Chem.* **2010**, *31*, 1603–1612.
- (39) Hirao, H.; Morokuma, K. *J. Phys. Chem. Lett.* **2010**, *1*, 901–906.
- (40) Sumowski, C. V.; Ochsenfeld, C. *J. Phys. Chem. A* **2009**, *113*, 11734–11741.
- (41) Sumowski, C. V.; Schmitt, B. B. T.; Schweizer, S.; Ochsenfeld, C. *Angew. Chem., Int. Ed.* **2010**, *49*, 9951–9955.
- (42) Hu, L. H.; Eliasson, J.; Heimdal, J.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 11793–11800.
- (43) Hu, L. H.; Söderhjelm, P.; Ryde, U. *J. Chem. Theory Comput.* **2011**, *7*, 761–777.
- (44) Seiffert, G. B.; Ullmann, G. M.; Messerschmidt, A.; Schink, B.; Kroneck, P. M. H.; Einsle, O. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 3073–3077.
- (45) Liao, R. Z.; Yu, J. G.; Himo, F. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22523–22527.
- (46) Liao, R. Z.; Himo, F. *ACS Catal.* **2011**, *1*, 937–944.
- (47) Liu, Y.-F.; Liao, R.-Z.; Ding, W.-J.; Yu, J.-G.; Liu, R.-Z. *J. Biol. Inorg. Chem.* **2011**, *16*, 745–752.
- (48) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714–2723.
- (49) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525–537.
- (50) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (51) Metz, S.; Wang, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, *131*, 4628–4640.
- (52) Metz, S.; Thiel, W. *J. Am. Chem. Soc.* **2009**, *131*, 14885–14902.
- (53) Metz, S.; Thiel, W. *J. Phys. Chem. B* **2010**, *114*, 1506–1517.
- (54) Rogers, D. W.; Dagdagan, O. A.; Allinger, N. L. *J. Am. Chem. Soc.* **1979**, *101*, 671–676.
- (55) Rosner, B. M.; Schink, B. *J. Bacteriol.* **1995**, *177*, 5767–5772.
- (56) MacKerell, A. D. Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L. Jr.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-MacCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E. III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (57) Brooks, C. L.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312–6325.
- (58) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1–28.
- (59) *Chemshell, a Computational Chemistry Shell*; Science & Technology Facilities Council: Swindon, U.K.; www.chemshell.org. (accessed February 29, 2012).
- (60) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (61) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136–141.
- (62) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.
- (63) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (64) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (65) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.
- (66) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.
- (67) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
- (68) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- (69) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (70) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (71) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*, Revision B.01; Gaussian, Inc.: Wallingford, CT, 2009.
- (72) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (73) Roy, L. E.; Hay, P. J.; Martin, R. L. *J. Chem. Theory Comput.* **2008**, *4*, 1029–1031.
- (74) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.